



HAL
open science

Estimation paramétrique d'une diffusion ergodique observée à temps discret

Sandie Souchet

► **To cite this version:**

Sandie Souchet. Estimation paramétrique d'une diffusion ergodique observée à temps discret. Mathématiques [math]. Université Panthéon-Sorbonne - Paris I, 1999. Français. NNT : . tel-00276933

HAL Id: tel-00276933

<https://theses.hal.science/tel-00276933v1>

Submitted on 2 May 2008

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.

UNIVERSITÉ PARIS I - PANTHÉON-SORBONNE

THÈSE DE DOCTORAT

Spécialité :
Mathématiques Appliquées

Présentée par

Sandie SOUCHET

Pour obtenir le grade de

Docteur en Sciences

(Arrêté du 30 mars 1992)

**Titre : Estimation des paramètres d'une diffusion ergodique
observée à temps discret.**

Sous la direction de

Xavier GUYON

Soutenue publiquement le mercredi 13 janvier 1999 devant le jury composé de :

Mesdames	Laure ELIE	
	Danielle FLORENS-LANDAIS	Rapporteur
	Valentine GENON-CATALOT	
	Catherine LAREDO	Rapporteur
Messieurs	Peter BROCKWELL	Rapporteur
	Christian GOURIEROUX	
	Xavier GUYON	

Contents

Introduction	3
A Estimation paramétrique d'une diffusion ergodique discrétisée	3
A.1 Estimation par fonction d'estimation	3
A.2 Estimation par minimum de contraste	4
A.2.1 Contraste basé sur des schémas de discrétisation anticipatifs	5
A.2.2 Contraste associé à un schéma de discrétisation adapté à l'ordre p	6
A.3 Estimation par minimum de contraste pour un pas de discrétisation tendant vers 0 et une dérive irrégulière	6
A.4 Estimation d'un CAR(p) incomplètement observé à partir des équations de Yule-Walker	8
A.5 Etudes expérimentales	10
B Estimation de la précision associée à un échantillonnage systématique et méthodes transitives	11
Bibliographie	14
Estimation paramétrique d'une diffusion ergodique discrétisée	18
1 Schémas de discrétisation anticipatifs et estimation du paramètre de dérive d'une diffusion	18
1.1 Introduction	18
1.2 Présentation des schémas d'approximation : trapèze, Simpson	20
1.3 Le cas d'une dérive linéaire en θ :	
Schéma du trapèze et Estimation par variable instrumentale	22
1.3.1 La méthode d'estimation par variables instrumentales	23
1.3.2 Résultats asymptotiques	24
1.4 Le cas d'une dérive générale :	
Schéma du trapèze et Estimation des Moments généralisés	27
1.4.1 Propriétés asymptotiques	28
1.4.2 Adaptation de la méthode au schéma de Simpson	36
1.5 Etude expérimentale	37
1.5.1 Processus de Ornstein-Uhlenbeck	38
1.5.2 Processus de Cox-Ingersoll-Ross	39
1.5.3 Cas d'une dérive non linéaire en x : diffusion Log-Normale	43

Bibliographie	45
Annexes	47
A Démonstration de la propriété 2 : schéma d'approximation de Simpson	47
B Démonstration du lemme 4	51
C Démonstration du lemme 5	53
D Démonstration du lemme 7	57
E Démonstration du lemme 8 : développement de $\Gamma_\delta(\theta_\delta)$	58
F Schéma d'approximation d'Euler	62
2 Schéma d'approximation adapté à l'ordre p et estimation du paramètre de dérive d'une diffusion.	64
2.1 Introduction	64
2.2 Schéma d'approximation à l'ordre p pour une fonctionnelle de la solution	65
2.3 Résultats asymptotiques	67
2.3.1 Construction du contraste	69
2.3.2 Calcul de l'écart entre $\theta_{p,\delta}$ et θ_0	72
2.3.3 Convergence de $\widehat{\theta}_{p,n}^\delta$ vers $\theta_{p,\delta}$	73
2.3.4 Normalité asymptotique de $\widehat{\theta}_{p,n}^\delta$	73
2.4 Etude expérimentale	76
2.4.1 Processus de Ornstein-Uhlenbeck	76
2.4.2 Processus de Cox-Ingersoll-Ross	79
Bibliographie	81
Annexes	83
A Démonstration du lemme 5	83
B Démonstration du lemme 7 :	87
C Démonstration du lemme 8 : développement de $K^p(\theta_{p,\delta})$ en fonction de δ	88
C.1 Développement de $T^p(\theta_{p,\delta})$	88
C.2 Majoration de $S^p(\theta_{p,\delta})$	90
3 Estimation du paramètre de dérive d'une diffusion sous des conditions d'irrégularité de la dérive.	92
3.1 Introduction	92
3.2 Résultat d'approximation d'Euler pour les modèles à seuil	93
3.3 Hypothèses et résultats préliminaires	95
3.4 Etude de la consistance	97
3.5 Vitesse de convergence et normalité asymptotique	101
3.6 Applications : Modèle CTAR(1)	109
3.7 Estimation du paramètre de seuil pour les chaînes de Markov	112
Bibliographie	116
Annexes	118
A Démonstration du lemme 2	118
B Démonstration de la propriété 1	120
C Démonstration de la propriété 2	121

D	Démonstration du lemme 8	122
E	Démonstration du lemme 9	124
F	Démonstration du lemme 10	126
G	Démonstration du théorème 4	127
H	Démonstration de la propriété 3	129
I	Schéma de la démonstration du théorème 7	130
I.1	Préliminaires	130
I.2	Consistance	131
I.3	Normalité asymptotique	131
4	Estimation d'un CAR(p) incomplètement observé à partir des équations de YULE-WALKER	134
4.1	Introduction :	134
4.2	Rappels et résultats préliminaires	136
4.3	Estimation du paramètre à partir des équations de Yule-Walker	138
4.4	Estimation des paramètres d'un CAR(2) à partir du schéma du trapèze.	142
4.5	Estimation du coefficient de diffusion σ	145
4.6	Etude Expérimentale	146
	Bibliographie	150
	Annexes	151
A	Démonstration du lemme 2	151
B	Démonstration du théorème 1	153
C	Démonstration du lemme 8	157
D	Démonstration du lemme 9	159
E	Démonstration du lemme 3 :	161
F	Démonstration du lemme 4	163
G	Démonstration du lemme 7	167
H	Démonstration du théorème 3	169
I	Démonstration du théorème 4 :	172
	Estimation de la précision associée à un échantillonnage systématique et méthodes transitives	175
5	Precision of systematic sampling and transitive methods	175
5.1	Introduction	175
5.2	Unbiased estimator of integral characteristics	177
5.3	Mean squared error expansion	180
5.4	Discrete approximations of the extension term	183
5.5	Extension term estimation	185
5.6	Robust extension term estimation	186
5.7	Discussion	188

Introduction

A Estimation paramétrique d'une diffusion ergodique discrétisée

On considère l'équation différentielle stochastique suivante :

$$dY_t = f(\theta_0, Y_t) dt + \sigma(Y_t) dW_t \quad (1)$$

où $(W_t)_{t \geq 0}$ est un mouvement Brownien standard issu de 0 et $\theta_0 \in \overset{\circ}{\Theta}$, Θ étant un compact de \mathbb{R}^d . La fonction de dérive f et le coefficient de diffusion σ sont supposés connus et vérifient les conditions usuelles qui assurent, pour toute condition initiale, l'existence d'une unique solution forte *stationnaire et ergodique* $Y = (Y_t)_{t \in [0, T]}$. Le paramètre de dérive θ_0 est inconnu et l'on désire l'estimer.

On note par ailleurs P_{θ_0} la loi du processus sous la loi stationnaire μ_{θ_0} et E_{θ_0} l'espérance sous P_{θ_0} .

Dans le cas où l'on dispose d'une observation continue de la trajectoire de Y sur $[0, T]$, les principaux résultats ont été donnés par Lipster et Shiryaev [34] et Kutoyants [33]. Considérant la log-vraisemblance en temps continu,

$$\int_0^T \frac{f(\theta, Y_t)}{\sigma^2(Y_t)} dY_t - \frac{1}{2} \int_0^T \frac{f^2(\theta, Y_t)}{\sigma^2(Y_t)} dt \quad (2)$$

ils montrent que, lorsque $T \rightarrow +\infty$, l'estimateur du maximum de vraisemblance est consistant et asymptotiquement normal et efficace avec une vitesse en \sqrt{T} .

Cependant, il paraît peu réaliste de pouvoir disposer d'une observation continue du processus solution. En conséquence, on supposera dans tout ce travail que l'estimation du paramètre de dérive repose sur une observation discrète de $Y : (Y_{k\delta})_{k=0, n}$ avec $\delta > 0$ petit mais fixé (Chapitres 1, 2 et 4), où encore $(Y_{kh_n})_{k=0, n}$ avec $h_n \rightarrow 0$ lorsque $n \rightarrow +\infty$ (Chapitre 3).

Dans le cas d'une diffusion unidimensionnelle, Dacunha-Castelle et Florens-Zmirou [12] montrent que l'estimateur du maximum de vraisemblance associé à l'observation $(Y_{k\delta})_{k=0, n}$ est consistant et asymptotiquement normal et efficace avec la vitesse $\sqrt{n\delta}$, lorsque $n \rightarrow +\infty$. Toutefois, les densités de transition ne sont en général pas explicites et la méthode d'estimation est difficile à mettre en oeuvre.

A.1 Estimation par fonction d'estimation

Dans le cas où le coefficient de diffusion dépend également de θ_0 , des auteurs tels que Kessler, Sorensen et Bibby proposent des procédures d'estimation basées sur les *fonctions d'estimation* $G_n(\theta) = \sum_{k=1}^n g(\theta, Y_{(k-1)\delta}, Y_{k\delta})$ et l'estimateur associé $\hat{\theta}_n$ est alors la solution de l'équation $G_n(\theta) = 0$.

Dans [6], Bibby et Sorensen considèrent des fonctions d'estimation martingales de la forme $G_n(\theta) = \sum_{k=1}^n g(\theta, Y_{(k-1)\delta}) [Y_{k\delta} - m(\theta, Y_{(k-1)\delta})]$, où $m(\theta, x)$ est l'espérance conditionnelle de Y_δ sachant $Y_0 = x$ sous θ . Différents critères d'optimalité développés par Godambe et Heyde [17] conditionnent le choix de g . Bibby [5] et Kessler [28] étudient des fonctions d'estimation martingales plus générales mais qui font également intervenir des quantités conditionnelles en général non explicites et qui doivent en conséquence être estimées. Dans [31], Kessler et Sorensen proposent des fonctions d'estimation martingales basées sur les fonctions propres du générateur de la diffusion. Dans ces différents cas, $\hat{\theta}_n$ est consistant et asymptotiquement normal avec une vitesse de l'ordre de \sqrt{n} .

Dans [30], Kessler propose des fonctions d'estimation *non martingales* de la forme $F_n(\theta) = \sum_{k=1}^n h(\theta, Y_{(k-1)\delta})$, où $h(\theta_0, \cdot)$ est d'intégrale nulle par rapport à la mesure invariante de la diffusion μ_{θ_0} . Là encore, l'estimateur obtenu est consistant et asymptotiquement normal à la vitesse \sqrt{n} .

A.2 Estimation par minimum de contraste

La méthode d'estimation la plus courante demeure l'estimation par minimum de contraste (définition 3.2.7. de [11]). Les différentes approches visant à la construction de contrastes $U_n(\theta)$ reposent en général sur une approximation de la log-vraisemblance discrète exacte.

Dans [36] et [37], Pedersen propose une estimation consistante $p_N(x, \delta, \cdot; \theta)$ de la probabilité de transition à pas δ , $p(x, \delta, \cdot; \theta)$, lorsque $N \rightarrow +\infty$. Il construit ainsi une suite de contrastes $(U_{n,N}(\theta, \delta))_N$ qui converge à n fixé vers la log-vraisemblance discrète du modèle. L'estimateur associé est asymptotiquement efficace. Pour un pas d'observation δ_n tendant vers 0, Prakasa-Rao [39] et Yoshida [46] obtiennent, à partir de la discrétisation de la vraisemblance continue (2), des estimateurs de θ_0 consistants lorsque $n\delta_n \rightarrow +\infty$ et asymptotiquement efficaces lorsque $n\delta_n^3 \rightarrow 0$. Dans [29], Kessler approxime la vraisemblance exacte par celle d'un processus gaussien dont les moments sont des développements en δ_n à l'ordre p des moments conditionnels, basés sur les itérés du générateur infinitésimal de la diffusion. Si $n\delta_n \rightarrow +\infty$ et $n\delta_n^p \rightarrow 0$, l'estimation associée est alors optimale.

La construction de contrastes peut également résulter de l'utilisation de schémas d'approximation de la diffusion. Dans [16], Florens-Zmirou utilise le schéma d'approximation d'Euler qui couplé aux Moindres Carrés, donne le contraste :

$$U_n(\theta, \delta) = \frac{1}{n\delta} \sum_{k=1}^n \left(\frac{Y_{k\delta} - Y_{(k-1)\delta} - \delta f(\theta, Y_{(k-1)\delta})}{\sqrt{\delta} \sigma(Y_{(k-1)\delta})} \right)^2 \quad (3)$$

L'estimateur du minimum de contraste $\hat{\theta}_n = \inf_{\theta \in \Theta} U_n(\theta, \delta_n)$ est alors consistant, lorsque $n\delta_n \rightarrow +\infty$ et $\delta_n \rightarrow 0$, et asymptotiquement normal et efficace, si de plus $n\delta_n^3 \rightarrow 0$.

Le cas où le pas d'observation est fixe est aussi étudié. L'estimateur est également convergent et asymptotiquement normal à la vitesse $\sqrt{n\delta}$, non pas pour θ_0 mais pour θ_δ , où θ_δ est l'unique minimum de la fonction de contraste $U(\theta, \delta)$ associée à $U_n(\theta, \delta)$. De plus, θ_δ approche θ_0 à un biais explicite en δ et la variance asymptotique est l'inverse de l'information de Fisher à un facteur $(1 + o(1))$ près.

A.2.1 Contraste basé sur des schémas de discrétisation anticipatifs

Nous adoptons une démarche identique dans le Chapitre 1, **Schémas de discrétisation anticipatifs et estimation du paramètre de dérive d'une diffusion**[44]. La méthode d'estimation est basée sur l'utilisation de schémas d'approximation *anticipatifs* : le *schéma du trapèze* et le *schéma de Simpson*. Le schéma discret exact du trapèze s'écrit :

$$\forall t \geq 0, Y_{t+\delta} = Y_t + \frac{\delta}{2} (f(\theta_0, Y_t) + f(\theta_0, Y_{t+\delta})) + \eta_t + \tilde{W}_t \quad (4)$$

où η_t est un $O_{P_{\theta_0}}(\delta^3)$ et $(\tilde{W}_t)_{t \geq 0}$ est une suite de variables de carré intégrable telles que $E_{\theta_0}(\tilde{W}_t | \mathcal{F}_t) \stackrel{\text{P.S.}}{=} 0$, $(\mathcal{F}_t)_{t \geq 0}$ étant la filtration naturelle associée à Y (propriété 1, p.20).

Pour le schéma du trapèze, et contrairement au schéma d'Euler, la variable explicative $(f(\theta, Y_t) + f(\theta, Y_{t+\delta}))$ est corrélée au bruit \tilde{W}_t . L'utilisation des Moindres Carrés engendre alors un biais systématique dans la procédure d'estimation. Pour lever cette difficulté, une solution classique utilisée en économétrie consiste à remplacer la variable explicative par un instrument, en l'occurrence $\frac{\partial}{\partial \theta} f(\theta, Y_t)$, qui a la caractéristique d'être décorrélé du bruit tout en étant lié à la variable explicative. Dans le cas où la fonction f est linéaire par rapport à θ , $f(\theta, \cdot) = \theta \cdot f(\cdot)$. cette méthode d'estimation est appelée méthode par *Variable instrumentale*. Bergstrom [3] et Sargan [42] sont les premiers à l'utiliser, dans le cadre d'une diffusion gaussienne et multidimensionnelle. La *méthode d'estimation des moments généralisés* étend la méthode par variable instrumentale dans le cas où l'hypothèse de linéarité n'est pas vérifiée (cf. [20] [22]; [19]).

Le contraste obtenu en couplant le schéma d'approximation du trapèze à la méthode des moments généralisés est :

$$U_n(\theta, \delta) = \left[\frac{1}{n\delta} \sum_{k=1}^n \frac{\partial}{\partial \theta} f(\theta, Y_{(k-1)\delta}) \left(\frac{Y_{k\delta} - Y_{(k-1)\delta} - \frac{\delta}{2} (f(\theta, Y_{(k-1)\delta}) + f(\theta, Y_{k\delta}))}{\delta \sigma^2(Y_{(k-1)\delta})} \right) \right]^2 \quad (5)$$

Nous appuyant sur la loi des grands nombres et le théorème central limite énoncé par Florens-Zmirou pour les diffusions ergodiques (lemme 2 et théorème 1 de [16]), nous montrons que l'estimateur du minimum de contraste converge vers θ_δ (théorème 2, p.32) et est asymptotiquement normal avec une vitesse de l'ordre de $\sqrt{n\delta}$ (théorème 4, p.34), θ_δ étant l'unique minimum de la fonction de contraste U associée à U_n . Nous montrons que θ_δ approche θ_0 avec un biais explicite de l'ordre de δ^2 (théorème 3, p.34). De plus, la variance asymptotique est, à un facteur $(1 + O(\delta))$ près, l'inverse de l'information de Fisher.

Une démarche identique est mise en oeuvre pour le schéma de Simpson :

$$\forall t \geq 0, Y_{t+2\delta} = Y_t + \frac{\delta}{6} (f(\theta_0, Y_t) + 4f(\theta_0, Y_{t+\delta}) + f(\theta_0, Y_{t+2\delta})) + \eta_t + \tilde{W}_t^1 + \tilde{W}_t^2 \quad (6)$$

où η_t est un $O_{P_{\theta_0}}(\delta^5)$ et $(\tilde{W}_t^1)_{t \geq 0}$ et $(\tilde{W}_t^2)_{t \geq 0}$ sont deux suites de variables de carré intégrable telles que : $E_{\theta_0}(\tilde{W}_t^1 | \mathcal{F}_t) \stackrel{\text{P.S.}}{=} 0$ et $E_{\theta_0}(\tilde{W}_t^2 | \mathcal{F}_{t+\delta}) \stackrel{\text{P.S.}}{=} 0$ (propriété 2, p.22).

L'estimateur obtenu est convergent avec un biais explicite de l'ordre de δ^4 , l'efficacité asymptotique étant préservée à un facteur $(1 + O(\delta))$ près (théorème 5, p.37).

A.2.2 Contraste associé à un schéma de discrétisation adapté à l'ordre p

La méthode d'estimation mise en oeuvre dans le Chapitre 2, **Schéma d'approximation adapté à l'ordre p et estimation du paramètre de dérive d'une diffusion**, résulte cette fois de l'association des Moindres Carrés et d'un schéma d'approximation adapté à l'ordre p de la diffusion. Ce schéma repose sur une approximation de l'intégrale $\int_{k\delta}^{(k+1)\delta} f(\theta, Y_s) ds$ basée sur les itérés du générateur de la diffusion. Si on note A_{θ_0} le générateur infinitésimal de la diffusion et $A_{\theta_0}^l$ son $l^{\text{ème}}$ itéré, le schéma (exact) considéré est donné par :

$$\forall t \geq 0, \quad Y_{t+\delta} - Y_t = \sum_{l=0}^{p-1} \frac{\delta^{l+1}}{(l+1)!} A_{\theta_0}^l f(\theta_0, Y_t) + \eta_t^p + \xi_t^p \quad (7)$$

où η_t est une variable d'ordre δ^{p+1} et ξ_t^p est une variable de carré intégrable telle que $E[\xi_t^p | \mathcal{F}_t] \stackrel{p.s.}{=} 0$. Le contraste est alors :

$$U_n^p(\theta, \delta) = \frac{1}{n\delta} \sum_{k=1}^n \left(\frac{Y_{k\delta} - Y_{(k-1)\delta} - M_p(Y_{(k-1)\delta}, \theta, \delta)}{\delta \sigma^2(Y_{(k-1)\delta})} \right)^2 \quad (8)$$

$$M_p(x, \theta, \delta) = \sum_{l=0}^{p-1} \frac{\delta^{l+1}}{(l+1)!} A_{\theta}^l f(\theta, x)$$

Pour δ dans un voisinage de 0, la fonction de contraste associée à $U_n^p(\cdot, \delta)$ admet un unique minimum $\theta_{p,\delta}$ qui approche θ_0 à δ^p près (lemme 1, p.72). Lorsque $n \rightarrow +\infty$, l'estimateur du minimum de contraste converge vers $\theta_{p,\delta}$ (théorème 2, p.73) et est asymptotiquement efficace à un facteur $(1 + O(\delta))$ près (théorème 3, p.74).

Il faut signaler que la procédure d'estimation exposée précédemment rejoint celle développée par Kessler dans [29] pour un développement à l'ordre δ de la variance conditionnelle.

A.3 Estimation par minimum de contraste pour un pas de discrétisation tendant vers 0 et une dérive irrégulière

Dans le Chapitre 3, **Estimation du paramètre de dérive d'une diffusion sous des conditions d'irrégularité de la dérive**, nous considérons une fonction de dérive de la forme $f(\theta_0, x) = f(x - \theta_0)$, où f est une fonction de classe C^2 sur $R - \{0\}$. Nous limitant au cas où le coefficient de diffusion est constant, on suppose que f' et f'' admettent des limites finies à gauche et à droite en 0 avec $f'(0^+) \neq f'(0^-)$: on dit alors que f est un élément de C_0^2 . Sous cette hypothèse, le modèle étudié appartient à la classe plus générale des modèles réguliers : de dérive $x \rightarrow f(\theta, x)$ continue et différentiable en moyenne quadratique par rapport à θ . Dans [33], Kutoyants montre que, pour un modèle régulier, l'estimateur du maximum de vraisemblance associé à une observation continue du processus stationnaire ergodique sur $[0, T]$ est optimal ($T \rightarrow +\infty$). Notre but est également de fournir une estimation optimale de θ_0 pour f dans C_0^2 mais basée sur une observation discrétisée à pas δ_n ($\delta_n \rightarrow 0$) du processus. Nous utilisons pour cela le contraste dérivé des moindres carrés appliqués au schéma d'approximation d'Euler, défini en (3). Dans le cas où f est de classe C^2 sur R , les résultats obtenus par Florens-Zmirou [16] impliquent que, si $n\delta_n \rightarrow +\infty$ et $n\delta_n^3 \rightarrow 0$, l'estimateur du minimum de contraste est consistant, asymptotiquement normal et efficace.

Nous montrons que ces résultats restent valables lorsque f est dans C_0^2 , si l'on suppose de plus que f est lipschitzienne.

La vérification des conditions classiques, qui assurent la consistance de l'estimateur lorsque $n\delta_n \rightarrow 0$ et $\delta_n \rightarrow 0$ (théorème 3.2.8., [11]), utilise une loi forte des grands nombres (lemme 3, p.96) basée sur l'hypothèse suivante de contractivité du générateur de la diffusion : si $P_{\theta_0}^t(x, dy) = P_{\theta_0}(Y_t \in dy \mid Y_0 = x)$, il existe $\lambda > 0$ telle que, pour toute fonction h vérifiant $\mu_{\theta_0}(h) = 0$ et $\mu_{\theta_0}(h^2) < +\infty$, on a :

$$\|P_{\theta_0}^t(h)\|_{L^2(\mu_{\theta_0})} \leq \exp(-\lambda t) \|h\|_{L^2(\mu_{\theta_0})} \quad (9)$$

Du fait du caractère Lipschitz de f , des conditions suffisantes simples assurent cette hypothèse de contractivité (cf. proposition 9, [21], p.801).

On montre également que si $n\delta_n^3 \rightarrow 0$, $\hat{\theta}_n$ est asymptotiquement normal avec une vitesse de l'ordre $\sqrt{n\delta_n}$ (théorème 2, p.101). Ce résultat ne peut être obtenu par la méthode classique qui consiste à étudier le développement de Taylor de la dérivée du contraste (cf. § 3.3.4. de [11]) :

$$\sqrt{n\delta_n} \left(\frac{\partial}{\partial \theta} U_n(\hat{\theta}_n) - \frac{\partial}{\partial \theta} U_n(\theta_0) \right)$$

Nous allons ainsi travailler directement sur le contraste en nous inspirant des méthodes développées par Ibragimov et Has'minskii [25]. Plus précisément, nous allons établir un développement local du contraste autour de θ_0 , qui satisfait une condition LAN (local asymptotic normality). Le schéma de démonstration comporte deux étapes.

- (i) Nous commençons par déterminer la vitesse de convergence de $\hat{\theta}_n$ vers θ_0 . Reprenant des arguments développés par Chan dans le cadre d'un modèle à seuil à temps discret [9], on montre que, si $n\delta_n \rightarrow 0$ et $(n\delta_n)$ est bornée, la suite $\left(\sqrt{n\delta_n} (\hat{\theta}_n - \theta_0) \right)_n$ est tendue.

On étudie alors le contraste localement en $\theta_0 + \frac{u}{\sqrt{n\delta_n}}$. Dans ce but, on introduit la variable Z_n définie par $Z_n(u) = n\delta_n \left(U_n \left(\theta_0 + \frac{u}{\sqrt{n\delta_n}} \right) - U_n(\theta_0) \right)$, dont nous exhibons une décomposition de la forme :

$$Z_n(u) = 2I(0)u\Delta_n + I(0)u^2 + \Psi_n(u)$$

Dans une situation standard de dérivabilité, $I(0)\Delta_n$ serait identifiée à $\sqrt{n\delta_n} \frac{\partial}{\partial \theta} U_n(\theta_0)$.

- (ii) Nous montrons alors que Z_n satisfait une condition LAN.

En effet, nous appuyant sur le théorème central limite pour les suites triangulaires (théorème 2.8.43. [11]), nous montrons que, lorsque $n\delta_n \rightarrow +\infty$ et $\delta_n \rightarrow 0$, Δ_n converge en loi vers la loi Normale $\mathcal{N} \left(0, \frac{1}{I(0)} \right)$ avec $I(0) = \int (f')^2(x - \theta_0) \mu_{\theta_0}(dx)$.

Nous montrons de plus que Ψ_n est une fonction continue presque sûrement telle que, lorsque $n\delta_n \rightarrow +\infty$ et $n\delta_n^3 \rightarrow 0$, on a pour tout $B > 0$:

$$\sup_{|u| \leq B} |\Psi_n(u)| \xrightarrow{P_{\theta_0}} 0$$

La normalité asymptotique de $\widehat{\theta}_n$ résulte alors d'une adaptation du théorème 1.2 de Ibragimov et Has'minskii [25] (théorème 4, p.103). Ce théorème établit en effet la normalité de l'estimateur du maximum de vraisemblance lorsque la vraisemblance du modèle considéré satisfait une condition LAN. On obtient ainsi le résultat de convergence en loi de $\widehat{\theta}_n$:

$$\sqrt{n\delta_n} (\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}(P_{\theta_0})} \mathcal{N} \left(0, \frac{1}{I(0)} \right)$$

Dans ce même chapitre, nous traitons également le cas des modèles autorégressifs à temps discret de la forme :

$$X_n = f(X_{n-1} - \theta_0) + e_n$$

où f est continue et présente le même type d'irrégularité que la dérive dans le contexte précédent. $(e_n)_n$ est une suite de variables indépendantes, identiquement distribuées et centrées, telle que e_n est indépendante de $\mathcal{F}_{n-1} = \sigma(X_0, X_1, \dots, X_{n-1})$. Ici, f n'est plus supposée lipschitzienne mais sous-linéaire : il existe $\lambda \in]0, 1[$ et $K > 0$ telles que $|f(x)| \leq \lambda|x| + K$. Sous des conditions supplémentaires qui assurent que la chaîne de Markov est récurrente positive, l'estimateur des Moindres Carrés Conditionnels est consistant et asymptotiquement normal avec une vitesse standard (théorème 7, p.115). Ceci répond à la question posée par Chan ([9], p.524) : " Another problem is to work out the limiting distribution of the CLSE when the autoregressive function is continuous". Rappelons que le contexte étudié par Chan est différent puisque θ_0 est aussi un point de discontinuité pour la fonction d'autorégression : dans ce cas, la vitesse et la loi asymptotique ne sont plus standards.

A.4 Estimation d'un CAR(p) incomplètement observé à partir des équations de Yule-Walker

Dans le Chapitre 4, la classe de processus considérée est celle des CAR(p) ([7],[23]). Un CAR(p) de paramètre $\theta = (\alpha_0, \dots, \alpha_{p-1}; \sigma)$ est un processus $X = (X(t))_{t \geq 0}$ dérivable à l'ordre $(p-1)$ tel que, si on note $X^{(j)}$ sa dérivée à l'ordre j , $Y(t) = {}^t(X(t), \dots, X^{(p-1)}(t))$ est l'unique solution stationnaire de l'équation différentielle stochastique p -dimensionnelle suivante :

$$dY(t) = AY(t) dt + \sigma b dW(t) \quad (10)$$

où $(W(t))_{t \geq 0}$ est un mouvement brownien standard unidimensionnel, $Y(0)$ est indépendant de W , et A et b sont respectivement une matrice de dimensions $p \times p$ et un vecteur de R^p définis par :

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & \cdots & -\alpha_{p-1} \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Si toutes les racines du polynôme caractéristique de A , $P(z) = z^p + \alpha_{p-1}z^{p-1} + \dots + \alpha_0$, admettent une partie réelle négative, (10) admet une unique solution stationnaire et on a :

$$dX^{(p-1)}(t) + [\alpha_0 X(t) + \dots + \alpha_{p-1} X^{(p-1)}(t)] dt = \sigma dW(t) \quad (11)$$

Le problème de l'estimation du paramètre $\alpha = (\alpha_0, \dots, \alpha_{p-1})$ à partir d'une observation continue et complète de Y sur $[0, T]$, c'est-à-dire une observation continue de X et de ses dérivées jusqu'à l'ordre $(p-1)$, est étudié par Bartlett [1] [2], Dzhaparidze [15], Dinh Tuan [13], Priestley [40]. Le cas où le processus est observé à temps discret est étudié par de nombreux auteurs, signalons Bartlett [1], Durbin [14], Phillips [38], Bergstrom [4], Jones [27] et Robinson [41].

Dans [24], Hyndman propose une méthode d'estimation basée sur les équations de Yule-Walker associées au modèle. Dans le cas des modèles autorégressifs à temps continu, ces équations font intervenir les fonctions de covariance dérivées $(D_{j,k})$ qui sont définies par :

$$\forall (j, k) \in \{0, 1, \dots, p-1\}, \quad D_{j,k}(h) = E_{\theta_0} \left[X^{(j)}(h) X^{(k)}(0) \right]$$

$$D_{j,p}(h) \stackrel{L^2(P_\theta)}{=} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T X^{(j)}(t+h) dX^{(p-1)}(t)$$

Les équations de Yule-Walker sont obtenues en multipliant l'équation (11) par $X^{(j)}(t+h)$ et en passant à l'espérance sous P_θ . On obtient alors, pour tout $j = 0, \dots, p-1$ et tout $h \in \mathbb{R}$:

$$D_{j,p}(h) + \alpha_{p-1} D_{j,p-1}(h) + \dots + \alpha_0 D_{j,0}(h) = 0 \quad (12)$$

La paramètre α satisfait ainsi le système constitué des $p-1$ équations précédentes. Pour h quelconque et fixé, il suffit donc d'estimer correctement les $(D_{j,k}(h))$ pour obtenir une estimation satisfaisante de α .

Pour une observation continue et complète de Y sur $[0, T]$, Hyndman propose d'estimer $D_{j,k}(0)$ par :

$$\forall (j, k) \in \{0, 1, \dots, p-1\}, \quad \bar{D}_{j,k} = \frac{1}{T} \int_0^T X^{(j)}(t) X^{(k)}(t) dt$$

$$\bar{D}_{j,p}(0) = \frac{1}{T} \int_0^T X^{(j)}(t) dX^{(p-1)}(t)$$

Ces estimateurs étant consistants, il obtient par l'intermédiaire des équations de Yule-Walker en $h = 0$, une estimation asymptotiquement normale de α à la vitesse \sqrt{T} .

Lorsque X , la première composante de Y , est seule observée et cela à des temps discrets équidistants de δ , les dérivés successives du CAR(p) doivent être estimés. Si on note $\Delta X(t+\delta) = X(t+\delta) - X(t)$ et $\Delta^{(j)}$ le $j^{\text{ième}}$ itéré de cet opérateur, une estimation naturelle de $X^{(j)}(t)$ est $\delta^{-j} \Delta^{(j)} X(t+j\delta)$. Les estimateurs des $(D_{j,k}(0))$ sont alors de la forme :

$$\hat{D}_{j,k} = \frac{\delta^{-(j+k)}}{n_p} \sum_{i=0}^{n_p-1} \Delta^{(j)} X(ip\delta + j\delta) \Delta^{(k)} X(ip\delta + k\delta), \quad n_p = \left\lceil \frac{n}{p} \right\rceil$$

Cependant, comme le remarque Hyndman [24] dans le cas d'un CAR(2) et comme nous le montrons plus généralement pour un CAR(p), $\hat{D}_{p-1,p}$ présente un biais systématique. Plus précisément, nous montrons que :

$$\hat{D}_{p-1,p} \xrightarrow{P_{\alpha,\sigma}} C(p) D_{p-1,p}(0)$$

où $C(p) \neq 1$ est une constante indépendante de δ et de (α, σ) (lemme 3, p.138).

Nous remédions à ce problème en estimant les équations de Yule-Walker, non pas en $h = 0$ mais en $h = -(p-1)\delta$. Les estimateurs empiriques des $(D_{j,k}(-(p-1)\delta))$ sont donnés par :

$$\tilde{D}_{j,k} = \frac{\delta^{-(j+k)}}{n_{2p-1}} \sum_{i=0}^{n_{2p-1}-1} \Delta^{(j)} X(i(2p-1)\delta + j\delta) \Delta^{(k)} X(i(2p-1)\delta + (p-1)\delta + k\delta), \quad n_{2p-1} = \left\lfloor \frac{n}{2p-1} \right\rfloor$$

et l'estimateur α^n de α est défini comme la solution du système d'équations suivant :

$$\forall j = 0, \dots, p-1, \quad \tilde{D}_{j,p} + \alpha_{p-1}^n \tilde{D}_{j,p-1} + \dots + \alpha_0^n \tilde{D}_{j,0} = 0$$

Nous montrons que cet estimateur converge vers α^δ qui approche α à un biais explicite de l'ordre de δ (lemme 5, p.140, théorème 2, p.141).

Dans le cas d'un CAR(2), nous proposons une méthode qui permet de réduire le biais d'estimation associé à α . Cette méthode d'estimation est basée sur l'utilisation du schéma anticipatif bi-dimensionnel du trapèze associé au CAR(2). α est alors estimé à δ^2 près (théorème 4, p.145).

Nous nous intéressons enfin à l'estimation du paramètre σ^2 . La méthode d'estimation repose sur le résultat suivant :

$$D_{p-1,p} = -\frac{\sigma^2}{2}$$

Il suffit donc de débiaiser $\hat{D}_{p-1,p}$ (l'estimateur naturel de $D_{p-1,p}(0)$) par le facteur explicite $C(p)$ pour obtenir une estimation de σ^2 à δ près (théorème 5, p.146).

A.5 Etudes expérimentales

Chacun des chapitres précédents se terminent par une étude expérimentale par simulations. Les méthodes d'estimation sont mises en oeuvre pour des modèles de diffusion classiques tels les modèles de Ornstein-Uhlenbeck et Cox-Ingersoll-Ross pour les deux premiers chapitres, mais aussi pour une diffusion log-normale dans le Chapitre 1. Dans le Chapitre 3, les applications numériques portent sur un CTAR(1) (cf. [8],[7] et [23]) à raccordement continu. Pour le Chapitre 4, les résultats théoriques sont testés pour un CAR(2).

La diffusion est simulée sur $[0, T]$ à partir d'un schéma d'Euler de pas fin 0.005 ou 0.0001. On calcule alors indépendamment N estimateurs ($N = 100$ ou $N = 200$). La moyenne et la variance empiriques ainsi obtenues sont comparées à leur valeur théorique. Cette procédure est effectuée pour différentes valeurs de T ($T = 200, 400, 600$).

B Estimation de la précision associée à un échantillonnage systématique et méthodes transitives

Dans le Chapitre 5 **Precision of systematic sampling and transitive methods**¹, on considère Q une quantité déterministe inconnue qui s'écrit sous la forme :

$$Q = \int_R f(x) dx$$

où f est une fonction appelée *fonction de mesure*. En outre, f est supposée (m, p) -régulière par morceaux, c'est-à-dire que f est à support borné et est dérivable par morceaux jusqu'à l'ordre $m + p$ telle que :

- pour tout $k = 0, \dots, m - 1$, $f^{(k)}$ est continue sur R
- pour tout $k = m, \dots, m + p$, $f^{(k)}$ peut être discontinue sur $D_{f^{(k)}}$ qui est un ensemble fini de points, avec pour tout $x \in D_{f^{(k)}}$, un saut $S_{f^{(k)}}(x) = f^{(k)}(x^+) - f^{(k)}(x^-)$ fini.

Q est estimée empiriquement à partir de mesures effectuées selon un mode opératoire systématique et uniforme. On dispose ainsi de données de la forme $\{f(UT + kT), k \in Z\}$, où $T > 0$ est fixé et U est une variable uniformément distribuée sur $[0, 1[$. L'estimateur empirique, $\hat{Q}_T = T \sum_{k \in Z} f((U + k)T)$, est alors sans biais.

L'estimation basée sur l'échantillonnage systématique est particulièrement utilisée dans le domaine de la stéréologie. Un problème classique est l'estimation du volume Q d'un objet tridimensionnel à partir de *coupes sériées* : l'objet est débité en sections équidistantes et parallèles selon une orientation déterminée. Dans ce cas, $f(x)$ représente l'aire de la coupe de hauteur x et \hat{Q}_T est l'estimateur dit de *Cavalieri* ([45],[10]). Lorsque l'objet considéré est, par exemple, un ellipsoïde, la fonction de mesure est alors $(1, \infty)$ -régulière par morceaux. Si Q représente maintenant le volume d'un tétraèdre dont une des faces est parallèle aux plans de coupe, f est dans ce cas $(0, \infty)$ -régulière par morceaux.

L'inconvénient principal attaché à cette méthodologie réside dans la corrélation des données obtenues qui rend difficile l'évaluation de la précision de l'estimateur \hat{Q}_T .

Une première approche consiste à modéliser la structure de cette corrélation. Une méthode

¹A paraitre dans " Journal of Statistical Planning and Inference". En collaboration avec Kiên Kiêu, Unité de Biométrie, INRA, F78026, Saint-Cyr, et Jacques Istas, Unité de Biométrie, INRA, F78352 Jouy-en-Josas.

alternative a été développée par Matheron sous le nom de *Méthodes transitives* [35]. Cette méthode consiste à approximer l'écart quadratique moyen $\text{MSE}(\widehat{\mathbf{Q}}_T)$ en se basant sur le comportement du *covariogramme* g de f au voisinage de 0, g étant défini par $g(y) = \int_{\mathbb{R}} f(x+y)f(x)dx$. Matheron montre en effet que l'écart quadratique moyen peut être exprimé en fonction de g (équation (5.7), p.181) :

$$\text{MSE}[\widehat{\mathbf{Q}}_T] = T \sum_{k \in \mathbb{Z}} g(kT) - \int_{\mathbb{R}} g(y) dy \quad (13)$$

La méthode consiste alors à développer l'expression précédente en fonction de T et à approximer $\text{MSE}[\widehat{\mathbf{Q}}_T]$ par la partie contributive du développement obtenu. On utilise pour cela une généralisation de la formule d'*Euler-Mac Laurin* pour les fonctions régulières par morceaux ([43], [32]). En particulier, si g est $(2m+1, 1)$ -régulière par morceaux, $\text{MSE}[\widehat{\mathbf{Q}}_T]$ s'écrit alors, à un $o(T^{2m+2})$ près, comme la somme d'un terme d'extension $E(T)$ (équation (5.11), p.182) et d'un terme oscillant $Z(T)$ aussi appelé *Zitterbewegung* (cf.[35]). Ces termes sont définis par :

$$E(T) = -2T^{2m+2} P_{2m+2}(0) g^{(2m+1)}(0^+) \quad (14)$$

$$Z(T) = -T^{2m+2} \sum_{\substack{c \in D_{g^{(2m+1)}} \\ c \neq 0}} P_{2m+2}\left(\frac{c}{T} - \left\lfloor \frac{c}{T} \right\rfloor\right) s_{g^{(2m+1)}}(c) \quad (15)$$

où P_{2m+2} est le *polynôme de Bernoulli* d'ordre $(2m+2)$. Les expressions précédentes font donc intervenir le degré de régularité et les sauts du covariogramme qui n'est pas directement observé. Nous montrons cependant que, si la fonction de mesure f est (m, p) -régulière par morceaux, alors g est $(2m+1, p)$ -régulière par morceaux. Les sauts de g peuvent de plus être exprimés en fonction de ceux de f (cf. (5.9)-(5.10), p.182)

Le *Zitterbewegung* est un terme qui oscille autour de 0. Dans un cas simple de fonction de mesure aléatoire, Kiêu [32] a montré que $Z(T)$ est négligeable devant $E(T)$. En pratique, on néglige le *Zitterbewegung*. Ainsi, en première approximation

$$\text{MSE}[\widehat{\mathbf{Q}}_T] \simeq -2T^{2m+2} P_{2m+2}(0) g^{(2m+1)}(0^+)$$

Cependant, $g^{(2m+1)}(0^+)$ est inconnu et va être approximé par une combinaison linéaire des $(g(iT))_{i=0, n_{2m+1}}$, où n_{2m+1} dépend du degré de régularité du covariogramme. Ceci ne fait que déplacer le problème dans la mesure où les $g(iT)$ sont également inconnus. Toutefois, $g(iT)$ peut être estimé de manière satisfaisante par $\widehat{\mathbf{g}}_i = T \sum_{k \in \mathbb{Z}} f(UT + kT)(UT + (k+i)T)$. L'estimateur obtenu $\widehat{\mathbf{E}}_m(T)$ est de la forme (cf. § 6.2 de [32]) :

$$\widehat{\mathbf{E}}_m(T) = T \sum_{i=0}^{n_{2m+1}} \lambda_i \widehat{\mathbf{g}}_i$$

La procédure décrite précédemment permet donc d'obtenir une estimation du terme d'extension $E(T)$ défini en (14) qui est la composante principale de l'écart quadratique moyen de $\widehat{\mathbf{Q}}_T$.

Cette méthode présente cependant un inconvénient majeur puisqu'elle suppose que le degré de régularité m de la fonction de mesure est connu. Nous proposons donc de lui associer

une méthode d'estimation de m . Dans ce but, nous adaptons une méthode développée par Istas et Lang [26] pour estimer le degré de régularité de processus gaussiens discrètement observés.

La procédure d'estimation de m repose sur la propriété suivante : si M est un entier fixé a priori tel que $m \leq M$ alors $E\left(\widehat{\mathbf{E}}_M(T)\right) = c_m E(T)$ où c_m est une constante indépendante de T . Donc, même si l'estimateur est biaisé, on conserve un développement de l'ordre de T^{2m+2} . Si on note $\widehat{\mathbf{E}}_M(2T) = 2T \sum_{i=0}^{n_{2m+1}} \lambda_i \widehat{\mathbf{g}}_i$, on a ainsi : $E\left(\widehat{\mathbf{E}}_M(2T)\right) \simeq 2^{2m+2} E\left(\widehat{\mathbf{E}}_M(T)\right)$. Un estimateur naturel associé à la relation précédente est alors :

$$\widehat{\mathbf{m}}_T = \left[\frac{1}{2 \ln(2)} \ln \left(\frac{\widehat{\mathbf{E}}_M(2T)}{\widehat{\mathbf{E}}_M(T)} \right) - \frac{1}{2} \right]$$

Si m est inférieur à l'entier que l'on s'est fixé a priori, l'estimateur obtenu converge en probabilité vers m lorsque T tend vers 0.

Bibliography

- [1] Bartlett, M. S. (1946) *On the theoretical specification and sampling properties of auto-correlated time-series.*- J. Roy. Statist. Soc. Ser. B, 8, 27-41.
- [2] Bartlett, M. S. (1955) *An introduction to stochastic processes with special reference to methods and applications.*- Cambridge University Press, London.
- [3] Bergstrom, A. R. (1976) *Statistical inference in Continuous Time Series.*- In *Statistical inference in Continuous Time Economic Models.*, ed. Bergstrom, North Holland.
- [4] Bergstrom, A. R. (1984) *Continuous time stochastic models and issues of aggregation over time.*- In *Handbook of econometrics*, Vol. 2, ed. Griliches, Z. and Intriligator, M.D., Elsevier Science Publishers : BV, 1146-1212.
- [5] Bibby, B. M. (1994) *Optimal combination of martingale estimating functions for discretely observed diffusion processes.*- Research Reports 298, Department of theoretical statistics, University of Aarhus.
- [6] Bibby, B. M. & Sorensen, M. (1995) *Martingale Estimation Functions for Discretely Observed Diffusion Processes.* - Bernoulli 1, 17-39.
- [7] Brockwell, P. J. & Hyndman, R. J. (1992) *On continuous-time threshold autoregression.*- International Journal of Forecasting 8, North-Holland, 157-173.
- [8] Brockwell, P. J., Hyndman, R. J. & G. K. Grunwald (1991) *Continuous time threshold autoregressive models.*- Statistica Sinica, Vol. 1, No. 2, 401-410.
- [9] Chan, K.S. (1993) *Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model.*- The Annals of Statistics, Vol. 21, 1, 520-533.
- [10] Cruz-Orive, L. M. (1987) *Stereology: historical notes and recent evolution.*- In *The comemorative-Memorial Volume. Twenty-Five Years of Stereology*, ed. R. E. Miles, Acta Stereol., 6, Suppl. II, 43-56.
- [11] Dacunha-Castelle, D. & Dufflo, M. (1993) *Probabilité et Statistiques. Tome 2- 2^{eme} ed.* Masson.
- [12] Dacunha-Castelle, D. & Florens-Zmirou, D. (1986) *Estimation of the coefficient of a diffusion from discrete observations.*- Stochastics 19, 263-284.
- [13] Dinh Tuan, P. (1977) *Estimation of parameters of a continuous time gaussian stationary process with rational spectral density.*- Biometrika 64, 285-299.

- [14] Durbin, J. (1961) *Efficient fitting of linear models for continuous stationary time series from discrete data.*- Bull. Int. Statist. Inst., 38, 273-281.
- [15] Dzhaparidze, K. O. (1971) *On methods for obtaining asymptotically efficient spectral parameter estimates for stationary Gaussian process with rational spectral density.*- Theory Prob. Appl., 16, 550-554.
- [16] Florens-Zmirou, D. (1989) *Approximate discrete schemes for statistics of diffusion processes.*- Statistics 20, 547-557.
- [17] Godambe, V. P. & Heyde, C. C. (1987) *Quasi likelihood and optimal estimation.*- Int. Statist. Rev. 55, 231-244.
- [18] Genon-Catalot, V. (1990) *Maximum contrast estimation for diffusion processes from discrete observations.*- Statistics 21, 99-116.
- [19] Gourieroux, C. & Monfort, A. (1989) *Statistique et Modèles Econométriques. Tome 1- Economica.*
- [20] Hansen, L. (1982) *Large Sample Properties of Generalized Method of Moments Estimators.*- Econometrica, 50, 4, 1029-1054.
- [21] Hansen, L. P. & Scheinkman, J.A. (1995) *Back to the future : generating moment implications for continuous-time markov processes.*- Econometrica, Vol. 63, 4, 767-804.
- [22] Hansen, L. & Singleton, K. (1982) *Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models.*- Econometrica, Vol. 50, 5, 1269-1286.
- [23] Hyndman, R. J. (1992) *Continuous time threshold autoregressive models.*- PhD thesis, University of Melbourne, Department of statistics.
- [24] Hyndman, R. J. (1993) *Yule-Walker estimates for continuous-time autoregressive models.* - Journal of Time Series Analysis, Vol. 14, 3, 281-296.
- [25] Ibragimov, I. A. & Has'minskii, R. Z. (1981) *Statistical Estimation- Asymptotic Theory.*- Springer-Verlag.
- [26] Istas, J. & Lang, G. (1997) *Quadratic variations and estimation of the local Hölder index of a Gaussian process.*- A paraitre dans *Annales de l'instut Henri Poincaré.*
- [27] Jones, R. H. (1981) *Fitting a continuous time autoregression to discrete data.*- In *Applied Time Series Analysis*, Vol. II, ed. D. F. Findlay, Academic Press, N.Y., 651-682.
- [28] Kessler, M. (1995) *Martingale estimating functions for a Markov chain.*- Prépublication du laboratoire de Probabilités de l'Université de PARIS VI.
- [29] Kessler, M. (1997) *Estimation of an ergodic diffusion from discrete observations.*- Scand. J. Stat. 24, No. 2, 211-229.
- [30] Kessler, M. (1996) *Simple and Explicit Estimating Functions for a Discretely Observed Diffusion Process.*- Research Reports 336, Department of theoretical statistics, University of Aarhus.

- [31] Kessler, M. & Sorensen, M. (1995) *Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process.*- Research Reports 332, Department of theoretical statistics, University of Aarhus.
- [32] Kiêu, K. (1997) *Three lectures on systematic geometric sampling.* Memoirs 13, Department of Theoretical Statistics, University of Aarhus.
- [33] Kutoyants, A. Y. (1984) *Parameter estimation for stochastic processes.*- Research and Exposition in math. 6, Heldermann Verlag, Berlin.
- [34] Lipster, R. S. & Shiryaev, A. N. (1977) *Statistics of random processes. Tome 1,2.*- Springer-Verlag, N. Y.
- [35] Matheron, G. (1965) *Les variables régionalisées et leur estimation.*- Masson, Paris.
- [36] Pedersen, A. R. (1995) *Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes.*- Bernoulli 1, 257-279.
- [37] Pedersen, A. R. (1995) *A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations.*- Scand. J. Statist. 22, 55-71.
- [38] Phillips, P. C. B. (1972) *The structural estimation of a stochastic differential equation system.*- Econometrica, 40, 1021-1041.
- [39] Prakasa-Rao, B. L. S. (1983) *Asymptotic theory for non linear least squares estimator for diffusion processes.*- Math. Operationsforsch. u. Statist, Ser. Statist 14, 2, 195-209.
- [40] Priestley, M. B. (1981) *Spectral analysis and time series.*- Vol. 1, Academic Press, London.
- [41] Robinson, P. M. (1977) *Estimation of a time series model from unequally spaced data.*- Stoch. Proc. Appl., 6, 9-24.
- [42] Sargan, J. D. (1976) *Some discrete approximations to continuous times stochastic models* - In *Statistical inference in Continuous Time Economic Models.*, ed. Bergstrom, North Holland, 27-80.
- [43] Souchet, S. (1995) *Précision de l'estimateur de Cavalieri.*- Rapport de stage, DEA de statistique et modèles aléatoires appliqués à la finance, Université Paris VII. Laboratoire de biométrie, INRA-Versailles.
- [44] Souchet, S. (1998) *Schéma de discrétisation anticipatif et estimation du paramètre de dérive d'une diffusion.*-A paraître comme note au CRAS.
- [45] Sterio, D. C. (1984) *The unbiased estimation of number and sizes of arbitrary particles using the disector.*- J. Microsc., 134, 127-136.
- [46] Yoshida, N. (1992) *Estimation for diffusion processes from discrete observations.*- J. Multivariate Anal. 41, 220-242.

Estimation paramétrique d'une diffusion ergodique discrétisée

Chapter 1

Schémas de discrétisation anticipatifs et estimation du paramètre de dérive d'une diffusion

1.1 Introduction

Nous abordons dans ce travail le problème de l'estimation du paramètre de dérive d'une diffusion ergodique observée à des instants équidistants de pas $\delta > 0$ constant et petit.

Le problème d'estimation paramétrique pour les diffusions a fait l'objet de nombreux travaux, que l'on peut classer en fonction du type d'observations disponibles.

Si l'on note $[0, T]$ l'intervalle d'étude et si l'on dispose d'une observation continue de la trajectoire du processus, l'estimateur du maximum de vraisemblance est consistant, asymptotiquement normal et efficace avec la vitesse \sqrt{T} , lorsque T tend vers l'infini ([16], [15]).

Cependant, le recours à une observation discrète du processus paraît plus réaliste. Là encore, on distingue le cas d'un pas d'observation constant $\delta > 0$ de celui où les temps d'observation sont asymptotiquement denses dans R^+ .

Dans un cadre asymptotique du type $T = nh_n \rightarrow \infty$, $h_n \rightarrow 0$, $nh_n^p \rightarrow 0$, Florens-Zmirou [5], Prakasa-Rao [18] et Yoshida [23] proposent des procédures d'estimation optimales issues de la discrétisation de la vraisemblance continue; Kessler [11] donne une approximation gaussienne de la vraisemblance. Pedersen propose une méthode d'évaluation numérique des densités de transition tout en contrôlant les propriétés théoriques de l'estimation ([19], [20]).

Dans le cas où le pas d'observation δ est constant, Dacunha-Castelle et Florens-Zmirou [4] ont prouvé que l'estimateur du maximum de vraisemblance est consistant et asymptotiquement efficace à la vitesse $\sqrt{n\delta}$, quand $n \rightarrow \infty$. Mais on ne dispose pas, en général, de la forme analytique des densités de transition du processus : un autre point de vue est de considérer des estimateurs issus de fonctions d'estimation (Bibby et Sorensen [2]; Kessler [12] [13]; [22]). Dans [5], Florens-Zmirou construit un contraste à partir du schéma d'Euler approchant le schéma discret exact. Nous adopterons une démarche identique mais en utilisant des schémas anticipatifs.

Soit (Ω, \mathcal{A}, P) , un espace de probabilité et W un mouvement brownien unidimensionnel standard issu de 0 défini sur cet espace. On munit l'espace précédent de la filtration naturelle

$\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}$ associée à W . On considère alors l'équation différentielle stochastique :

$$dY_t = f(\theta_0, Y_t) dt + \sigma(Y_t) dW_t, \quad Y_0 = x \quad (1.1)$$

où $f : R^p \times R \rightarrow R$ et $\sigma : R \rightarrow R$ sont des fonctions connues et $\theta_0 \in R^p$ est le paramètre à estimer.

Supposons que (1.1) admette une unique solution forte de condition initiale x , notée $Y = (Y_t)_{t \geq 0}$ (Karatzas et Shreve [10]). La relation entre deux observations successives est :

$$Y_{(k+1)\delta} = Y_{k\delta} + \int_{k\delta}^{(k+1)\delta} f(\theta_0, Y_t) dt + \int_{k\delta}^{(k+1)\delta} \sigma(Y_t) dW_t$$

Florens-Zmirou utilise l'approximation $\delta f(\theta_0, Y_{k\delta})$ de l'intégrale $\int_{k\delta}^{(k+1)\delta} f(\theta_0, Y_t) dt$ et obtient ainsi le schéma d'Euler :

$$Y_{(k+1)\delta} = Y_{k\delta} + \delta f(\theta_0, Y_{k\delta}) + \eta_{k\delta} + \tilde{W}_{k\delta}$$

Sous certaines hypothèses, $(\tilde{W}_{k\delta})_{k \geq 0}$ est une suite de variables centrées, de carré intégrable, et $\eta_{k\delta}$ est une variable en $O(\delta^2)$ qui correspond à l'erreur commise en approchant l'intégrale par $\delta f(\theta_0, Y_{k\delta})$. La méthode d'estimation des moindres carrés associée au schéma d'approximation d'Euler à pas δ donne une estimation de θ_0 avec un biais de l'ordre de δ ([5] et annexe F).

Cependant, des résultats d'analyse classique nous montrent qu'il existe des méthodes plus fines permettant de mieux approcher l'intégrale d'une fonction déterministe, méthodes où l'erreur commise est inférieure à un $O(\delta^2)$. On peut citer la méthode du trapèze (méthode à deux points), la méthode de Simpson (méthode à trois points) ou encore la méthode de Bode (méthode à cinq points). Elles donnent respectivement, sous de bonnes conditions de régularité pour f (cf. [17]) :

$$\begin{aligned} \int_a^{a+\delta} f(x) dx &= \frac{\delta}{2} (f(a) + f(a+\delta)) + O(\delta^3) \\ \int_a^{a+2\delta} f(x) dx &= \frac{\delta}{3} (f(a) + 4f(a+\delta) + f(a+2\delta)) + O(\delta^5) \\ \int_a^{a+4\delta} f(x) dx &= \frac{\delta}{45} (14f(a) + 64f(a+\delta) + 24f(a+2\delta) + 64f(a+3\delta) + 14f(a+4\delta)) + O(\delta^7) \end{aligned}$$

D'où l'idée de transposer ces approximations à notre contexte, à savoir construire de nouveaux schémas d'approximation en les couplant à une méthode d'estimation adaptée afin d'obtenir un biais d'estimation d'un ordre supérieur à δ et ceci, sans perte d'efficacité.

La paternité de cette idée revient aux économètres (cf. Bergstrom [1], Sargan [21]): utilisant le schéma du trapèze couplé avec la méthode d'estimation par variable instrumentale (V.I.), ils obtiennent un contrôle du biais en δ^2 pour une diffusion vectorielle gaussienne ergodique Y , de dérive linéaire en θ : $f(\theta, Y_t) = \theta Y_t$, θ étant une matrice de dimensions $r \times r$ et Y_t un vecteur de R^r . Dans le cas d'un modèle de diffusion unidimensionnel (1.1), nous montrons que, sous des conditions très générales, le schéma anticipatif du trapèze (resp. de Simpson) couplé à la méthode d'estimation des moments généralisés (G.M.M.) (Hansen [8],[9]; Gourieroux et Monfort [7]) permet d'affiner le contrôle du biais qui est alors en δ^2 (resp. δ^4), cela sans perte d'efficacité.

Dans le paragraphe 2, nous présentons les schémas exacts obtenus en utilisant l'approximation du trapèze et l'approximation de Simpson pour l'intégrale $\int_{k\delta}^{(k+1)\delta} f(\theta, Y_v) dv$. La section 3 est consacrée à l'étude des propriétés asymptotiques de l'estimateur issu de ces schémas d'approximation couplés à la méthode d'estimation par variable instrumentale, dans le cas particulier où la dérive $f(\theta_0, y)$ est linéaire en θ_0 , $f(\theta_0, y) = \sum_{i=1}^p f_i(y) \theta_0^i$. Dans le paragraphe 4, utilisant la méthode des moments généralisés, nous étendons ces résultats pour une dérive $f(\theta_0, y)$ non nécessairement linéaire en θ_0 . La section 5 est une étude expérimentale qui illustre les résultats précédents pour trois modèles de diffusion particuliers.

1.2 Présentation des schémas d'approximation : trapèze, Simpson

Soit l'équation différentielle stochastique définie par :

$$dY_t = f(Y_t) dt + \sigma(Y_t) dW_t, \quad Y_0 = y \quad (1.2)$$

Supposons que (Y, W) , (Ω, \mathcal{F}, P) , $\{\mathcal{F}_t\}_{t \geq 0}$ soit une solution faible de (1.2) [10].

Commençons par introduire quelques notations. Pour tout $p \in \mathbb{N}^*$, $C^p(R)$ représente l'espace des fonctions p fois continument dérivables sur R . Pour toute fonction g dans $C^p(R)$, g' , g'' et $g^{(k)}$ sont respectivement les dérivées à l'ordre 1, 2 et k de g en x . Notons A le générateur infinitésimal associé à (1.2) défini par : $\forall h \in C^2(R)$, $Ah = fh' + \frac{1}{2}\sigma^2 h''$; A^l représente alors le $l^{\text{ième}}$ itéré de A . La notation abrégée, $l = m, r$, se lit $l \in \{m, \dots, r\}$. Si $B = (B_k^i)_{k,i}$ est une matrice, k représente l'indice des lignes et i celui des colonnes; ${}^t B$ est la matrice transposée de B .

Propriété 1 : Schéma du trapèze

Supposons que $f \in C^1(R)$ et $\sigma \in C^2(R)$. Alors :

$$\forall t \geq 0, \quad Y_{t+\delta} = Y_t + \frac{\delta}{2} (f(Y_t) + f(Y_{t+\delta})) + \eta_t + \tilde{W}_t \quad (1.3)$$

avec :

$$\begin{aligned} \eta_t &= \frac{1}{2} \int_t^{t+\delta} (t+\delta-v)(t-v) A^2 f(Y_v) dv \\ \tilde{W}_t &= \int_t^{t+\delta} \sigma(Y_v) \left[1 + \frac{1}{2} (t+\delta-v)(t-v) (Af)'(Y_v) + \left(\frac{\delta}{2} + t-v \right) f'(Y_v) \right] dW_v \end{aligned}$$

Si de plus, pour $l = 0, 1$, $E \left(\int_0^t [(A^l f)' \sigma]^2(Y_v) dv \right) < \infty$ et $E \left(\int_0^t \sigma^2(Y_v) dv \right) < \infty$, alors $(\tilde{W}_t)_{t \geq 0}$ est une suite de variables de carré intégrable telles que $E \left(\tilde{W}_t \mid \mathcal{F}_t \right) \stackrel{p.s.}{=} 0$.

Démonstration:

La démonstration repose sur le résultat suivant d'interversion de l'ordre d'intégration entre intégrale de Riemann et intégrale stochastique : soit $\left(\int_0^t J_v dW_v \right)_{t \geq 0}$ une martingale locale

continue et g une fonction continue sur R ; alors, on obtient, par application de la formule de Ito pour les martingales locales [10] :

$$\int_t^{t+\delta} \left(\int_v^{t+\delta} g(u) du \right) J_v dW_v = \int_t^{t+\delta} \left(\int_t^u J_v dW_v \right) g(u) du \quad (1.4)$$

$$\int_t^{t+\delta} \left(\int_t^v g(u) du \right) J_v dW_v = \int_t^{t+\delta} \left(\int_u^{t+\delta} J_v dW_v \right) g(u) du \quad (1.5)$$

Posons :

$$\begin{aligned} R_t &= \int_t^{t+\delta} f(Y_v) dv - \frac{\delta}{2} (f(Y_t) + f(Y_{t+\delta})) \\ &= \frac{1}{2} \int_t^{t+\delta} [f(Y_v) - f(Y_t)] dv - \frac{1}{2} \int_t^{t+\delta} [f(Y_{t+\delta}) - f(Y_v)] dv \end{aligned}$$

Par application de la formule de Ito, on obtient :

$$\begin{aligned} R_t &= \frac{1}{2} \int_t^{t+\delta} \left[\int_t^v Af(Y_u) du \right] dv + \frac{1}{2} \int_t^{t+\delta} \left[\int_t^v (f' \sigma)(Y_u) dW_u \right] dv \\ &\quad - \frac{1}{2} \int_t^{t+\delta} \left[\int_v^{t+\delta} Af(Y_u) du \right] dv - \frac{1}{2} \int_t^{t+\delta} \left[\int_v^{t+\delta} (f' \sigma)(Y_u) dW_u \right] dv \end{aligned}$$

Par (1.4) et (1.5), on a :

$$R_t = \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) Af(Y_u) du + \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) (f' \sigma)(Y_u) dW_u$$

Or :

$$\int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) du = 0$$

Donc :

$$\begin{aligned} \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) Af(Y_u) du &= \frac{1}{2} \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) [Af(Y_u) - Af(Y_t)] du \\ &\quad - \frac{1}{2} \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) [Af(Y_{t+\delta}) - Af(Y_u)] du \end{aligned}$$

De manière similaire, on obtient :

$$\begin{aligned} \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) Af(Y_u) du &= \frac{1}{2} \int_t^{t+\delta} (t + \delta - v)(t - v) A^2 f(Y_v) dv \\ &\quad + \frac{1}{2} \int_t^{t+\delta} (t + \delta - v)(t - v) \left((Af)' \sigma \right)(Y_v) dW_v \end{aligned}$$

Ceci permet d'écrire :

$$Y_{t+\delta} = Y_t + \frac{\delta}{2} (f(Y_t) + f(Y_{t+\delta})) + R_t + \int_t^{t+\delta} \sigma(Y_v) dW_v$$

ce qui n'est autre que le schéma du trapèze (1.3).

On constate que l'erreur η_t est en $O(\delta^3)$, ce qui est conforme à l'approximation d'une intégrale ordinaire par la méthode du trapèze. De même, la transposition de la méthode de Simpson à notre problème donne un résultat conforme à celui attendu où l'erreur η_t est un $O(\delta^5)$.

Propriété 2 : Schéma de Simpson

Supposons que $f \in C^8(R)$ et $\sigma \in C^6(R)$. Alors :

$$\forall t \geq 0, Y_{t+2\delta} = Y_t + \frac{\delta}{3} (f(Y_t) + 4f(Y_{t+\delta}) + f(Y_{t+2\delta})) + \eta_t + \tilde{W}_t \quad (1.6)$$

avec:

$$\begin{aligned} \eta_t &= \frac{1}{4!} \int_t^{t+\delta} (v-t)^3 \left(v-t - \frac{4\delta}{3}\right) A^4 f(Y_v) dv \\ &+ \frac{1}{4!} \int_{t+\delta}^{t+2\delta} (t+2\delta-v)^3 \left(t + \frac{2\delta}{3} - v\right) A^4 f(Y_v) dv \\ \tilde{W}_t &= \tilde{W}_t^1 + \tilde{W}_t^2 \\ \tilde{W}_t^1 &= \int_t^{t+\delta} \sigma(Y_v) \left[1 + \sum_{k=0}^3 \frac{(-1)^k (v-t)^k}{(k+1)!} \left(t + \frac{(k+1)\delta}{3} - v\right) (A^k f)'(Y_v) \right] dW_v \\ \tilde{W}_t^2 &= \int_{t+\delta}^{t+2\delta} \sigma(Y_v) \left[1 + \sum_{k=0}^3 \frac{(t+2\delta-v)^k}{(k+1)!} \left(t + 2\delta - \frac{(k+1)\delta}{3} - v\right) (A^k f)'(Y_v) \right] dW_v \end{aligned}$$

Si de plus, pour $l = 0, 3$, $E \left(\int_0^t [(A^l f)' \sigma]^2(Y_v) dv \right) < \infty$ et $E \left(\int_0^t \sigma^2(Y_v) dv \right) < \infty$, alors

$(\tilde{W}_t^1)_{t \geq 0}$ et $(\tilde{W}_t^2)_{t \geq 0}$ sont des suites de variables de carré intégrable qui vérifient $E(\tilde{W}_t^1 | \mathcal{F}_t) \stackrel{p.s.}{=} 0$ et $E(\tilde{W}_t^2 | \mathcal{F}_{t+\delta}) \stackrel{p.s.}{=} 0$

La démonstration de cette propriété est donnée en annexe A.

1.3 Le cas d'une dérive linéaire en θ :

Schéma du trapèze et Estimation par variable instrumentale

Dans la suite, on supposera que la fonction de dérive dépend d'un paramètre $\theta_0 \in R^p$. Pour l'estimation de θ_0 basée sur le schéma d'approximation d'Euler (cf. [5] et annexe F), le contraste employé est celui des moindres carrés. Si nous utilisons le schéma d'approximation du trapèze, le caractère anticipatif de ce schéma conduit à un biais systématique dans la procédure d'estimation par moindres carrés.

Si le modèle est linéaire en θ , une alternative classique utilisée en économétrie est l'estimation par *variable instrumentale* ([1], [7]) : pour un bon choix de *l'instrument*, instrument que l'on substitue alors à la variable explicative corrélée au résidu, elle fournit une estimation convergente là où les moindres carrés conduisent à un biais systématique.

Pour estimer le paramètre d'une diffusion vectorielle gaussienne, $dY_t = \theta_0 Y_t dt + dW_t$, Bergstrom et Sargan ([1], [21]) ont eu l'idée d'utiliser le schéma du trapèze associé à une méthode d'estimation par V.I. Dans ce cas, la dérive $f(\theta, y) = \theta y$ est doublement linéaire, en θ et en y . Nous examinerons ici le cas d'une diffusion pour laquelle la dérive s'écrit $\theta \cdot f(y)$, sans hypothèses de linéarité en y . Il est intéressant de mettre l'accent sur cet exemple à plusieurs titres :

- Il permet, d'une part, de comprendre le choix du contraste qui sera retenu dans le cas d'une dérive générale $f(\theta, y)$, sans condition de linéarité ni en θ ni en y . Comme nous le verrons au paragraphe 4, ce choix (estimation des moments généralisés, G.M.M.) coïncidera avec celui de l'estimation par V.I. dans le cas linéaire en θ .
- D'autre part, dans le cas linéaire, les hypothèses (notées H 1) assurant de bonnes propriétés asymptotiques sont simplifiées par rapport au cas général.
- Enfin, la linéarité permet d'obtenir une expression explicite et linéaire dans les observations de l'estimateur par V.I. Ceci rend alors la méthode numériquement très simple.

1.3.1 La méthode d'estimation par variables instrumentales

Considérons l'e.d.s. définie par :

$$dY_t = \theta_0 \cdot f(Y_t) dt + \sigma(Y_t) dW_t, \quad Y_0 = x \quad (1.7)$$

avec $\theta_0 \cdot f(Y_t) = \sum_{i=1}^p f_i(Y_t) \theta_0^i$. On suppose que f_1, \dots, f_p et σ sont des fonctions connues, définies de R dans R , σ^2 étant de plus strictement positive sur R . Soit Y une solution de (1.7). Pour $i = 1, p$ et $k = 1, n$, on note :

$$X_k^i = \frac{f_i(Y_{k\delta}) + f_i(Y_{(k-1)\delta})}{2\sigma(Y_{(k-1)\delta})}, \quad X_k = (X_k^1, \dots, X_k^p), \quad X = (X_k^i)_{k,i}$$

$$Y_k = \frac{Y_{k\delta} - Y_{(k-1)\delta}}{\delta\sigma(Y_{(k-1)\delta})}, \quad Y = {}^t(Y_1, \dots, Y_n)$$

Dans le schéma d'approximation du trapèze dérivé du schéma exact (1.3), le fait que la variable explicative X_k et le bruit blanc $\tilde{W}_{(k-1)\delta}$ soient corrélés conduit à un biais systématique dans l'estimation des moindres carrés : l'estimateur, $\theta_n = ({}^t X X)^{-1} ({}^t X Y)$, doit être adapté. La méthode d'estimation par V.I. consiste à remplacer la variable explicative X_k par un *instrument* \tilde{X}_k décorrélé de $\tilde{W}_{(k-1)\delta}$, mais bien lié à X_k . Dans le cas présent, il est naturel de choisir $\tilde{X}_k = (\tilde{X}_k^1, \dots, \tilde{X}_k^p)$ avec $\tilde{X}_k^i = \frac{f_i(Y_{(k-1)\delta})}{\sigma(Y_{(k-1)\delta})}$. Comme on le verra, ce choix est bon car il permet de préserver l'efficacité asymptotique à un facteur $(I_p + O(\delta))$ près. L'estimateur par V.I. est défini par (cf. [1], [7]) :

$$\theta_n^I = ({}^t \tilde{X} X)^{-1} ({}^t \tilde{X} Y)$$

avec :

$$\begin{aligned}\tilde{X} &= \left(\tilde{X}_k^i \right)_{k,i} \\ {}^t \tilde{X} X &= \left(\sum_{k=1}^n \tilde{X}_k^i X_k^j \right)_{1 \leq i, j \leq p}, \quad \tilde{X}_k^i X_k^j = \frac{f_i(Y_{(k-1)\delta})}{2\sigma^2(Y_{(k-1)\delta})} [f_j(Y_{(k-1)\delta}) + f_j(Y_{k\delta})] \\ {}^t \tilde{X} Y &= {}^t \left(\sum_{k=1}^n \tilde{X}_k^1 Y_k, \dots, \sum_{k=1}^n \tilde{X}_k^p Y_k \right), \quad \tilde{X}_k^i Y_k = \frac{f_i(Y_{(k-1)\delta})}{\delta\sigma^2(Y_{(k-1)\delta})} [Y_{k\delta} - Y_{(k-1)\delta}]\end{aligned}$$

1.3.2 Résultats asymptotiques

Nous donnerons ici une preuve succincte des résultats du théorème 1. Ceux-ci seront démontrés dans le cadre général au paragraphe 4. On note, $s(y, \theta_0) = \exp\left(-2 \int_{x_0}^y \frac{\theta_0 \cdot f(z)}{\sigma^2(z)} dz\right)$, la dérivée de la fonction d'échelle associée à l'équation (1.7).

Hypothèses H 1 : Schéma du trapèze, dérive linéaire en θ

H 1.1 Pour tout $i = 1, p$, f_i est de classe C^4 sur R et σ est de classe C^2 sur R et vérifie :

$$\forall x \in R, \sigma^2(x) > 0$$

Il existe de plus $K > 0$ telle que : $\forall i = 1, p, \forall x \in R, |f_i(x)| + |\sigma(x)| \leq K(1 + |x|)$. Sous ces conditions, (1.7) admet une unique solution forte $(Y_t)_{t \geq 0}$ non explosive pour toute condition initiale $Y_0 = x$ [10].

H 1.2

$$\int_0^\infty s(x, \theta_0) dx = \int_{-\infty}^0 s(x, \theta_0) dx = \infty \text{ et } \int_{-\infty}^\infty [s(x, \theta_0) \sigma^2(x)]^{-1} dx = C(\theta_0) < \infty$$

$(Y_t)_{t \geq 0}$ est alors récurrente positive sur R pour $\theta = \theta_0$, de loi invariante :

$$\mu_{\theta_0}(dy) = [C(\theta_0) s(y, \theta_0) \sigma^2(y)]^{-1} dy$$

On note P_{θ_0} la loi du processus stationnaire et E_{θ_0} l'espérance sous P_{θ_0} . Si $\pi_{\theta_0}^t(x, dy) = P_{\theta_0}(Y_t \in dy \mid Y_0 = x)$, $Q_{\theta_0}^t = \mu_{\theta_0} \otimes \pi_{\theta_0}^t$ est alors la loi de (Y_0, Y_t) en régime stationnaire.

H 1.3 On note $c_{\theta_0}(x) = \frac{(\theta_0 \cdot f)^2(x)}{2\sigma^2(x)} + \frac{(\theta_0 \cdot f)'(x)}{2} - \frac{\sigma'(x)(\theta_0 \cdot f)(x)}{\sigma(x)} + \frac{(\sigma')^2(x)}{8}$.
On suppose que :

$$\min \left\{ \lim_{x \rightarrow +\infty} c_{\theta_0}(x), \lim_{x \rightarrow -\infty} c_{\theta_0}(x) \right\} = c_{\theta_0} > 0$$

H 1.4 La loi invariante μ_{θ_0} admet des moments de tous ordres et vérifie :

$$\forall r \in N, \int \frac{|x|^r}{\sigma^8(x)} d\mu_{\theta_0}(x) < \infty$$

H 1.5 Pour tout $i = 1, p$, les dérivées successives de f_i jusqu'à l'ordre 4 sont à croissance polynomiale, ainsi que σ' et σ'' .

H 1.6 $H_{\theta_0} = \left(E_{\theta_0} \left[\frac{f_i(Y_0)f_j(Y_0)}{\sigma^2(Y_0)} \right] \right)_{1 \leq i, j \leq p}$ est inversible.

H 1.7 Si, pour tout $i = 1, p$, $E_{\theta_0} \left[f_i \left(\frac{(\theta_0 \cdot f)(Y_0) - (\theta \cdot f)(Y_0)}{\sigma^2(Y_0)} \right) \right] = 0$, alors $\theta = \theta_0$.

Quelques remarques s'imposent :

- Les conditions (H 1.2), (H 1.3) et (H 1.4) ne sont pas spécifiques au caractère linéaire en θ de la fonction de dérive. Elles seront reprises dans l'étude du cas général au paragraphe 4.
- Notons $L_0^2(\mu_{\theta_0}) = \{f \in L^2(\mu_{\theta_0}), \mu_{\theta_0}(f) = 0\}$. D'après [12], sous (H 1.1), (H 1.2) et (H 1.3), il existe $\lambda > 0$ telle que :
$$\forall f \in L_0^2, \|\pi_{\theta_0}^t(f)\|_{L^2(\mu_{\theta_0})} \leq \exp(-\lambda t) \|f\|_{L^2(\mu_{\theta_0})} \quad (1.8)$$
- Si μ_{θ_0} admet des moments de tous ordres, une condition suffisante, pour que l'hypothèse (H 1.4) soit satisfaite, est que $\inf_{x \in R} |\sigma(x)| > 0$ (modèle de Ornstein-Uhlenbeck). Comme nous le verrons lors de l'étude expérimentale, (H 1.4) est vérifiée pour le modèle de Cox-Ingersoll-Ross sous certaines conditions sur θ_0 .
- En fait, il est suffisant que (H 1.4) soit satisfaite pour un entier r assez grand : (H 1.4) et (H 1.5) doivent assurer certaines conditions d'intégrabilité par rapport à Q_{θ_0} indispensables à l'établissement des résultats asymptotiques.

Pour tout $i = 1, p$ et tout $j = 1, p$, on définit les fonctions suivantes :

$$\begin{aligned} M_{i,j}(v) &= E_{\theta_0} \left[\frac{f_i(Y_0) f_j(Y_v)}{\sigma^2(Y_0)} \right], \quad N_i(v) = E_{\theta_0} \left[f_i(Y_0) \frac{[A_{\theta_0}^2(\theta_0 \cdot f)](Y_v)}{\sigma^2(Y_0)} \right] \\ Q_{i,j}(v) &= E_{\theta_0} \left[(f_i f_j)(Y_0) \frac{[\sigma^2(\theta_0 \cdot f)]'(Y_v)}{\sigma^4(Y_0)} \right], \quad P_{i,j}(v) = E_{\theta_0} \left[(f_i f_j)(Y_0) \frac{\sigma^2(Y_v)}{\sigma^4(Y_0)} \right] \end{aligned}$$

Lemme 1 Sous (H 1), N_i est continue sur R^+ et $M_{i,j}$, $Q_{i,j}$ et $P_{i,j}$ sont de classe C^1 sur R_*^+ . On a en particulier :

$$\forall v \in R^+, M'_{i,j}(v) = E_{\theta_0} \left[f_i(Y_0) \frac{(A_{\theta_0} f_j)(Y_v)}{\sigma^2(Y_0)} \right], \quad P'_{i,j}(v) = E_{\theta_0} \left[(f_i f_j)(Y_0) \frac{(A_{\theta_0} \sigma^2)(Y_v)}{\sigma^4(Y_0)} \right]$$

De plus, $M_{i,j}$, $Q_{i,j}$ et $P_{i,j}$ sont dérivables à droite en 0 avec :

$$\lim_{h \rightarrow 0} M'_{i,j}(h) = M'_{i,j}(0^+), \quad \lim_{h \rightarrow 0} P'_{i,j}(h) = P'_{i,j}(0^+), \quad \lim_{h \rightarrow 0} Q'_{i,j}(h) = Q'_{i,j}(0^+).$$

Ce résultat est important car il conditionne l'ensemble des développements effectués par la suite. Nous en donnerons une justification dans le cas général.

On note :

$$M_{\theta_0} = (M'_{i,j}(0^+))_{i,j} = \left(E_{\theta_0} \left[f_i(Y_0) \frac{(A_{\theta_0} f_j)(Y_0)}{\sigma^2(Y_0)} \right] \right)_{i,j} \quad (1.9)$$

$$P_{\theta_0} = (P'_{i,j}(0^+))_{i,j} = \left(E_{\theta_0} \left[(f_i f_j)(Y_0) \frac{(A_{\theta_0} \sigma^2)(Y_0)}{\sigma^4(Y_0)} \right] \right)_{i,j} \quad (1.10)$$

$${}^t N_{\theta_0} = (N_i(0))_i = \left(E_{\theta_0} \left[f_i(Y_0) \frac{[A_{\theta_0}^2(\theta_0 \cdot f)](Y_0)}{\sigma^2(Y_0)} \right] \right)_i \quad (1.11)$$

Sous (H 1), à δ fixé et proche de 0, on va voir que l'estimation par V.I. converge vers une valeur θ_δ , où θ_δ est une fonction explicite de θ_0 et de δ . Pour cela, on considère le système linéaire suivant :

$$\left[{}^t \tilde{X} Y \right]^\delta - \left[{}^t \tilde{X} X \right]^\delta \theta = 0 \quad (1.12)$$

avec :

$$\left[{}^t \tilde{X} Y \right]^\delta = \left(E_{\theta_0} \left[f_i(Y_0) \left(\frac{Y_\delta - Y_0}{\delta \sigma^2(Y_0)} \right) \right] \right)_i \quad \text{et} \quad \left[{}^t \tilde{X} X \right]^\delta = \left(E_{\theta_0} \left[\frac{M_{i,j}(\delta) + M_{i,j}(0)}{2} \right] \right)_{i,j}$$

Lemme 2 *Sous (H 1), il existe $\delta_0 > 0$ tel que pour tout δ , $0 < \delta \leq \delta_0$, (1.12) admet une unique solution :*

$$\theta_\delta = \left(\left[{}^t \tilde{X} X \right]^\delta \right)^{-1} \left[{}^t \tilde{X} Y \right]^\delta$$

Nous démontrerons ce résultat dans le cas général traité dans le paragraphe suivant.

Les propriétés asymptotiques de l'estimateur par V.I. reposent enfin sur les résultats d'ergodicité et de normalité asymptotique suivants (cf. lemme 3.2 de [2]; [5]) :

Lemme 3 *Sous (H 1.1), (H 1.2) et (H 1.3), si $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ vérifie $Q_{\theta_0}^\delta(g^2) < \infty$, alors :*

$$\frac{1}{n} \sum_{k=1}^n g(Y_{k\delta}, Y_{(k-1)\delta}) \xrightarrow{L^2(P_{\theta_0})} Q_{\theta_0}^\delta(g)$$

Si de plus, $Q_{\theta_0}^\delta(g) = 0$, on a :

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n g(Y_{k\delta}, Y_{(k-1)\delta}) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}(0, V_{\theta_0}^\delta)$$

où $V_{\theta_0}^\delta = \int [g^2(x, y) + 2g(x, y) [U_\delta(g)](y)] Q_{\theta_0}^\delta(dx, dy)$ avec $U_\delta(g)(x) = \sum_{k=1}^{+\infty} \pi_{\theta_0}^{k\delta}(g)(x)$.

Théorème 1 *Sous (H 1), pour H_{θ_0} définie en (H 1.6), P_{θ_0} en (1.10) et N_{θ_0} en (1.11), et pour tout δ , $0 < \delta \leq \delta_0$, on a lorsque $n \rightarrow +\infty$:*

$$\theta_n^I = \left({}^t \tilde{X} X \right)^{-1} \left({}^t \tilde{X} Y \right) \xrightarrow{P_{\theta_0}} \theta_\delta \quad \text{et} \quad \sqrt{n\delta} (\theta_n^I - \theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, V_\delta(\theta_0))$$

avec :

$$\begin{aligned} \theta_\delta - \theta_0 &= -\frac{\delta^2}{12} H_{\theta_0}^{-1} N_{\theta_0} + o(\delta^2) \\ V_\delta(\theta_0) &= H_{\theta_0}^{-1} \left(I_p + \frac{\delta}{2} P_{\theta_0} H_{\theta_0}^{-1} + o(\delta^2) \right) \end{aligned}$$

Comme nous l'avons souligné, H_{θ_0} étant l'information de Fisher, l'efficacité asymptotique est préservée à un facteur en $(I_p + O(\delta))$ près.

Démonstration:

1. Par application du lemme 3, lorsque $n \rightarrow \infty$, $\frac{1}{n} {}^t \tilde{X} X \xrightarrow{L^2(P_{\theta_0})} [{}^t \tilde{X} X]^\delta$. Or, d'après le lemme 2, on sait que pour $\delta \in]0, \delta_0]$, $[{}^t \tilde{X} X]^\delta$ est inversible. On en déduit que pour n assez grand, ${}^t \tilde{X} X$ est inversible et donc que $\theta_n^I = ({}^t \tilde{X} X)^{-1} ({}^t \tilde{X} Y)$ existe. Pour montrer la convergence de θ_n^I vers θ_δ , il suffit alors d'appliquer à nouveau le lemme 3 : $\frac{1}{n} {}^t \tilde{X} Y \xrightarrow{L^2(P_{\theta_0})} [{}^t \tilde{X} Y]^\delta$. On obtient alors :

$$\theta_n^I \xrightarrow{P_{\theta_0}} \left([{}^t \tilde{X} X]^\delta \right)^{-1} \left([{}^t \tilde{X} Y]^\delta \right) = \theta_\delta$$

2. L'écart entre θ_δ et θ_0 se calcule en utilisant la décomposition du trapèze (1.3) :

$$\begin{aligned} \theta_\delta &= \left([{}^t \tilde{X} X]^\delta \right)^{-1} \left[[{}^t \tilde{X} X]^\delta \theta_0 + B_\delta \right] = \theta_0 + \left([{}^t \tilde{X} X]^\delta \right)^{-1} B_\delta \\ \text{avec } {}^t B_\delta &= \left(E_{\theta_0} \left[f_i(Y_0) \frac{\eta_0}{\delta \sigma^2(Y_0)} \right] \right)_i \end{aligned}$$

Par application du lemme 1, $B_\delta = -\frac{\delta^2}{12} N_{\theta_0} + o(\delta^2)$ et $[{}^t \tilde{X} X]^\delta = H_{\theta_0} + \frac{\delta}{2} M_{\theta_0} + o(\delta)$, M_{θ_0} étant définie en (1.9). Le développement de B_δ sera donné dans le cas général lors de la démonstration du théorème 3, celui de $[{}^t \tilde{X} X]^\delta$ dans le lemme 7. On a donc le résultat annoncé pour le biais asymptotique.

3. On utilise à nouveau le lemme 3 pour prouver la normalité asymptotique. Celle-ci résulte directement des résultats suivants :

$$\sqrt{n\delta} (\theta_n^I - \theta_\delta) = \left(\frac{1}{n} {}^t \tilde{X} X \right)^{-1} \sqrt{\frac{\delta}{n}} ({}^t \tilde{X} Y - {}^t \tilde{X} X \theta_\delta)$$

avec :

$$\sqrt{\frac{\delta}{n}} ({}^t \tilde{X} Y - {}^t \tilde{X} X \theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, \Gamma_\delta(\theta_\delta)) \text{ et } \frac{1}{n} {}^t \tilde{X} X \xrightarrow{P_{\theta_0}} [{}^t \tilde{X} X]^\delta$$

On obtient donc :

$$\sqrt{n\delta} (\theta_n^I - \theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, V_\delta(\theta_0)), \quad V_\delta(\theta_0) = \left([{}^t \tilde{X} X]^\delta \right)^{-1} \Gamma_\delta(\theta_\delta) \left([{}^t \tilde{X} X]^\delta \right)^{-1}$$

Or, $\Gamma_\delta(\theta_\delta) = H_{\theta_0} + \frac{\delta}{2} P_{\theta_0} + o(\delta)$ (cf. lemme 8) et $[{}^t \tilde{X} X]^\delta = H_{\theta_0} + \frac{\delta}{2} M_{\theta_0} + o(\delta)$. On a ainsi le développement annoncé pour $V_\delta(\theta_0)$.

1.4 Le cas d'une dérive générale :

Schéma du trapèze et Estimation des Moments généralisés

Ce paragraphe est consacré à l'étude du cas général, sans nécessité de linéarité en θ pour la fonction de dérive. Dans la suite, on notera $f_\theta(x) = f(\theta, x)$.

Soient U un ouvert de R^p et V un ouvert de R . On dit que $(\theta, x) \rightarrow g(\theta, x)$ est de classe $C^{m,k}$ sur $U \times V$ si :

1. $(\theta, x) \rightarrow g(\theta, x)$ est continue sur $U \times V$.
2. Pour tout $x \in V$, $\theta \rightarrow g(\theta, x)$ admet des dérivées partielles jusqu'à l'ordre m en tout point de U et ses dérivées partielles sont continues sur $U \times V$. Pour tout $i = 1, p$, on note $\frac{\partial}{\partial \theta^i} g(\theta, x)$ la $i^{\text{ème}}$ dérivée partielle de $g(\cdot, x)$ au point θ .
3. Pour tout $\theta \in U$, $x \rightarrow g(\theta, x)$ est dérivable jusqu'à l'ordre k sur V et ses dérivées successives sont continues sur $U \times V$. On note alors $g'(\theta, x)$, $g''(\theta, x)$ et $g^{(l)}(\theta, x)$, les dérivées à l'ordre 1, 2 et l de $g(\theta, \cdot)$ au point x .

On dit aussi que $(\theta, x) \rightarrow g(\theta, x)$ est de classe $C^{m,1}$ sur $U \times R^+$ si :

- Les points 1 et 2 de la définition précédente sont vérifiés en remplaçant V par R^+ .
- Pour tout $\theta \in U$, $x \rightarrow g(\theta, x)$ est dérivable sur R_*^+ et dérivable à droite en 0. La fonction $(\theta, x) \rightarrow g'(\theta, x)$ est continue sur $U \times R_*^+$. De plus, pour tout $\theta \in U$ et toute suite $((\theta_n, x_n))_n \subset U \times R_*^+$ telle que $(\theta_n, x_n) \rightarrow (\theta, 0)$, on a : $\lim_{n \rightarrow \infty} g'(\theta_n, x_n) = g'(\theta, 0^+)$, $g'(\theta, 0^+)$ étant la dérivée à droite de $g(\theta, \cdot)$ en 0.

On va utiliser la méthode d'estimation des moments généralisés (G.M.M.) ([9],[8]; [7]). Cette méthode généralise au cas non linéaire la méthode des V.I.

Notons :

$$\mathcal{V}_\delta = {}^t (\mathcal{V}_\delta^1, \dots, \mathcal{V}_\delta^p), \quad \mathcal{V}_\delta^i(y, x, \theta) = \frac{\partial}{\partial \theta^i} f_\theta(x) \left(\frac{y - x - \frac{\delta}{2} (f_\theta(x) + f_\theta(y))}{\delta \sigma^2(x)} \right) \quad (1.13)$$

$$V_n(\theta) = \frac{1}{n} \sum_{k=1}^n \mathcal{V}_\delta(Y_{k\delta}, Y_{(k-1)\delta}, \theta), \quad V_n^i(\theta) = \frac{1}{n} \mathcal{V}_\delta^i(Y_{k\delta}, Y_{(k-1)\delta}, \theta) \quad (1.14)$$

L'estimateur des moments généralisés, $\hat{\theta}_n$, associé à \mathcal{V}_δ et à la matrice de poids l'identité de dimension p , est défini par :

$$\hat{\theta}_n = \inf_{\theta \in \Theta} U_n(\theta), \quad U_n(\theta) = {}^t V_n(\theta) I_p V_n(\theta) = \|V_n(\theta)\|_2^2$$

On vérifie aisément que θ_n^I et $\hat{\theta}_n$ coïncident dans le cas d'un modèle linéaire puisqu'alors $U_n(\theta_n^I) = 0$. Nous allons montrer que $\hat{\theta}_n$ converge vers une limite θ_δ que l'on peut caractériser, et est asymptotiquement normal et efficace à un facteur $(I_p + O(\delta))$ près. Dans le cas des modèles linéaires, pour δ petit, θ_δ est l'unique solution explicite d'un certain système linéaire. Dans le cas général, nous allons montrer que, lorsque δ est petit, θ_δ est l'unique solution de l'équation $E_{\theta_0} [\mathcal{V}_\delta(Y_\delta, Y_0, \theta)] = 0$.

1.4.1 Propriétés asymptotiques

Soit Θ un sous ensemble **compact** de R^p tel que $\theta_0 \in \overset{\circ}{\Theta}$. Soit $U(\Theta)$ un ouvert de R^p tel que $\Theta \subset U(\Theta)$.

Hypothèses H 2 : Schéma du trapèze et G.M.M.

H 2.1 f_{θ_0} est de classe C^4 sur R , σ est de classe C^2 sur R et vérifie :

$$\forall x \in R, \sigma^2(x) > 0$$

De plus, il existe $K > 0$ telle que : $\forall x \in R, |f_{\theta_0}(x)| + |\sigma(x)| \leq K(1 + |x|)$.

H 2.2 -3 -4 On maintient **(H 1.2)**, **(H 1.3)** et **(H 1.4)** en remplaçant le cas échéant $\theta_0 \cdot f$ par f_{θ_0} .

H 2.5 1. Pour tout $x \in R$, $\theta \rightarrow f_{\theta}(x)$ est de classe C^4 sur $U(\Theta)$. Cette fonction et ses dérivées partielles successives jusqu'à l'ordre 4 sont à croissance polynomiale en x , uniformément en θ sur $U(\Theta)$.

De plus, $f_{\theta}(x)$, $\frac{\partial}{\partial \theta^i} f_{\theta}(x)$ et $\frac{\partial^2}{\partial \theta^i \partial \theta^j} f_{\theta}(x)$ sont continues sur $U(\Theta) \times R$.

2. Pour $\theta \in U(\Theta)$, $x \rightarrow f_{\theta}(x)$ est deux fois dérivable sur R . La fonction $(\theta, x) \rightarrow A_{\theta_0} f_{\theta}(x)$ est continue sur $U(\Theta) \times R$ et est à croissance polynomiale en x , uniformément en θ sur $U(\Theta)$.

De plus, pour tout $x \in R$, $\theta \rightarrow A_{\theta_0} f_{\theta}(x)$ est de classe C^1 sur $U(\Theta)$.

3. Pour tout $i = 1, p$ et tout $\theta \in U(\Theta)$, $x \rightarrow \frac{\partial}{\partial \theta^i} f_{\theta_0}(x)$ est de classe C^2 sur R .

De plus, $x \rightarrow [A_{\theta_0} (\frac{\partial}{\partial \theta^i} f_{\theta_0})](x)$ est à croissance polynomiale en x .

4. Les fonctions f'_{θ_0} , $(A_{\theta_0} f_{\theta_0})'$, $A_{\theta_0}^2 f_{\theta_0}$ et $A_{\theta_0}(\sigma^2 f'_{\theta_0})$ sont à croissance polynomiale en x , ainsi que σ' et σ'' .

H 2.6 -7 (H 1.6) et **(H 1.7)** sont maintenues en remplaçant dans la définition de H_{θ_0} , f_i par $\frac{\partial}{\partial \theta^i} f_{\theta_0}$.

Par souci de simplicité, nous avons supposé que (H 2.1) est vérifiée sur R . Il se peut cependant que ces hypothèses ne soient satisfaites que sur un intervalle de R , $I =]l, r[$, c'est le cas du modèle de Cox-Ingersoll-Ross dont nous détaillons l'étude dans le paragraphe 5.

Dans tous les cas, la continuité des fonctions et le fait que $\sigma^2 > 0$ assurent, pour toute condition initiale $x \in I$, l'existence et l'unicité en loi d'une solution de (1.1) avant un temps d'explosion ([10], théorème 5.15, p.341). Dans le cas où $I = R$, l'hypothèse de croissance polynomiale faite sur f_{θ_0} et σ rend la solution faible non explosive ([10], problème 5.3.15). Par contre, si I est strictement contenu dans R , ce résultat s'obtient en supposant que $\lim_{y \rightarrow l^+} \int_y^{x_0} s(x, \theta_0) dx = -\infty$ et $\lim_{y \rightarrow r^-} \int_{x_0}^y s(x, \theta_0) dx = +\infty$ ([10], proposition 5.5.22).

Bien que l'existence et l'unicité d'une solution faible non explosive soient suffisantes, nous obtenons l'existence et l'unicité d'une solution forte non explosive. Pour établir le schéma du trapèze, nous supposons en effet que f_{θ_0} et σ sont respectivement de classe C^4 et C^2 sur R : elles vérifient donc une condition de lipschitz locale sur R . En se référant au théorème 5.2.5 et au corollaire 5.3.23 de [10], (1.1) admet alors une unique solution forte pour toute condition initiale $x \in R$.

On note pour tout $i = 1, p$:

$$\forall \delta > 0, V_i(\theta, \delta) = E_{\theta_0} [\mathcal{V}_{\delta}^i(Y_{\delta}, Y_0, \theta)] \quad (1.15)$$

$$V_i(\theta, 0) = E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] \quad (1.16)$$

Pour tout $i = 1, p$ et tout $j = 1, p$, on redéfinit les fonctions $M_{i,j}$, $Q_{i,j}$, $P_{i,j}$ et N_i introduites lors du paragraphe précédent en remplaçant $\theta_0 \cdot f$ par f_{θ_0} et f_i par $\frac{\partial}{\partial \theta^i} f_{\theta_0}$:

$$\begin{aligned} M_{i,j}(v) &= E_{\theta_0} \left[\frac{\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_v)}{\sigma^2(Y_0)} \right], \quad N_i(v) = E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_v)}{\sigma^2(Y_0)} \right] \\ Q_{i,j}(v) &= E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \frac{(\sigma^2 f'_{\theta_0})(Y_v)}{\sigma^4(Y_0)} \right] \\ P_{i,j}(v) &= E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \frac{\sigma^2(Y_v)}{\sigma^4(Y_0)} \right] \end{aligned}$$

Nous commencerons par énoncer un lemme technique qui permet d'établir la régularité de ces fonctions (cf. lemme 1) et celle de $(\theta, \delta) \rightarrow V_i(\theta, \delta)$ pour tout $i = 1, p$ (cf. lemme 5). Le contrôle de la régularité des V_i est indispensable pour la caractérisation de θ_δ : en utilisant le lemme 5, nous allons montrer que pour δ fixé et petit, il existe un unique θ_δ dans $\overset{\circ}{\Theta}$ tel que pour tout $i = 1, p$, $V_i(\theta_\delta, \delta) = 0$ (cf. lemme 6).

Lemme 4 Soient $(\theta, x) \rightarrow h(\theta, x)$ et $(\theta, x) \rightarrow g(\theta, x)$ deux fonctions définies sur $U(\Theta) \times R$, à croissance polynomiale en x uniformément en θ sur $U(\Theta)$ et telles que : g est continue sur $U(\Theta) \times R$ et, pour tout $x \in R$, $\theta \rightarrow h(\theta, x)$ est continue sur $U(\Theta)$.

Sous (H 2.1), (H 2.2) et si μ_{θ_0} admet des moments de tous ordres, la fonction définie sur $U(\Theta) \times R^+$ par $(\theta, \delta) \rightarrow m(\theta, \delta) = E_{\theta_0} [h(\theta, Y_0) g(\theta, Y_\delta)]$ est continue sur $U(\Theta) \times R^+$.

Si on suppose de plus que, pour tout $x \in R$, $\theta \rightarrow h(\theta, x)$ est de classe C^1 sur $U(\Theta)$, que g est de classe $C^{1,2}$ sur $U(\Theta) \times R$ et que leurs dérivées partielles, ainsi que $A_{\theta_0} g$, sont à croissance polynomiale en x uniformément en θ sur $U(\Theta)$, alors, $(\theta, \delta) \rightarrow m(\theta, \delta)$ est de classe $C^{1,1}$ sur $U(\Theta) \times R^+$ avec :

$$\begin{aligned} \frac{\partial}{\partial \theta^i} m(\theta, \delta) &= E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} h(\theta, Y_0) g(\theta, Y_\delta) + h(\theta, Y_0) \frac{\partial}{\partial \theta^i} g(\theta, Y_\delta) \right] \\ m'(\theta, \delta) &= E_{\theta_0} [h(\theta, Y_0) (A_{\theta_0} g)(\theta, Y_\delta)] \end{aligned}$$

La démonstration de ce résultat est donnée en annexe B. Le lemme 1 est donc une conséquence de ce résultat. Notons :

$$M_{\theta_0} = \left(E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{(A_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] \right)_{i,j} \quad (1.17)$$

$$P_{\theta_0} = \left(E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \frac{(A_{\theta_0} \sigma^2)(Y_0)}{\sigma^4(Y_0)} \right] \right)_{i,j} \quad (1.18)$$

$${}^t N_{\theta_0} = \left(E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] \right)_i \quad (1.19)$$

Outre la régularité de $(\theta, \delta) \rightarrow V_i(\theta, \delta)$, le lemme qui suit permet d'établir deux développements qui seront utilisés ultérieurement.

Lemme 5 Sous (H 2), pour tout $i = 1, p$, $(\theta, \delta) \rightarrow V_i(\theta, \delta)$ est de classe $C^{2,1}$ sur $U(\Theta) \times R^+$ et on a :

$$V_i(\theta_0, \delta) = -\frac{\delta^2}{12} N_i(0) + o(\delta^2) \quad (1.20)$$

$$\frac{\partial}{\partial \theta^j} V_i(\theta_0, \delta) = -\frac{1}{2} (M_{i,j}(\delta) + M_{i,j}(0)) + O(\delta^2) \quad (1.21)$$

De plus, $\theta \rightarrow V_i'(\theta, 0^+)$ est de classe C^1 sur $U(\Theta)$.

La démonstration de ce lemme est donnée en annexe C. Montrons alors que le cadre choisi s'inscrit dans celui de l'estimation par minimum de contraste comme défini dans [3] (définition 3.2.7.).

Lemme 6 *Sous (H 2), il existe $\delta_0 > 0$ tel que pour tout δ , $0 < \delta \leq \delta_0$, $U(\theta, \delta) = \sum_{i=1}^p (V_i(\theta, \delta))^2$ admet un unique minimum $\theta_\delta = \phi(\delta) \in \overset{\circ}{\Theta}$ qui vérifie $U(\theta_\delta, \delta) = 0$ et $\lim_{\delta \rightarrow 0} \phi(\delta) = \theta_0$. $U(\theta, \delta)$ est la fonction de contraste associée au processus de contraste $U_n(\theta)$.*

Dans le cas linéaire, cette caractérisation de θ_δ correspond à celle du lemme 2. Il suffit en effet de remarquer que pour une dérive linéaire en θ , on a :

$$V(\theta, \delta) = {}^t(V_1(\theta, \delta), \dots, V_p(\theta, \delta)) = [{}^t\tilde{X}Y]^\delta - [{}^t\tilde{X}X]^\delta \theta$$

Démonstration :

1. Pour tout $\theta \in U(\Theta)$, $U_n(\theta)$ est $\mathcal{F}_{n\delta}$ -mesurable et $U(\theta, \delta)$ est positive.
2. Montrons que, pour tout $\theta \in U(\Theta)$, $U_n(\theta) \xrightarrow{P_{\theta_0}} U(\theta, \delta)$. Les hypothèses (H 2.1), (H 2.4) et (H 2.5) impliquent que $\mathcal{V}_\delta^i(y, x, \theta) \in L^2(Q_{\theta_0}^\delta)$. Sous (H 2), par application du lemme 3, on obtient pour tout $i = 1, p$:

$$\forall \theta \in U(\Theta), V_n^i(\theta) \xrightarrow{L^2(P_{\theta_0})} V_i(\theta, \delta)$$

D'où :

$$\forall \theta \in \Theta, U_n(\theta) \xrightarrow{L^1(P_{\theta_0})} U(\theta, \delta)$$

3. Il nous reste à montrer que pour δ petit, $U(\cdot, \delta)$ admet un unique minimum θ_δ dans $\overset{\circ}{\Theta}$. Soit, pour tout $i = 1, p$, \bar{V}_i la fonction définie sur $U(\Theta) \times R$ par :

$$\begin{aligned} \forall \theta \in U(\Theta), \text{ si } \delta > 0, \bar{V}_i(\theta, \delta) &= V_i(\theta, \delta) \\ \text{si } \delta \leq 0, \bar{V}_i(\theta, \delta) &= V_i(\theta, 0) + \delta V_i'(\theta, 0^+) \end{aligned}$$

En utilisant le lemme 5, on montre que pour tout $i = 1, p$, \bar{V}_i est de classe C^1 sur $U(\Theta) \times R$. Notons \bar{V} la fonction définie de $U(\Theta) \times R$ dans R^p par $\bar{V}(\theta, \delta) = {}^t(\bar{V}_1(\theta, \delta), \dots, \bar{V}_p(\theta, \delta))$: \bar{V} est de classe C^1 sur $\overset{\circ}{\Theta} \times R$ avec $\bar{V}(\theta_0, 0) = V(\theta_0, 0) = 0$, θ_0 étant unique (H 2.7). De plus, si on note $D_\theta \bar{V}_0$ la différentielle de $\theta \rightarrow \bar{V}(\theta, 0)$, on a : $D_\theta \bar{V}_0(\theta_0) = -H_{\theta_0}$. Or, d'après (H 2.6), H_{θ_0} est inversible. Par application du *théorème des fonctions implicites*, il existe $\varepsilon > 0$ et $\delta_0 > 0$ tels que pour tout δ , $|\delta| < \delta_0$, il existe un unique θ_δ dans la boule ouverte, $B(\theta_0, \varepsilon) \subset \overset{\circ}{\Theta}$, qui vérifie $\bar{V}(\theta_\delta, \delta) = 0$ et $\theta_\delta = \phi(\delta)$. ϕ est de plus dérivable en 0. En particulier pour tout δ , $0 < \delta < \delta_0$, il existe un unique $\theta_\delta \in \overset{\circ}{\Theta}$ tel que $U(\theta_\delta, \delta) = 0$ et $\theta_\delta = \phi(\delta)$ avec $\lim_{\delta \rightarrow 0} \phi(\delta) = \theta_0$.

$U_n(\theta)$ et $U(\theta, \delta)$ vérifient donc bien les conditions de la définition 3.2.7. de [3].

Etude de la convergence de $\widehat{\theta}_n$

Théorème 2 Sous (H 2), pour tout δ , $0 < \delta \leq \delta_0$, $\widehat{\theta}_n \xrightarrow{P_{\theta_0}} \theta_\delta$

Démonstration :

La démonstration repose sur la vérification des hypothèses du théorème 3.2.8. de [3].

Sous (H 2), à δ fixé, $0 < \delta \leq \delta_0$, $\theta \rightarrow V_i(\theta, \delta)$ et $\theta \rightarrow V_n^i(\theta)$ sont de classe C^2 sur $U(\Theta)$; il en est donc de même pour $U_n(\theta)$ et $U(\theta, \delta)$.

Une condition suffisante pour obtenir le point 2 du théorème 3.2.8. est qu'il existe une suite de variables aléatoires (Z_n) telle que :

1. pour tout $\eta > 0$, pour tout $(\theta, \beta) \in \Theta^2$ tels $\|\theta - \beta\|_2 \leq \eta$,
 $|U_n(\theta) - U_n(\beta)| \leq Z_n \eta$.
2. $Z_n \xrightarrow{P_{\theta_0}} Z$, Z étant une constante indépendante de η .

Montrons que tel est le cas. Pour tout (θ, β) tels que $\|\theta - \beta\|_2 \leq \eta$, on a :

$$|U_n(\theta) - U_n(\beta)| \leq \sum_{i=1}^p \left| (V_n^i(\theta))^2 - (V_n^i(\beta))^2 \right| \leq 2p \sum_{i=1}^p \left[\left(\sum_{j=1}^p D_n^{i,j} \right) C_n^i \right] \eta$$

avec :

$$D_n^{i,j} = \frac{1}{n} \sum_{k=1}^n \sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta^j} \mathcal{V}_\delta^i(Y_{k\delta}, Y_{(k-1)\delta}, \theta) \right| \quad \text{et} \quad C_n^i = \frac{1}{n} \sum_{k=1}^n \sup_{\theta \in \Theta} |\mathcal{V}_\delta^i(Y_{k\delta}, Y_{(k-1)\delta}, \theta)|$$

Or, sous (H 2.1), (H 2.4) et (H 2.5), on a :

$$(y, x) \rightarrow \sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta^j} \mathcal{V}_\delta^i(y, x, \theta) \right| \in L^2(Q_{\theta_0}^\delta) \quad \text{et} \quad (y, x) \rightarrow \sup_{\theta \in \Theta} |\mathcal{V}_\delta^i(y, x, \theta)| \in L^2(Q_{\theta_0}^\delta)$$

Par application du lemme 3, on obtient :

$$D_n^{i,j} \xrightarrow{L^2(P_{\theta_0})} D^{i,j} = E_{\theta_0} \left[\sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta^j} \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta) \right| \right]$$

$$C_n^i \xrightarrow{L^2(P_{\theta_0})} C^i = E_{\theta_0} \left[\sup_{\theta \in \Theta} |\mathcal{V}_\delta^i(Y_\delta, Y_0, \theta)| \right]$$

où C^i et $D^{i,j}$ sont des constantes indépendantes de θ . Pour Z_n et Z définies par :

$$Z_n = 2p \sum_{i=1}^p \left[\left(\sum_{j=1}^p D_n^{i,j} \right) C_n^i \right] \xrightarrow{L^1(P_{\theta_0})} Z = 2p \sum_{i=1}^p \left[\left(\sum_{j=1}^p D^{i,j} \right) C^i \right]$$

et Θ étant supposé compact, les conditions du théorème 3.2.8. sont donc bien satisfaites.

Evaluation de l'écart $\theta_\delta - \theta_0$

Théorème 3 Sous (H 2), pour H_{θ_0} définie en (H 2.6) et N_{θ_0} en (1.19) et pour δ , $0 < \delta \leq \delta_0$, on a :

$$\theta_\delta = \theta_0 + B_{\theta_0} \delta^2 + o(\delta^2) \text{ avec } B_{\theta_0} = -\frac{1}{12} H_{\theta_0}^{-1} N_{\theta_0}$$

Démonstration :

Sous (H 2), à δ fixé, $0 < \delta \leq \delta_0$, $\theta \rightarrow V_i(\theta, \delta)$ est de classe C^2 sur $U(\Theta)$. De plus $V_i(\theta_\delta, \delta) = 0$. Par application du théorème des accroissements finis, il existe $t_i \in]0, 1[$ tel que $\theta_{\delta,i} = t_i \theta_0 + (1 - t_i) \theta_\delta$ vérifie :

$$V_i(\theta_\delta, \delta) = 0 = V_i(\theta_0, \delta) + \sum_{j=1}^p \frac{\partial}{\partial \theta^j} V_i(\theta_{\delta,i}, \delta) (\theta_\delta^j - \theta_0^j) \quad (1.22)$$

D'après le lemme 6, lorsque $\delta \rightarrow 0$, $\theta_\delta \rightarrow \theta_0$ et donc $\theta_{\delta,i} \rightarrow \theta_0$. En utilisant le lemme 5, on a lorsque $\delta \rightarrow 0$:

$$\frac{\partial}{\partial \theta^j} V_i(\theta_{\delta,i}, \delta) \rightarrow \frac{\partial}{\partial \theta^j} V_i(\theta_0, 0) = -M_{i,j}(0)$$

On peut donc écrire :

$$\frac{\partial}{\partial \theta^j} V_i(\theta_{\delta,i}, \delta) = -M_{i,j}(0) + o(1)$$

D'après (1.20), (1.22) s'écrit alors :

$$0 = -\frac{\delta^2}{12} N_{\theta_0} - (H_{\theta_0} + o(1)) (\theta_\delta - \theta_0) + o(\delta^2)$$

L'hypothèse (H 2.6) conduit alors au résultat annoncé.

Normalité asymptotique de $\hat{\theta}_n$

On note pour tout δ , $0 < \delta \leq \delta_0$:

$$H_{\theta_0}^\delta = \left(\frac{\partial}{\partial \theta^j} V_i(\theta_\delta, \delta) \right)_{i,j} \quad (1.23)$$

$$\Gamma_\delta(\theta_\delta) = T_\delta(\theta_\delta) + 2R_\delta(\theta_\delta) \quad (1.24)$$

avec :

$$T_\delta(\theta_\delta) = \left(E_{\theta_0} \left[\delta \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta_\delta) \mathcal{V}_\delta^j(Y_\delta, Y_0, \theta_\delta) \right] \right)_{i,j}$$

et, si on note $\mathcal{V}_\delta^j(\theta_\delta) = \mathcal{V}_\delta^j(\cdot, \cdot, \theta_\delta)$

$$R_\delta(\theta_\delta) = \left(E_{\theta_0} \left[\delta \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta_\delta) G_\delta^j(Y_\delta) \right] \right)_{i,j}, \quad G_\delta^j = U_\delta \left(\mathcal{V}_\delta^j(\theta_\delta) \right)$$

U_δ étant définie dans le lemme 3.

Notons que dans le cas des modèles linéaires en θ , $H_{\theta_0}^\delta$ et $-\left[{}^t \tilde{X} X \right]^\delta$ coïncident.

Théorème 4 Sous (H 2), pour H_{θ_0} définie en (H 2.6) et P_{θ_0} en (1.18), il existe δ_1 , $0 < \delta_1 \leq \delta_0$, tel que pour tout δ , $0 < \delta \leq \delta_1$, on a lorsque $n \rightarrow \infty$:

$$\sqrt{n\delta} \left(\hat{\theta}_n - \theta_\delta \right) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, V_\delta(\theta_0)), \quad V_\delta(\theta_0) = H_{\theta_0}^{-1} \left(I_p + \frac{\delta}{2} P_{\theta_0} H_{\theta_0}^{-1} + o(\delta) \right)$$

H_{θ_0} étant l'information de Fisher asymptotique, $\hat{\theta}_n$ est un estimateur asymptotiquement efficace de θ_δ à un facteur $(I_p + O(\delta))$ près.

Démonstration:

Sous (H 2), pour δ fixé, $0 < \delta \leq \delta_0$, on sait que $\theta_\delta \in \overset{\circ}{\Theta}$ et que $U_n(\theta)$ est de classe C^3 sur $U(\Theta)$. Notons pour tout $\theta \in \Theta$:

$$DU_n(\theta) = 2 {}^t(DV_n(\theta)) V_n(\theta), \quad DV_n(\theta) = \left(\frac{\partial}{\partial \theta^j} V_n^i(\theta) \right)_{i,j}$$

$$D^2U_n(\theta) = \left(2 \sum_{i=1}^p \left(\frac{\partial^2}{\partial \theta^k \partial \theta^j} V_n^i(\theta) \right) V_n^i(\theta) + 2 \sum_{i=1}^p \left(\frac{\partial}{\partial \theta^j} V_n^i(\theta) \right) \left(\frac{\partial}{\partial \theta^k} V_n^i(\theta) \right) \right)_{j,k}$$

$$R_n = \int_0^1 \left(D^2U_n(\theta_\delta + s(\hat{\theta}_n - \theta_\delta)) - D^2U_n(\theta_\delta) \right) ds$$

L'application de la formule de Taylor avec reste intégral donne :

$$\sqrt{n\delta} DU_n(\hat{\theta}_n) = 0 = \sqrt{n\delta} DU_n(\theta_\delta) + (D^2U_n(\theta_\delta) + R_n) \sqrt{n\delta} (\hat{\theta}_n - \theta_\delta)$$

Pour $H_{\theta_0}^\delta$ définie en (1.23) et $\Gamma_\delta(\theta_\delta)$ en (1.24), la normalité asymptotique de $\hat{\theta}_n$ résultera des trois propriétés suivantes :

$$\sqrt{n\delta} DU_n(\theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, K(\theta_\delta)), \quad K(\theta_\delta) = 4 {}^t H_{\theta_0}^\delta \Gamma_\delta(\theta_\delta) H_{\theta_0}^\delta \quad (1.25)$$

$$D^2U_n(\theta_\delta) \xrightarrow{P_{\theta_0}} 2 {}^t H_{\theta_0}^\delta H_{\theta_0}^\delta \quad (1.26)$$

où $H_{\theta_0}^\delta$ est inversible, et :

$$R_n \xrightarrow{P_{\theta_0}} 0 \quad (1.27)$$

1. Montrons (1.25). Il suffit de vérifier que :

$$\sqrt{n\delta} V_n(\theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, \Gamma_\delta(\theta_\delta)) \quad \text{et} \quad {}^t DV_n(\theta_\delta) \xrightarrow{L^2(P_{\theta_0})} {}^t H_{\theta_0}^\delta$$

Sous (H 2.1), (H 2.4) et (H 2.5), pour tout $T \in R^p$, ${}^t T \mathcal{V}_\delta(y, x, \theta_\delta) \in L^2(Q_{\theta_0}^\delta)$. Par application, du lemme 3, on obtient :

$${}^t T \sqrt{n\delta} V_n(\theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}(0, {}^t T \Gamma_\delta(\theta_\delta) T)$$

Donc :

$$\sqrt{n\delta}V_n(\theta_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, \Gamma_\delta(\theta_\delta))$$

La convergence en probabilité de ${}^tDV_n(\theta_\delta) = \left(\frac{1}{n} \sum_{k=1}^n \frac{\partial}{\partial \theta^j} \mathcal{V}_\delta^i(Y_{k\delta}, Y_{(k-1)\delta}, \theta_\delta)\right)_{j,i}$ est obtenue en appliquant le lemme 3. En effet, sous (H 2.1), (H 2.4) et (H 2.5), $(y, x, \theta_\delta) \rightarrow \frac{\partial}{\partial \theta^j} \mathcal{V}_\delta^i(y, x, \theta_\delta) \in L^2(Q_{\theta_0}^\delta)$. On a alors :

$${}^tDV_n(\theta_\delta) \xrightarrow{L^2(P_{\theta_0})} \left(\frac{\partial}{\partial \theta^j} V_i(\theta_\delta, \delta)\right)_{j,i} = {}^tH_{\theta_0}^\delta$$

2. Montrons (1.26). On déduit de ce qui précède que :

$${}^t(DV_n(\theta_\delta))DV_n(\theta_\delta) = \left(\sum_{i=1}^p \left(\frac{\partial}{\partial \theta^j} V_n^i \frac{\partial}{\partial \theta^k} V_n^i\right)(\theta_\delta)\right)_{j,k} \xrightarrow{L^1(P_{\theta_0})} {}^tH_{\theta_0}^\delta H_{\theta_0}^\delta \quad (1.28)$$

D'autre part, (H 2.1) (H 2.4) et (H 2.5) impliquent que, pour tout (j, k) , $(y, x) \rightarrow \frac{\partial^2}{\partial \theta^k \partial \theta^j} \mathcal{V}_\delta^i(y, x, \theta_\delta) \in L^2(Q_{\theta_0}^\delta)$. Appliquant le lemme 3, on obtient :

$$\frac{\partial^2}{\partial \theta^k \partial \theta^j} V_n^i(\theta_\delta) \xrightarrow{L^2(P_{\theta_0})} E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^k \partial \theta^j} \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta_\delta) \right] < \infty$$

En outre, $V_n^i(\theta_\delta) \xrightarrow{L^2(P_{\theta_0})} V_i(\theta_\delta) = 0$. On a donc :

$$\sum_{i=1}^p \left(\frac{\partial^2}{\partial \theta^k \partial \theta^j} V_n^i(\theta_\delta)\right) V_n^i(\theta_\delta) \xrightarrow{L^1(P_{\theta_0})} 0 \quad (1.29)$$

(1.26) est alors une conséquence de (1.28) et (1.29). De plus :

Lemme 7 *Sous (H 2), pour H_{θ_0} définie en (H 2.6) et M_{θ_0} définie en (1.17), on a :*

$$\forall \delta, 0 < \delta \leq \delta_0, H_{\theta_0}^\delta = -H_{\theta_0} - \frac{\delta}{2} M_{\theta_0} + o(\delta)$$

De plus, il existe $\delta_1, 0 < \delta_1 \leq \delta_0$, tel que pour tout $\delta, 0 < \delta \leq \delta_1$, $H_{\theta_0}^\delta$ est inversible.

La démonstration de ce lemme est donnée en annexe D.

3. Il nous reste à établir (1.27). Par application de la formule de Rolle, on a pour tout i et tout j :

$$\left| \int_0^1 \left[\frac{\partial^2}{\partial \theta^i \partial \theta^j} U_n(\theta_\delta + s(\hat{\theta}_n - \theta_\delta)) - \frac{\partial^2}{\partial \theta^i \partial \theta^j} U_n(\theta_\delta) \right] ds \right| \leq \sum_{k=1}^p D_n^{i,j,k} \left| \hat{\theta}_n^k - \theta_\delta^k \right|$$

avec :

$$D_n^{i,j,k} = \sup_{\theta \in \Theta} \left| \frac{\partial^3}{\partial \theta^k \partial \theta^j \partial \theta^i} U_n(\theta) \right|$$

Sous (H 2.1) (H 2.4) et (H 2.5), $D_n^{i,j,k}$ est une variable positive, indépendante de θ qui, d'après le lemme 3, converge en probabilité vers une constante finie. La convergence en probabilité de $\hat{\theta}_n$ vers θ_δ entraîne celle de R_n vers 0.

Ces résultats impliquent que pour tout δ , $0 < \delta \leq \delta_1$, δ_1 étant défini au lemme 7 :

$$\sqrt{n\delta} \left(\hat{\theta}_n - \theta_\delta \right) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p \left(0, V_\delta(\theta_0) \right), \quad V_\delta(\theta_0) = \left(H_{\theta_0}^\delta \right)^{-1} \Gamma_\delta(\theta_0) \left({}^t H_{\theta_0}^\delta \right)^{-1}$$

Or :

Lemme 8 *Sous (H 2), pour tout δ , $0 < \delta \leq \delta_0$, on a :*

$$\Gamma_\delta(\theta_\delta) = H_{\theta_0} + \frac{\delta}{2} P_{\theta_0} + o(\delta)$$

Ce lemme est démontré en annexe E. On obtient ainsi :

$$\begin{aligned} V_\delta(\theta_0) &= (-H_{\theta_0})^{-1} \left[H_{\theta_0} + \frac{\delta}{2} P_{\theta_0} \right] (-H_{\theta_0})^{-1} \\ &+ \left(\left[I_p + (H_{\theta_0})^{-1} H_{\theta_0}^\delta \right] \left(H_{\theta_0}^\delta \right)^{-1} \right) \left[H_{\theta_0} + \frac{\delta}{2} P_{\theta_0} \right] {}^t \left(\left[I_p + (H_{\theta_0})^{-1} H_{\theta_0}^\delta \right] \left(H_{\theta_0}^\delta \right)^{-1} \right) + o(\delta) \end{aligned}$$

Or, d'après lemme 7, on a :

$$\lim_{\delta \rightarrow 0} H_{\theta_0}^\delta = -H_{\theta_0} \quad \text{et} \quad \frac{1}{\delta} \left[I_p + (H_{\theta_0})^{-1} H_{\theta_0}^\delta \right] = -\frac{1}{2} H_{\theta_0}^{-1} M_{\theta_0} + o(1)$$

On obtient donc :

$$V_\delta(\theta_0) = H_{\theta_0}^{-1} \left(I_p + \frac{\delta}{2} P_{\theta_0} H_{\theta_0}^{-1} + o(\delta) \right)$$

1.4.2 Adaptation de la méthode au schéma de Simpson

La méthode d'estimation des moments généralisés peut être adaptée au schéma d'approximation de Simpson associé à la décomposition (1.6). On définit pour tout $i = 1, p$:

$$\begin{aligned} \tilde{V}_n^i(\theta) &= \frac{2}{n} \sum_{k=1}^{\frac{n}{2}} \tilde{\mathcal{V}}_\delta^i(Y_{2k\delta}, Y_{2k\delta-1}, Y_{2(k-1)\delta}, \theta) \\ \tilde{\mathcal{V}}_\delta^i(z, y, x, \theta) &= \frac{\partial}{\partial \theta^i} f_\theta(x) \left(\frac{z - x - \frac{\delta}{3} (f_\theta(x) + 4f_\theta(y) + f_\theta(z))}{2\delta\sigma^2(x)} \right) \end{aligned}$$

Soit $\tilde{U}_n(\theta) = \sum_{i=1}^p \left(\tilde{V}_n^i(\theta) \right)^2$ et $\tilde{\theta}_n = \inf_{\theta \in \Theta} \tilde{U}_n(\theta)$ l'estimateur des moments généralisés. Moyennant quelques adaptations dans les hypothèses, on peut montrer que cet estimateur converge vers θ_0 , à un biais en $O(\delta^4)$, et qu'il est asymptotiquement normal et efficace à un facteur $(I_p + O(\delta))$ près.

Hypothèses H 3 : Schéma de Simpson et G.M.M.

H 3.1 (H 2.1) *est maintenue avec f_{θ_0} et σ respectivement de classe C^8 et C^6 sur R .*

H 3.2 -3 -4 -6 -7 On reprend à l'identique (H 2.2), (H 2.3), (H 2.4), (H 2.6) et (H 2.7).

H 3.5 On maintient (H 2.5.1) (H 2.5.2) et (H 2.5.3). (H 2.5.4) est remplacée par (H 3.5.4) : Pour tout $l = 0, 3$, $(A_{\theta_0}^l f_{\theta_0})'$ et $A_{\theta_0}^4 f_{\theta_0}$ sont à croissance polynomiale en x , ainsi que celles de σ jusqu'à l'ordre 6.

Pour tout i et tout j , $M_{i,j}$, $Q_{i,j}$ et $P_{i,j}$ sont définies comme dans le paragraphe précédent. On remplace par contre N_i par \tilde{N}_i et N_{θ_0} par \tilde{N}_{θ_0} avec :

$$\tilde{N}_i(v) = E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \frac{(A_{\theta_0}^4 f_{\theta_0})(Y_v)}{\sigma^2(Y_0)} \right]$$

$${}^t \tilde{N}_{\theta_0} = \left(E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{(A_{\theta_0}^4 f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] \right)_i \quad (1.30)$$

On définit de plus pour tout $i = 1, p$:

$$\forall \delta > 0, \tilde{V}_i(\theta, \delta) = E_{\theta_0} \left[\tilde{V}_\delta^i(Y_{2\delta}, Y_\delta, Y_0, \theta) \right] \quad (1.31)$$

$$\tilde{V}_i(\theta, 0) = E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_\theta(Y_0)}{\sigma^2(Y_0)} \right) \right] \quad (1.32)$$

Théorème 5 Sous (H 3), pour H_{θ_0} définie en (H 3.6), P_{θ_0} en (1.18), \tilde{N}_{θ_0} en (1.30) et \tilde{V}_i en (1.31)-(1.32), il existe $\delta_1 > 0$ tel que pour tout δ , $0 < \delta \leq \delta_1$:

1. Il existe un unique $\tilde{\theta}_\delta \in \mathring{\Theta}$ qui vérifie, pour tout $i = 1, p$, $\tilde{V}_i(\tilde{\theta}_\delta, \delta) = 0$ et

$$\tilde{\theta}_\delta - \theta_0 = -\frac{\delta^4}{180} H_{\theta_0}^{-1} \tilde{N}_{\theta_0} + o(\delta^4)$$

2. Lorsque $n \rightarrow \infty$, on a :

$$\tilde{\theta}_n \xrightarrow{P_{\theta_0}} \tilde{\theta}_\delta$$

$$\sqrt{n\delta} (\tilde{\theta}_n - \tilde{\theta}_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p \left(0, \tilde{V}_\delta(\theta_0) \right), \quad \tilde{V}_\delta(\theta_0) = H_{\theta_0}^{-1} \left(I_p + \delta P_{\theta_0} H_{\theta_0}^{-1} + o(\delta) \right)$$

1.5 Etude expérimentale

Ce paragraphe est consacré à la mise en oeuvre des méthodes décrites précédemment et ceci pour trois modèles de diffusion. Pour les deux premiers (Ornstein-Uhlenbeck, (O.U.) et Cox-Ingersoll-Ross, (C.I.R.)) la dérive est linéaire dans le paramètre et affine en x . Pour le troisième (L.N.), la fonction de dérive $f_\theta(x) = \theta x \ln(x) + \frac{1}{2}x$ n'est plus linéaire en x . D'autre part, outre les méthodes du trapèze et de Simpson, il est possible d'utiliser le schéma de Bode à cinq points (cf. §1). Bien que nous n'ayons pas étudié l'estimateur des moments généralisés

associé, son comportement asymptotique empirique est notifié pour les trois modèles choisis. Nous comparerons les résultats associés aux schémas anticipatifs à ceux obtenus par la méthode des moindres carrés et le schéma d'Euler (cf. annexe F).

Nous indiquerons par e , t , s et b , les quantités se rapportant respectivement au schéma d'Euler, du trapeze, de Simpson et de Bode. On va illustrer expérimentalement quelques unes des propriétés annoncées pour $\hat{\theta}_n$: convergence vers θ_δ , écart entre $\hat{\theta}_n$ et θ_0 , biais théorique et empirique, variance asymptotique. Pour cela, on simule la diffusion sur un intervalle $[0, T]$ en utilisant un schéma d'Euler sur une grille fine de pas 0.005. Puis, on estime le paramètre par la méthode des moments généralisés pour les schémas t , s , b et par les moindres carrés pour e , ceci pour un choix (δ, n) , $n\delta = T$ avec $T = 40$ et $\delta \in \{0.01, 0.1\}$, $T = 400$ et $\delta \in \{0.1, 1\}$, $T = 4000$ et $\delta = 1$. On repète 100 fois cette expérience afin de calculer la moyenne et l'écart quadratique moyen empiriques notées respectivement $m(\hat{\theta}_n) = \frac{1}{n} \sum_{i=1}^{100} \hat{\theta}_n^i$ et $eqm(\hat{\theta}_n) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_n^i - \theta_0)^2$.

1.5.1 Processus de Ornstein-Uhlenbeck

On considère l'e.d.s. suivante :

$$dY_t = -\theta_0 Y_t dt + dW_t, Y_0 = x \quad (1.33)$$

(1.33) admet une unique solution forte, à trajectoires continues :

$$Y_t = \exp(-\theta_0 t) \left(x + \int_0^t \exp(\theta_0 s) dW_s \right)$$

Si $\theta_0 > 0$, $(Y_t)_{t \geq 0}$ est récurrente positive sur R , de loi invariante $\mu_{\theta_0} = \mathcal{N}\left(0, \frac{1}{2\theta_0}\right)$. On vérifie facilement (H 1), (H 3) et (H 4) (cf. annexe F). On montre de plus que : $A_{\theta_0}^l f_{\theta_0}(x) = (-\theta_0)^{l+1} x$ et

$$\begin{aligned} M(v) &= \frac{\exp(-\theta_0 v)}{2\theta_0}, P(v) = -Q(v) = \frac{1}{2\theta_0} \\ \check{N}(v) &= -\frac{\theta_0}{2} \exp(-\theta_0 v), N(v) = \frac{\theta_0^2}{2} \exp(-\theta_0 v), \tilde{N}(v) = \frac{\theta_0^4}{2} \exp(-\theta_0 v) \end{aligned}$$

La variance asymptotique théorique de $\sqrt{T}(\hat{\theta}_n - \theta_\delta)$ est encore notée $V_\delta(\theta_0)$. Pour le schéma d'Euler (cf. annexe F), on a :

$$\begin{aligned} \hat{\theta}_n^e &= -\frac{1}{\delta} \frac{\sum_{k=1}^n Y_{(k-1)\delta} (Y_{k\delta} - Y_{(k-1)\delta})}{\sum_{k=0}^{n-1} (Y_{k\delta})^2} \\ \theta_\delta^e &= -\frac{1}{\delta} (\exp(-\theta_0 \delta) - 1) = \theta_0 - \frac{\theta_0^2}{2} \delta + o(\delta) \\ V_\delta^e(\theta_0) &= 2\theta_0 (1 - \theta_0 \delta + (\delta)) \end{aligned}$$

Pour le schéma du trapèze, on a :

$$\begin{aligned}\widehat{\theta}_n^t &= -\frac{2 \sum_{k=1}^n Y_{(k-1)\delta} (Y_{k\delta} - Y_{(k-1)\delta})}{\delta \sum_{k=1}^n Y_{(k-1)\delta} (Y_{k\delta} + Y_{(k-1)\delta})} \\ \theta_\delta^t &= -\frac{2}{\delta} \left(\frac{\exp(-\theta_0\delta) - 1}{1 + \exp(-\theta_0\delta)} \right) = \theta_0 - \frac{\theta_0^3}{12} \delta^2 + o(\delta^2) \\ V_\delta^t(\theta_0) &= 2\theta_0 (1 + o(\delta))\end{aligned}$$

Pour le schéma de Simpson :

$$\begin{aligned}\widehat{\theta}_n^s &= -\frac{3 \sum_{k=1}^{\frac{n}{2}} Y_{2(k-1)\delta} (Y_{2k\delta} - Y_{2(k-1)\delta})}{\delta \sum_{k=1}^{\frac{n}{2}} Y_{2(k-1)\delta} (Y_{2(k-1)\delta} + 4Y_{2k\delta-1} + Y_{2k\delta})} \\ \theta_\delta^s &= -\frac{3}{\delta} \left(\frac{\exp(-2\delta\theta_0) - 1}{\exp(-2\delta\theta_0) + 4\exp(-\delta\theta_0) + 1} \right) = \theta_0 - \frac{\theta_0^5}{180} \delta^4 + o(\delta^4) \\ V_\delta^s(\theta_0) &= 2\theta_0 (1 + o(\delta))\end{aligned}$$

Pour le schéma de Bode :

$$\begin{aligned}\widehat{\theta}_n^b &= -\frac{45 \sum_{k=1}^{\frac{n}{4}} Y_{4(k-1)\delta} (Y_{4k\delta} - Y_{4(k-1)\delta})}{\delta \sum_{k=0}^{\frac{n}{4}} Y_{4(k-1)\delta} (14Y_{4(k-1)\delta} + 64Y_{(4k-3)\delta} + 24Y_{(4k-2)\delta} + 64Y_{(4k-1)\delta} + 14Y_{4k\delta})} \\ \theta_\delta^b &= -\frac{45}{\delta} \left(\frac{\exp(-4\theta_0\delta) - 1}{14 + 64\exp(-\theta_0\delta) + 24\exp(-2\theta_0\delta) + 64\exp(-3\theta_0\delta) + 14\exp(-4\theta_0\delta)} \right) \\ &= \theta_0 - \frac{2\theta_0^7}{945} \delta^6 + o(\delta^6)\end{aligned}$$

On choisit : $Y_0 = 0$, $\theta_0 = 1$. On constate que les estimateurs associés aux schémas anticipatifs améliorent le biais empirique et l'écart quadratique moyen obtenus avec le schéma d'Euler pour l'ensemble des partitions de T , (δ, n) , sauf lorsque $n = 4000$ et $\delta = 0.01$.

Lorsque $\delta = 0.1$ et $n = 400$, c'est la valeur relativement faible de T , $T = 40$, qui détériore le biais empirique. Dans les cas où $\delta = 1$ ($T = 400$ ou $T = 4000$), le biais vient de la différence entre θ_δ et θ_0 : pour les schémas d'Euler et du trapèze, on constate que l'écart entre l'estimation et θ_0 coïncide avec $\theta_\delta - \theta_0$. L'influence de δ et du type de schéma de discrétisation apparaît bien dès que T est suffisamment grand.

On remarque enfin que les résultats obtenus pour le schéma de Bode, dans le cas où $\delta = 1$, sont moins bons que ceux associés aux deux autres schémas anticipatifs.

1.5.2 Processus de Cox-Ingersoll-Ross

On considère l'e.d.s. suivante :

$$dY_t = b_0 - a_0 Y_t dt + \sqrt{Y_t} dW_t, \quad Y_0 = x$$

$$\theta = (b, a), \quad \sigma(x) = \sqrt{x} \text{ et } f = {}^t(f_1, f_2) \text{ avec } f_1(x) = -x \text{ et } f_2(x) = 1.$$

f est C^∞ sur R et lipschitzienne, σ est C^∞ sur R_+^* et h"olderienne de rapport $\frac{1}{2}$. L'e.d.s. admet donc une solution forte unique, à trajectoires continues. Si $b_0 \geq \frac{1}{2}$ et $a_0 > 0$, la

δ	θ_δ^e	θ_δ^t	θ_δ^s	θ_δ^b
0.01	0.995017	0.999992	1	1
0.1	0.951626	0.999167	0.999999	1
1	0.632121	0.92434	0.995067	0.998659

Table 1.1: Modèle de O.U. : comparaison des valeurs θ_δ pour les quatre schémas.

	δ	n	$m(\widehat{\theta}_n^e)$	$m(\widehat{\theta}_n^t)$	$m(\widehat{\theta}_n^s)$	$m(\widehat{\theta}_n^b)$
$T = 40$	0.1	400	0.846248	0.884295	0.928478	0.919213
	0.01	4000	1.012309	1.017524	1.019922	1.012849
$T = 400$	1	400	0.638412	0.940662	1.007697	1.063251
	0.1	4000	0.956130	1.004378	1.004885	1.001454
$T = 4000$	1	4000	0.634981	0.930612	1.004823	1.015734

Table 1.2: Modèle de O.U. : comparaison des moyennes empiriques pour 100 réalisations.

	δ	n	$eqm(\widehat{\theta}_n^e)$	$eqm(\widehat{\theta}_n^t)$	$eqm(\widehat{\theta}_n^s)$	$eqm(\widehat{\theta}_n^b)$
$T = 40$	0.1	400	0.03527	0.02707	0.01996	0.02
	0.01	4000	0.01295	0.01336	0.01591	0.02165
$T = 400$	1	400	0.1325	0.01223	0.03111	0.1359
	0.1	4000	0.00613	0.00515	0.00552	0.0065
$T = 4000$	1	4000	0.13339	0.005546	0.00276	0.0118

Table 1.3: Modèle de O.U. : comparaison des écarts quadratiques moyens pour 100 réalisations.

diffusion est récurrente positive sur R_+^* , de loi invariante $\mu_{\theta_0} = \Gamma(2b_0, 2a_0)$, admettant des moments de tous ordres. On vérifie sans difficulté que, dès que $b_0 > 2$, (H 1), (H 3) et (H 4) sont satisfaites et que :

$$\begin{aligned} M_{1,1}(v) &= \frac{b_0}{a_0}, \quad M_{1,2}(v) = -1, \quad M_{2,1}(v) = \frac{\exp(-a_0v) - 2b_0}{2b_0 - 1}, \quad M_{2,2}(v) = \frac{2a_0}{2b_0 - 1} \\ N_1(v) &= 0, \quad N_2(v) = \frac{a_0^3}{2b_0 - 1} \exp(-a_0v) \\ P_{1,1}(v) &= \frac{b_0}{a_0}, \quad P_{1,2}(v) = P_{2,1}(v) = \frac{\exp(-a_0v) - 2b_0}{2b_0 - 1}, \quad P_{2,2}(v) = 2 \frac{a_0 b_0 (b_0 - \exp(-a_0v))}{(2b_0 - 1)(b_0 - 1)} \\ Q_{i,j}(v) &= -a_0 P_{i,j}(v) \end{aligned}$$

Pour le schéma d'Euler, on a :

$$\begin{aligned} a_\delta^\epsilon &= -\frac{1}{\delta} (\exp(-a_0\delta) - 1) = a_0 - \frac{a_0^2}{2} \delta + o(\delta) \\ b_\delta^\epsilon &= \frac{b_0}{a_0} a_\delta = b_0 - \frac{b_0 a_0}{2} \delta + o(\delta) \end{aligned}$$

Pour le schéma du trapèze :

$$a_\delta^t = -\frac{\delta \exp(-a_0\delta) - 1}{2 \exp(-a_0\delta) + 1} = a_0 - \frac{a_0^3}{12} \delta^2 + o(\delta^2), \quad b_\delta^t = \frac{b_0}{a_0} a_\delta^t = b_0 - \frac{b_0 a_0^2}{12} \delta^2 + o(\delta^2)$$

Pour le schéma de Simpson :

$$\begin{aligned} a_\delta^s &= -\frac{3}{\delta} \left(\frac{\exp(-2\delta a_0) - 1}{\exp(-2\delta a_0) + 4 \exp(-\delta a_0) + 1} \right) = a_0 - \frac{a_0^5}{180} \delta^4 + o(\delta^4) \\ b_\delta^s &= \frac{b_0}{a_0} a_\delta = b_0 - \frac{b_0 a_0^4}{180} \delta^4 + o(\delta^4) \end{aligned}$$

Enfin, pour le schéma de Bode, on a :

$$\begin{aligned} a_\delta^b &= -\frac{45}{\delta} \left(\frac{\exp(-4a_0\delta) - 1}{14 + 64 \exp(-a_0\delta) + 24 \exp(-2a_0\delta) + 64 \exp(-3a_0\delta) + 14 \exp(-4a_0\delta)} \right) \\ &= a_0 - \frac{2a_0^7}{945} \delta^6 + o(\delta^6) \\ b_\delta^b &= \frac{b_0}{a_0} a_\delta = b_0 - \frac{2b_0 a_0^6}{945} \delta^6 + o(\delta^6) \end{aligned}$$

Bien qu'il soit possible de calculer le développement à l'ordre 1 en δ des différentes variances asymptotiques, nous nous contenterons de donner leur terme principal commun :

$$H_{\theta_0}^{-1} = (2b_0 - 1) \begin{pmatrix} \frac{2a_0}{2b_0 - 1} & 1 \\ 1 & \frac{b_0}{a_0} \end{pmatrix}.$$

On choisit : $Y_0 = 3$, $(a_0, b_0) = (1, 3)$. L'amélioration apportée par le schéma du trapèze par rapport au schéma d'Euler est particulièrement nette pour $T = 400$ avec $\delta = 1$ et $T = 4000$. Lorsque $T = 40$, l'imprécision des résultats obtenus s'expliquent par la faible valeur de T .

δ	$a_\delta (a_0 = 1)$				$b_\delta (b_0 = 3)$			
	a_δ^e	a_δ^t	a_δ^s	a_δ^b	b_δ^e	b_δ^t	b_δ^s	b_δ^b
0.01	0.995	0.999	1	1	2.985	2.999	3	3
0.1	0.951	0.999	0.999	1	2.854	2.997	2.999	3
1	0.632	0.924	0.995	0.998	1.896	2.772	2.985	2.995

Table 1.4: Modèle de C.I.R. : comparaison des valeurs (a_δ, b_δ) pour les quatre schémas.

	δ	n	$a_n (a_0 = 1)$				$b_n (b_0 = 3)$			
			$m(a_n^e)$	$m(a_n^t)$	$m(a_n^s)$	$m(a_n^b)$	$m(b_n^e)$	$m(b_n^t)$	$m(b_n^s)$	$m(b_n^b)$
$T = 40$	0.1	400	1.097	1.1659	1.171	1.177	3.235	3.429	3.4447	3.464
	0.01	4000	1.0759	1.102	1.099	1.101	3.208	3.279	3.268	3.274
$T = 400$	1	400	0.644	0.957	1.052	1.175	1.932	2.867	3.151	3.521
	0.1	4000	0.967	1.016	1.016	1.007	2.903	3.050	3.049	3.024
$T = 4000$	1	4000	0.633	0.926	1.006	1.011	1.897	2.776	3.015	3.030

Table 1.5: Modèle de C.I.R. : comparaison des moyennes empiriques pour 100 réalisations.

	δ	n	$a_n (a_0 = 1)$			
			$eqm(a_n^e)$	$eqm(a_n^t)$	$eqm(a_n^s)$	$eqm(a_n^b)$
$T = 40$	0.1	400	0.0734	0.0938	0.1	0.128
	0.01	4000	0.0869	0.0846	0.086	0.0884
$T = 400$	1	400	0.129	0.0147	0.045	0.29
	0.1	4000	0.0065	0.0062	0.0082	0.0089
$T = 4000$	1	4000	0.135	0.0064	0.003	0.0169

Table 1.6: Modèle de C.I.R. : comparaison des écarts quadratiques moyens pour le paramètre a_0 pour 100 réalisations.

	δ	n	$b_n (b_0 = 3)$			
			$eqm(b_n^e)$	$eqm(b_n^t)$	$eqm(b_n^s)$	$eqm(b_n^b)$
$T = 40$	0.1	400	0.581	0.7	0.78	1.02
	0.01	4000	0.1125	0.72	0.73	0.747
$T = 400$	1	400	1.16	0.134	0.412	2.65
	0.1	4000	0.053	0.054	0.065	0.074
$T = 4000$	1	4000	1.219	0.061	0.032	0.145

Table 1.7: Modèle de C.I.R. : comparaison des écarts quadratiques moyens pour le paramètre b_0 pour 100 réalisations.

L'influence de δ sur la variance asymptotique est d'autant plus forte que la discrétisation associée au schéma anticipatif est importante et que la valeur de T est faible. Ceci explique en particulier que l'écart quadratique moyen et le biais empirique soient moins satisfaisants pour le schéma de Simpson et le schéma de Bode que pour le schéma du trapèze lorsque $\delta = 1$ et $T = 400$.

1.5.3 Cas d'une dérive non linéaire en x : diffusion Log-Normale

Soit l'e.d.s. :

$$dY_t = \left(-\theta_0 \ln(Y_t) Y_t + \frac{1}{2} Y_t \right) dt + Y_t dW_t, \quad Y_0 = x$$

Cette équation admet une unique solution forte à trajectoires continues (cf. [14], p.125) :

$$Y_t = \exp \left(\exp(-\theta_0 t) \ln(x) + \exp(-\theta_0 t) \int_0^t \exp(\theta_0 s) dW_s \right)$$

On montre que $\pi_{\theta_0}^t(x, \cdot)$ est une loi log-normale de paramètre $\left(\ln(x) \exp(-\theta_0 t), \frac{1 - \exp(-2\theta_0 t)}{2\theta_0} \right)$. Si $\theta_0 > 0$, la diffusion est récurrente positive de loi invariante μ_{θ_0} log-normale de paramètres $\left(0, \frac{1}{2\theta_0} \right)$. Dans ce cas, $f_{\theta}(x) = -\theta_0 \ln(x) x + \frac{1}{2} x$ ne vérifie pas l'intégralité de (H 2), (H 3) et (H 4) : en particulier, l'hypothèse de croissance polynomiale faite sur f_{θ_0} n'est pas satisfaite. Cependant, puisque la transition $\pi_{\theta_0}^t$ est explicite, on peut prouver l'ensemble des régularités requises pour l'obtention des résultats asymptotiques. On a ainsi, en notant $\alpha(v) = \frac{1 - \exp(-\theta_0 v)}{2\theta_0}$:

$$\begin{aligned} M(v) &= \left(\frac{\exp(-\theta_0 v)}{2\theta_0} - \alpha^2(v) \right) \exp(\alpha(v)) \\ P(v) &= \left(\frac{1}{2\theta_0} + 4\alpha^2(v) \right) \exp(4\alpha(v)) \\ R(v) &= \alpha(v) \exp(\alpha(v)) \end{aligned}$$

Les estimateurs associés aux quatre schémas d'approximation vont converger respectivement vers :

$$\begin{aligned} \theta_{\delta}^c &= \frac{1}{\delta} \frac{R(\delta)}{M(0)} = \theta_0 + \frac{1}{2} (\theta_0 - \theta_0^2) \delta + o(\delta) \\ \theta_{\delta}^t &= \frac{2}{\delta} \left(1 - \frac{\delta}{4} \right) \frac{R(\delta)}{M(0) + M(\delta)} = \theta_0 + \left(-\frac{\theta_0^3}{12} + \frac{\theta_0^2}{4} - \frac{\theta_0}{16} \right) \delta^2 + o(\delta^2) \\ \theta_{\delta}^s &= \frac{\left(\frac{3}{\delta} - \frac{1}{2} \right) R(2\delta) - 2R(\delta)}{M(0) + 4M(\delta) + M(2\delta)} = \theta_0 + \frac{1}{12} \left(-\frac{\theta_0^5}{15} + \theta_0^4 - \frac{65\theta_0^3}{4} + \frac{13\theta_0^2}{3} - \frac{\theta_0}{48} \right) \delta^4 + o(\delta^4) \\ \theta_{\delta}^b &= \frac{\left(\frac{45}{\delta} - 7 \right) R(4\delta) - 32R(3\delta) - 12R(2\delta) - 32R(\delta)}{14M(0) + 64M(\delta) + 24M(\delta^2) + 64M(\delta^3) + 14M(\delta^4)} \\ &= \theta_0 + \frac{\theta_0}{30240} (-7 + 252\theta_0 - 2800\theta_0^2 + 11200\theta_0^3 - 14448\theta_0^4 + 4032\theta_0^5 - 64\theta_0^6) \delta^6 + o(\delta^6) \end{aligned}$$

On choisit : $Y_0 = \exp\left(\frac{1}{4}\right)$ et $\theta_0 = 1$. Si, comme pour le modèle de C.I.R. les schémas anticipatifs n'améliorent pas sensiblement les résultats obtenus par le schéma d'Euler pour des valeurs importantes de δ et faibles de T , le schéma du trapèze semble un bon choix en situation de discrétisation assez générale.

δ	θ_δ^e	θ_δ^t	θ_δ^s	θ_δ^b
0.01	0.999979	1	1	1
0.1	0.998	1.001	1	1
1	0.867	1.0569	1.0268	0.990

Table 1.8: Diffusion Log-Normale : comparaison des valeurs θ_δ pour les quatre schémas.

	δ	n	$m(\widehat{\theta}_n^e)$	$m(\widehat{\theta}_n^t)$	$m(\widehat{\theta}_n^s)$	$m(\widehat{\theta}_n^b)$
$T = 40$	0.1	400	1.100	1.107	1.117	1.215
	0.01	4000	1.094	1.094	1.102	1.10
$T = 400$	1	400	0.863	1.068	0.682	0.734
	0.1	4000	1.005	1.008	1.007	1.008
$T = 4000$	1	4000	0.872	1.067	1.104	1.136

Table 1.9: Diffusion Log-Normale : comparaison des moyennes empiriques pour 100 réalisations.

	δ	n	$eqm(\widehat{\theta}_n^e)$	$eqm(\widehat{\theta}_n^t)$	$eqm(\widehat{\theta}_n^s)$	$eqm(\widehat{\theta}_n^b)$
$T = 40$	0.1	400	0.048	0.054	0.048	0.116
	0.01	4000	0.0638	0.064	0.0735	0.0692
$T = 400$	1	400	0.03	0.0765	4.85	28.55
	0.1	4000	0.0043	0.0048	0.00635	0.0114
$T = 4000$	1	4000	0.0173	0.011	0.08	2.678

Table 1.10: Diffusion Log-Normale : comparaison des écarts quadratiques moyens pour 100 réalisations.

Bibliography

- [1] Bergstrom, A. R. (1976) *Statistical inference in Continuous Time Series*.- In *Statistical inference in Continuous Time Economic Models*., Ed. Bergstrom, North Holland.
- [2] Bibby, B.M. & Sorensen, M. (1995) *Martingale Estimation Functions for Discretely Observed Diffusion Processes*. - Bernoulli 1, 17-39.
- [3] Dacunha-Castelle, D. & Duflo, M. (1993) *Probabilité et Statistiques. Tome 2- 2^{eme} Ed.* Masson.
- [4] Dacunha-Castelle, D. & Florens-Zmirou, D. (1986) *Estimation of the coefficient of a diffusion from discrete observations*.- Stochastics 19, 263-284.
- [5] Florens-Zmirou, D. (1989) *Approximate discrete schemes for statistics of diffusion processes*.- Statistics 20, 547-557.
- [6] Genon-Catalot, V. (1990) *Maximum contrast estimation for diffusion processes from discrete observations*.- Statistics 21, 99-116.
- [7] Gouriéroux, C. & Monfort, A. *Statistique et Modèles Econométriques. Tome 1*- Economica.
- [8] Hansen, L (1982) *Large Sample Properties of Generalized Method of Moments Estimators*.- Econometrica, 50 (4), 1029-1054.
- [9] Hansen, L & Singleton, K (1982) *Generalized Instrumental Variables Estimation of Non-linear Rational Expectations Models*.- Econometrica, 50 (5), 1269-1286.
- [10] Karatzas, I. & Shreve S.E. (1996) *Brownian Motion and Stochastic Calculus*.- 2nd Ed. Springer.
- [11] Kessler, M. (1997) *Estimation of an ergodic diffusion from discrete observations*.- Scand. J. Stat. 24, No. 2, 211-229.
- [12] Kessler, M. (1996) *Simple and Explicit Estimating Functions for a Discretely Observed Diffusion Process*.- Research Reports 336, Department of theoretical statistics, University of Aarhus.
- [13] Kessler, M. & Sorensen, M. (1995) *Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process*.- Research Reports 332, Department of theoretical statistics, University of Aarhus.

- [14] Kloeden, P.E. & Platen, E. (1995) *Numerical Solution of Stochastic Differential Equations*.- Springer.
- [15] Kutoyants, Yu. A. (1984) *Parameter estimation for stochastic processes*.- Research and Exposition in math. 6, Heldermann Verlag, Berlin.
- [16] Lipster, R.S. & Shiryaev, A.N. (1977) *Statistics of random processes. Tome 1,2*- Springer-Verlag.
- [17] Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. (1988) *Numerical Recipes in C- 2nd Ed.* Cambridge University Press, 132-133.
- [18] Prakasa-Rao, B. L. S. (1983) *Asymptotic theory for non linear least squares estimator for diffusion proceses*.- Math. Operationsforsch. Statist Ser. Berlin 14, 2, 195-209.
- [19] Pedersen, A. R. (1995) *Consistency and asymptotic normality of an approximate maximum likelihood estimator for dicretely observed diffusion processes*.- Bernoulli 1, 257-279.
- [20] Pedersen, A. R. (1995) *A new approch to maximum likelihood estimation for stochastic differential equations based on discrete observations*.- Scand. J. Statist. 22, 55-71.
- [21] Sargan, J.D. (1976) *Some discrete approximations to continuous times stochastics models* - In *Statistical inference in Continuous Time Economic Models*., Ed. Bergstrom, North Holand, 27-80.
- [22] Sorensen, M. (1996) *Estimating functions for discretely observed diffusions: A review*.- Research Reports 348, Department of theoretical statistics, University of Aarhus.
- [23] Yoshida, N. (1992) *Estimation for diffusion processes from discrete observations*.- J. Multivariate Anal. 41, 220-242.

Annexes

A Démonstration de la propriété 2 : schéma d'approximation de Simpson

Comme pour le schéma du trapèze, cette démonstration repose sur les résultats (1.4)-(1.5) d'intervention de l'ordre d'intégration des intégrales de Riemann et des intégrales stochastique.

- Posons:

$$\begin{aligned} A_t &= \int_t^{t+\delta} f(Y_v) dv - \frac{\delta}{3} [f(Y_t) + 2f(Y_{t+\delta})] + \int_{t+\delta}^{t+2\delta} f(Y_v) dv - \frac{\delta}{3} [f(Y_{t+2\delta}) + 2f(Y_{t+\delta})] \\ &= A_t^1 + A_t^2 \end{aligned}$$

$$A_t^1 = \frac{1}{3} \int_t^{t+\delta} [f(Y_v) - f(Y_t)] dv - \frac{2}{3} \int_t^{t+\delta} [f(Y_{t+\delta}) - f(Y_v)] dv$$

Par application de la formule de Ito, on obtient:

$$\begin{aligned} A_t^1 &= \frac{1}{3} \int_t^{t+\delta} \left[\int_t^v Af(Y_u) du \right] dv + \frac{1}{3} \int_t^{t+\delta} \left[\int_t^v (\sigma f')(Y_u) dW_u \right] dv \\ &\quad - \frac{2}{3} \int_t^{t+\delta} \left[\int_v^{t+\delta} Af(Y_u) du \right] dv - \frac{2}{3} \int_t^{t+\delta} \left[\int_v^{t+\delta} (\sigma f')(Y_u) dW_u \right] dv \end{aligned}$$

(1.4) et (1.5) impliquent que :

$$\begin{aligned} A_t^1 &= \frac{1}{3} \int_t^{t+\delta} (t + \delta - u) Af(Y_u) du + \frac{1}{3} \int_t^{t+\delta} (t + \delta - u) (\sigma f')(Y_u) dW_u \\ &\quad - \frac{2}{3} \int_t^{t+\delta} (u - t) Af(Y_u) du - \frac{2}{3} \int_t^{t+\delta} (u - t) (\sigma f')(Y_u) dW_u \\ &= \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) Af(Y_u) du + \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) (\sigma f')(Y_u) dW_u \end{aligned}$$

En procédant de même que pour A_t^1 , on montre que :

$$\begin{aligned} A_t^2 &= \frac{2}{3} \int_{t+\delta}^{t+2\delta} [f(Y_v) - f(Y_{t+\delta})] dv - \frac{1}{3} \int_{t+\delta}^{t+2\delta} [f(Y_{t+2\delta}) - f(Y_v)] dv \\ &= \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) Af(Y_u) du + \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) (\sigma f')(Y_u) dW_u \end{aligned}$$

Donc :

$$\begin{aligned} A_t &= \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) Af(Y_u) du + \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) Af(Y_u) du \\ &+ \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) (\sigma f')(Y_u) dW_u + \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) (\sigma f')(Y_u) dW_u \end{aligned}$$

Or :

$$\int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) du = -\frac{\delta^2}{6}, \quad \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) du = \frac{\delta^2}{6}$$

Donc, si on note :

$$B_t = \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) Af(Y_u) du + \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) Af(Y_u) du$$

on a :

$$\begin{aligned} B_t &= \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) [Af(Y_u) - A(Y_{t+\delta})] du \\ &+ \int_{t+\delta}^{t+2\delta} \left(t + \frac{5\delta}{3} - u \right) [Af(Y_u) - Af(Y_{t+\delta})] du \\ &= B_t^1 + B_t^2 \end{aligned}$$

- (1.4) et (1.5) entraînent que :

$$\begin{aligned} B_t^1 &= - \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) \left[\int_u^{t+\delta} A^2 f(Y_v) dv \right] du \\ &- \int_t^{t+\delta} \left(t + \frac{\delta}{3} - u \right) \left[\int_u^{t+\delta} [\sigma(Af)'](Y_v) dW_v \right] du \\ &= - \int_t^{t+\delta} \left[\int_t^v \left(t + \frac{\delta}{3} - u \right) du \right] A^2 f(Y_v) dv \\ &- \int_t^{t+\delta} \left[\int_t^v \left(t + \frac{\delta}{3} - u \right) du \right] [\sigma(Af)'](Y_v) dW_v \\ &= \frac{1}{2} \int_t^{t+\delta} (v-t) \left(v-t - \frac{2\delta}{3} \right) A^2 f(Y_v) dv \\ &+ \frac{1}{2} \int_t^{t+\delta} (v-t) \left(v-t - \frac{2\delta}{3} \right) [\sigma(Af)'](Y_v) dW_v \end{aligned}$$

De même :

$$\begin{aligned} B_t^2 &= \frac{1}{2} \int_{t+\delta}^{t+2\delta} (t+2\delta-v) \left(t + \frac{4\delta}{3} - v \right) A^2 f(Y_v) dv \\ &+ \frac{1}{2} \int_{t+\delta}^{t+2\delta} (t+2\delta-v) \left(t + \frac{4\delta}{3} - v \right) [\sigma(Af)'](Y_v) dW_v \end{aligned}$$

Or :

$$\frac{1}{2} \int_t^{t+\delta} (v-t) \left(v-t-\frac{2\delta}{3}\right) dv = 0 \text{ et } \frac{1}{2} \int_{t+\delta}^{t+2\delta} (t+2\delta-v) \left(t+\frac{4\delta}{3}-v\right) dv = 0$$

Donc si on note :

$$\begin{aligned} C_t &= \frac{1}{2} \int_t^{t+\delta} (v-t) \left(v-t-\frac{2\delta}{3}\right) A^2 f(Y_v) dv \\ &\quad + \frac{1}{2} \int_{t+\delta}^{t+2\delta} (t+2\delta-v) \left(t+\frac{4\delta}{3}-v\right) A^2 f(Y_v) dv \end{aligned}$$

On a :

$$\begin{aligned} C_t &= \frac{1}{2} \int_t^{t+\delta} (v-t) \left(v-t-\frac{2\delta}{3}\right) [A^2 f(Y_v) - A^2 f(Y_{t+\delta})] dv \\ &\quad + \frac{1}{2} \int_{t+\delta}^{t+2\delta} (t+2\delta-v) \left(t+\frac{4\delta}{3}-v\right) [A^2 f(Y_v) - A^2 f(Y_{t+\delta})] dv \\ &= C_t^1 + C_t^2 \end{aligned}$$

• Par (1.4) et (1.5), on obtient :

$$\begin{aligned} C_t^1 &= -\frac{1}{2} \int_t^{t+\delta} (u-t) \left(u-t-\frac{2\delta}{3}\right) \left[\int_u^{t+\delta} A^3 f(Y_v) dv \right] du \\ &\quad - \frac{1}{2} \int_t^{t+\delta} (u-t) \left(u-t-\frac{2\delta}{3}\right) \left[\int_u^{t+\delta} [\sigma(A^2 f)'] (Y_v) dW_v \right] du \\ &= -\frac{1}{2} \int_t^{t+\delta} \left[\int_t^v (u-t) \left(u-t-\frac{2\delta}{3}\right) du \right] A^3 f(Y_v) dv \\ &\quad - \frac{1}{2} \int_t^{t+\delta} \left[\int_t^v (u-t) \left(u-t-\frac{2\delta}{3}\right) du \right] [\sigma(A^2 f)'] (Y_v) dW_v \\ &= \frac{1}{6} \int_t^{t+\delta} (v-t)^2 (t+\delta-v) A^3 f(Y_v) dv \\ &\quad + \frac{1}{6} \int_t^{t+\delta} (v-t)^2 (t+\delta-v) [\sigma(A^2 f)'] (Y_v) dW_v \end{aligned}$$

De même :

$$\begin{aligned} C_t^2 &= \frac{1}{6} \int_{t+\delta}^{t+2\delta} (v-t-2\delta)^2 (t+\delta-v) A^3 f(Y_v) dv \\ &\quad + \frac{1}{6} \int_{t+\delta}^{t+2\delta} (v-t-2\delta)^2 (t+\delta-v) [\sigma(A^2 f)'] (Y_v) dW_v \end{aligned}$$

Or :

$$\int_t^{t+\delta} (v-t)^2 (t+\delta-v) dv = \frac{\delta^4}{12}, \quad \int_{t+\delta}^{t+2\delta} (v-t-2\delta)^2 (t+\delta-v) dv = -\frac{\delta^4}{12}$$

Donc, si on pose :

$$\begin{aligned} D_t &= \frac{1}{6} \int_t^{t+\delta} (v-t)^2 (t+\delta-v) A^3 f(Y_v) dv \\ &+ \frac{1}{6} \int_{t+\delta}^{t+2\delta} (v-t-2\delta)^2 (t+\delta-v) A^3 f(Y_v) dv \end{aligned}$$

On a :

$$\begin{aligned} D_t &= \frac{1}{6} \int_t^{t+\delta} (v-t)^2 (t+\delta-v) [A^3 f(Y_v) - A^3 f(Y_{t+\delta})] dv \\ &+ \frac{1}{6} \int_{t+\delta}^{t+2\delta} (v-t-2\delta)^2 (t+\delta-v) [A^3 f(Y_v) - A^3 f(Y_{t+\delta})] dv \\ &= D_t^1 + D_t^2 \end{aligned}$$

• D'après (1.4) et (1.5) :

$$\begin{aligned} D_t^1 &= -\frac{1}{6} \int_t^{t+\delta} (u-t)^2 (t+\delta-u) \left[\int_u^{t+\delta} A^4 f(Y_v) dv \right] du \\ &- \frac{1}{6} \int_t^{t+\delta} (u-t)^2 (t+\delta-u) \left[\int_u^{t+\delta} [\sigma(A^3 f)'] (Y_v) dW_v \right] du \\ &= -\frac{1}{6} \int_t^{t+\delta} \left[\int_t^v (u-t)^2 (t+\delta-u) du \right] A^4 f(Y_v) dv \\ &- \frac{1}{6} \int_t^{t+\delta} \left[\int_t^v (u-t)^2 (t+\delta-u) du \right] [\sigma(A^3 f)'] (Y_v) dW_v \\ &= \frac{1}{24} \int_t^{t+\delta} (v-t)^3 \left(v-t-\frac{4\delta}{3} \right) A^4 f(Y_v) dv \\ &+ \frac{1}{24} \int_t^{t+\delta} (v-t)^3 \left(v-t-\frac{4\delta}{3} \right) [\sigma(A^3 f)'] (Y_v) dW_v \end{aligned}$$

De même :

$$\begin{aligned} D_t^2 &= \frac{1}{24} \int_{t+\delta}^{t+2\delta} (t+2\delta-v)^3 \left(t+\frac{2\delta}{3}-v \right) A^4 f(Y_v) dv \\ &+ \frac{1}{24} \int_{t+\delta}^{t+2\delta} (t+2\delta-v)^3 \left(t+\frac{2\delta}{3}-v \right) [\sigma(A^3 f)'] (Y_v) dW_v \end{aligned}$$

avec :

$$\begin{aligned} \int_t^{t+\delta} (v-t)^3 \left(v-t-\frac{4\delta}{3} \right) dv &= \int_{t+\delta}^{t+2\delta} (t+2\delta-v)^3 \left(t+\frac{2\delta}{3}-v \right) dv \\ &= -\frac{2\delta^5}{15} \end{aligned}$$

Le développement s'arrête donc à la quatrième itération. Si l'on rassemble l'ensemble des résultats obtenus, on a :

$$A_t + \int_t^{t+2\delta} \sigma(Y_v) dW_v = \eta_t + \tilde{W}_t$$

où η_t et \tilde{W}_t sont les variables définies dans la propriété 2.

B Démonstration du lemme 4

La démonstration repose sur le résultat du problème 3.15 de [10], p.306 :

Lemme 9 Soient f_{θ_0} et σ deux fonctions mesurables de $(R, \mathcal{B}(R))$ dans $(R, \mathcal{B}(R))$. On suppose qu'il existe une constante $K > 0$ telle que :

$$\forall x \in R, |f_{\theta_0}(x)| + |\sigma(x)| \leq K(1 + |x|)$$

Si (Y, W) , (Ω, \mathcal{F}, P) , $\{\mathcal{F}_t\}_{t \geq 0}$ est une solution faible de (1.1) telle que, pour $m \geq 1$, $E[|Y_0|^{2m}] < \infty$, alors pour tout $T > 0$, on a :

$$\forall t, 0 \leq t \leq T, E \left[\max_{0 \leq s \leq t} |Y_s|^{2m} \right] \leq C \left(1 + E[|Y_0|^{2m}] \right) \exp(Ct)$$

où C est une constante strictement positive qui ne dépend que de m , T et K .

Sous (H 2.1), (H 2.2) et si l'on suppose que la loi invariante μ_{θ_0} admet des moments de tous ordres, le processus stationnaire associé à (1.1) satisfait les hypothèses de ce lemme. On a donc pour tout $m \in N^*$ et tout $T > 0$:

$$\forall t, 0 \leq t \leq T, E_{\theta_0} \left[\max_{0 \leq s \leq t} |Y_s|^{2m} \right] \leq C \left(1 + E_{\theta_0}[|Y_0|^{2m}] \right) \exp(Ct) \quad (1.34)$$

1. La justification de la continuité de m sur $U(\Theta) \times R^+$ repose sur l'application conjointe du théorème de convergence dominée de Lebesgue et de (1.34). Soit $((\theta_n, \delta_n))_n$ une suite de $U(\Theta) \times R^+$ convergeant vers (θ, δ) , également dans $U(\Theta) \times R^+$. Par continuité de h et de g sur $U(\Theta) \times R$ et celle en t du processus stationnaire solution de (1.1), on a :

$$h(\theta_n, Y_0) g(\theta_n, Y_{\delta_n}) \xrightarrow{P_{\theta_0-p.s.}} h(\theta, Y_0) g(\theta, Y_\delta)$$

De plus, d'après l'hypothèse de croissance polynomiale, on sait qu'il existe $M > 0$ et $\alpha > 0$ telles que :

$$\forall n \in N, |h(\theta_n, Y_0) g(\theta_n, Y_{\delta_n})| \leq M(1 + |Y_0|)^\alpha \left(1 + \max_{n \in N} |Y_{\delta_n}| \right)^\alpha$$

Or, d'après (1.34), $M(1 + |Y_0|)^\alpha (1 + \max_{n \in N} |Y_{\delta_n}|)^\alpha \in L^1(Q_{\theta_0})$. Par application du théorème de Lebesgue, on obtient, lorsque $n \rightarrow \infty$:

$$m(\theta_n, \delta_n) \xrightarrow{L^1(P_{\theta_0})} m(\theta, \delta)$$

2. On montre en utilisant les mêmes arguments que, pour tout $\delta \geq 0$, $\theta \rightarrow m(\theta, \delta)$ admet des dérivées partielles en tout point de $U(\Theta)$. Elles sont obtenues par dérivation sous l'espérance et sont continues sur $U(\Theta) \times R^+$.

3. Montrons que pour tout $\theta \in U(\Theta)$, $\delta \rightarrow m(\theta, \delta)$ est dérivable à droite en tout point de R^+ . Pour tout $h > 0$:

$$m(\theta, \delta + h) - m(\theta, \delta) = E_{\theta_0} [h(\theta, Y_0) [g(\theta, Y_{\delta+h}) - g(\theta, Y_\delta)]]$$

Par application de la formule de Ito, on a :

$$g(\theta, Y_{\delta+h}) - g(\theta, Y_\delta) = \int_\delta^{\delta+h} (\sigma g')(\theta, Y_u) dW_u + \int_\delta^{\delta+h} (A_{\theta_0} g)(\theta, Y_u) du$$

où, pour tout $\delta \geq 0$, $\int_\delta^{\delta+h} (\sigma g')(\theta, Y_u) dW_u$ est décorrélée de Y_0 . On obtient donc :

$$m(\theta, \delta + h) - m(\theta, \delta) = \int_\delta^{\delta+h} E_{\theta_0} [h(\theta, Y_0) (A_{\theta_0} g)(\theta, Y_u)] du$$

Or, d'après ce qui précède, on sait que $(\theta, u) \rightarrow E_{\theta_0} [h(\theta, Y_0) (A_{\theta_0} g)(\theta, Y_u)]$ est continue sur $U(\Theta) \times R^+$. Par conséquent :

$$\lim_{h \rightarrow 0} \frac{1}{h} [m(\theta, \delta + h) - m(\theta, \delta)] = E_{\theta_0} [h(\theta, Y_0) (A_{\theta_0} g)(\theta, Y_\delta)]$$

On procède de même pour étudier la limite de $\frac{1}{h} [m(\theta, \delta) - m(\theta, \delta - h)]$ pour tout $\delta > 0$. On montre ainsi que, pour tout $\theta \in U(\Theta)$, $\delta \rightarrow m(\theta, \delta)$ est dérivable sur R_*^+ et est dérivable à droite en 0 avec :

$$\begin{aligned} \forall \delta, \delta > 0, m'(\theta, \delta) &= E_{\theta_0} [h(\theta, Y_0) (A_{\theta_0} g)(\theta, Y_\delta)] \\ m'(\theta, 0^+) &= E_{\theta_0} [h(\theta, Y_0) (A_{\theta_0} g)(\theta, Y_0)] \end{aligned}$$

D'après 1., on sait que $(\theta, \delta) \rightarrow m'(\theta, \delta)$ est continue sur $U(\Theta) \times R_*^+$ et que, pour tout $\theta \in U(\Theta)$ et toute suite $((\theta_n, \delta_n))_n \subset U(\Theta) \times R_*^+$ telle que $(\theta_n, \delta_n) \rightarrow (\theta, 0)$, on a : $\lim_{n \rightarrow \infty} m'(\theta_n, \delta_n) = m'(\theta, 0^+)$. On a obtenu le résultat annoncé.

C Démonstration du lemme 5

1. Nous commencerons par montrer la continuité de V_i sur $U(\Theta) \times R^+$. Par application du lemme 4, on sait que sous (H 2), pour tout $i = 1, p$, $(\theta, \delta) \rightarrow V_i(\theta, \delta)$ est continue sur $U(\Theta) \times R_*^+$. Pour obtenir la continuité de la fonction sur $U(\Theta) \times R^+$, il nous faut montrer que V_i est également continue en tout point de la forme $(\theta, 0)$, $\theta \in U(\Theta)$. En utilisant la décomposition du trapèze (1.3), on a pour tout $\theta \in U(\Theta)$ et tout $\delta > 0$:

$$\begin{aligned}
V_i(\theta, \delta) &= \frac{1}{\delta} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{\eta_0}{\sigma^2(Y_0)} \right] \\
&+ \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] \\
&+ \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_{\delta}) - f_{\theta}(Y_{\delta})}{\sigma^2(Y_0)} \right) \right] \\
&= -\frac{\delta^2}{2} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{u\delta})}{\sigma^2(Y_0)} \right] du \\
&+ \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] \\
&+ \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_{\delta}) - f_{\theta}(Y_{\delta})}{\sigma^2(Y_0)} \right) \right]
\end{aligned}$$

Or, d'après le lemme 4, les trois fonctions précédentes sont continues sur $U(\Theta) \times R^+$. Pour toute suite $((\theta_n, \delta_n))_n \subset U(\Theta) \times R_*^+$ convergeant vers $(\theta, 0)$, on obtient donc :

$$\lim_{n \rightarrow \infty} V_i(\theta_n, \delta_n) = E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] = V_i(\theta, 0)$$

On en déduit que V_i est continue sur $U(\Theta) \times R^+$. De plus, en utilisant le développement précédent, on a :

$$\begin{aligned}
V_i(\theta_0, \delta) &= -\frac{\delta^2}{2} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{u\delta})}{\sigma^2(Y_0)} \right] du \\
&= -\frac{\delta^2}{2} \int_0^1 (1-u) u N_i(u\delta) du
\end{aligned}$$

Par continuité de $v \rightarrow N_i(v)$, on a : $V_i(\theta_0, \delta) = -\frac{\delta^2}{12} N_i(0) + o(\delta^2)$. L'équation (1.20) est donc démontrée.

2. D'après le lemme 4, pour tout $\delta \geq 0$, $\theta \rightarrow V_i(\theta, \delta)$ admet des dérivées partielles sur $U(\Theta)$, obtenues par dérivation sous l'espérance. On a, pour tout $j = 1, p$ et tout $\theta \in U(\Theta)$:

$$\begin{aligned}
\frac{\partial}{\partial \theta^j} V_i(\theta, 0) &= E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] \\
&- E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{\frac{\partial}{\partial \theta^j} f_{\theta}(Y_0) + \frac{\partial}{\partial \theta^j} f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right]
\end{aligned}$$

$$\begin{aligned} \forall \delta > 0, \quad \frac{\partial}{\partial \theta^j} V_i(\theta, \delta) &= E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \theta^i} f_\theta(Y_0) \left(\frac{Y_\delta - Y_0 - \frac{\delta}{2} (f_\theta(Y_0) + f_\theta(Y_\delta))}{\delta \sigma^2(Y_0)} \right) \right] \\ &\quad - \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \left(\frac{\frac{\partial}{\partial \theta^j} f_\theta(Y_0) + \frac{\partial}{\partial \theta^j} f_\theta(Y_\delta)}{\sigma^2(Y_0)} \right) \right] \end{aligned}$$

On sait de plus que pour tout $j = 1, p$, $(\theta, \delta) \rightarrow \frac{\partial}{\partial \theta^j} V_i(\theta, \delta)$ est continue sur $U(\Theta) \times R_*^+$. Pour prolonger ce résultat sur $U(\Theta) \times R^+$, il nous faut montrer que $\frac{\partial}{\partial \theta^j} V_i$ est continue en tout point de la forme $(\theta, 0)$, $\theta \in U(\Theta)$. En utilisant le schéma du trapèze (1.3), on obtient :

$$\begin{aligned} \forall \delta > 0, \quad \frac{\partial}{\partial \theta^j} V_i(\theta, \delta) &= -\frac{\delta^2}{2} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \theta^i} f_\theta(Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{\delta u})}{\sigma^2(Y_0)} \right] du \\ &\quad + \frac{1}{2} E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \theta^i} f_\theta(Y_0) \left(\frac{f_{\theta_0}(Y_\delta) - f_\theta(Y_\delta)}{\sigma^2(Y_0)} \right) \right] du \\ &\quad + \frac{1}{2} E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \theta^i} f_\theta(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_\theta(Y_0)}{\sigma^2(Y_0)} \right) \right] du \\ &\quad - \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \left(\frac{\frac{\partial}{\partial \theta^j} f_\theta(Y_0) + \frac{\partial}{\partial \theta^j} f_\theta(Y_\delta)}{\sigma^2(Y_0)} \right) \right] \end{aligned}$$

D'après le lemme 4, les fonctions qui interviennent dans le développement précédent sont continues sur $U(\Theta) \times R^+$. Si $((\theta_n, \delta_n))_n \subset U(\Theta) \times R_*^+$ est une suite qui converge vers $(\theta, 0)$, on a donc :

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\partial}{\partial \theta^j} V_i(\theta_n, \delta_n) &= E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \theta^i} f_\theta(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_\theta(Y_0)}{\sigma^2(Y_0)} \right) \right] du \\ &\quad - \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \left(\frac{\frac{\partial}{\partial \theta^j} f_\theta(Y_0) + \frac{\partial}{\partial \theta^j} f_\theta(Y_\delta)}{\sigma^2(Y_0)} \right) \right] \\ &= \frac{\partial}{\partial \theta^j} V_i(\theta, 0) \end{aligned}$$

On a montré que, pour tout $\delta \geq 0$, $\theta \rightarrow V_i(\theta, \delta)$ admet p dérivées partielles continues sur $U(\Theta) \times R^+$.

On montrerait de manière similaire que sous (H 2), pour tout $\delta \geq 0$ et tout couple (i, j) , $\theta \rightarrow \frac{\partial}{\partial \theta^j} V_i(\theta, \delta)$ admet p dérivées partielles en tout point de $U(\Theta)$ et que ces dérivées partielles sont continues sur $U(\Theta) \times R^+$.

Pour établir (1.21), il suffit de reprendre le développement précédemment établi pour $\frac{\partial}{\partial \theta^j} V_i(\theta, \delta)$ lorsque $\delta > 0$. On a ainsi pour $\theta = \theta_0$:

$$\begin{aligned} \forall \delta > 0, \quad \frac{\partial}{\partial \theta^j} V_i(\theta_0, \delta) &= -\frac{\delta^2}{2} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \theta^i} f_{\theta_0}(Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{\delta u})}{\sigma^2(Y_0)} \right] du \\ &\quad - \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \left(\frac{\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) + \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_\delta)}{\sigma^2(Y_0)} \right) \right] \end{aligned}$$

$$\text{Or, } \lim_{\delta \rightarrow 0} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^i \partial \theta^j} f_{\theta_0} (Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{\delta u})}{\sigma^2(Y_0)} \right] du = \frac{1}{6} E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^i \partial \theta^j} f_{\theta_0} (Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right].$$

Donc :

$$\forall \delta > 0, \frac{\partial}{\partial \theta^j} V_i (\theta_0, \delta) = -\frac{1}{2} (M_{i,j} (\delta) + M_{i,j} (0)) + O(\delta^2)$$

3. D'après le lemme 4, on sait que, pour tout $\theta \in U(\Theta)$, $\delta \rightarrow V_i(\theta, \delta)$ est dérivable sur R_*^+ avec :

$$\begin{aligned} \forall \delta > 0, V_i'(\theta, \delta) &= -\frac{1}{\delta^2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(Y_{\delta} - Y_0)}{\sigma^2(Y_0)} \right] \\ &+ \frac{1}{\delta} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{f_{\theta_0}(Y_{\delta})}{\sigma^2(Y_0)} \right] \\ &- \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(A_{\theta_0} f_{\theta})(Y_{\delta})}{\sigma^2(Y_0)} \right] \end{aligned}$$

V_i' étant de plus continue sur $U(\Theta) \times R_*^+$.

Montrons que, pour tout $\theta \in U(\Theta)$, $\delta \rightarrow V_i(\theta, \delta)$ est également dérivable à droite en 0. Utilisant (1.3) et la formule de Ito, on obtient :

$$\begin{aligned} \frac{1}{\delta} (V_i(\theta, \delta) - V_i(\theta, 0)) &= -\frac{\delta}{2} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{u\delta})}{\sigma^2(Y_0)} \right] du \\ &+ \frac{1}{2\delta} \int_0^{\delta} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{A_{\theta_0}(f_{\theta_0} - f_{\theta})(Y_u)}{\sigma^2(Y_0)} \right] du \end{aligned}$$

Par le lemme 4, les fonctions précédentes sont continues sur R^+ . On obtient donc :

$$\lim_{h \rightarrow 0} \frac{1}{\delta} (V_i(\theta, \delta) - V_i(\theta, 0)) = \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{A_{\theta_0}(f_{\theta_0} - f_{\theta})(Y_0)}{\sigma^2(Y_0)} \right]$$

On en déduit que $V_i(\theta, 0^+)$ existe et est égale à $\frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{A_{\theta_0}(f_{\theta_0} - f_{\theta})(Y_0)}{\sigma^2(Y_0)} \right]$.

Par la décomposition du trapèze, on a pour tout $(\theta, \delta) \in U(\Theta) \times R_*^+$:

$$\begin{aligned} V_i'(\theta, \delta) &= \frac{1}{2\delta} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(f_{\theta_0}(Y_{\delta}) - f_{\theta_0}(Y_0))}{\sigma^2(Y_0)} \right] \\ &- \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(A_{\theta_0} f_{\theta})(Y_{\delta})}{\sigma^2(Y_0)} \right] - \frac{1}{\delta^2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{\eta_0}{\sigma^2(Y_0)} \right] \\ &= \frac{1}{2\delta} \int_0^{\delta} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(A_{\theta_0} f_{\theta_0})(Y_u)}{\sigma^2(Y_0)} \right] du \\ &- \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(A_{\theta_0} f_{\theta})(Y_{\delta})}{\sigma^2(Y_0)} \right] \\ &- \frac{\delta}{2} \int_0^1 (1-u) u E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta} (Y_0) \frac{(A_{\theta_0}^2 f_{\theta_0})(Y_{u\delta})}{\sigma^2(Y_0)} \right] du \end{aligned}$$

D'après le lemme 4, les fonctions qui interviennent dans ce développement sont continues sur $U(\Theta) \times \mathbb{R}^+$. Donc, Si $((\theta_n, \delta_n))_n \subset U(\Theta) \times \mathbb{R}_*^+$ est une suite qui converge vers $(\theta, 0)$, on obtient que :

$$\begin{aligned} \lim_{n \rightarrow \infty} V_i'(\theta_n, \delta_n) &= \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(A_{\theta_0} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] - \frac{1}{2} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(A_{\theta_0} f_{\theta})(Y_0)}{\sigma^2(Y_0)} \right] \\ &= V_i'(\theta, 0^+) \end{aligned}$$

De 1. 2. et 3., il résulte que V_i est de classe $C^{2,1}$ sur $U(\Theta) \times \mathbb{R}^+$.

4. D'après le lemme 4 et l'hypothèse (H 2.5.2), $\theta \rightarrow V_i'(\theta, 0^+)$ est de classe C^1 sur $U(\Theta)$.

D Démonstration du lemme 7

Utilisant le lemme 5, on sait que, pour tout $\delta > 0$ fixé et tout couple $(i, j) \in \{1, \dots, p\}^2$, $\theta \rightarrow \frac{\partial}{\partial \theta^j} V_i(\theta, \delta)$ est de classe C^1 sur $U(\Theta)$. Par application de la formule de Taylor-Young, on a :

$$\frac{\partial}{\partial \theta^j} V_i(\theta_\delta, \delta) = \frac{\partial}{\partial \theta^j} V_i(\theta_0, \delta) + \sum_{k=1}^p \frac{\partial^2}{\partial \theta^k \partial \theta^j} V_i(\theta_0, \delta) (\theta_\delta^k - \theta_0^k) + \varepsilon(\theta_\delta - \theta_0) \|\theta_\delta - \theta_0\|_2 \quad (1.35)$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

D'après (1.21), on a :

$$\frac{\partial}{\partial \theta^j} V_i(\theta_0, \delta) = -\frac{1}{2} (M_{i,j}(\delta) + M_{i,j}(0)) + O(\delta^2)$$

Or, d'après le lemme 1, on peut faire un développement limité de $M_{i,j}$ au voisinage de 0 :

$$\frac{\partial}{\partial \theta^j} V_i(\theta_0, \delta) = -M_{i,j}(0) - \frac{\delta}{2} M'_{i,j}(0^+) + o(\delta) \quad (1.36)$$

Par le lemme 5, on sait que $\delta \rightarrow \frac{\partial^2}{\partial \theta^k \partial \theta^j} V_i(\theta_0, \delta)$ est continue sur R^+ . On en déduit que :

$$\lim_{\delta \rightarrow 0} \frac{\partial^2}{\partial \theta^k \partial \theta^j} V_i(\theta_0, \delta) = \frac{\partial^2}{\partial \theta^k \partial \theta^j} V_i(\theta_0, 0) \quad (1.37)$$

On sait enfin que $\theta_\delta - \theta_0 = O(\delta^2)$. D'après (1.35), (1.36) et (1.37), on a :

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} \left[\frac{\partial}{\partial \theta^j} V_i(\theta_\delta, \delta) + M_{i,j}(0) \right] = -\frac{1}{2} M'_{i,j}(0^+)$$

Donc :

$$H_{\theta_0}^\delta = -H_{\theta_0} - \frac{\delta}{2} M_{\theta_0} + o(\delta)$$

D'après (H 2.6), H_{θ_0} est inversible. Or, le développement précédent montre en particulier que $\delta \rightarrow H_{\theta_0}^\delta$ est continue en 0. On en déduit donc qu'il existe δ_1 , $0 < \delta_1 \leq \delta_0$, tel que pour tout δ , $0 < \delta \leq \delta_1$, $H_{\theta_0}^\delta$ est inversible.

E Démonstration du lemme 8 : développement de $\Gamma_\delta(\theta_\delta)$

Pour tout $0 < \delta \leq \delta_0$ et pour tout couple $(i, j) \in \{1, \dots, p\}$, $[\Gamma_\delta(\theta_\delta)]_i^j$ est définie par :

$$[\Gamma_\delta(\theta_\delta)]_i^j = [T_\delta(\theta_\delta)]_i^j + 2[R_\delta(\theta_\delta)]_i^j$$

avec :

$$[T_\delta(\theta_\delta)]_i^j = E_{\theta_0} \left[\delta \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta_\delta) \mathcal{V}_\delta^j(Y_\delta, Y_0, \theta_\delta) \right]$$

et, si on note $\mathcal{V}_\delta^j(\theta_\delta) = \mathcal{V}_\delta^j(\cdot, \cdot, \theta_\delta)$

$$\begin{aligned} [R_\delta(\theta_\delta)]_i^j &= E_{\theta_0} \left[\delta \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta_\delta) G_\delta^j(Y_\delta) \right] \\ G_\delta^j &= U_\delta \left(\mathcal{V}_\delta^j(\theta_\delta) \right) \end{aligned}$$

U_δ étant définie dans le lemme 3.

Nous commencerons par montrer que $T_\delta(\theta_\delta) = H_{\theta_0} + \frac{\delta}{2} P_{\theta_0} + o(\delta)$. Par application du lemme 4, on montre que $\theta \rightarrow T_\delta(\theta)$ est de classe C^3 sur $\bar{U}(\Theta)$, ses dérivées étant obtenues par dérivation sous le signe somme. Par application de la formule de Taylor-Young, on obtient:

$$[T_\delta(\theta_\delta)]_i^j = [T_\delta(\theta_0)]_i^j + \sum_{k=1}^p \left(\frac{\partial}{\partial \theta^k} [T_\delta(\theta_0)]_i^j \right) (\theta_\delta^k - \theta_0^k) + \varepsilon(\theta_\delta - \theta_0) \|\theta_\delta - \theta_0\|_2$$

où $\varepsilon(h) \rightarrow 0$ lorsque $h \rightarrow 0$. Si on montre que :

1. $T_\delta(\theta_0) = H_{\theta_0} + \frac{\delta}{2} P_{\theta_0} + o(\delta)$.
2. Pour $k = 1, p$:

$$\begin{aligned} \frac{\partial}{\partial \theta^k} [T_\delta(\theta_0)]_i^j &= E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^k \partial \theta^i} f_{\theta_0}(Y_0) \frac{\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right] \\ &+ E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^k \partial \theta^j} f_{\theta_0}(Y_0) \frac{\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right] + o(1) \end{aligned}$$

et, compte tenu du fait que $\theta_\delta - \theta_0 = O(\delta^2)$, nous aurons le développement annoncé. La démonstration de 2. reposant sur le même type d'argument que 1., nous nous contenterons de montrer 1. D'après la décomposition (1.3), on a :

$$\begin{aligned} [T_\delta(\theta_0)]_i^j &= \frac{1}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \left(\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \right) \left(\frac{\tilde{W}_0}{\sigma^2(Y_0)} \right)^2 \right] \\ &+ \frac{1}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \left(\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \right) \left(\frac{\eta_0}{\sigma^2(Y_0)} \right)^2 \right] \\ &+ \frac{2}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \left(\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \right) \frac{\eta_0 \tilde{W}_0}{\sigma^2(Y_0)} \right] \end{aligned}$$

Etudions chacun des termes intervenant dans la somme précédente.

• Montrons que :

$$\begin{aligned}
& \frac{1}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \left(\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \right) \left(\frac{\tilde{W}_0}{\sigma^2(Y_0)} \right)^2 \right] = M_{i,j}(0) + \frac{\delta}{2} P'_{i,j}(0^+) + o(\delta) \\
& E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \left(\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \right) \left(\frac{\tilde{W}_0}{\sigma^2(Y_0)} \right)^2 \right] \\
& = \int_0^\delta P_{i,j}(v) dv + 2 \int_0^\delta \left(\frac{\delta}{2} - v \right) Q_{i,j}(v) dv \\
& + \delta^3 \int_0^1 \left(\frac{1}{2} - v \right)^2 E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\sigma f'_{\theta_0})^2(Y_{v\delta})}{\sigma^4(Y_0)} \right] dv \\
& - \delta^3 \int_0^1 (1-v)v E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\sigma^2(A_{\theta_0} f_{\theta_0}))'(Y_{v\delta})}{\sigma^4(Y_0)} \right] dv \\
& - \frac{\delta^4}{2} \int_0^1 v E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\sigma^2 f'_{\theta_0}(A_{\theta_0} f_{\theta_0}))'(Y_{v\delta})}{\sigma^4(Y_0)} \right] dv \\
& + \frac{\delta^4}{2} \int_0^1 (3v^2 - 2v^3) E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\sigma^2 f'_{\theta_0}(A_{\theta_0} f_{\theta_0}))'(Y_{v\delta})}{\sigma^4(Y_0)} \right] dv \\
& + \frac{\delta^5}{4} \int_0^1 (1-v)^2 v^2 E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\sigma(A_{\theta_0} f_{\theta_0}))'(Y_{v\delta})^2}{\sigma^4(Y_0)} \right] dv
\end{aligned}$$

Le lemme 1 implique que l'on peut faire un développement de $t \rightarrow \int_0^t P_{i,j}(v) dv$ au voisinage de 0. On obtient ainsi :

$$\begin{aligned}
\int_0^\delta P_{i,j}(v) dv &= \delta P_{i,j}(0) + \frac{\delta^2}{2} P'_{i,j}(0^+) + o(\delta^2) \\
&= \delta M_{i,j}(0) + \frac{\delta^2}{2} P'_{i,j}(0^+) + o(\delta^2)
\end{aligned}$$

Par les mêmes arguments, on a aussi :

$$\int_0^\delta \left(\frac{\delta}{2} - v \right) Q_{i,j}(v) dv = o(\delta^2)$$

De plus, les fonctions qui interviennent dans les autres intégrales sont continues sur R^+ et donc bornées sur $[0, \delta_0]$. On en déduit que le reste du développement de $T_\delta(\theta_0)$ est en $o(\delta^2)$. On obtient donc le résultat annoncé.

- Par deux applications successives de l'inégalité de Schwarz, on montre que :

$$\begin{aligned} & \left| E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right) \left(\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \right) \left(\frac{\eta_0}{\sigma^2(Y_0)} \right)^2 \right] \right| \\ & \leq \delta^6 C^i C^j \sup_{t \in [0, \delta_0]} \left(E_{\theta_0} \left[\left(\frac{|A_{\theta_0}^2 f_{\theta_0}(Y_t)|}{\sigma(Y_0)} \right)^4 \right] \right)^{\frac{1}{2}} \end{aligned}$$

$$\text{où } C^i = \left(E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \right)^4 / \sigma^4(Y_0) \right] \right)^{\frac{1}{4}}$$

- Enfin, on montre aisément grâce à l'inégalité de Schwarz et aux résultats précédents que le double produit restant est un $o(\delta^3)$.

On a donc le développement attendu pour $[T_\delta(\theta_0)]_i^j$.

Pour obtenir le résultat annoncé, il nous reste à montrer que : $R_\delta(\theta_\delta) = o(\delta)$.

Notons $v_\delta^j(\theta_\delta)(x) = E_{\theta_0} [\mathcal{V}_\delta^j(Y_\delta, Y_0, \theta_\delta) \mid Y_0 = x]$. D'après la propriété 1, on a :

$$\begin{aligned} \mathcal{V}_\delta^j(Y_\delta, Y_0, \theta_\delta) &= \frac{\partial}{\partial \theta^j} f_{\theta_\delta}(Y_0) \left[\frac{f_{\theta_0}(Y_0) - f_{\theta_\delta}(Y_0)}{\sigma^2(Y_0)} + \frac{f_{\theta_0}(Y_\delta) - f_{\theta_\delta}(Y_\delta)}{\sigma^2(Y_0)} \right] \\ &+ \frac{\partial}{\partial \theta^j} f_{\theta_\delta}(Y_0) \left[\frac{\eta_0}{\delta \sigma^2(Y_0)} + \frac{\tilde{W}_0}{\delta \sigma^2(Y_0)} \right] \end{aligned}$$

Notons $D_\theta f_{\theta_0}$ le gradient de f par rapport à θ calculé en θ_0 . Compte tenu de l'écart entre θ_δ et θ_0 , on obtient par application de la formule de Taylor-Young :

$$\begin{aligned} \mathcal{V}_\delta^j(Y_\delta, Y_0, \theta_\delta) &= \delta^2 \frac{\partial}{\partial \theta^j} f_{\theta_\delta}(Y_0) \left[\frac{B_{\theta_0} \cdot D_\theta f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} + \frac{B_{\theta_0} \cdot D_\theta f_{\theta_0}(Y_\delta)}{\sigma^2(Y_0)} \right] + o(\delta^2) \\ &+ \frac{\partial}{\partial \theta^j} f_{\theta_\delta}(Y_0) \left[\frac{\eta_0}{\delta \sigma^2(Y_0)} + \frac{\tilde{W}_0}{\delta \sigma^2(Y_0)} \right] \end{aligned}$$

D'où :

$$\begin{aligned} v_\delta^j(\theta_\delta)(x) &= \delta^2 \frac{\partial}{\partial \theta^j} f_{\theta_\delta}(x) \left[\frac{B_{\theta_0} \cdot D_\theta f_{\theta_0}(x)}{\sigma^2(x)} + \frac{B_{\theta_0} \cdot \pi_{\theta_0}^\delta(D_\theta f_{\theta_0})(x)}{\sigma^2(x)} \right] + o(\delta^2) \\ &- \delta^2 \frac{\partial}{\partial \theta^j} f_{\theta_\delta}(x) \frac{\int_0^1 (1-u) u \pi_{\theta_0}^{u\delta}(A_{\theta_0}^2 f_{\theta_0})(x) du}{2\sigma^2(x)} \end{aligned}$$

Il existe donc une constante K finie telle que : $\|v_\delta^j(\theta_\delta)\|_{L^2(\mu_{\theta_0})} \leq K\delta^2$.

De plus, par définition de θ_δ , on a : $\mu_{\theta_0}(v_\delta^j(\theta_\delta)) = 0$. D'après la propriété de mélange (1.8), on obtient :

$$\|G_\delta^j\|_{L^2(\mu_{\theta_0})} \leq \sum_{k=0}^{+\infty} \|\pi_{\theta_0}^{k\delta}(v_\delta^j(\theta_\delta))\|_{L^2(\mu_{\theta_0})} \leq K \frac{\delta^2}{1 - \exp(-\lambda\delta)}$$

On a donc :

$$\left\| G_\delta^j \right\|_{L^2(\mu_{\theta_0})} \leq O(\delta) \quad (1.38)$$

De plus, d'après la première partie de la démonstration, on sait que :

$$\sqrt{\delta} \left\| \mathcal{V}_\delta^i(\theta_\delta) \right\|_{L^2(\mu_{\theta_0})} = E_{\theta_0} \left[\delta \mathcal{V}_\delta^i(Y_\delta, Y_0, \theta_\delta)^2 \right]^{\frac{1}{2}} = \sqrt{M_{i,i}(0) + O(\delta)} \quad (1.39)$$

Or : $[R_\delta]_i^j \leq \sqrt{\delta} \sqrt{\delta} \left\| \mathcal{V}_\delta^i(\theta_\delta) \right\|_{L^2(\mu_{\theta_0})} \left\| G_\delta^j \right\|_{L^2(\mu_{\theta_0})}$. Utilisant (1.38) et (1.39), on déduit que $[R_\delta]_i^j \leq O\left(\delta^{\frac{3}{2}}\right)$ et donc le résultat annoncé.

F Schéma d'approximation d'Euler

Présentons brièvement les résultats sur l'estimation déduite du schéma d'Euler. Ces résultats étendent ceux de [5]. Rappelons tout d'abord la propriété d'approximation du schéma d'Euler pour $Y = (Y_t)_{t \geq 0}$, solution de (1.1)

Propriété 3

Si f est de classe C^2 sur R et σ est continue sur R . Alors :

$$\forall t \geq 0, Y_{t+\delta} = Y_t + \delta f(Y_t) + \eta_t + \tilde{W}_t \quad (1.40)$$

avec :

$$\begin{aligned} \eta_t &= \int_t^{t+\delta} (t + \delta - v) Af(Y_v) dv \\ \tilde{W}_t &= \int_t^{t+\delta} \sigma(Y_v) \left[1 + (t + \delta - v) f'(Y_v) \right] dW_v \end{aligned}$$

Si de plus, $E \left(\int_0^t [f' \sigma]^2(Y_v) dv \right) < +\infty$ et $E \left(\int_0^t \sigma^2(Y_v) dv \right) < \infty$, alors $(\tilde{W}_t)_{t \geq 0}$ est une suite de variables de carré intégrable telles que $E(\tilde{W}_t | \mathcal{F}_t) \stackrel{p.s.}{=} 0$

L'estimateur des moindres carrés associés au schéma d'approximation d'Euler est donné par :

$$\check{\theta}_n = \inf_{\theta \in \Theta} K_n(\theta), \quad K_n(\theta) = \frac{1}{n} \sum_{k=1}^n \left(\frac{Y_{k\delta} - Y_{(k-1)\delta} - \delta f_\theta(Y_{(k-1)\delta})}{\delta \sigma(Y_{(k-1)\delta})} \right)^2$$

On note : $K_\delta(\theta) = E_{\theta_0} \left[\left(\frac{Y_\delta - Y_0 - \delta f_\theta(Y_0)}{\delta \sigma(Y_0)} \right)^2 \right]$. Les conditions d'obtention des propriétés asymptotiques de $\check{\theta}_n$ sont les suivantes :

Hypothèses H 4 : Schéma d'Euler et Moindres Carrés

H 4.1 (H 2.1) est maintenue avec f_{θ_0} et σ de classe C^2 sur R .

H 4.2 -3 -4 -6 -7 On reprend à l'identique les hypothèses **(H 2.2)**, **(H 2.3)**, **(H 2.4)**, **(H 2.6)** et **(H 2.7)**.

H 4.5 (H 2.5) est remplacée par les hypothèses suivantes :

1. Pour tout x , $\theta \rightarrow f_\theta(x)$ est de classe C^3 sur Θ et cette fonction, ainsi que ses dérivées partielles successives jusqu'à l'ordre 3 sont à croissance polynomiale en x uniformément en θ sur Θ .
2. f'_{θ_0} , $A_{\theta_0} f_{\theta_0}$, σ' et σ'' sont à croissance polynomiale en x .

Sous (H 4), K_δ sont de classe C^3 sur Θ , ses dérivées partielles étant obtenues par dérivation sous E_{θ_0} . On note ${}^t \dot{K}_\delta$ le gardient de K_δ et \ddot{K}_δ sa différentielle à l'ordre 2 :

$${}^t \dot{K}_\delta(\theta) = \left(-2E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \left(\frac{Y_\delta - Y_0 - \delta f_\theta(Y_0)}{\delta \sigma^2(Y_0)} \right) \right] \right)_i \quad (1.41)$$

On définit de plus :

$${}^t\check{N}_{\theta_0} = \left(E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{A_{\theta_0} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right] \right)_i \quad (1.42)$$

Théorème 6 *Sous (H 4), pour H_{θ_0} définie en (H 4.6) et \check{N}_{θ_0} en (1.42), il existe $\delta_1 > 0$ tel que pour tout δ , $0 < \delta \leq \delta_1$:*

1. *Il existe un unique $\check{\theta}_\delta \in \overset{\circ}{\Theta}$ qui vérifie $\dot{K}_\delta(\check{\theta}_\delta) = 0$ et*

$$\check{\theta}_\delta - \theta_0 = \frac{\delta}{2} H_{\theta_0}^{-1} \check{N}_{\theta_0} + o(\delta), \quad \ddot{K}_\delta(\check{\theta}_\delta) = H_{\theta_0} + O(\delta)$$

2. *Lorsque $n \rightarrow +\infty$, on a :*

$$\check{\theta}_n \xrightarrow{P_{\theta_0}} \check{\theta}_\delta$$

$$\sqrt{n\delta} (\check{\theta}_n - \check{\theta}_\delta) \xrightarrow{\mathcal{D}(P_{\theta_0})} \mathcal{N}_p(0, \check{V}_\delta(\theta_0)), \quad \check{V}_\delta(\theta_0) = H_{\theta_0}^{-1} [I_p + o(1)]$$

Ce théorème étend les résultats obtenus par Florens-Zmirou [5] : pour δ petit, on prouve en effet l'existence et l'unicité de θ_δ . Pour le montrer, on procède comme au § 4.1. Pour tout $i = 1, p$, on définit \overline{K}_i sur $U(\Theta) \times R$ par :

$$\begin{aligned} \forall \theta \in U(\Theta), \text{ si } \delta > 0, \quad \overline{K}_i(\theta, \delta) &= \dot{K}_\delta^i(\theta) \\ \overline{K}_i(\theta, 0) &= -2E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_\theta(Y_0)}{\sigma^2(Y_0)} \right) \right] \\ \text{si } \delta < 0, \quad \overline{K}_i(\theta, \delta) &= \overline{K}_i(\theta, 0) - \delta E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \frac{A_{\theta_0} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right] \end{aligned}$$

On montre alors que $\overline{K} = {}^t(\overline{K}_1, \dots, \overline{K}_p)$ est de classe C^1 sur $U(\Theta) \times R$. Il suffit alors d'appliquer le théorème des fonctions implicites avec les hypothèses (H 4.6) et (H 4.7).

Chapter 2

Schéma d'approximation adapté à l'ordre p et estimation du paramètre de dérive d'une diffusion.

2.1 Introduction

Considérons l'équation différentielle stochastique suivante :

$$dY_t = f(\theta_0, Y_t) dt + \sigma(Y_t) dW_t, \quad Y_0 = x \quad (2.1)$$

où f et σ sont connues et à valeurs dans R et θ_0 est un paramètre inconnu de R^d . L'estimation du paramètre de dérive à partir d'une observation discrète à pas $\delta > 0$ fixé du processus solution $(Y_t)_{t \geq 0}$ peut être abordée sous des angles aussi divers que les fonctions d'estimations ([1], [9], [13], [8]), la construction de contrastes basés sur des schémas d'approximation de la diffusion ([4], [14]), la maximisation de la log-vraisemblance ([10]) ou encore d'une discrétisation de celle-ci ([5], [15]). Bien que l'estimateur du maximum de vraisemblance possède en théorie de bonnes propriétés asymptotiques (consistance et efficacité asymptotique à la vitesse $\sqrt{n\delta}$, lorsque $n \rightarrow \infty$, voir [3]), les densités de transition ne sont en général pas explicites et cette méthode d'estimation ne peut alors être mise en pratique. Signalons cependant une procédure numérique d'obtention de ces transitions, mise en oeuvre par Pedersen et l'étude théorique des estimateurs associés ([11], [12]).

Une solution consiste à approximer les densités de transition par des densités gaussiennes d'espérance $E_\theta [Y_{k\delta} | Y_{(k-1)\delta}]$ et de variance $Var_\theta [Y_{k\delta} | Y_{(k-1)\delta}]$, ce qui ne fait que déplacer le problème dans la mesure où ces quantités ne sont pas explicites. Cependant, pour une régularité à l'ordre $2p$ (dans un sens que l'on précisera) de f et σ , on peut écrire un schéma de discrétisation adapté à l'ordre p pour $h(Y_t)$ dès que h est $2(p+1)$ fois continûment différentiable. En utilisant ce résultat avec $h(x) = x$ et $h(x) = x^2$, on approche, sous certaines conditions, les moyenne et variance conditionnelles à un $o(\delta^p)$ près.

Cette méthodologie est celle adoptée par M. Kessler dans [7] pour un pas d'observation δ_n non constant. Par minimisation du contraste gaussien, il obtient une estimation consistante et asymptotiquement efficace de θ_0 (mais aussi du paramètre de diffusion σ_0), lorsque $n\delta_n \rightarrow \infty$ et $n\delta_n^2 \rightarrow 0$. Dans notre démarche, le pas d'observation $\delta > 0$ est proche de 0, mais fixé. La diffusion étant étudiée sur $[0, T]$, on dispose ainsi d'un vecteur d'observations $(Y_{k\delta})_{k=0, n}$ avec $n\delta = T$.

Le paragraphe 2 sera consacré à la présentation du schéma de discrétisation à l'ordre p de $h(Y_t)$. Dans une troisième partie, nous déduirons une approximation à l'ordre p de l'espérance conditionnelle sur laquelle nous nous baserons pour la construction du contraste. Pour simplifier, l'approximation requise pour la variance conditionnelle est d'ordre 1. Nous montrerons alors que l'estimateur du minimum de contraste associé est asymptotiquement efficace pour $\theta_{p,\delta}$ à un facteur $1 + O(\delta)$, $\theta_{p,\delta}$ approchant θ_0 avec un biais explicite en δ^p . Dans le paragraphe 4, une étude expérimentale par simulation nous permettra de vérifier ces résultats pour deux modèles de diffusion particuliers : les modèles de Ornstein-Uhlenbeck et de Cox-Ingersoll-Ross.

2.2 Schéma d'approximation à l'ordre p pour une fonctionnelle de la solution

Soit l'e.d.s.

$$dY_t = f(Y_t) dt + \sigma(Y_t) dW_t \quad (2.2)$$

Notons $\mathcal{A} = f \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2}$ le générateur associé à l'e.d.s. (2.2) et \mathcal{A}^l le $l^{\text{ième}}$ itéré de \mathcal{A} . Soient $M^{c,loc}$ l'espace des martingales locales continues et $(\mathcal{F}_t)_{t \geq 0}$ la filtration naturelle associée à (2.2) [6]. De plus, $l = r, p$ se lit $l \in \{r, \dots, p\}$ et C^m est l'ensemble des fonctions m fois continûment différentiables.

Propriété 1

Soient h une fonction de R dans R et p un entier naturel non nul. Supposons que :

1. $h \in C^{2(p+1)}$, f et $\sigma \in C^{2p}$.
2. $\forall l = 1, p, \left(\int_0^t [(\mathcal{A}^l h)' \sigma](Y_v) dW_v \right)_{t \geq 0} \in M^{c,loc}$

Alors :

$$\forall t \geq 0, \quad h(Y_{t+\delta}) = h(Y_t) + \sum_{l=1}^p \frac{\delta^l}{l!} \mathcal{A}^l h(Y_t) + \eta_t^p + \xi_t^p$$

avec :

$$\begin{aligned} \eta_t^p &= \int_t^{t+\delta} \frac{(t+\delta-v)^p}{p!} \mathcal{A}^{p+1} h(Y_v) dv \\ \xi_t^p &= \int_t^{t+\delta} \sigma(Y_v) \left[\sum_{l=0}^p \frac{(t+\delta-v)^l}{l!} (\mathcal{A}^l h)'(Y_v) \right] dW_v \end{aligned}$$

3. Si de plus $\forall l = 0, p, E \left(\int_0^t [(\mathcal{A}^l h)' \sigma]^2(Y_v) dv \right) < +\infty$, $(\xi_t^p)_{t \geq 0}$ sont des variables de carré intégrable telles que $E[\xi_t^p | \mathcal{F}_t] \stackrel{p.s.}{=} 0$.

Démonstration :

Soient $\left(\int_0^t J_v dW_v \right)_{t \geq 0}$ un élément de $M^{c,loc}$ et f une fonction continue sur R . On montre par application de la formule de Ito que :

$$\int_t^{t+\delta} \left(\int_v^{t+\delta} f(u) du \right) J_v dW_v = \int_t^{t+\delta} \left(\int_t^u J_v dW_v \right) f(u) du$$

On peut ainsi intervertir l'ordre d'intégration entre l'intégrale de Riemann et l'intégrale stochastique.

Par application de la formule de Ito, on a :

$$h(Y_{t+\delta}) = h(Y_t) + \int_t^{t+\delta} \mathcal{A}h(Y_v) dv + \int_t^{t+\delta} [h'\sigma](Y_v) dW_v$$

- Montrons la propriété pour $p = 1$:

On pose :

$$\begin{aligned} R_t &= \int_t^{t+\delta} \mathcal{A}h(Y_v) dv - \delta \mathcal{A}h(Y_t) \\ &= \int_t^{t+\delta} \left(\int_t^v \mathcal{A}^2 h(Y_u) du \right) dv + \int_t^{t+\delta} \left(\int_t^v [(\mathcal{A}h)'\sigma](Y_u) dW_u \right) dv \end{aligned}$$

Par interversion de l'ordre d'intégration, on obtient :

$$R_t = \int_t^{t+\delta} (t + \delta - u) \mathcal{A}^2 h(Y_u) du + \int_t^{t+\delta} (t + \delta - u) [(\mathcal{A}h)'\sigma](Y_u) dW_u$$

D'où :

$$\begin{aligned} h(Y_{t+\delta}) - h(Y_t) &= \delta \mathcal{A}h(Y_t) + R_t + \int_t^{t+\delta} [h'\sigma](Y_v) dW_v \\ &= \delta \mathcal{A}h(Y_t) + \eta_t^1 + \xi_t^1 \end{aligned}$$

où η_t^1 et ξ_t^1 sont donnés dans l'énoncé.

- Supposons que la propriété soit vraie au rang $p - 1$ et que les hypothèses sont vérifiées au rang p , on a donc :

$$h(Y_{t+\delta}) - h(Y_t) = \sum_{l=1}^{p-1} \frac{\delta^l}{l!} \mathcal{A}^l h(Y_t) + \eta_t^{p-1} + \xi_t^{p-1}$$

avec :

$$\begin{aligned} \eta_t^{p-1} &= \int_t^{t+\delta} \frac{(t + \delta - v)^{p-1}}{(p-1)!} \mathcal{A}^p h(Y_v) dv \\ \xi_t^{p-1} &= \int_t^{t+\delta} \sigma(Y_v) \left[\sum_{l=0}^{p-1} \frac{(t + \delta - v)^l}{l!} (\mathcal{A}^l h)'\sigma(Y_v) \right] dW_v \end{aligned}$$

Notons : $R_t = \eta_t^{p-1} - \frac{\delta^p}{p!} \mathcal{A}^p h(Y_t)$. On a :

$$\begin{aligned} R_t &= \int_t^{t+\delta} \frac{(t + \delta - v)^{p-1}}{(p-1)!} (\mathcal{A}^p h(Y_v) - \mathcal{A}^p h(Y_t)) dv \\ &= \int_t^{t+\delta} \frac{(t + \delta - v)^{p-1}}{(p-1)!} \left(\int_t^v \mathcal{A}^{p+1} h(Y_u) du \right) dv \\ &\quad + \int_t^{t+\delta} \frac{(t + \delta - v)^{p-1}}{(p-1)!} \left(\int_t^v [(\mathcal{A}^p h)'\sigma](Y_u) dW_u \right) dv \end{aligned}$$

Par interversion de l'ordre d'intégration, on obtient :

$$R_t = \int_t^{t+\delta} \frac{(t+\delta-u)^p}{p!} \mathcal{A}^{p+1} h(Y_u) du + \int_t^{t+\delta} \frac{(t+\delta-u)^p}{p!} [(\mathcal{A}^p h)' \sigma](Y_u) dW_u$$

Or :

$$Y_{t+\delta} - Y_t = \sum_{l=1}^p \frac{\delta^l}{l!} \mathcal{A}^l h(Y_t) + R_t + \xi_t^{p-1}$$

De ce qui précède, on déduit que :

$$Y_{t+\delta} - Y_t = \sum_{l=1}^p \frac{\delta^l}{l!} \mathcal{A}^l h(Y_t) + \eta_t^p + \xi_t^p$$

où η_t^p et ξ_t^p ont la même expression que celle donnée dans l'énoncé. .

On a en particulier pour $h(x) = x$:

$$\forall t \geq 0, \quad Y_{t+\delta} - Y_t = \sum_{l=0}^{p-1} \frac{\delta^{l+1}}{(l+1)!} \mathcal{A}^l f(Y_t) + \eta_t^p + \xi_t^p \quad (2.3)$$

avec :

$$\begin{aligned} \eta_t^p &= \int_t^{t+\delta} \frac{(t+\delta-v)^p}{p!} \mathcal{A}^p f(Y_v) dv \\ \xi_t^p &= \int_t^{t+\delta} \sigma(Y_v) \left[1 + \sum_{l=1}^p \frac{(t+\delta-v)^l}{l!} (\mathcal{A}^{l-1} f)'(Y_v) \right] dW_v \end{aligned}$$

Dans le paragraphe qui suit, nous utilisons ce schéma discret et adapté afin d'obtenir un développement en δ à l'ordre p de $E[Y_{t+\delta} | Y_t]$.

2.3 Résultats asymptotiques

Dans toute la suite, la fonction de dérive dépend d'un paramètre θ_0 et on note $f(\theta, \cdot) = f_\theta(\cdot)$. La dérivée de la fonction d'échelle associée à (2.1) est $s(x, \theta) = \exp \left[-2 \int^x \frac{f_\theta(u)}{\sigma^2(u)} du \right]$. Soient

Θ un sous-ensemble compact de R^p tel que $\theta_0 \in \overset{\circ}{\Theta}$ et $U(\Theta)$ un ouvert de R^d qui contient Θ . Si $M = (M_{i,j})_{i,j}$ est une matrice de dimensions $d \times d$, i représente l'indice des lignes et j celui des colonnes.

Par convention, U et V étant deux ouverts respectivement de R^p et de R , on dira que $(\theta, x) \rightarrow g(\theta, x)$ est de classe $C^{m,k}$ sur $U \times V$ si :

1. g est continue sur $U \times V$.
2. Pour tout $x \in V$, $\theta \rightarrow g(\theta, x)$ admet des dérivées partielles jusqu'à l'ordre m en tout point de U et ses dérivées partielles étant continues sur $U \times V$. Pour tout $i = 1, p$, $\frac{\partial}{\partial \theta^i} g(\theta, x)$ représente la $i^{\text{ème}}$ dérivée partielle de $g(\cdot, x)$ au point θ .

3. Pour tout $\theta \in U$, $x \rightarrow g(\theta, x)$ est dérivable jusqu'à l'ordre k sur V et ses dérivées successives sont continues sur $U \times V$. On note alors $g'(\theta, x)$, $g''(\theta, x)$ et $g^{(l)}(\theta, x)$, les dérivées à l'ordre 1, 2 et l de $g(\theta, \cdot)$ au point x .

Par extension, on dira que $(\theta, x) \rightarrow g(\theta, x)$ est de classe $C^{m,1}$ sur $U \times R^+$ si :

- Les conditions 1 et 2 de la définition précédente sont satisfaites sur $U \times R^+$.
- Pour tout $\theta \in U$, $x \rightarrow g(\theta, x)$ est dérivable sur R_*^+ et dérivable à droite en 0 : $g'(\theta, 0)$ représente alors sa dérivée à droite en 0. La fonction $(\theta, x) \rightarrow g'(\theta, x)$ est de plus continue sur $U \times R^+$.

On considère le jeu d'hypothèses général noté **H** :

H 1 f_{θ_0} et σ sont de classe C^{2p} sur R et $\forall x \in R, \sigma^2(x) > 0$.

De plus, il existe $K > 0$ telle que : $\forall x \in R, |f_{\theta_0}(x)| + |\sigma(x)| \leq K(1 + |x|)$.

Sous cette hypothèse, pour toute condition initiale $x \in R$, (2.1) admet une unique solution forte $(Y_t)_{t \geq 0}$.

H 2

$$\int_0^\infty s(x, \theta_0) dx = \int_{-\infty}^0 s(x, \theta_0) dx = \infty \text{ et } \int_{-\infty}^\infty [s(x, \theta_0) \sigma^2(x)]^{-1} dx = C(\theta_0) < \infty$$

Sous (H 1) et (H 2), $(Y_t)_{t \geq 0}$ est alors récurrente positive sur R de loi invariante :

$$\mu_{\theta_0}(dy) = [C(\theta_0) s(y, \theta_0) \sigma^2(y)]^{-1} dy$$

H 3 On note : $c_{\theta_0}(x) = \frac{f_{\theta_0}^2(x)}{\sigma^2(x)} + \frac{f_{\theta_0}'(x)}{2} - \frac{\sigma'(x)f_{\theta_0}(x)}{\sigma(x)} + \frac{(\sigma')^2(x)}{8}$.

On suppose que :

$$\min \left\{ \lim_{x \rightarrow +\infty} c_{\theta_0}(x), \lim_{x \rightarrow -\infty} c_{\theta_0}(x) \right\} = c_{\theta_0} > 0$$

H 4 La loi invariante μ_{θ_0} admet des moments de tous ordres et vérifie :

$$\forall p \geq 0, \int \frac{|x|^p}{\sigma^4(x)} d\mu_{\theta_0}(x) < \infty$$

H 5 - Pour tout $l = 0, p - 1$ et tout $x \in R$, $\theta \rightarrow \mathcal{A}_\theta^l f_\theta(x)$ est de classe C^3 sur Θ et à croissance polynomiale en x uniformément en θ , ainsi que ses dérivées partielles en θ jusqu'à l'ordre 3.

- Pour tout $l = 0, p - 1$, $(\mathcal{A}_{\theta_0}^l f_{\theta_0})'$ et $\mathcal{A}_{\theta_0}^p f_{\theta_0}$ sont à croissance polynomiale en x .

H 6 $H_{\theta_0} = (E_{\theta_0} [\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) / \sigma^2(Y_0)])_{i,j}$ est inversible.

H 7 Si $f_\theta = f_{\theta_0}$ μ_{θ_0} -p.s. alors $\theta = \theta_0$.

On note P_{x,θ_0} la loi du processus solution de (2.1) et de condition initiale x , $\Pi_{\theta_0}^t(x, dy) = P_{x,\theta_0}(Y_t \in dy)$ la probabilité de transition du processus issu de x et $P_{\theta_0}^t = \mu_{\theta_0} \otimes \Pi_{\theta_0}^t$ la loi de (Y_0, Y_t) lorsque le processus est stationnaire. P_{θ_0} est alors la loi du processus stationnaire solution de (2.1) et E_{θ_0} l'espérance sous P_{θ_0} .

L'ensemble des développements ultérieurs s'appuient sur le résultat d'ergodicité et normalité asymptotique qui suit et qui est directement dérivé du théorème 1 et du lemme 2 de Florenz-Zmirou [4]. (voir aussi lemme 3.2 de [1]).

Lemme 1 *Sous (H 1), (H 2) et (H 3), si $g : R^2 \rightarrow R$ vérifie $P_{\theta_0}^\delta(g^2) < \infty$, alors :*

$$\frac{1}{n} \sum_{k=1}^n g(Y_{k\delta}, Y_{(k-1)\delta}) \xrightarrow{L^2(P_{x,\theta_0})} P_{\theta_0}^\delta(g)$$

Si de plus, $P_{\theta_0}^\delta(g) = 0$, on a :

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n g(Y_{k\delta}, Y_{(k-1)\delta}) \xrightarrow{\mathcal{D}(P_{x,\theta_0})} N(0, P_{\theta_0}^\delta(G))$$

où $G(y, x) = g^2(y, x) + 2g(y, x)U_\delta(g)(y)$ et $U_\delta(g)(y) = \sum_{k=1}^{+\infty} \pi_{\theta_0}^{k\delta}(g)(y)$.

La démonstration de ce résultat est basée sur la propriété suivante :

Propriété 2 *Sous (H 2) et (H 3), il existe $\lambda > 0$ telle que, pour toute fonction g vérifiant $\mu_{\theta_0}(g) = 0$ et $\mu_{\theta_0}(g^2) < +\infty$, on a :*

$$\|\pi_{\theta_0}^t(g)\|_{L^2(\mu_{\theta_0})} \leq \|g\|_{L^2(\mu_{\theta_0})} \exp(-\lambda t)$$

Dans la suite, nous utiliserons également le résultat de régularité suivant (lemme 4 de [14]) :

Lemme 2 *Soient $(\theta, x) \rightarrow h(\theta, x)$ et $(\theta, x) \rightarrow g(\theta, x)$ deux fonctions définies sur $U(\Theta) \times R$, à croissance polynomiale en x uniformément en θ sur $U(\Theta)$ et telles que : g est continue sur $U(\Theta) \times R$ et, pour tout $x \in R$, $\theta \rightarrow h(\theta, x)$ est continue sur $U(\Theta)$.*

Sous (H 1), (H 2) et si μ_{θ_0} admet des moments de tous ordres, la fonction définie sur $U(\Theta) \times R^+$ par $(\theta, \delta) \rightarrow m(\theta, \delta) = E_{\theta_0}[h(\theta, Y_0)g(\theta, Y_\delta)]$ est continue sur $U(\Theta) \times R^+$.

Si on suppose de plus que, pour tout $x \in R$, $\theta \rightarrow h(\theta, x)$ est de classe C^1 sur $U(\Theta)$, que g est de classe $C^{1,2}$ sur $U(\Theta) \times R$ et que leurs dérivées partielles ainsi que $\mathcal{A}_{\theta_0}g$ sont à croissance polynomiale en x uniformément en θ sur $U(\Theta)$, alors $(\theta, \delta) \rightarrow m(\theta, \delta)$ est de classe $C^{1,1}$ sur $U(\Theta) \times R^+$ avec :

$$\begin{aligned} \frac{\partial}{\partial \theta^i} m(\theta, \delta) &= E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} h(\theta, Y_0) g(\theta, Y_\delta) + h(\theta, Y_0) \frac{\partial}{\partial \theta^i} g(\theta, Y_\delta) \right] \\ m'(\theta, \delta) &= E_{\theta_0} [h(\theta, Y_0) (\mathcal{A}_{\theta_0}g)(\theta, Y_\delta)] \end{aligned}$$

2.3.1 Construction du contraste

Donnons tout d'abord les résultats d'approximation sur lesquels nous nous basons pour construire la fonction d'estimation. On note :

$$M_p(x, \theta, \delta) = \sum_{l=0}^{p-1} \frac{\delta^{l+1}}{(l+1)!} \mathcal{A}_\theta^l f_\theta(x) \quad (2.4)$$

Lemme 3 Sous (H 1), (H 2), (H 4) et (H 5), pour tout $t > 0$ et tout $\delta > 0$, on a :

$$\begin{aligned} E_{\theta_0} [Y_{t+\delta} | Y_t] &= Y_t + M_p(Y_t, \theta, \delta) + O_{P_{x, \theta_0}}(\delta^{p+1}) \\ \text{Var}_{\theta_0} [Y_{t+\delta} | Y_t] &= \delta \sigma^2(Y_t) + O_{P_{x, \theta_0}}(\delta^2) \end{aligned}$$

Démonstration :

L'approximation à l'ordre p de $E_{\theta_0} [Y_{t+\delta} | Y_t]$ est obtenue en utilisant le schéma (2.3). Compte tenu des hypothèses d'intégrabilité et de croissance polynomiale, on obtient par passage à l'espérance conditionnelle :

$$E_{\theta_0} [Y_{t+\delta} | Y_t] = Y_t + \sum_{l=0}^{p-1} \frac{\delta^{l+1}}{(l+1)!} \mathcal{A}_{\theta_0}^l f_{\theta_0}(Y_t) + E_{\theta_0} [\eta_t^p | Y_t]$$

avec $\eta_t^p = \int_t^{t+\delta} \frac{(t+\delta-v)^p}{p!} \left(\mathcal{A}_{\theta_0}^p f_{\theta_0} \right) (Y_v) dv$ et $\lim_{\delta \rightarrow 0} \frac{1}{\delta^{p+1}} E_{\theta_0} [\eta_t^p | Y_t] = \frac{1}{(p+1)!} \left(\mathcal{A}_{\theta_0}^p f_{\theta_0} \right) (Y_t)$ P_{x, θ_0} p.s..

Pour l'approximation de la variance, on utilise la propriété 1 avec $h(x) = x^2$ et $p = 1$. On obtient ainsi :

$$E_{\theta_0} [Y_{t+\delta}^2 | Y_t] = Y_t^2 + 2\delta Y_t f_{\theta_0}(Y_t) + \delta \sigma^2(Y_t) + O_{P_{x, \theta_0}}(\delta^2)$$

Il suffit alors de remplacer $(E_{\theta_0} [Y_{t+\delta} | Y_t])^2$ par son approximation pour obtenir le résultat annoncé.

Nous nous limitons à une approximation à l'ordre 1 pour la variance conditionnelle dans la mesure où une approximation plus fine n'améliore pas les propriétés asymptotiques de l'estimateur du paramètre de dérive. Le contraste gaussien associé aux approximations précédentes s'écrit alors :

$$U_{p,n}(\theta, \delta) = \frac{1}{n\delta^2} \sum_{k=1}^n \Phi_p(Y_{k\delta}, Y_{(k-1)\delta}, \theta, \delta), \quad \Phi_p(y, x, \theta, \delta) = \frac{(y - x - M_p(x, \theta, \delta))^2}{\sigma^2(x)}$$

Pour montrer que notre cadre d'étude s'inscrit bien dans celui de l'estimation par minimum de contraste, nous allons vérifier successivement les conditions de la définition 3.2.7. de [2]. Notons :

$$U_p(\theta, \delta) = \frac{1}{\delta^2} E_{\theta_0} [\Phi_p(Y_\delta, Y_0, \theta, \delta)]$$

Lemme 4 Sous (H), pour tout $\delta > 0$, $U_{p,n}(\cdot, \delta)$ et $U_p(\cdot, \delta)$ sont de classe C^3 sur $U(\Theta)$, les dérivées partielles successives de $U_p(\theta, \delta)$ étant obtenues par dérivation sous l'espérance. Pour tout $\theta \in U(\Theta)$, $U_{p,n}(\theta, \delta)$ est $\mathcal{F}_{n\delta}$ mesurable et on a :

$$U_{p,n}(\theta, \delta) \xrightarrow{L^2(P_{x, \theta_0})} U_p(\theta, \delta)$$

Démonstration :

(H 4) et (H 5) impliquent que pour $\delta > 0$ et $\theta \in U(\Theta)$, $(y, x) \rightarrow \Phi_p(y, x, \theta, \delta) \in L^2(P_{\theta_0}^\delta)$. Donc, par application du lemme 1, on a :

$$U_{p,n}(\theta, \delta) \xrightarrow{L^2(P_{x, \theta_0})} \frac{1}{\delta^2} E_{\theta_0} [\Phi_p(Y_\delta, Y_0, \theta, \delta)] = U_p(\theta, \delta)$$

Pour tout $\delta > 0$, $U_{p,n}(\theta, \delta)$ est $\mathcal{F}_{n\delta}$ mesurable car $(y, x) \rightarrow \Phi_p(y, x, \theta, \delta)$ est continue et $(Y_t)_{t \geq 0}$ est adapté à la filtration $(\mathcal{F})_t$. $\theta \rightarrow U_{p,n}(\theta, \delta)$ est de classe C^3 sur $U(\Theta)$ car, pour tout $x \in \mathbb{R}$, $\theta \rightarrow M_p(x, \theta, \delta)$ est de classe C^3 sur $U(\Theta)$. On sait de plus que M_p et ses dérivées sont à croissance polynomiale en x uniformément en θ sur $U(\Theta)$. Par application du théorème de Lebesgue, on montre que $\theta \rightarrow U_p(\theta, \delta)$ est de classe C^3 sur $U(\Theta)$, ses dérivées partielles étant obtenues par dérivation sous l'espérance.

Les deux lemmes qui suivent vont nous permettre de prouver que pour $\delta > 0$ dans un voisinage de 0, $\theta \rightarrow U_p(\theta, \delta)$ admet un unique minimum $\theta_{p,\delta} \in \overset{\circ}{\Theta}$. Dans ce but, on définit pour tout $i = 1, d$, la fonction V_i^p telle que :

$$\begin{aligned} \delta > 0, \quad V_i^p(\theta, \delta) &= \frac{\partial}{\partial \theta^i} U_p(\theta, \delta) \\ V_i^p(\theta, 0) &= -2E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2} \right) \right] \\ \delta < 0, \quad V_i^p(\theta, \delta) &= V_i^p(\theta, 0) + \delta R_i^p(\theta) \end{aligned}$$

avec, pour $p \geq 1$:

$$\begin{aligned} R_i^p(\theta) &= -E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] \\ &+ E_{\theta_0} \left[\frac{\frac{\partial}{\partial \theta^i} (\mathcal{A}_{\theta} f_{\theta})(Y_0) f_{\theta}(Y_0) + \frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) (\mathcal{A}_{\theta} f_{\theta})(Y_0)}{\sigma^2(Y_0)} \right] \end{aligned}$$

et :

$$R_i^1(\theta) = -E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right]$$

Lemme 5 *Sous (H), pour tout $i = 1, d$, $(\theta, \delta) \rightarrow V_i^p(\theta, \delta)$ est de classe $C^{2,1}$ sur $U(\Theta) \times \mathbb{R}$. De plus, pour $\delta > 0$:*

$$\frac{\partial}{\partial \theta^i} U_p(\theta_0, \delta) = -\frac{2\delta^p}{(p+1)!} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{(\mathcal{A}_{\theta_0}^p f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] + o(\delta^p) \quad (2.5)$$

$$\frac{\partial^2}{\partial \theta^j \partial \theta^i} U_p(\theta_0, \delta) = 2E_{\theta_0} \left[\frac{\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right] + O(\delta) \quad (2.6)$$

$$\frac{\partial^3}{\partial \theta^k \partial \theta^j \partial \theta^i} U_p(\theta_0, \delta) = 2E_{\theta_0} \left[\frac{\left(\frac{\partial^2}{\partial \theta^k \partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} + \frac{\partial^2}{\partial \theta^k \partial \theta^j} f_{\theta_0} \frac{\partial}{\partial \theta^i} f_{\theta_0} + \frac{\partial^2}{\partial \theta^j \partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^k} f_{\theta_0} \right) (Y_0)}{\sigma^2(Y_0)} \right] + O(\delta) \quad (2.7)$$

La démonstration de ce résultat est donnée en annexe A. L'existence et l'unicité d'un minimum $\theta_{p,\delta}$ pour $U_p(\theta, \delta)$, lorsque δ est dans un voisinage de 0, est alors une conséquence directe du lemme précédent.

Lemme 6 *Sous (H), il existe $\delta_0 > 0$ tel que pour δ , $0 < \delta \leq \delta_0$, $\theta \rightarrow U_p(\theta, \delta)$ admet un unique minimum $\theta_{p,\delta} = \psi(\delta)$ qui vérifie $\lim_{\delta \rightarrow 0} \psi(\delta) = \theta_0$.*

Démonstration :

Utilisant le lemme 5, on sait que $(\theta, \delta) \rightarrow V^p(\theta, \delta) = (V_i^p(\theta, \delta))_{i=1,d}$ est de classe C^1 sur $\overset{\circ}{\Theta} \times R$.

D'après l'hypothèse (H 7), on sait que θ_0 est l'unique point de $\overset{\circ}{\Theta}$ tel que $V^p(\theta_0, 0) = 0$. De plus, le gradient de $\theta \rightarrow V^p(\theta, 0)$ en θ_0 est égale à $2H_{\theta_0}$ et est donc inversible (H 6). Par application du théorème des fonctions implicites, il existe un $\varepsilon > 0$ et $\delta_0 > 0$ tels que pour δ , $|\delta| \leq \delta_0$, il existe un unique $\theta_{p,\delta}$, dans la boule ouverte de centre θ_0 et de rayon ε , vérifiant $V^p(\theta_{p,\delta}, \delta) = 0$ et $\theta_{p,\delta} = \psi(\delta)$. ψ étant dérivable en 0, on a en particulier $\lim_{\delta \rightarrow 0} \psi(\delta) = \theta_0$.

Donc, si δ est tel que $0 < \delta \leq \delta_0$, il existe un unique $\theta_{p,\delta} \in \overset{\circ}{\Theta}$ qui satisfait pour tout $i = 1, d$, $\frac{\partial}{\partial \theta^i} U_p(\theta_{p,\delta}, \delta) = 0$ et $\theta_{p,\delta} = \psi(\delta)$ avec $\lim_{\delta \rightarrow 0} \psi(\delta) = \theta_0$.

D'après les lemmes 4 et 6, pour $\delta \in]0, \delta_0]$, $U_{p,n}(\cdot, \delta)$ est bien un processus de contraste relatif à la fonction de contraste $U_p(\cdot, \delta)$. On définit alors l'estimateur du minimum de contraste $\widehat{\theta}_{p,n}^\delta$ par :

$$\widehat{\theta}_{p,n}^\delta = \inf_{\theta \in \overset{\circ}{\Theta}} U_{p,n}(\theta, \delta)$$

Avant de montrer que $\widehat{\theta}_{p,n}^\delta$ est un estimateur convergent et asymptotiquement normal pour $\theta_{p,\delta}$, nous allons établir que $\theta_{p,\delta}$ approche θ_0 à δ^p près.

2.3.2 Calcul de l'écart entre $\theta_{p,\delta}$ et θ_0

On note :

$$\forall i = 1, d, N_i^p(v) = E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\mathcal{A}_{\theta_0}^p f_{\theta_0}(Y_v)}{\sigma^2(Y_0)} \right], \quad N_{\theta_0}^p = (N_i^p(0))_i \quad (2.8)$$

Théorème 1 *Sous (H), pour H_{θ_0} définie en (H 6) et pour δ , $0 < \delta \leq \delta_0$, on a :*

$$\theta_{p,\delta} = \theta_0 + \frac{\delta^p}{(p+1)!} H_{\theta_0}^{-1} N_{\theta_0}^p + o(\delta^p)$$

Démonstration :

Compte tenu du lemme 5 et par application du théorème des accroissements finis, pour tout $i = 1, d$, il existe $u_i \in]0, 1[$ tel que pour $\tilde{\theta}_\delta^{(i)} = u_i \theta_0 + (1 - u_i) \theta_{p,\delta}$, on a :

$$\frac{\partial}{\partial \theta^i} U_p(\theta_{p,\delta}, \delta) = 0 = \frac{\partial}{\partial \theta^i} U_p(\theta_0, \delta) + \sum_{j=1}^d \frac{\partial^2}{\partial \theta^j \partial \theta^i} U_p(\tilde{\theta}_\delta^{(i)}, \delta) (\theta_{p,\delta}^j - \theta_0^j)$$

Lorsque $\delta \rightarrow 0$, $\theta_{p,\delta} \rightarrow \theta_0$ (lemme 6) et donc $\tilde{\theta}_\delta^{(i)} \rightarrow \theta_0$. Utilisant la continuité de $\frac{\partial^2}{\partial \theta^j \partial \theta^i} U_p(\theta, \delta)$ et (2.6), on obtient :

$$\lim_{\delta \rightarrow 0} \frac{\partial^2}{\partial \theta^j \partial \theta^i} U_p(\tilde{\theta}_\delta^{(i)}, \delta) = 2E_{\theta_0} \left[\frac{\frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right]$$

De plus, par (2.5) :

$$\frac{\partial}{\partial \theta^i} U_p(\theta_0, \delta) = -2 \frac{\delta^p}{(p+1)!} N_i^p(0) + o(\delta^p)$$

On a donc :

$$[H_{\theta_0} + o(1)](\theta_{p,\delta} - \theta_0) = \frac{\delta^p}{(p+1)!} N_{\theta_0}^p + o(\delta^p)$$

Utilisant (H 6), on obtient le résultat annoncé.

2.3.3 Convergence de $\widehat{\theta}_{p,n}^\delta$ vers $\theta_{p,\delta}$

Théorème 2 Sous (H), pour δ , $0 < \delta \leq \delta_0$: $\widehat{\theta}_{p,n}^\delta \xrightarrow{P_{x,\theta_0}} \theta_{p,\delta}$

Démonstration :

La démonstration de cette convergence repose sur la vérification des hypothèses du théorème 3.2.8. de [2]. Par hypothèse, Θ est compact et, d'après le lemme 4, $U_{p,n}(\theta, \delta)$ et $U_p(\theta, \delta)$ sont de classe C^3 sur Θ . Il nous reste donc à montrer que la condition suivante est satisfaite; il existe deux suites $(\eta_k)_k$ et $(\varepsilon_k)_k$ réelles et décroissant vers 0, telles que, pour tout $k \in N$:

$$\lim_{n \rightarrow \infty} P_{x,\theta_0} \left(\sup_{\|\theta - \tilde{\theta}\|_2 \leq \eta_k} |U_{p,n}(\theta, \delta) - U_{p,n}(\tilde{\theta}, \delta)| > \varepsilon_k \right) = 0$$

Cette condition est vérifiée s'il existe D_n telle que $D_n \xrightarrow{P_{x,\theta_0}} D$ où D est une constante strictement positive et que :

$$\forall \eta > 0, \quad \sup_{\|\theta - \tilde{\theta}\|_2 \leq \eta} |U_{p,n}(\theta, \delta) - U_{p,n}(\tilde{\theta}, \delta)| \leq \eta D_n$$

Posons pour tout couple (y, x) et tout $i = 1, d$, $\phi_p^{(i)}(y, x, \delta) = \sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta^i} \Phi_p(y, x, \theta, \delta) \right|$. Sous (H 4) et (H 5), $\phi_p^{(i)} \in L^2(P_{\theta_0}^\delta)$. De plus, par application de la formule de Rolles, on a :

$$\sup_{\|\theta - \tilde{\theta}\|_2 \leq \eta} |U_{p,n}(\theta, \delta) - U_{p,n}(\tilde{\theta}, \delta)| \leq \sum_{i=1}^d \left(\frac{1}{n} \sum_{k=1}^n \phi_p^{(i)}(Y_{k\delta}, Y_{(k-1)\delta}, \delta) \right) \frac{d}{\delta^2} \eta$$

Par application du lemme 1, on obtient :

$$D_n = \frac{d}{n} \sum_{k=1}^n \left[\frac{1}{\delta^2} \sum_{i=1}^d \phi_p^{(i)}(Y_{k\delta}, Y_{(k-1)\delta}, \delta) \right] \xrightarrow{L^2(P_{x,\theta_0}^\delta)} D = \frac{d}{\delta^2} \sum_{i=1}^d E_{\theta_0} \left[\phi_p^{(i)}(Y_\delta, Y_0, \delta) \right] < \infty$$

On a donc la convergence annoncée.

2.3.4 Normalité asymptotique de $\widehat{\theta}_{p,n}^\delta$

On note pour $i = 1, d$ et $j = 1, d$:

$$P_{i,j}(v) = E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{\sigma^2(Y_v)}{\sigma^4(Y_0)} \right] \quad (2.9)$$

$$P'_{i,j}(v) = E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\mathcal{A}_{\theta_0} \sigma^2)(Y_v)}{\sigma^4(Y_0)} \right], \quad D_{\theta_0} = (P'_{i,j}(0^+))_{i,j} \quad (2.10)$$

Sous (H), $P_{i,j}$ est en effet dérivable sur R_*^+ et à droite en 0 (lemme 2). On définit également :

$$Q_{i,j}(v) = E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\sigma^2 f'_{\theta_0})(Y_v)}{\sigma^4(Y_0)} \right], \quad Q_{\theta_0} = (Q_{i,j}(0))_{i,j} \quad (2.11)$$

$$C_{i,j}(v) = E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} (\mathcal{A}_{\theta_0} f_{\theta_0}) \frac{\partial}{\partial \theta^j} f_{\theta_0} + \frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} (\mathcal{A}_{\theta_0} f_{\theta_0}) \right) (Y_0) \frac{\sigma^2(Y_v)}{\sigma^4(Y_0)} \right], \quad C_{\theta_0} = (C_{i,j}(0))_{i,j} \quad (2.12)$$

Comme le montre le théorème qui suit, $\hat{\theta}_{p,n}^\delta$ est asymptotiquement normal et sa variance asymptotique $V_\delta^p(\theta_0)$ est égale à l'inverse de l'information de Fisher $(H_{\theta_0})^{-1}$ à un facteur $I_d + o(1)$ près :

$$\forall p \geq 1, \quad V_\delta^p(\theta_0) = (H_{\theta_0})^{-1} [I_d + o(1)] \quad (2.13)$$

Cependant, si $p \geq 2$, l'expression de $V_\delta^p(\theta_0)$ peut être affinée et on a :

$$V_\delta^p(\theta_0) = (H_{\theta_0})^{-1} \left[I_d + \delta \left(\frac{1}{2} D_{\theta_0} + Q_{\theta_0} + \frac{1}{2} C_{\theta_0} \right) + o(\delta) \right] \quad (2.14)$$

Théorème 3 *Sous (H), pour H_{θ_0} définie en (H 6), il existe $\delta_1 > 0$ tel que pour δ , $0 < \delta \leq \delta_1$, et $V_\delta^p(\theta_0)$ définie en (2.13) et (2.14), on a :*

$$\sqrt{T} \left(\hat{\theta}_{p,n}^\delta - \theta_{p,\delta} \right) \xrightarrow{\mathcal{D}(P_{\theta_0})} N_d(0, V_\delta^p(\theta_0))$$

Démonstration :

$U_{p,n}(\theta, \delta)$ est de classe C^3 sur Θ . On note alors :

$$DU_{p,n}(\theta, \delta) = \left(\frac{\partial}{\partial \theta^i} U_{p,n}(\theta, \delta) \right)_i \quad \text{et} \quad D^2 U_{p,n}(\theta, \delta) = \left(\frac{\partial^2}{\partial \theta^j \partial \theta^i} U_{p,n}(\theta, \delta) \right)_{i,j}$$

On a par application de la formule de Taylor avec reste intégral :

$$\sqrt{n\delta} DU_{p,n}(\hat{\theta}_{p,n}^\delta, \delta) = 0 = \sqrt{n\delta} DU_{p,n}(\theta_{p,\delta}, \delta) + (D^2 U_{p,n}(\theta_{p,\delta}, \delta) + R_{p,n}) \sqrt{n\delta} (\hat{\theta}_{p,n}^\delta - \theta_{p,\delta})$$

avec :

$$R_{p,n} = \int_0^1 \left[D^2 U_{p,n}(\theta_{p,\delta} + s(\hat{\theta}_{p,n}^\delta - \theta_{p,\delta}), \delta) - U_{p,n}(\theta_{p,\delta}, \delta) \right] ds$$

Si l'on prouve que :

1. $\sqrt{n\delta} DU_{p,n}(\theta_{p,\delta}, \delta) \xrightarrow{\mathcal{D}(P_{x,\theta_0})} N_d(0, K^p(\theta_{p,\delta}))$.
2. $D^2 U_{p,n}(\theta_{p,\delta}, \delta) \xrightarrow{P_{x,\theta_0}} H_\delta^p$, H_δ^p étant symétrique et inversible.
3. $R_{p,n} \xrightarrow{P_{x,\theta_0}} 0$.

on aura la normalité asymptotique annoncée pour $\hat{\theta}_{p,n}^\delta$ avec $V_\delta^p(\theta_0) = (H_\delta^p)^{-1} K^p(\theta_{p,\delta}) (H_\delta^p)^{-1}$.

1. Sous (H 4) et (H 5), $\frac{\partial}{\partial \theta^i} \Phi_p(y, x, \theta_{p,\delta}, \delta) \in L^2(P_{\theta_0}^\delta)$. De plus, d'après les lemmes 4 et 6, on a pour tout $i = 1, d$, $E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} \Phi_p(Y_\delta, Y_0, \theta_{p,\delta}, \delta) \right] = \delta^2 \frac{\partial}{\partial \theta^i} U_p(\theta_{p,\delta}, \delta) = 0$. Utilisant le lemme 1, on obtient :

$$\sqrt{n\delta} DU_{p,n}(\theta_{p,\delta}, \delta) \xrightarrow{\mathcal{D}(P_{\theta_0}^\delta)} N_d(0, K^p(\theta_{p,\delta})), \quad K^p(\theta_{p,\delta}) = T^p(\theta_{p,\delta}) + S^p(\theta_{p,\delta})$$

avec :

$$T^p(\theta_{p,\delta}) = \left(\frac{1}{\delta^3} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} \Phi_p(Y_\delta, Y_0, \theta_{p,\delta}, \delta) \frac{\partial}{\partial \theta^j} \Phi_p(Y_\delta, Y_0, \theta_{p,\delta}, \delta) \right] \right)_{i,j} \quad (2.15)$$

$$S^p(\theta_{p,\delta}) = \left(\frac{2}{\delta^3} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} \Phi_p(Y_\delta, Y_0, \theta_{p,\delta}, \delta) U_\delta \left(\frac{\partial}{\partial \theta^j} \Phi_p \right) (Y_\delta) \right] \right)_{i,j} \quad (2.16)$$

où U_δ est défini dans lemme 1. Le premier résultat de convergence faible est donc vérifié.

2. Pour tout $i, j = 1, d$, $\frac{\partial^2}{\partial \theta^j \partial \theta^i} \Phi_p(y, x, \theta_{p,\delta}, \delta) \in L^2(P_{\theta_0}^\delta)$. Par application du lemme 1, on obtient :

$$D^2 U_{p,n}(\theta_{p,\delta}, \delta) \xrightarrow{L^2(P_{x,\theta_0})} H_\delta^p = \left(\frac{1}{\delta^2} E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \partial \theta^i} \Phi_p(Y_\delta, Y_0, \theta_{p,\delta}, \delta) \right] \right)_{i,j}$$

Etant la matrice heissienne de $\theta \rightarrow U_p(\theta, \delta)$, H_δ^p est symétrique. De plus :

Lemme 7 Sous (H), pour H_{θ_0} définie en (H 6), on a pour δ , $0 < \delta \leq \delta_0$:

$$H_\delta^p = 2H_{\theta_0} + O(\delta)$$

De plus, il existe δ_1 , $0 < \delta_1 \leq \delta_0$, tel que H_δ^p est inversible.

Ce résultat est démontré en annexe B. On a donc la convergence annoncée.

3. Montrons la convergence de $R_{p,n}$ vers 0. Pour $i, j, k = 1, d$ et tout $(y, x) \in R^2$, on définit $\phi_p^{(k,j,i)}(y, x, \delta) = \sup_{\theta \in \Theta} \left| \frac{\partial^3}{\partial \theta^k \partial \theta^j \partial \theta^i} \Phi_p(y, x, \theta, \delta) \right|$. D'après la formule de Rolles, on a :

$$\left| \int_0^1 \left[\frac{\partial^2}{\partial \theta^j \partial \theta^i} U_{p,n}(\theta_{p,\delta} + s(\widehat{\theta}_{p,n}^\delta - \theta_{p,\delta}), \delta) - \frac{\partial^2}{\partial \theta^j \partial \theta^i} U_{p,n}(\theta_{p,\delta}, \delta) \right] ds \right| \leq \sum_{k=1}^d D_n^{k,j,i} \left| (\widehat{\theta}_{p,n}^\delta)^k - \theta_{p,\delta}^k \right|$$

avec $D_n^{i,j,l} = \frac{1}{n\delta^2} \sum_{k=1}^n \phi_p^{(k,j,i)}(Y_{k\delta}, Y_{(k-1)\delta}, \delta)$. Or, sous (H 4) et (H 5), $\phi_p^{(k,j,i)}(y, x, \delta) \in L^2(P_{\theta_0}^\delta)$. Par application du lemme 1, $D_n^{i,j,k}$ converge en probabilité vers

$\frac{1}{\delta^2} E_{\theta_0} \left[\phi_p^{(k,j,i)}(Y_\delta, Y_0, \delta) \right]$ qui est une constante finie. Compte tenu de la convergence de

$\widehat{\theta}_{p,n}^\delta$ vers $\theta_{p,\delta}$, on obtient $R_{p,n} \xrightarrow{P_{x,\theta_0}} 0$.

Il nous reste à établir (2.13) et (2.14). On utilise pour cela le lemme 7 énoncé en 2 et le résultat suivant, dont la démonstration est donnée en annexe C.

Lemme 8 Sous (H), pour H_{θ_0} définie en (H 6), D_{θ_0} en (2.10), Q_{θ_0} en (2.11) et C_{θ_0} en (2.12), on a :

$$\begin{aligned} K^1(\theta_{p,\delta}) &= 4H_{\theta_0} + o(1) \\ \forall p \geq 2, K^p(\theta_{p,\delta}) &= 4H_{\theta_0} + 4\delta \left[\frac{1}{2}D_{\theta_0} + Q_{\theta_0} + \frac{1}{2}C_{\theta_0} \right] + o(\delta) \end{aligned}$$

Pour $p \geq 2$, notons M_{θ_0} la partie en δ du développement précédent. On obtient :

$$\begin{aligned} V_{\delta}^p(\theta_0) - (2H_{\theta_0})^{-1} \left(4H_{\theta_0} + 4\delta M_{\theta_0}^p \right) (2H_{\theta_0})^{-1} \\ = (H_{\delta}^p)^{-1} \left(\left[I_d - H_{\delta}^p (2H_{\theta_0})^{-1} \right] \left[4H_{\theta_0} + 4\delta M_{\theta_0}^p \right] \left[I_d - (2H_{\theta_0})^{-1} H_{\delta}^p \right] \right) (H_{\delta}^p)^{-1} + o(\delta) \end{aligned}$$

Or : $\lim_{\delta \rightarrow 0} H_{\delta}^p = 2H_{\theta_0}$ et $I_d - H_{\delta}^p (2H_{\theta_0})^{-1} = O(\delta)$ (lemme 7). On obtient :

$$V_{\delta}^p(\theta_0) = (H_{\theta_0})^{-1} \left[I_d + \delta M_{\theta_0}^p (H_{\theta_0})^{-1} \right] + o(\delta)$$

On procède de manière similaire lorsque $p = 1$.

2.4 Etude expérimentale

Dans ce paragraphe, nous vérifierons expérimentalement quelques unes des propriétés annoncées pour $\hat{\theta}_{p,n}^{\delta}$ (convergence vers $\theta_{p,\delta}$, écart entre $\hat{\theta}_{p,n}^{\delta}$ et θ_0 , variance asymptotique) pour deux modèles de diffusion : les processus de Ornstein-Uhlenbeck (O.U.) et de Cox-Ingersoll-Ross (C.I.R.). Pour cela, on simule la diffusion sur un intervalle $[0, T]$, en utilisant un schéma d'Euler au pas fin 0.01. Pour un choix (δ, n) , avec $n\delta = T$, pour $\delta \in \{0.05, 0.1, 0.5\}$ et $T = 200$, le paramètre est alors estimé par les moindres carrés (E.M.C.) pour différents ordres p du développement l'espérance conditionnelle. On répète 1000 fois cette expérience afin de calculer la moyenne et la variance empiriques, notées respectivement $m(\hat{\theta}_{p,n}^{\delta})$ et $S^2(\hat{\theta}_{p,n}^{\delta})$.

2.4.1 Processus de Ornstein-Uhlenbeck

On considère l'e.d.s. suivante:

$$dY_t = -\theta_0 Y_t dt + dW_t, Y_0 = x \quad (2.17)$$

(2.17) admet une unique solution forte, à trajectoires continues,

$$Y_t = \exp(-\theta_0 t) \left(x + \int_0^t \exp(\theta_0 s) dW_s \right)$$

$\pi_{\theta_0}^t(x, \cdot)$ est donc une loi normale $N\left(x \exp(-\theta_0 t), \frac{1 - \exp(-2\theta_0 t)}{2\theta_0}\right)$.

Si $\theta_0 > 0$, $(Y_t)_{t \geq 0}$ est récurrente positive sur R , de loi invariante $\mu_{\theta_0} = N\left(0, \frac{1}{2\theta_0}\right)$. En outre, $\lim_{|x| \rightarrow +\infty} c_{\theta_0} = +\infty$, (H 3) est vérifiée. Enfin, on a :

$$\mathcal{A}_{\theta_0}^l f_{\theta_0}(x) = (-\theta_0)^{l+1} x$$

$$P(v) = \frac{1}{2\theta_0}, \quad N^p(v) = \frac{(-\theta_0)^p}{2} \exp(-\theta_0 v), \quad H_{\theta_0} = \frac{1}{2\theta_0}$$

Notons :

$$Q_p(x) = S_n, \quad S_n = \sum_{l=0}^{p-1} \frac{(-x)^{l+1}}{(l+1)!} \quad (2.18)$$

L'estimateur associé à l'approximation au rang p de l'espérance conditionnelle vérifie :

$$Q_p\left(\delta \widehat{\theta}_{p,n}^\delta\right) = \frac{\sum_{k=1}^n Y_{(k-1)\delta} (Y_{k\delta} - Y_{(k-1)\delta})}{\sum_{k=1}^n Y_{(k-1)\delta}^2}$$

De plus, $\theta_{p,\delta}$ est solution de l'équation :

$$Q_p(\delta\theta) = \exp(-\delta\theta) - 1$$

Pour $p = 1$ et $p = 3$ et pour tout $y \leq 0$, l'équation $Q_p(x) = y$ admet une unique solution positive. Pour $p = 2$ et $y \in [-\frac{1}{2}, 0]$, cette équation admet une unique solution dans $[0, 1]$. Enfin, si $p = \infty$ et $y \in]-1, 0]$, l'unique solution est $x = -\ln(1 + y)$. Pour δ suffisamment petit, on a en particulier pour $p = 1$, $p = 2$ et $p = \infty$:

$$\begin{aligned} \theta_{1,\delta} &= \frac{1}{\delta} (1 - \exp(-\delta\theta_0)) = \theta_0 - \frac{\theta_0^2}{2}\delta + o(\delta) \\ \theta_{2,\delta} &= \frac{1}{\delta} \left(1 - \sqrt{2 \exp(-\delta\theta_0) - 1}\right) = \theta_0 + \frac{\theta_0^3}{6}\delta^2 + o(\delta^2) \\ \theta_{\infty,\delta} &= \theta_0 \end{aligned}$$

On prend comme valeur initiale pour les simulations $Y_0 = 0$ et $\theta_0 = 1$.

On constate que le biais empirique est d'autant plus proche de 0 que δ est petit et que l'ordre p est élevé. Notons également que, dans le cas de ce modèle pour lequel les quantités conditionnelles sont explicites, lorsque $p = \infty$, $\widehat{\theta}_{p,n}^\delta$ correspond à l'estimateur associé au contraste gaussien basé sur les espérances conditionnelles exactes.

		$\widehat{\theta}_{p,n}^\delta (\theta_0 = 1)$					
		δ	n	$m(\widehat{\theta}_{p,n}^\delta)$	$S^2(\widehat{\theta}_{p,n}^\delta)$	$\theta_{p,\delta}$	$\frac{H_{\theta_0}^{-1}}{T}$
$p = 1$	0.05	4000	0.9878	0.0098	0.9754	0.01	
	0.1	2000	0.964525	0.008976	0.951626		
	0.5	400	0.7999	0.0068	0.7869		

Table 2.1: Processus de O.U. et approximation à l'ordre 1 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.

		$\widehat{\theta}_{p,n}^\delta (\theta_0 = 1)$					
		δ	n	$m(\widehat{\theta}_{p,n}^\delta)$	$S^2(\widehat{\theta}_{p,n}^\delta)$	$\theta_{p,\delta}$	$\frac{H_{\theta_0}^{-1}}{T}$
$p = 2$	0.05	4000	1.0138	0.0109	1.0004	0.01	
	0.1	2000	1.016774	0.011170	1.001807		
	0.5	400	1.1172	0.0344	1.076		

Table 2.2: Processus de O.U. et approximation à l'ordre 2 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.. Pour $\delta = 0.5$, nous avons obtenu 987 réalisations de $\widehat{\theta}_{p,n}^\delta$; dans les autres cas, les réalisations de S_n sont en dehors de $[-\frac{1}{2}, 0]$.

		$\widehat{\theta}_{p,n}^\delta (\theta_0 = 1)$					
		δ	n	$m(\widehat{\theta}_{p,n}^\delta)$	$S^2(\widehat{\theta}_{p,n}^\delta)$	$\theta_{p,\delta}$	$\frac{H_{\theta_0}^{-1}}{T}$
$p = 3$	0.05	4000	1.01335	0.01089	0.999756	0.01	
	0.1	2000	1.014716	0.011030	1.000366		
	0.5	400	1.0167	0.0179	0.9925		

Table 2.3: Processus de O.U. et approximation à l'ordre 3 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.

		$\widehat{\theta}_{p,n}^\delta (\theta_0 = 1)$					
		δ	n	$m(\widehat{\theta}_{p,n}^\delta)$	$S^2(\widehat{\theta}_{p,n}^\delta)$	$\theta_{p,\delta}$	$\frac{H_{\theta_0}^{-1}}{T}$
$p = \infty$	0.05	4000	1.013373	0.010912	1	0.01	
	0.1	2000	1.014816	0.011036	1		
	0.5	400	1.0264	0.0196	1		

Table 2.4: Processus de O.U. et approximation à l'ordre ∞ : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.

2.4.2 Processus de Cox-Ingersoll-Ross

On considère l'e.d.s. suivante:

$$dY_t = b_0 - a_0 Y_t dt + \sqrt{Y_t} dW_t, \quad Y_0 = x$$

On a donc :

$$\theta = (b, a), \quad \sigma(x) = \sqrt{x}, \quad f = {}^t(f_1, f_2), \quad f_1(x) = -x \text{ et } f_2(x) = 1$$

f est C^∞ sur R et lipschitzienne, σ est C^∞ sur R_+^* et h"olderienne de rapport $\frac{1}{2}$. L'e.d.s. admet donc une solution forte unique, à trajectoires continues. Si $b \geq \frac{1}{2}$ et $a > 0$, la diffusion est r"ecurrente positive sur R_+^* , de loi invariante $\mu_\theta = \Gamma(2b, 2a)$. Cette loi admet des moments de tous ordres.

De plus, si $b > 2$, $\int \frac{1}{\sigma^s(x)} d\mu_{\theta_0}(x) < \infty$ et donc (H 4) est v"erifi"ee. (H 3) est v"erifi"ee, car $\lim_{x \rightarrow 0^+} c_{\theta_0}(x) = +\infty$ et $\lim_{x \rightarrow +\infty} c_{\theta_0}(x) = +\infty$. Enfin, on a pour tout $l \leq 0$, $\mathcal{A}_\theta^l f_\theta(x) = (-a)^l f_\theta(x)$, et on montre que :

$$\begin{aligned} N_1^p(v) &= 0, \quad N_2^p(v) = -\frac{(-a_0)^{p+1}}{2b_0 - 1} \exp(-a_0 v) \\ P_{1,1}(v) &= \frac{b_0}{a_0}, \quad P_{1,2}(v) = P_{2,1}(v) = \frac{\exp(-a_0 v) - 2b_0}{2b_0 - 1}, \quad P_{2,2}(v) = 2 \frac{a_0 b_0 (b_0 - \exp(-a_0 v))}{(2b_0 - 1)(b_0 - 1)} \end{aligned}$$

$$H_{\theta_0}^{-1} = (2b_0 - 1) \begin{pmatrix} \frac{2a_0}{2b_0 - 1} & 1 \\ 1 & \frac{b_0}{a_0} \end{pmatrix}.$$

Q_p "etant donn"e par (2.18), l'estimateur du minimum de contraste $(\hat{a}_{p,n}^\delta, \hat{b}_{p,n}^\delta)$ v"erifie le syst"eme suivant :

$$\begin{aligned} Q_p(\delta \hat{a}_{p,n}^\delta) &= S_n, \quad S_n = \frac{\left(\frac{1}{n} \sum_{k=1}^n \frac{1}{Y_{(k-1)\delta}}\right) \left(\frac{1}{n} \sum_{k=1}^n (Y_{k\delta} - Y_{(k-1)\delta})\right) - \frac{1}{n} \sum_{k=1}^n \frac{(Y_{k\delta} - Y_{(k-1)\delta})}{Y_{(k-1)\delta}}}{\left(\frac{1}{n} \sum_{k=1}^n Y_{(k-1)\delta}\right) \left(\frac{1}{n} \sum_{k=1}^n \frac{1}{Y_{(k-1)\delta}}\right) - 1} \\ \hat{b}_{p,n}^\delta &= \frac{\hat{a}_{p,n}^\delta}{S_n} \left[S_n \frac{1}{n} \sum_{k=1}^n Y_{(k-1)\delta} - \frac{1}{n} \sum_{k=1}^n \frac{(Y_{k\delta} - Y_{(k-1)\delta})}{Y_{(k-1)\delta}} \right] \end{aligned}$$

De plus, $(a_{p,\delta}, b_{p,\delta})$ est solution de :

$$\begin{aligned} Q_p(\delta a_{p,\delta}) &= \exp(-\delta a_0) - 1 \\ b_{p,\delta} &= \frac{b_0}{a_0} a_{p,\delta} \end{aligned}$$

Il apparait que $a_{p,\delta}$ v"erifie la m"eme "equation que $\theta_{p,\delta}$ pour le mod"ele O.U.

Pour $p = 1$, $p = 2$ et $p = \infty$, lorsque δ est proche de 0 :

$$\begin{aligned} b_{1,\delta} &= b_0 - \frac{b_0 a_0}{2} \delta + o(\delta) \\ b_{2,\delta} &= b_0 + \frac{b_0 a_0^2}{6} \delta^2 + o(\delta^2) \\ b_{\infty,\delta} &= b_0 \end{aligned}$$

On prend $Y_0 = 1$ et $(a_0, b_0) = (1, 3)$.

		$a_n (a_0 = 1)$				$\widehat{b}_{p,n}^\delta (b_0 = 3)$				
	δ	n	$m(\widehat{a}_{p,n}^\delta)$	$S^2(\widehat{a}_{p,n}^\delta)$	$a_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1$	$m(\widehat{b}_{p,n}^\delta)$	$S^2(\widehat{b}_{p,n}^\delta)$	$b_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2$
$p = 1$	0.05	4000	1.0008	0.01002	0.975	0.01	2.980	0.0718	2.926	0.075
	0.1	2000	0.96	0.010	0.951		2.885	0.0764	2.854	
	0.5	400	0.802	0.0066	0.786		2.411	0.062	2.360	

Table 2.5: Processus de C.I.R. et approximation à l'ordre 1 : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.

		$\widehat{a}_{p,n}^\delta (a_0 = 1)$				$\widehat{b}_{p,n}^\delta (b_0 = 3)$				
	δ	n	$m(\widehat{a}_{p,n}^\delta)$	$S^2(\widehat{a}_{p,n}^\delta)$	$a_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1$	$m(\widehat{b}_{p,n}^\delta)$	$S^2(\widehat{b}_{p,n}^\delta)$	$b_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2$
$p = 2$	0.05	4000	1.027	0.011	1.0004	0.01	3.059	0.079	3.001	0.075
	0.1	2000	1.012	0.012	1.0018		3.041	0.095	3.005	
	0.5	400	1.133	0.039	1.076		3.404	0.349	3.23	

Table 2.6: Processus de C.I.R. et approximation à l'ordre 2 : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.

		$\widehat{a}_{p,n}^\delta (a_0 = 1)$				$\widehat{b}_{p,n}^\delta (b_0 = 3)$				
	δ	n	$m(\widehat{a}_{p,n}^\delta)$	$S^2(\widehat{a}_{p,n}^\delta)$	$a_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1$	$m(\widehat{b}_{p,n}^\delta)$	$S^2(\widehat{b}_{p,n}^\delta)$	$b_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2$
$p = 3$	0.05	4000	1.027	0.011	0.999	0.01	3.058	0.079	2.999	0.075
	0.1	2000	1.010	0.012	1.0003		3.035	0.094	3.001	
	0.5	400	1.020	0.017	0.992		3.066	0.155	2.977	

Table 2.7: Processus de C.I.R. et approximation à l'ordre 3 : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.

		$\widehat{a}_{p,n}^\delta (a_0 = 1)$				$\widehat{b}_{p,n}^\delta (b_0 = 3)$				
	δ	n	$m(\widehat{a}_{p,n}^\delta)$	$S^2(\widehat{a}_{p,n}^\delta)$	$a_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1$	$m(\widehat{b}_{p,n}^\delta)$	$S^2(\widehat{b}_{p,n}^\delta)$	$b_{p,\delta}$	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2$
$p = \infty$	0.05	4000	1.027	0.011	1	0.01	3.058	0.079	3	0.075
	0.1	2000	1.010	0.012	1		3.035	0.094	3	
	0.5	400	1.030	0.018	1		3.095	0.167	3	

Table 2.8: Processus de C.I.R. et approximation à l'ordre ∞ : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.

Bibliography

- [1] Bibby, B.M. & Sorensen, M. (1995) *Martingale Estimation Functions for Discretely Observed Diffusion Processes.* - Bernoulli 1, 17-39.
- [2] Dacunha-Castelle, D. & Duflo, M. (1993) *Probabilité et Statistiques. Tome 2- 2^{eme} Ed.* Masson.
- [3] Dacunha-Castelle, D. & Florens-Zmirou, D. (1986) *Estimation of the coefficient of a diffusion from discrete observations.-* Stochastics 19, 263-284.
- [4] Florens-Zmirou, D. (1989) *Approximate discrete schemes for stastics of diffusion processes.-* Statistics 20, 547-557.
- [5] Genon-Catalot, V. (1990) *Maximum contrast estimation for diffusion processes from discrete observations.-* Statistics 21, 99-116.
- [6] Karatzas, I. & Shreve S.E. (1996) *Brownian Motion and Stochastic Calculus.- 2nd Ed.* Springer.
- [7] Kessler, M. (1995) *Estimation of an ergodic diffusion from discrete observations.-* Scand. J. Stat. 24, No 2, 211-229.
- [8] Kessler, M. (1996) *Simple and Explicit Estimating Functions for a Discretely Observed Diffusion Process.-* Research Reports 336, Department of theoretical statistics, University of Aarhus.
- [9] Kessler, M. & Sorensen, M. (1995) *Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process.-* Research Reports 332, Department of theoretical statistics, University of Aarhus.
- [10] Lipster, R.S. & Shiryaev, A.N. (1977) *Statistics of random processes. Tome 1,2-* Springer-Verlag.
- [11] Pedersen, A. R. (1995) *Consistency and asymptotic normality of an approximate maximum likelihood estimator for dicretely observed diffusion processes.-* Bernoulli 1, 257-279.
- [12] Pedersen, A. R. (1995) *A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations.-* Scand. J. Statist. 22, 55-71.
- [13] Sorensen, M. (1996) *Estimating functions for discretely observed diffusions: A review.-* Research Reports 348, Department of theoretical statistics, University of Aarhus.

- [14] Souchet, S. (1998) *Schéma de discrétisation anticipatif et estimation du paramètre de dérive d'une diffusion.*-A paraître comme note au CRAS.
- [15] Yoshida, N. (1992) *Estimation for diffusion processes from discrete observations.*- J. Multivariate Anal. 41, 220-242.

Annexes

A Démonstration du lemme 5

Pour simplifier la démonstration nous montrons ce résultat dans le cas où $p = 1$. On a alors : $M_1(\theta, \delta, x) = \delta f_\theta(x)$ et $\eta_t^1 = \int_t^{t+\delta} \frac{(t+\delta-v)}{2} (\mathcal{A}_{\theta_0} f_{\theta_0})(Y_0) dv$.

Montrons que V_i^1 est de classe $C^{1,2}$ sur $U(\Theta) \times R$.

- Nous commencerons par montrer que $(\theta, \delta) \rightarrow V_i^1(\theta, \delta)$ est de classe $C^{2,1}$ sur $U(\Theta) \times R_*^+$. Pour tout $\delta > 0$, on a :

$$V_i^1(\theta, \delta) = \frac{\partial}{\partial \theta^i} U_1(\theta, \delta) = -2E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_\theta(Y_0) \frac{(Y_\delta - Y_0 - \delta f_\theta(Y_0))}{\delta \sigma^2(Y_0)} \right] \quad (2.19)$$

La régularité de V_i^1 sur $U(\Theta) \times R_*^+$ est alors une simple conséquence du lemme 2. On a en particulier pour tout $(\theta, \delta) \in U(\Theta) \times R_*^+$:

$$\begin{aligned} \frac{\partial}{\partial \theta^j} V_i^1(\theta, \delta) &= -\frac{2}{\delta} E_{\theta_0} \left[\left(\frac{\partial^2}{\partial \theta^j \partial \theta^i} f_\theta \right) (Y_0) \left(\frac{Y_\delta - Y_0 - f_\theta(Y_0)}{\sigma^2(Y_0)} \right) \right] \\ &+ 2E_{\theta_0} \left[\frac{\left(\frac{\partial}{\partial \theta^i} f_\theta \right) (Y_0) \left(\frac{\partial}{\partial \theta^j} f_\theta \right) (Y_0)}{\sigma^2(Y_0)} \right] \end{aligned} \quad (2.20)$$

$$\begin{aligned} \frac{\partial^2}{\partial \theta^k \partial \theta^j} V_i^1(\theta, \delta) &= -\frac{2}{\delta} E_{\theta_0} \left[\frac{\partial^3}{\partial \theta^k \partial \theta^j \partial \theta^i} f_\theta(Y_0) \left(\frac{Y_\delta - Y_0 - \delta f_\theta(Y_0)}{\sigma^2(Y_0)} \right) \right] \\ &+ 2E_{\theta_0} \left[\frac{\left(\frac{\partial^2}{\partial \theta^k \partial \theta^i} f_\theta \frac{\partial}{\partial \theta^j} f_\theta + \frac{\partial^2}{\partial \theta^k \partial \theta^j} f_\theta \frac{\partial}{\partial \theta^i} f_\theta + \frac{\partial^2}{\partial \theta^j \partial \theta^i} f_\theta \frac{\partial}{\partial \theta^k} f_\theta \right) (Y_0)}{\sigma^2(Y_0)} \right] \end{aligned} \quad (2.21)$$

et :

$$(V_i^1)'(\theta, \delta) = \frac{2}{\delta^2} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_\theta \right) (Y_0) \left(\frac{Y_\delta - Y_0}{\sigma^2(Y_0)} \right) \right] - \frac{2}{\delta} E_{\theta_0} \left[\frac{\left(\frac{\partial}{\partial \theta^i} f_\theta \right) (Y_0) f_{\theta_0}(Y_\delta)}{\sigma^2(Y_0)} \right] \quad (2.22)$$

- Utilisant à nouveau le lemme 2, on montre que $\theta \rightarrow V_i^1(\theta, 0)$ et $\theta \rightarrow R_i^1(\theta)$ sont de classe C^2 sur $U(\Theta)$. On déduit en particulier que V_i^1 est de classe C^2 sur $U(\Theta) \times R_*^-$.
- Pour achever cette démonstration, il nous reste à montrer que :

A V_i^1 et ses dérivées partielles successives par rapport à θ sont continues en tout point de la forme $(\theta, 0)$ avec $\theta \in U(\Theta)$.

B Pour tout $\theta \in U(\Theta)$, $\delta \rightarrow V_i^1(\theta, \delta)$ est dérivable en 0 et sa dérivée $(\theta, \delta) \rightarrow (V_i^1)'(\theta, \delta)$ est continue en tout point $(\theta, 0)$.

Dans les deux cas, nous nous basons sur l'équation suivante qui est obtenue en utilisant le schéma (2.3) avec $p = 1$ dans (2.19).

$$\begin{aligned} V_i^1(\theta, \delta) &= -2\delta \int_0^1 (1-v) E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_{\delta v})}{\sigma^2(Y_0)} \right] dv \\ &\quad - 2E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] \end{aligned} \quad (2.23)$$

Montrons A.

Nous nous contenterons en fait de prouver la continuité de V_i^1 , la continuité de ses dérivées partielles étant obtenue de manière similaire. Soient $\theta \in U(\Theta)$ et $((\theta_n, \delta_n))_n$ une suite de $U(\Theta) \times R_*^+$ qui converge vers $(\theta, 0)$. Utilisant (2.23), on a :

$$\begin{aligned} V_i^1(\theta_n, \delta_n) &= -2\delta_n \int_0^1 (1-v) E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_n}(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_{\delta_n v})}{\sigma^2(Y_0)} \right] dv \\ &\quad - 2E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_n} \right)(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta_n}(Y_0)}{\sigma^2(Y_0)} \right) \right] \end{aligned}$$

D'après le lemme 2, on sait que, sous (H), les fonctions qui interviennent dans ce développement sont continues respectivement sur $U(\Theta) \times R^+$ et $U(\Theta)$. On en déduit que :

$$\begin{aligned} \lim_{n \rightarrow \infty} V_i^1(\theta_n, \delta_n) &= -2E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta} \right)(Y_0) \left(\frac{f_{\theta_0}(Y_0) - f_{\theta}(Y_0)}{\sigma^2(Y_0)} \right) \right] \\ &= V_i^1(\theta, 0) \end{aligned}$$

De même :

$$\lim_{n \rightarrow \infty} V_i^1(\theta_n, -\delta_n) = V_i^1(\theta, 0)$$

car, $\theta \rightarrow V_i^1(\theta, 0)$ et $\theta \rightarrow R_i^1(\theta)$ sont continues sur $U(\Theta)$. La continuité de V_i^1 sur tout point de la forme $(\theta, 0)$ est donc démontrée.

Montrons B.

Utilisant l'équation (2.23), on a pour $\delta > 0$:

$$\frac{V_i^1(\theta, \delta) - V_i^1(\theta, 0)}{\delta} = -2 \int_0^1 (1-v) E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_{\delta v})}{\sigma^2(Y_0)} \right] dv$$

Par continuité, on a que la dérivée à droite est :

$$\begin{aligned} (V_i^1)'(\theta, 0) &= -E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta}(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] \\ &= R_i^1(\theta) \end{aligned}$$

Or, la limite à gauche est identique, on en déduit la dérivabilité en 0.

Pour montrer la continuité de $(V_i^1)'$, on considère $((\theta_n, \delta_n))_n$ une suite de points de $U(\Theta) \times R_*^+$ qui converge vers $(\theta, 0)$. D'après (2.22) et (2.3), on a :

$$\begin{aligned} (V_i^1)'(\theta_n, \delta_n) &= 2 \int_0^1 (1-v) E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_n}(Y_0) \left(\frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_{\delta_n v})}{\sigma^2(Y_0)} \right) \right] dv \\ &\quad - \frac{2}{\delta_n} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_n}(Y_0) \left(\frac{f_{\theta_0}(Y_\delta) - f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right) \right] \end{aligned}$$

Or, par application de la formule de Ito, on a :

$$-\frac{2}{\delta_n} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_n}(Y_0) \left(\frac{f_{\theta_0}(Y_\delta) - f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right) \right] = -\frac{2}{\delta_n} \int_0^{\delta_n} E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_n}(Y_0) \left(\frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_v)}{\sigma^2(Y_0)} \right) \right] dv$$

Par continuité des fonctions qui dans le développement de $(V_i^1)'$ (lemme 2), on obtient :

$$\lim_{n \rightarrow \infty} (V_i^1)'(\theta_n, \delta_n) = -E_{\theta_0} \left[\frac{(\frac{\partial}{\partial \theta^i} f_{\theta_0})(Y_0) (\mathcal{A}_{\theta_0} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] = R_i^1(\theta)$$

Or :

$$\lim_{n \rightarrow \infty} (V_i^1)'(\theta_n, -\delta_n) = \lim_{n \rightarrow \infty} R_i^1(\theta_n) = R_i^1(\theta)$$

La fonction $(\theta, \delta) \rightarrow (V_i^1)'(\theta, \delta)$ est donc continue en tout point de la forme $(\theta, 0)$.

On a donc le résultat annoncé : $(\theta, \delta) \rightarrow V_i^1(\theta, \delta)$ est de classe $C^{2,1}$ sur $U(\Theta) \times R$.

Montrons (2.5), (2.6) et (2.7).

L'équation (2.5) est obtenue en utilisant (2.23). Pour $\delta > 0$ et $\theta = \theta_0$, on a :

$$V_i^1(\theta_0, \delta) = -2\delta \int_0^1 (1-v) E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \right)(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_{\delta v})}{\sigma^2(Y_0)} \right] dv$$

Par application du lemme 2, on sait que $\delta \rightarrow E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \right)(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_{\delta v})}{\sigma^2(Y_0)} \right]$ est continue.

D'où :

$$\begin{aligned} \frac{\partial}{\partial \theta^i} U_1(\theta_0, \delta) &= V_i^1(\theta_0, \delta) \\ &= -\delta E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \right)(Y_0) \frac{(\mathcal{A}_{\theta_0} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] + o(\delta) \end{aligned}$$

Par les mêmes arguments, en utilisant (2.20), (2.21) et la décomposition (2.3), on obtient pour $\delta > 0$:

$$\begin{aligned} \frac{\partial^2}{\partial \theta^j \partial \theta^i} U_1(\theta_0, \delta) &= \frac{\partial}{\partial \theta^j} V_i^1(\theta_0, \delta) \\ &= 2E_{\theta_0} \left[\frac{(\frac{\partial}{\partial \theta^i} f_{\theta_0})(Y_0) (\frac{\partial}{\partial \theta^j} f_{\theta_0})(Y_0)}{\sigma^2(Y_0)} \right] + O(\delta) \end{aligned}$$

et :

$$\begin{aligned} \frac{\partial^3}{\partial \theta^k \partial \theta^j \partial \theta^i} U(\theta_0, \delta) &= \frac{\partial^2}{\partial \theta^k \partial \theta^j} V_i^1(\theta_0, \delta) \\ &= 2E_{\theta_0} \left[\frac{\left(\frac{\partial^2}{\partial \theta^k \partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} + \frac{\partial^2}{\partial \theta^k \partial \theta^j} f_{\theta_0} \frac{\partial}{\partial \theta^i} f_{\theta_0} + \frac{\partial^2}{\partial \theta^j \partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^k} f_{\theta_0} \right) (Y_0)}{\sigma^2(Y_0)} \right] + O(\delta) \end{aligned}$$

(2.6) et (2.7) sont donc vérifiées.

B Démonstration du lemme 7 :

Pour $i, j = 1, d$, d'après le lemme 5, pour $\delta > 0$, $\theta \rightarrow \frac{\partial^2}{\partial\theta^j\partial\theta^i}U_p(\theta, \delta)$ est de classe C^1 sur $\overset{\circ}{\Theta}$. Utilisant la formule de Taylor-Young, on obtient pour δ , $0 < \delta \leq \delta_0$:

$$\frac{\partial^2}{\partial\theta^j\partial\theta^i}U_p(\theta_{p,\delta}, \delta) = \frac{\partial^2}{\partial\theta^j\partial\theta^i}U_p(\theta_0, \delta) + \sum_{k=1}^d \frac{\partial^3}{\partial\theta^k\partial\theta^j\partial\theta^i}U(\theta_0, \delta) \left(\theta_{p,\delta}^k - \theta_0^k\right) + \varepsilon(\theta_{p,\delta} - \theta_0) \|\theta_{p,\delta} - \theta_0\|_2$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. D'après (2.6), (2.7) et l'écart calculé entre $\theta_{p,\delta}$ et θ_0 , on a :

$$\frac{\partial^2}{\partial\theta^j\partial\theta^i}U_p(\theta_{p,\delta}, \delta) = 2E_{\theta_0} \left[\frac{\left(\frac{\partial}{\partial\theta^i}f_{\theta_0} \frac{\partial}{\partial\theta^j}f_{\theta_0}\right)(Y_0)}{\sigma^2(Y_0)} \right] + O(\delta)$$

et, donc : $H_\delta^p = 2H_{\theta_0} + O(\delta)$.

Il nous reste à montrer que H_δ^p est inversible. Le développement précédent implique que $\lim_{\delta \rightarrow 0} H_\delta^p = 2H_{\theta_0}$. H_{θ_0} étant inversible par (H 6), il existe en particulier δ_1 , $0 < \delta_1 \leq \delta_0$, tel que H_δ^p est inversible.

C Démonstration du lemme 8 : développement de $K^p(\theta_{p,\delta})$ en fonction de δ

Pour tout $p \in N^*$, on a :

$$K^p(\theta_{p,\delta}) = T^p(\theta_{p,\delta}) + S^p(\theta_{p,\delta})$$

où $T^p(\theta_{p,\delta})$ et $S^p(\theta_{p,\delta})$ sont en (2.15)-(2.16). Nous commencerons par établir que, pour tout $p \geq 1$:

$$T^p(\theta_{p,\delta}) = 4H_{\theta_0} + 4\delta M_{\theta_0}^p + o(\delta)$$

où, pour tout $p \geq 2$, $M_{\theta_0}^p = \frac{1}{2}D_{\theta_0} + Q_{\theta_0} + \frac{1}{2}C_{\theta_0}$.

Dans une seconde partie, nous montrerons que, pour tout $p \geq 1$, il existe $C_p > 0$ finie telle que :

$$\forall i = 1, d, \forall j = 1, d, \left| S_{i,j}^p(\theta_{p,\delta}) \right| \leq C_p \delta^{p-\frac{1}{2}}$$

On aura alors le développement annoncé pour $K^p(\theta_{p,\delta})$.

C.1 Développement de $T^p(\theta_{p,\delta})$

Sous (H), par application du lemme 2, $\theta \rightarrow T^p(\theta)$ est de classe C^2 sur $\overset{\circ}{\Theta}$ par dérivation sous l'espérance. Pour tout $i, j = 1, d$, la formule de Taylor-Young donne :

$$T_{i,j}^p(\theta_{p,\delta}) = T_{i,j}^p(\theta_0) + \sum_{k=1}^d \frac{\partial}{\partial \theta^k} T_{i,j}^p(\theta_0) (\theta_{p,\delta} - \theta_0) + \varepsilon(\theta_{p,\delta} - \theta_0) \|\theta_{p,\delta} - \theta_0\|_2 \quad (2.24)$$

Nous allons établir un développement en δ pour les termes $T_{i,j}^p(\theta_0)$ et $\frac{\partial}{\partial \theta^k} T_{i,j}^p(\theta_0)$. Pour simplifier, nous mènerons les calculs pour $p = 1$ et donnerons les résultats pour $p \geq 2$.

- Du fait de (2.3), on a :

$$T_{i,j}^1(\theta_0) = \frac{4}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\eta_0^1 + \xi_0^1)^2}{\sigma^4(Y_0)} \right]$$

$$1. \text{ Etudions } A(\delta) = \frac{4}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0) \frac{(\xi_0^1)^2}{\sigma^4(Y_0)} \right]$$

$$\begin{aligned} A(\delta) &= \frac{4}{\delta} \int_0^\delta P_{i,j}(v) dv + \frac{8}{\delta} \int_0^\delta (\delta - v) Q_{i,j}(v) dv \\ &+ \int_0^\delta (\delta - v)^2 E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0} (Y_0) \frac{\sigma^2(Y_v) (f'_{\theta_0})^2(Y_v)}{\sigma^4(Y_0)} \right] \end{aligned}$$

Compte tenu de la régularité des fonctions qui interviennent dans le développement précédent (lemme 2), on a :

$$\begin{aligned}\int_0^\delta P_{i,j}(v) dv &= \delta P_{i,j}(0) + \frac{\delta^2}{2} P'_{i,j}(0) + o(\delta^2) \\ \int_0^\delta (\delta - v) Q_{i,j}(v) dv &= \frac{\delta^2}{2} Q_{i,j}(0) + o(\delta^2) \\ \int_0^\delta (\delta - v)^2 E_{\theta_0} \left[\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\sigma^2(Y_v) (f'_{\theta_0})^2(Y_v)}{\sigma^4(Y_0)} \right] &= o(\delta^2)\end{aligned}$$

De ce qui précède, on déduit que :

$$A(\delta) = 4P_{i,j}(0) + 2\delta P'_{i,j}(0) + 4\delta Q_{i,j}(0) + o(\delta)$$

2. Soit $B(\delta) = \frac{4}{\delta} E_{\theta_0} \left[\left(\frac{\partial}{\partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} \right) (Y_0), \frac{(\eta_0^1)^2}{\sigma^4(Y_0)} \right]$. Par une double application de l'inégalité de Schwarz, on montre que :

$$|B(\delta)| \leq D_i D_j M \delta^3$$

avec :

$$D_i = E_{\theta_0} \left[\left(\frac{\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right)^4 \right]^{\frac{1}{4}}, \quad M = E_{\theta_0} \left[(\mathcal{A}_{\theta_0} f_{\theta_0})^4(Y_0) \right]^{\frac{1}{2}}$$

3. Majorant le double produit restant grâce à l'inégalité de Schwarz, on montre qu'il est de l'ordre d'un $o(\delta)$.

On a donc :

$$T_{i,j}^1(\theta_0) = 4P_{i,j}(0) + 2\delta P'_{i,j}(0) + 4\delta Q_{i,j}(0) + o(\delta)$$

Pour $p \geq 2$, on voit apparaître le terme supplémentaire $C_{i,j}$ défini en (2.12) :

$$T_{i,j}^1(\theta_0) = 4P_{i,j}(0) + 2\delta P'_{i,j}(0) + 4\delta Q_{i,j}(0) + 2\delta C_{i,j}(0) + o(\delta)$$

- Traitons les termes $\frac{\partial}{\partial \theta^k} T_{i,j}^p(\theta_0)$ ($k = 1, d$). Par le même type d'arguments, on montre que pour tout $p \geq 1$ et tout $k = 1, d$:

$$\frac{\partial}{\partial \theta^k} T_{i,j}^p(\theta_0) = E_{\theta_0} \left[\frac{\left(\frac{\partial^2}{\partial \theta^k \partial \theta^i} f_{\theta_0} \frac{\partial}{\partial \theta^j} f_{\theta_0} + \frac{\partial^2}{\partial \theta^k \partial \theta^j} f_{\theta_0} \frac{\partial}{\partial \theta^i} f_{\theta_0} \right) (Y_0)}{\sigma^2(Y_0)} \right] + o(1)$$

Pour tout $p \geq 2$, l'écart entre $\theta_{p,\delta}$ et θ_0 est $o(\delta)$. L'équation (2.24) s'écrit alors :

$$T_{i,j}^p(\theta_{p,\delta}) = 4P_{i,j}(0) + 2\delta P'_{i,j}(0) + 4\delta Q_{i,j}(0) + 2C_{i,j}(0) + o(\delta)$$

et :

$$T^p(\theta_{p,\delta}) = 4H_{\theta_0} + 2\delta D_{\theta_0} + 4\delta Q_{\theta_0} + 2\delta C_{\theta_0} + o(\delta)$$

Pour $p = 1$, $\theta_{1,\delta} - \theta_0 = \frac{1}{2}H_{\theta_0}^{-1}N_{\theta_0}^1\delta + o(\delta)$. L'expression de (2.24) est alors plus compliquée. Notons, pour $i, j = 1, d$:

$$B_{i,j} = \left(E_{\theta_0} \left[\frac{\left(\frac{\partial^2}{\partial\theta^k\partial\theta^i} f_{\theta_0} \frac{\partial}{\partial\theta^j} f_{\theta_0} + \frac{\partial^2}{\partial\theta^k\partial\theta^j} f_{\theta_0} \frac{\partial}{\partial\theta^i} f_{\theta_0} \right) (Y_0)}{\sigma^2(Y_0)} \right] \right)_{k=1,d}$$

et :

$$B_{\theta_0} = \left({}^t B_{i,j} H_{\theta_0}^{-1} N_{\theta_0}^1 \right)_{i,j}$$

On a ainsi :

$$T^1(\theta_{p,\delta}) = 4H_{\theta_0} + 2\delta D_{\theta_0} + 4\delta Q_{\theta_0} + 2\delta B_{\theta_0} + o(\delta)$$

On obtient ainsi le développement annoncé.

C.2 Majoration de $S^p(\theta_{p,\delta})$

Pour tout $p \geq 1$, par application de l'inégalité de Schwarz, pour $i = 1, d$ et $j = 1, d$, on a :

$$\begin{aligned} |S_{i,j}^p| &= \left| \frac{2}{\delta^3} E_{\theta_0} \left[\frac{\partial}{\partial\theta^i} \Phi_p(Y_\delta, Y_0, \theta_{p,\delta}, \delta) U_\delta \left(\frac{\partial}{\partial\theta^j} \Phi_p \right) (Y_\delta) \right] \right| \\ &\leq \frac{2}{\delta^{\frac{3}{2}}} \sqrt{T_{i,i}^p(\theta_{p,\delta})} \left\| U_\delta \left(\frac{\partial}{\partial\theta^j} \Phi_p \right) \right\|_{L^2(\mu_{\theta_0})} \end{aligned}$$

D'après C.1, $T_{i,i}^p(\theta_{p,\delta})$ est majoré par une constante finie. Il nous reste donc à majorer $\left\| U_\delta \left(\frac{\partial}{\partial\theta^j} \Phi_p \right) \right\|_{L^2(\mu_{\theta_0})}$. Notons : $\Psi_p^j(x) = \pi_{\theta_0}^\delta \left(\frac{\partial}{\partial\theta^j} \Phi_p \right) (x)$.

On a : $\mu_{\theta_0} \left(\Psi_p^j \right) = \delta^2 \frac{\partial}{\partial\theta^j} U_p(\theta_{p,\delta}, \delta) = 0$ (lemme 6). De plus, utilisant (2.3) avec $h(x) = x$, on obtient pour M_p définie en (2.4) :

$$\Psi_p^j(x) = -2 \left(\frac{\partial}{\partial\theta^j} M_p \right) (x, \theta_{p,\delta}, \delta) \left[M_p(x, \theta_{p,\delta}, \delta) - M_p(x, \theta_0, \delta) + \int_0^\delta \frac{(\delta-v)^p}{p!} \pi_{\theta_0}^\delta \left(\mathcal{A}_{\theta_0}^p f_{\theta_0} \right) (x) \right]$$

Compte tenu de l'écart entre $\theta_{p,\delta}$ et θ_0 , on déduit de ce qui précède et de la propriété 2 qu'il existe D_j une constante positive finie telle que :

$$\left\| \Psi_p^j \right\|_{L^2(\mu_{\theta_0})} \leq D_j \delta^{p+2}$$

Donc :

$$\begin{aligned} \left\| U_\delta \left(\frac{\partial}{\partial\theta^j} \Phi_p \right) \right\|_{L^2(\mu_{\theta_0})} &\leq \sum_{k=0}^{\infty} \left\| \pi_{\theta_0}^{k\delta} \left(\Psi_p^j \right) \right\|_{L^2(\mu_{\theta_0})} \\ &\leq D_j \sum_{k=0}^{\infty} \exp(-\lambda k\delta) \delta^{p+2} \\ &\leq D_j \frac{\delta}{1 - \exp(-\lambda\delta)} \delta^{p+1} = \tilde{D}_j \delta^{p+1} \end{aligned}$$

où \tilde{D}_j est une constante finie.

En conclusion, pour tout $i = 1, d$ et tout $j = 1, d$, il existe une constante finie $D_{i,j}$ telle que :

$$\left| S_{i,j}^p \right| \leq D_{i,j} \delta^{p-\frac{1}{2}}$$

On prend alors $C_p = \max_{i,j} D_{i,j}$ et on obtient la majoration annoncée. Lorsque $p \geq 2$, $S_{i,j}^p(\theta_{p,\delta})$ est un $o(\delta)$, $T_{i,j}^p(\theta_{p,\delta})$ étant le terme dominant dans le développement en δ de $K_{i,j}^p(\theta_{p,\delta})$. Par contre, pour $p = 1$, $S_{i,j}^1(\theta_{p,\delta}) = o(1)$ et $K_{i,j}^1(\theta_{p,\delta})$ se développe en $4P_{i,j}(0) + o(1)$.

Chapter 3

Estimation du paramètre de dérive d'une diffusion sous des conditions d'irrégularité de la dérive.

3.1 Introduction

Soit (Ω, \mathcal{F}, P) un espace de probabilité, $B = (B_t)_{t \geq 0}$ un mouvement brownien standard issu de 0 et ξ une variable indépendante de B , définis sur l'espace précédent. On munit cet espace de la filtration $\{\mathcal{F}_t\}_{t \geq 0}$, où $\mathcal{F}_t = \sigma(\mathcal{G}_t \cup \mathcal{N})$ est la tribu $\mathcal{G}_t = \sigma(\xi, B_s; 0 \leq s \leq t)$ complétée par $\mathcal{N} = \{N \subseteq \Omega; \exists G \in \mathcal{G}_\infty, N \subseteq G, P(G) = 0\}$.

On considère l'équation différentielle stochastique (e.d.s.) définie par :

$$\begin{aligned} dX_t &= b(X_t - r_0) dt + dB_t, \quad 0 \leq t \leq T \\ X_0 &= \xi \end{aligned} \tag{3.1}$$

où b est une fonction connue et r_0 est un paramètre inconnu que l'on désire estimer.

On définit, pour tout réel a , $C_{\{a\}}^2$, l'ensemble des fonctions f continues sur R , de classe C^2 sur $R - \{a\}$ telles que, si l'on note f' et f'' les dérivées à l'ordre 1 et 2, $\lim_{x \rightarrow a^+} f'(x) = f'(a^+)$ et $\lim_{x \rightarrow a^+} f''(x) = f''(a^+)$ existent et sont finies avec $f'(a^+) \neq f'(a^-)$. Dans la suite, nous considérons une fonction b appartenant à $C_{\{0\}}^2$ et nous nommerons **diffusion à seuil** un tel modèle; un prototype est :

$$dX_t = b_1(X_t - r_0) I_{(X_t > r_0)} + b_2(X_t - r_0) I_{(X_t \leq r_0)} dt + dB_t \tag{3.2}$$

où b_1 et b_2 sont de classe C^2 sur R , telles que $b_1(0) = b_2(0)$ et $b_1'(0) \neq b_2'(0)$.

Sous l'hypothèse $b \in C_{\{0\}}^2$, le modèle étudié appartient à la classe plus générale des modèles réguliers : de dérive $x \rightarrow b(x, r)$ continue et différentiable en moyenne quadratique par rapport à r . Dans [13], Kutoyants montre que, pour un modèle régulier, l'estimateur du maximum de vraisemblance associé à une observation continue du processus stationnaire ergodique sur $[0, T]$ est optimal ($T \rightarrow +\infty$). Notre but est également de fournir une estimation optimale de r_0 dans le cas d'une diffusion à seuil mais basée sur une observation discrétisée à pas fin h_n ($h_n \rightarrow 0$) du processus : $(X_{kh_n})_{k=0, n}$ avec $nh_n = T$. Nous nous basons pour cela sur l'estimateur des Moindres Carrés (M.C.) associés au schéma d'approximation d'Euler.

Lorsque $(X_t)_{t \geq 0}$ est ergodique et la fonction de dérive est de classe C^2 sur R (modèle C^2), les résultats de Florens-Zmirou [8] impliquent la consistance de cet estimateur sous la condition $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$. Si de plus $nh_n^3 \rightarrow 0$, il est asymptotiquement normal et efficace avec une vitesse de l'ordre de $\sqrt{nh_n}$. Nous allons montrer que cet estimateur conserve ses propriétés asymptotiques sous des conditions analogues lorsque $b \in C_{\{0\}}^2$, si l'on impose que b soit **lipschitzienne**.

Dans une deuxième partie, nous présentons le schéma d'approximation d'Euler pour les diffusions à seuil. Le § 3 regroupe les hypothèses et résultats techniques permettant l'étude du comportement asymptotique de l'estimateur dérivé des moindres carrés et du schéma d'Euler. Dans la section 4, nous prouvons la consistance de cet estimateur lorsque $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$. Le § 5 est consacré à l'étude de la loi asymptotique. Le fait que le contraste ne soit pas de classe C^2 sur tout R , nous oblige à recourir à un schéma de démonstration différent des techniques habituellement utilisées (dérivation du contraste au point d'estimation). Nous commençons par établir la vitesse de convergence de l'estimateur en imposant un premier renforcement de l'hypothèse de décroissance de la suite $(h_n)_n$ vers 0 : si $(nh_n^3)_n$ est bornée, nous obtenons une vitesse de convergence de l'estimateur de l'ordre de $\sqrt{nh_n}$. Nous montrons ensuite que l'estimateur est asymptotiquement normal et efficace, sous la condition supplémentaire $nh_n^3 \rightarrow 0$. Dans le § 6, nous expérimentons les résultats précédents pour un modèle de diffusion particulier du type de (3.2) : un CTAR(1) (modèle AR à seuil et à temps continu) à raccordement continu. Dans la dernière partie, nous nous intéressons aux modèles autorégressifs à l'ordre 1 et à temps discret dont la fonction d'autorégression est de la forme $b(x - r_0)$ où b sous linéaire et appartient à $C_{\{0\}}^2$. Nous transposons ainsi les techniques précédemment utilisées pour les diffusions à seuil afin d'étudier le comportement asymptotique de l'estimateur des moindres carrés conditionnels du paramètre de seuil.

3.2 Résultat d'approximation d'Euler pour les modèles à seuil

Considérons l' e.d.s. définie par :

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t \quad (3.3)$$

Soit (X, B) une solution faible de cette équation (cf. [12], p.300). Pour un modèle C^2 , le schéma d'approximation d'Euler à pas δ permet d'approcher le modèle discret exact avec une erreur de l'ordre $O_P(\delta^2)$ [16]. Pour les diffusions à seuil, le résultat est le suivant :

Lemme 1 (Schéma d'Euler pour les modèles à seuil)

Supposons que :

1. $b \in C_{\{a\}}^2$ et σ est continue sur R .
2. Pour tout $t \geq 0$, $E\left(\int_0^t \sigma^2(X_u) du\right) < \infty$ et $E\left(\int_0^t [b'\sigma]^2(X_u) du\right) < \infty$.

Alors :

$$\forall t \geq 0, X_{t+\delta} = X_t + \delta b(X_t) + \eta_t + \varepsilon_t$$

avec :

$$\eta_t = \int_t^{t+\delta} (t + \delta - s) h(X_s) ds + \frac{1}{2} [b'(a^+) - b'(a^-)] \int_t^{t+\delta} [L_s(a) - L_t(a)] ds$$

où $h(x) = b(x)b'(x) + \frac{1}{2}\sigma^2(x)b''(x)$ et $(L_t(a))_{t \geq 0}$ est le temps local de $(X_t)_{t \geq 0}$ en a ([12], p.218). De plus, $(\varepsilon_t)_{t \geq 0}$ est une suite de variables de carré intégrable, telles que $\bar{E}[\varepsilon_t | \mathcal{F}_t] = 0$ et définies par :

$$\varepsilon_t = \int_t^{t+\delta} \sigma(X_s) [1 + (t + \delta - s)b'(X_s)] dB_s$$

Commentaire :

La différence entre le cas C^2 et le cas à seuil réside dans l'ordre de grandeur de la variable d'erreur η_t . Dans le premier cas, elle est en $O_P(\delta^2)$. Ceci n'est plus vrai pour les diffusions à seuil du fait de l'apparition du temps local. Cependant, si b est lipschitzienne, on va montrer que dans les deux cas, $E(|\eta_t|) \leq C\delta^2$. Ainsi, pour les diffusions à seuil comme pour les modèles C^2 [8], si $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$, les M.C. associés au schéma d'Euler fournissent un estimateur consistant et asymptotiquement normal.

Démonstration du lemme 1 :

La démonstration de ce résultat repose sur le lemme suivant dont la démonstration est donnée en annexe A.

Lemme 2 Soit $(X_t)_{t \geq 0}$ une semi-martingale continue, $X_t = X_0 + M_t + V_t$, et f un élément de $C_{\{a\}}^2$. Alors, pour tout $t \geq 0$:

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle M \rangle_s + [f'(a^+) - f'(a^-)] L_t(a) \quad (3.4)$$

Pour tout $t \geq 0$, on a :

$$X_{t+\delta} = X_t + \delta b(X_t) + \int_t^{t+\delta} (b(X_s) - b(X_t)) ds + \int_t^{t+\delta} \sigma(X_s) dB_s$$

Par application du lemme 2, on a pour tout $s \in [t, t + \delta]$:

$$\begin{aligned} b(X_s) - b(X_t) &= \int_t^s (b'\sigma)(X_u) dB_u + \int_t^s \left(bb' + \frac{1}{2}\sigma^2 b'' \right) (X_u) du \\ &+ [b'(a^+) - b'(a^-)] [L_s(a) - L_t(a)] \end{aligned} \quad (3.5)$$

Donc :

$$db(X_t) = h(X_t) dt + (b'\sigma)(X_t) dB_t + [b'(a^+) - b'(a^-)] dL_t(a) \quad (3.6)$$

avec $h(x) = (bb' + \frac{1}{2}\sigma^2 b'')(x)$.

D'après (3.6) et la formule de Ito appliquée au processus $(sb(X_s))_{s \geq 0}$, on a :

$$\begin{aligned} (t + \delta) b(X_{t+\delta}) - tb(X_t) &= \int_t^{t+\delta} b(X_s) ds + \int_t^{t+\delta} s db(X_s) \\ &= \int_t^{t+\delta} b(X_s) ds + \int_t^{t+\delta} s h(X_s) ds \\ &+ \int_t^{t+\delta} s (\sigma b')(X_s) dB_s + [b'(a^+) - b'(a^-)] \int_t^{t+\delta} s dL_s(a) \end{aligned}$$

De plus, si on multiplie (3.5) par $(t + \delta)$, on obtient :

$$(t + \delta) b(X_{t+\delta}) - (t + \delta) b(X_t) = \int_t^{t+\delta} (t + \delta) (b'\sigma)(X_s) dB_s + \int_t^{t+\delta} (t + \delta) h(X_s) ds \\ + [b'(a^+) - b'(a^-)] [(t + \delta) L_{t+\delta}(a) - (t + \delta) L_t(a)]$$

En soustrayant membre à membre les deux égalités précédentes, on obtient :

$$\int_t^{t+\delta} (b(X_s) - b(X_t)) ds = \int_t^{t+\delta} (t + \delta - s) (b'\sigma)(X_s) dB_s + \int_t^{t+\delta} (t + \delta - s) h(X_s) ds \\ = [b'(a^+) - b'(a^-)] \left[(t + \delta) L_{t+\delta}(a) - (t + \delta) L_t(a) - \int_t^{t+\delta} s dL_s(a) \right]$$

Or, appliquant la formule de Ito au processus $(sL_s(a))_{s \geq 0}$, on a :

$$(t + \delta) L_{t+\delta}(a) - (t + \delta) L_t(a) - \int_t^{t+\delta} s dL_s(a) = \int_t^{t+\delta} L_s(a) ds - \delta L_t(a)$$

On obtient ainsi le résultat annoncé. .

3.3 Hypothèses et résultats préliminaires

On définit $L_0^2(\mu) = \{f : R \rightarrow R, \mu(f) = 0, \mu(f^2) < \infty\}$ et pour tout $p \in N^*$, $\|f\|_{p,\mu} = [\mu(|f|^p)]^{\frac{1}{p}}$.

Pour $k \in N^*$, $\Delta_{kh_n} B = B_{kh_n} - B_{(k-1)h_n}$ représente l'accroissement brownien entre kh_n et $(k+1)h_n$.

On note pour $s \in [0, T]$, $b_s(r) = b(X_s - r)$, et pour $(a, b) \in R^2$, $M_{a,b} = \max\{|a|, |b|\}$.

On suppose que $r_0 \in]a, b[$, où a et b sont deux réels tels que $a < b$.

Posons les conditions suivantes **(A)** :

(A 1) b appartient à $C_{\{0\}}^2$ et il existe $K > 0$ telle que :

$$\forall (x, y) \in R^2, |b(x) - b(y)| \leq K |x - y|$$

On suppose de plus que b'' est à croissance polynomiale.

Sous (A 1), pour toute condition initiale ξ indépendante de $(B_t)_{t \geq 0}$, (3.1) admet une unique solution forte notée $(X_t)_{t \geq 0}$.

(A 2) Notons $s_{r_0}(x) = \exp\left(-2 \int_{r_0}^x b(y - r_0) dy\right)$ la dérivée de la fonction d'échelle associée à (3.1) et supposons que :

$$\int_0^\infty s_{r_0}(x) dx = \int_{-\infty}^0 s_{r_0}(x) dx = \infty \text{ et } \int_{-\infty}^\infty [s_{r_0}(x)]^{-1} dx = C(r_0) < \infty$$

Sous (A 1) et (A 2), $(X_t)_{t \geq 0}$ est récurrente positive sur R , de loi invariante $\mu_{r_0}(dx) = [C(r_0) s_{r_0}(x)]^{-1} dx$. Notons q_{r_0} la densité de probabilité de μ_{r_0} .

(A 3) La loi de ξ est μ_{r_0} .

(A 4) Notons $P_{r_0}^t$ le semi-groupe associé à (3.1).

μ_{r_0} admet des moments de tous ordres et il existe $\delta > 0$ tel que pour tout $t \geq 0$, on a :

$$\forall f \in L_0^2(\mu_{r_0}), \quad \|P_{r_0}^t f\|_{2, \mu_{r_0}} \leq \exp(-\delta t) \|f\|_{2, \mu_{r_0}}$$

(A 5) $I(r_0) = \int_R \left[(b')^2 ((x - r_0)^+) I_{(x > r_0)} + (b')^2 ((x - r_0)^-) I_{(x \leq r_0)} \right] q_{r_0}(x) dx > 0$

On peut faire les observations suivantes sur (A) :

1. Hansen et Scheinkman donnent dans [9] (proposition 9, p.801) des conditions suffisantes pour vérifier (A 4). Sous (A 1), (A 2) et (A 3), il suffit de montrer que :

(a) $\liminf_{x \rightarrow +\infty} q_{r_0}(x) = 0$ et $\lim_{x \rightarrow +\infty} \frac{1}{2b(x-r_0)}$ existe et est finie.

(b) $\liminf_{x \rightarrow -\infty} q_{r_0}(x) = 0$ et $\lim_{x \rightarrow -\infty} \frac{1}{2b(x-r_0)}$ existe et est finie.

Cette hypothèse est par exemple vérifiée pour un CTAR(1) (cf. § 6).

2. On remarque que si $(X_t)_{t \geq 0}$ est l'unique solution forte, stationnaire stricte et ergodique de (3.1), il en est de même pour $(X_t - r_0)_{t \geq 0}$ relativement à l'e.d.s. :

$$dX_t = b(X_t) dt + dB_t, \quad 0 \leq t \leq T$$

On peut donc raisonner en prenant $r_0 = 0$. On note alors $P_0 = P_{\mu_0}$ la loi du processus stationnaire et E_0 l'espérance sous P_0 .

3. Sous (A 1), b' est bornée sur R par K .

Nous terminerons ce paragraphe en énonçant deux propriétés et un lemme qui nous servirons par la suite.

Propriété 1 Sous (A) et si $(h_n)_{n \geq 0}$ est bornée, il existe une constante¹ C telle que, pour tout $k \in N^*$ et tout n , on a :

$$E_0 \left[\left(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] \right)^4 \right] \leq Ch_n^2$$

Propriété 2 Sous (A), il existe C telle que, pour tout $k \in N^*$ et tout n , on a :

$$E_0 \left[\left| E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0) \right| \right] \leq Ch_n^2$$

Les démonstrations de ces propriétés sont données respectivement en annexes B et C.

Lemme 3 Sous (A), pour toute f dans $L_0^2(\mu_0)$, on a, si $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$:

$$\frac{1}{n} \sum_{k=1}^n f(X_{(k-1)h_n}) \xrightarrow{L^2(P_0)} 0$$

¹Dans la suite, l'ensemble des constantes introduites sont finies et strictement positives.

Démonstration :

Si $f \in L_0^2(\mu_0)$, on a, du fait de la stationnarité, $E_0 \left[\frac{1}{n} \sum_{k=1}^n f(X_{(k-1)h_n}) \right] = 0$. En outre, pour $k \leq j$, en utilisant l'inégalité de Schwarz et l'hypothèse (A 4), on a :

$$\begin{aligned} |E_0 [f(X_{kh_n}) f(X_{jh_n})]| &= |E_0 [f(X_{kh_n}) E_0 [f(X_{jh_n}) | \mathcal{F}_{kh_n}]]| \\ &\leq \|f\|_{2,\mu_0} \left\| P_{r_0}^{(j-k)h_n} f \right\|_{2,\mu_0} \\ &\leq \|f\|_{2,\mu_0}^2 \exp(-\delta(j-k)h_n) \end{aligned}$$

D'où :

$$\begin{aligned} \text{Var} \left[\frac{1}{n} \sum_{k=1}^n f(X_{(k-1)h_n}) \right] &\leq \frac{1}{n^2} \|f\|_{2,\mu_0}^2 \left(n + 2 \sum_{0 \leq k < j \leq n} \exp(-\delta(j-k)h_n) \right) \\ &\leq \frac{1}{n} \|f\|_{2,\mu_0}^2 \left(1 + 2 \frac{\exp(-\delta h_n)}{1 - \exp(-\delta h_n)} \right) \end{aligned}$$

Or, si $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, $n(1 - \exp(-\delta h_n)) \rightarrow \infty$. On a donc la convergence annoncée.

3.4 Etude de la consistance

La fonction d'estimation obtenue en couplant la méthode des M.C. au schéma d'approximation d'Euler est donnée par :

$$U_n(r) = \frac{1}{nh_n^2} \sum_{k=1}^n (X_{kh_n} - X_{(k-1)h_n} - h_n b(X_{(k-1)h_n} - r))^2$$

Nous allons commencer par vérifier que U_n est bien un contraste au sens de la définition 3.2.7 de [7]. U_n s'exprime en outre en fonction de la martingale centrée M_n et du reste R_n définis par :

$$M_n(r) = \sum_{k=1}^n b_{(k-1)h_n}(r) (X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])$$

$$R_n(r) = \sum_{k=1}^n (b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)) [E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0)]$$

Ces deux variables satisfont les lemmes suivants qui sont des conséquences directes respectivement des propriétés 1 et 2.

Lemme 4 *Sous (A) et si $(h_n)_{n \geq 0}$ est bornée, il existe C telle que pour tout couple $(r, s) \in [a, b]^2$ et tout $n \in N^*$:*

$$E_0 \left[(M_n(r))^2 \right] \leq Cnh_n \quad \text{et} \quad E_0 \left[(M_n(r) - M_n(s))^2 \right] \leq Cnh_n (r - s)^2$$

Démonstration du lemme 4 :

M_n est une martingale centrée. Utilisant la propriété de lipschitz de b , on obtient :

$$\begin{aligned} & E_0 \left[(M_n(v) - M_n(u))^2 \right] \\ &= \sum_{k=1}^n E_0 \left[(b_{(k-1)h_n}(v) - b_{(k-1)h_n}(u))^2 (X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^2 \right] \\ &\leq K^2 (v - u)^2 \sum_{k=1}^n E_0 \left[(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^2 \right] \end{aligned}$$

En utilisant la propriété 1, on a l'inégalité annoncée.

Pour montrer l'autre inégalité, il suffit d'appliquer l'inégalité de Schwarz et d'utiliser à nouveau la propriété 1.

Lemme 5 *Sous (A), il existe C telle que pour tout $n \in N^*$:*

$$\forall (v, u) \in [a, b]^2, E_0 [|R_n(v) - R_n(u)|] \leq C |v - u| nh_n^2 \quad (3.7)$$

$$\forall \eta, \eta \in [0, b - a], E_0 \left[\sup_{|v-u| \leq \eta, (v,u) \in [a,b]^2} |R_n(v) - R_n(u)| \right] \leq C \eta nh_n^2 \quad (3.8)$$

$$\forall \delta, \delta \in]0, M_{a,b}[, E_0 \left[\sup_{\delta \leq |r| \leq M_{a,b}, r \in [a,b]} \left| \frac{R_n(r)}{r} \right| \right] \leq C nh_n^2 \quad (3.9)$$

Démonstration du lemme 5 :

Montrons (3.7). Pour tout $(v, u) \in [a, b]^2$, le caractère lipschitz de b implique que :

$$\begin{aligned} & E_0 [|R_n(v) - R_n(u)|] \\ &\leq \sum_{k=1}^n E_0 [|b_{(k-1)h_n}(v) - b_{(k-1)h_n}(u)| |E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0)|] \\ &\leq K |v - u| \sum_{k=1}^n E_0 [|E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0)|] \end{aligned}$$

Il suffit alors d'utiliser la propriété 2 pour obtenir (3.7). (3.8) et (3.9) se montrent de manière similaire.

Utilisant ces deux lemmes, on montre que :

Lemme 6 *Sous (A), lorsque $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, on a :*

$$\forall r \in [a, b], U_n(r) - U_n(0) \xrightarrow{L^1(P_0)} U(r, 0) = \int (b(x-r) - b(x))^2 \mu_0(dx)$$

où $r \rightarrow U(r, 0)$ est une fonction continue et positive, qui admet un unique minimum en $r = 0$.

Démonstration du lemme 6 :

Pour tout $r \in [a, b]$, on a :

$$U_n(r) - U_n(0) = \frac{2}{nh_n} (M_n(0) - M_n(r)) + \frac{1}{n} \sum_{k=1}^n (b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r))^2 + \frac{2}{nh_n} R_n(r)$$

D'où, d'après les lemmes 5 et 4 :

$$\begin{aligned} E_0 [|U_n(r) - U_n(0) - U(r, 0)|] &\leq \frac{2}{nh_n} E_0 [|M_n(0) - M_n(r)|] + \frac{2}{nh_n} E_0 [|R_n(r)|] \\ &+ E_0 \left[\left| \frac{1}{n} \sum_{k=1}^n (b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r))^2 - U(r, 0) \right| \right] \\ &\leq C \left(\frac{1}{\sqrt{nh_n}} + h_n \right) M_{a,b} \\ &+ E_0 \left[\left| \frac{1}{n} \sum_{k=1}^n (b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r))^2 - U(r, 0) \right| \right] \end{aligned}$$

De plus, le lemme 3 implique que si $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, on a :

$$E_0 \left[\left| \frac{1}{n} \sum_{k=1}^n (b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r))^2 - U(r, 0) \right| \right] \rightarrow 0$$

Pour tout r dans $[a, b]$, on obtient la convergence dans $L^1(P_0)$ de $U_n(r) - U_n(0)$ vers $U(r, 0)$. Montrons que $r \rightarrow U(r, 0)$ est une fonction continue qui admet un unique minimum en $r = 0$. Par hypothèse, b est lipschitzienne et μ_0 admet des moments de tous ordres. Par application du théorème de Lebesgue, on obtient la continuité de U sur R .

Si l'on suppose qu'il existe \tilde{r} tel que $\tilde{r} \neq 0$ et $U(\tilde{r}, 0) = 0$, alors pour tout $x \in R$, $b(x - \tilde{r}) = b(x)$. On a ainsi : $b'(\tilde{r}^+) = \lim_{h \rightarrow 0^+} \frac{b(\tilde{r}+h) - b(\tilde{r})}{h} = \lim_{h \rightarrow 0^+} \frac{b(h) - b(0)}{h} = b'(0^+)$ et de même $b'(\tilde{r}^-) = b'(0^-)$. Or, $b'(\tilde{r}^+) = b'(\tilde{r}^-)$. On obtient donc $b'(0^+) = b'(0^-)$ ce qui contredit l'hypothèse $b \in C_{\{0\}}^2$.

Notons \hat{r}_n l'estimateur du minimum de contraste associé à U_n :

$$\hat{r}_n = \inf_{r \in [a, b]} U_n(r)$$

Avant de démontrer la consistance de \hat{r}_n , il nous faut énoncer une version simplifiée du théorème 19 de Ibragimov et Has'minskii ([11], annexe I, p.372) :

Lemme 7 Soient a et b deux réels tels que $a < b$. On considère $(Y_r)_{r \in [a, b]}$ un processus continu à valeurs dans R . On suppose qu'il existe trois constantes $\alpha \geq \gamma + 1 > 1$ et $\beta > 0$ telles que, pour tout couple $(r, u) \in [a, b]^2$ vérifiant $r + u \in [a, b]$, on a :

$$E [|Y_r|^\alpha] \leq \beta, \quad E [|Y_{r+u} - Y_r|^\alpha] \leq \beta |u|^{\gamma+1}$$

Alors, si on note $V(Y, L, \delta) = \sup \{|Y_r - Y_u|; a \leq u \leq r \leq b, |u| \leq L, |r| \leq L, r - u \leq \delta\}$, on a :

$$E [V(Y, L, \delta)] \leq C_{\alpha, \gamma} \beta^{\frac{1}{\alpha}} L \delta^{\frac{\gamma}{\alpha}}$$

où $C_{\alpha, \gamma}$ est une constante qui ne dépend que de α et γ .

Théorème 1 Sous (A), si $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, $\hat{r}_n \xrightarrow{P_0} 0$

Démonstration du théorème 1 :

La démonstration de ce résultat repose sur la vérification des hypothèses du théorème 3.2.8. de [7]. La compacité de l'espace du paramètre et les continuités de $r \rightarrow U_n(r)$ et $r \rightarrow U(r, 0)$ sont bien vérifiées. Pour montrer la consistance de \hat{r}_n , il suffit alors d'exhiber une variable $V_n(\eta)$ telle que, pour la fonction V définie au lemme 7 :

$$\forall \eta \leq \eta_0, V_n(\eta) \geq V(U_n, M_{a,b}, \eta) \text{ p.s. et } V_n(\eta) \xrightarrow{P_0} \phi(\eta)$$

où ϕ est une fonction de R^+ dans R^+ qui vérifie $\lim_{\eta \rightarrow 0} \phi(\eta) = 0$.

Pour tout (r, u) , on a :

$$\begin{aligned} U_n(r) - U_n(u) &= \frac{2}{nh_n} (M_n(u) - M_n(r)) + \frac{2}{nh_n} (R_n(r) - R_n(u)) \\ &+ \frac{1}{n} \sum_{k=1}^n [b_{(k-1)h_n}(u) - b_{(k-1)h_n}(r)] [b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)] \\ &+ \frac{1}{n} \sum_{k=1}^n [b_{(k-1)h_n}(u) - b_{(k-1)h_n}(r)] [b_{(k-1)h_n}(0) - b_{(k-1)h_n}(u)] \end{aligned}$$

Pour tout $0 < \eta \leq b - a$, on a donc :

$$V(U_n, M_{a,b}, \eta) \leq \frac{2}{nh_n} V(M_n, M_{a,b}, b - a) + \frac{2}{nh_n} V(R_n, M_{a,b}, b - a) + 2K^2(b - a)\eta$$

D'après le lemme 4, il existe une constante C indépendante de n telle que, pour tout couple $(r, u) \in [a, b]^2$ vérifiant $r + u \in [a, b]$, on a :

$$E_0 [M_n^2(r)] \leq Cnh_n \text{ et } E_0 [(M_n(r + u) - M_n(r))^2] \leq Cnh_n u^2$$

Par application du lemme 7, on obtient :

$$\frac{2}{nh_n} E_0 [V(M_n, M_{a,b}, b - a)] \leq C' \sqrt{b - a} \frac{M_{a,b}}{\sqrt{nh_n}}$$

où C' est indépendante de n . D'où :

$$\frac{2}{nh_n} V(M_n, M_{a,b}, b - a) \xrightarrow{L^1(P_0)} 0$$

De plus, le lemme 5 implique que :

$$\frac{2}{nh_n} V(R_n, M_{a,b}, b - a) \xrightarrow{L^1(P_0)} 0$$

Posons alors :

$$V_n(\eta) = \frac{2}{nh_n} V(M_n, M_{a,b}, b - a) + \frac{2}{nh_n} V(R_n, M_{a,b}, b - a) + 2K^2(b - a)\eta$$

Compte tenu de ce qui précède, pour tout $\eta \leq b - a$, on a :

$$V_n(\eta) \xrightarrow{L^1(P_0)} \phi(\eta) = 2K^2\eta(b - a)$$

3.5 Vitesse de convergence et normalité asymptotique

Enonçons le résultat de normalité asymptotique de \hat{r}_n :

Théorème 2 *Sous (A), si $(h_n)_n$ est telle que $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$, et pour $I(0)$ définie en (A 5), on a :*

$$\sqrt{nh_n}\hat{r}_n \xrightarrow{\mathcal{L}(P_0)} \mathcal{N}\left(0, \frac{1}{I(0)}\right)$$

Du fait de la non régularité du contraste, ce résultat ne peut être obtenu par la méthode classique (résultat § 3.3.4. de [7]), basée sur la dérivabilité à l'ordre 2 du contraste et qui consisterait à développer $U'_n(\hat{r}_n) - U'_n(0)$. Ainsi, après avoir identifié la vitesse de convergence de \hat{r}_n vers 0, nous développons le contraste autour de 0 en mettant en évidence une décomposition qui satisfait une propriété LAN (Local Asymptotique Normality, cf. [11]).

Comme le montre le théorème qui suit, l'identification de la vitesse de convergence de \hat{r}_n vers 0 nécessite un premier renforcement de l'hypothèse de décroissance de la suite $(h_n)_n$ vers 0 : si $(nh_n^3)_n$ est bornée, la suite $(\sqrt{nh_n}\hat{r}_n)_{n \geq 0}$ est alors tendue.

Théorème 3 *Sous (A), si $(h_n)_n$ est telle que $nh_n \rightarrow \infty$ et $(nh_n^3)_n$ est bornée, alors pour tout $\varepsilon > 0$, il existe $H_0 > 0$ et $n_{H_0} \in \mathbb{N}^*$ tels que pour $n \geq n_{H_0}$:*

$$P_0\left(|\hat{r}_n| \leq \frac{H_0}{\sqrt{nh_n}}\right) \leq 1 - \varepsilon$$

La démonstration de ce résultat repose sur la transposition d'arguments développés par Chan [4] dans le cadre de modèles autorégressifs à seuil à temps discret.

Démonstration du théorème 3 :

Pour tout $\Delta \in]0, b[$, on a :

$$P_0\left(\inf_{\Delta \geq |r| \geq \frac{H}{\sqrt{nh_n}}} (U_n(r) - U_n(0)) > 0\right) \leq P_0(|\hat{r}_n| > \Delta) + P_0\left(|\hat{r}_n| < \frac{H}{\sqrt{nh_n}}\right)$$

Du fait de la consistance de \hat{r}_n , pour tout $\Delta \in]0, b[$ et tout $\varepsilon > 0$, on sait qu'il existe $n_\Delta \in \mathbb{N}^*$ tel que $\forall n \geq n_\Delta$, $P_0(|\hat{r}_n| > \Delta) \leq \frac{\varepsilon}{2}$. Il nous suffit donc de trouver $\Delta_0 > 0$ tel que, pour tout $\varepsilon > 0$, il existe $H_0 > 0$ et $n_{H_0} \in \mathbb{N}^*$ qui satisfont pour $n \geq n_{H_0}$:

$$P_0\left(\inf_{\Delta_0 \geq |r| \geq \frac{H_0}{\sqrt{nh_n}}} (U_n(r) - U_n(0)) > 0\right) \geq 1 - \frac{\varepsilon}{2}$$

On aura alors le résultat attendu en prenant $n'_{H_0} = \max\{n_{\Delta_0}, n_{H_0}\}$.

Dans le cas où $r > 0$, on a :

$$\begin{aligned} U_n(r) - U_n(0) &= \frac{1}{n} \sum_{k=1}^n (b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r))^2 + \frac{2}{nh_n} (M_n(0) - M_n(r)) + \frac{2}{nh_n} R_n(r) \\ &\geq r^2 \inf_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{n} \sum_{k=1}^n \left(\frac{b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)}{r}\right)^2 \\ &\quad - 2r^2 \sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left|\frac{M_n(0) - M_n(r)}{r^2}\right| - 2r^2 \sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left|\frac{R_n(r)}{r^2}\right| \end{aligned}$$

Or :

Lemme 8 Sous (A), si $(h_n)_n$ est telle que $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, $\forall \Delta \in]0, b[$, $\forall \varepsilon > 0$, $\forall H > 0$ et $\forall \eta > 0$, il existe n_0 tel que, $\forall n \geq n_0$:

$$P_0 \left(\inf_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{n} \sum_{k=1}^n \left(\frac{b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)}{r} \right)^2 \geq \phi(\Delta) - (2 + \Delta)\eta \right) \geq 1 - \varepsilon$$

avec $\phi(\Delta) = \|b'\|_{2, \mu_0}^2 - 4K^2q(\Delta) - \Delta G(\Delta)$, où q et G sont des fonctions continues, croissantes, telles que $q(0) = 0$ et $G(0) < \infty$.

La démonstration de ce résultat est donnée en annexe D.

De plus, on a les deux résultats suivants, démontrés en annexes E et F :

Lemme 9 Sous (A), si $(h_n)_n$ est telle que $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, pour tout $\Delta \in]0, b[$, $\varepsilon > 0$ et tout $\eta > 0$, il existe $H_1 > 0$ telle que pour $H \geq H_1$, $\exists n_H^1 \in N^*$ tel que, $\forall n \geq n_H^1$:

$$P_0 \left(\sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left| \frac{M_n(0) - M_n(r)}{r^2} \right| > \eta \right) \leq \varepsilon \quad (3.10)$$

Lemme 10 Sous (A), si $(h_n)_n$ est telle que $(nh_n^3)_n$ est bornée, pour $\Delta \in]0, b[$, $\varepsilon > 0$ et tout $\eta > 0$, il existe $H_2 > 0$ telle que, $\forall H \geq H_2$, $\exists n_H^2 \in N^*$ tel que, $\forall n \geq n_H^2$:

$$P_0 \left(\sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left| \frac{R_n(r)}{r^2} \right| > \eta \right) \leq \varepsilon \quad (3.11)$$

On obtient donc que, pour tout $\Delta \in]0, b[$, tout $\varepsilon > 0$ et tout $\eta > 0$, il existe $H_0 = \max\{H_1, H_2\} > 0$ et $n_{H_0} = \max\{n_0, n_{H_0}^1, n_{H_0}^2\}$ tels que, $\forall n \geq n_{H_0}$:

$$P_0 \left(\forall r \in \left[\frac{H_0}{\sqrt{nh_n}}, \Delta \right], U_n(r) - U_n(0) \geq r^2 [\phi(\Delta) - (6 + \Delta)\eta] \right) \geq 1 - 3\varepsilon$$

La fonction $\Delta \rightarrow \phi(\Delta)$ est continue et décroissante avec $\phi(0) = \|b'\|_{2, \mu_0}^2 > 0$. Il existe donc $\Delta_0 > 0$ tel que $\phi(\Delta_0) > 0$. On choisit alors $\eta = \frac{\phi(\Delta_0)}{2(6 + \Delta_0)}$.

Le cas où $r \in \left[-\Delta, -\frac{H}{\sqrt{nh_n}} \right]$ se traite de manière analogue.

Notons que la condition imposée sur la vitesse de décroissance de $(h_n)_n$ vers 0 (lemme 10) résulte de la majoration $E_0 \left[\left| \eta_{(k-1)h_n} \right| \right] \leq Ch_n^2$ (cf. annexe C, (3.27), p.121) où $(\eta_{(k-1)h_n})_{k=1, n}$ sont les erreurs définies au lemme 1.

Comme nous l'avons précisé, pour établir la normalité asymptotique de \widehat{r}_n , nous adoptons le type de critère développé par Ibragimov et Has'minskii dans [11]. Le théorème 1.2 de [11] établit la normalité asymptotique de l'estimateur du maximum de vraisemblance lorsque la vraisemblance du modèle considéré satisfait une condition LAN. Nous allons adapter ce résultat à notre cadre d'étude (théorème 4).

La suite $(\sqrt{nh_n}\widehat{r}_n)$ étant bornée en probabilité, on définit :

$$Z_n(u) = nh_n \left(U_n \left(\frac{u}{\sqrt{nh_n}} \right) - U_n(0) \right) \quad (3.12)$$

Théorème 4 *Supposons que, pour tout u , on puisse écrire :*

$$Z_n(u) = 2I(0)u\Delta_n + I(0)u^2 + \Psi_n(u) \quad (3.13)$$

où $I(0)$ est une constante strictement positive. Si de plus :

1. Δ_n converge vers la loi normale centrée et de variance $\frac{1}{I(0)}$.
2. $u \rightarrow \Psi_n(u)$ est continue p.s. et telle que, pour tout $B > 0$, $\sup_{|u| \leq B} |\Psi_n(u)|$ tend en probabilité vers 0 lorsque $n \rightarrow +\infty$.

Alors :
$$\sqrt{nh_n}\widehat{r}_n \xrightarrow{\mathcal{L}(P_0)} \mathcal{N}\left(0, \frac{1}{I(0)}\right)$$

La démonstration du théorème 4 est donnée en annexe G.

Pour montrer le théorème 2, il nous suffit donc d'exhiber une décomposition de Z_n de la forme (3.13) telle que les trois termes de ce développement, $I(0)$, Δ_n et Ψ_n , satisfont les hypothèses du théorème 4.

Commençons par identifier chacun des trois termes qui constituent ce développement. Nous utilisons pour cela une généralisation de la formule de Taylor dans le cas d'une fonction de $C_{\{0\}}^2$.

Lemme 11 *Si b est un élément de $C_{\{0\}}^2$, on a pour tout x et tout r :*

$$\begin{aligned} b(x) - b(x-r) &= [b'(x^+)I_{(x>0)} + b'(x^-)I_{(x\leq 0)}]r \\ &\quad - [b'(0^+) - b'(0^-)] [(r-x)^+I_{(x>0)} + (x-r)^+I_{(x\leq 0)}] \\ &\quad - \int_0^r \int_0^v b''(x-u) dudv \end{aligned}$$

Ce résultat repose sur le même type d'arguments que ceux développés dans la démonstration du lemme 2 (annexe A-1).

Nous noterons par la suite :

$$Db(x) = [b'(x^+)I_{(x>0)} + b'(x^-)I_{(x\leq 0)}] \quad (3.14)$$

$$\begin{aligned} \Lambda(r, x) &= [b'(0^+) - b'(0^-)] [(r-x)^+I_{(x>0)} + (x-r)^+I_{(x\leq 0)}] \\ &\quad + \int_0^r \int_0^v b''(x-u) dudv \end{aligned} \quad (3.15)$$

Par application du lemme 11, on a pour $I(0)$ définie en (A 5) :

$$\begin{aligned} Z_n(u) &= 2u \frac{1}{\sqrt{nh_n}} \sum_{k=1}^n Db(X_{(k-1)h_n}) (X_{kh_n} - E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}]) \\ &\quad - 2 \sum_{k=1}^n \Lambda\left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n}\right) (X_{kh_n} - E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}]) \\ &\quad + u^2 \left(\frac{1}{n} \sum_{k=1}^n Db^2(X_{(k-1)h_n}) - I(0) \right) + u^2 I(0) \\ &\quad - 2h_n \frac{u}{\sqrt{nh_n}} \sum_{k=1}^n Db(X_{(k-1)h_n}) \Lambda\left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n}\right) \\ &\quad + h_n \sum_{k=1}^n \Lambda^2\left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n}\right) + 2R_n\left(\frac{u}{\sqrt{nh_n}}\right) \end{aligned}$$

Posons :

$$\Delta_n = \frac{D_n^{(n)}}{I(0)} \quad \text{où} \quad D_j^{(n)} = \frac{1}{\sqrt{nh_n}} \sum_{k=1}^j Db(X_{(k-1)h_n}) (X_{kh_n} - E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}]), \quad (j = 1, n)$$

$$D_0^{(n)} = 0$$

$$\begin{aligned} \Psi_n(u) &= -2 \sum_{k=1}^n \Lambda\left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n}\right) (X_{kh_n} - E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}]) + 2R_n\left(\frac{u}{\sqrt{nh_n}}\right) \\ &+ u^2 \left(\frac{1}{n} \sum_{k=1}^n Db^2(X_{(k-1)h_n}) - I(0) \right) + h_n \sum_{k=1}^n \Lambda^2\left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n}\right) \\ &- 2u \sqrt{\frac{h_n}{n}} \sum_{k=1}^n Db(X_{(k-1)h_n}) \Lambda\left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n}\right) \end{aligned}$$

Il nous reste à montrer que, sous **(A)**, les termes de cette décomposition (3.13) satisfont les deux conditions du théorème 4.

La loi asymptotique de Δ_n résulte directement du théorème suivant :

Théorème 5 *Sous (A), lorsque $nh_n \rightarrow \infty$ et $h_n \rightarrow 0$, on a :*

$$D_n^{(n)} \xrightarrow{\mathcal{L}(P_0)} \mathcal{N}(0, I(0))$$

Démonstration du théorème 5 :

Sous (A), $D^{(n)} = \left(D_k^{(n)}\right)_{0 \leq k \leq n}$ est une $(\mathcal{F}_{(k-1)h_n})_{k \geq 1}$ -martingale de carré intégrable. La démonstration du théorème repose alors sur la vérification des deux conditions du théorème 2.8.43 de [7] (p.82).

• Montrons que $\langle D \rangle_n^{(n)} = \sum_{k=1}^n E_0 \left[\left(D_k^{(n)} - D_{k-1}^{(n)} \right)^2 \mid \mathcal{F}_{(k-1)h_n} \right] \xrightarrow{L^1(P_0)} I(0)$. En effet :

$$\begin{aligned} \langle D \rangle_n^{(n)} &= \frac{1}{nh_n} \sum_{k=1}^n (Db)^2(X_{(k-1)h_n}) E_0 \left[(X_{kh_n} - E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^2 \mid \mathcal{F}_{(k-1)h_n} \right] \\ &= \frac{1}{n} \sum_{k=1}^n (Db)^2(X_{(k-1)h_n}) + C_n \end{aligned}$$

avec :

$$C_n = \frac{1}{nh_n} \sum_{k=1}^n (Db)^2(X_{(k-1)h_n}) E_0 \left[\left(\int_{(k-1)h_n}^{kh_n} b_s(0) - E_0[b_s(0) | \mathcal{F}_{(k-1)h_n}] ds \right)^2 \mid \mathcal{F}_{(k-1)h_n} \right]$$

Le lemme 3 implique que :

$$\frac{1}{n} \sum_{k=1}^n (Db)^2(X_{(k-1)h_n}) \xrightarrow{L^1(P_0)} I(0)$$

On sait de plus que b' est bornée sur R par K (remarque 3, § 3.3), on a donc :

$$E_0 [C_n] \leq \frac{K^2}{nh_n} \sum_{k=1}^n E_0 \left[\left(\int_{(k-1)h_n}^{kh_n} b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}] ds \right)^2 \right]$$

L'inégalité de Schwarz donne alors :

$$\begin{aligned} E_0 [C_n] &\leq \frac{K^2}{n} \sum_{k=1}^n \int_{(k-1)h_n}^{kh_n} E_0 \left[(b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}])^2 \right] ds \\ &\leq 2K^2 \|b\|_{2, \mu_0}^2 h_n \end{aligned}$$

On obtient donc : $C_n \xrightarrow{L^1(P_0)} 0$. D'où : $\langle D \rangle_n \xrightarrow{L^1(P_0)} I(0)$. La première hypothèse du théorème 2.8.43 est donc vérifiée avec $\nu(n) = n$.

• Il nous faut maintenant montrer que la condition de Lindeberg est satisfaite. Pour tout $\varepsilon > 0$, on obtient par application de l'inégalité de Schwarz :

$$\begin{aligned} &\sum_{k=1}^n E_0 \left[E_0 \left[\left| D_k^{(n)} - D_{k-1}^{(n)} \right|^2 I_{(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon)} | \mathcal{F}_{(k-1)h_n} \right] \right] \\ &\leq \frac{1}{nh_n} \sum_{k=1}^n E_0 \left[(Db)^4 (X_{(k-1)h_n}) I_{(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon)} \right]^{\frac{1}{2}} E_0 \left[(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^4 \right]^{\frac{1}{2}} \end{aligned}$$

Or, d'après la propriété 1, il existe C telle que :

$$E_0 \left[(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^4 \right]^{\frac{1}{2}} \leq Ch_n$$

De plus :

$$E_0 \left[(Db)^4 (X_{(k-1)h_n}) I_{(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon)} \right]^{\frac{1}{2}} \leq K^2 P_0 \left(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon \right)^{\frac{1}{2}}$$

Or, l'inégalité de Markov à l'ordre 1 et la propriété 1 impliquent que :

$$\begin{aligned} P_0 \left(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon \right) &\leq \frac{1}{\varepsilon} E_0 \left[|D_k^{(n)} - D_{k-1}^{(n)}| \right] \\ &\leq \frac{K}{\varepsilon \sqrt{nh_n}} E_0 \left[(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^2 \right]^{\frac{1}{2}} \\ &\leq \frac{CK}{\varepsilon \sqrt{n}} \end{aligned}$$

Il existe donc une constante C' telle que pour tout n :

$$\sum_{k=1}^n E_0 \left[E_0 \left[\left| D_k^{(n)} - D_{k-1}^{(n)} \right|^2 I_{(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon)} | \mathcal{F}_{(k-1)h_n} \right] \right] \leq \frac{C'}{\varepsilon^{\frac{1}{2}} n^{\frac{1}{4}}}$$

D'où :

$$\sum_{k=1}^n E_0 \left[\left| D_k^{(n)} - D_{k-1}^{(n)} \right|^2 I_{(|D_k^{(n)} - D_{k-1}^{(n)}| \geq \varepsilon)} | \mathcal{F}_{(k-1)h_n} \right] \xrightarrow{L^1(P_0)} 0$$

Puisque la condition de Lindeberg est vérifiée, on obtient la convergence en loi annoncée. L'hypothèse 1 du théorème 4 est donc satisfaite car $\Delta_n = (I(0))^{-1} D_n^{(n)}$.

La vérification de l'hypothèse 2 du théorème 4 repose entre autre sur l'obtention du résultat suivant : pour tout $B > 0$, $\sup_{|u| \leq B} \left| R_n \left(\frac{u}{\sqrt{nh_n}} \right) \right| \xrightarrow{P_0} 0$. Comme dans le théorème 3, c'est l'ordre de la majoration, $E_0 [\eta_{(k-1)h_n}] \leq Ch_n^2$, qui va imposer une vitesse de décroissance de (h_n) vers 0. On a ainsi un nouveau renforcement de l'hypothèse (nh_n^3) bornée.

Théorème 6 *Sous (A), si $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$, on a pour tout $B > 0$:*

$$\sup_{|u| \leq B} |\Psi_n(u)| \xrightarrow{L^1(P_0)} 0$$

Démonstration du théorème 6 :

On a :

$$\sup_{|u| \leq B} |\Psi_n(u)| \leq 2 \sup_{|u| \leq B} \left| \sum_{k=1}^n \Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) (X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}]) \right| \quad (3.16)$$

$$+ 2 \sup_{|u| \leq B} \left| R_n \left(\frac{u}{\sqrt{nh_n}} \right) \right| \quad (3.17)$$

$$+ B^2 \left| \frac{1}{n} \sum_{k=1}^n Db^2 (X_{(k-1)h_n}) - I(0) \right| \quad (3.18)$$

$$+ h_n \sum_{k=1}^n \sup_{|u| \leq B} \Lambda^2 \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \quad (3.19)$$

$$+ 2B \sqrt{\frac{h_n}{n}} \sum_{k=1}^n |Db (X_{(k-1)h_n})| \sup_{|u| \leq B} \left| \Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right| \quad (3.20)$$

Montrons que chacun de ces cinq termes, (3.16) à (3.20), tend vers 0 au sens L^1 sous P_0 . Pour cela, nous utiliserons la propriété suivante, dont la démonstration est donnée en annexe H.

Propriété 3 *Sous (A), pour tout $B > 0$, il existe $C_B > 0$ (qui ne dépend que de B) telle que, pour $|u| \leq B$, $|v| \leq B$, et pour n tel que $nh_n \geq 1$, on a :*

$$E_0 \left[\left(\Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) - \Lambda \left(\frac{v}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right)^4 \right] \leq C_B \left(\frac{r-u}{\sqrt{nh_n}} \right)^4 \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{(nh_n)^2} \right) \quad (3.21)$$

$$E_0 \left[\sup_{|u| \leq B} \Lambda^2 \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right] \leq C_B \frac{1}{nh_n} \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{nh_n} \right) \quad (3.22)$$

où Q est une application continue telle que $Q(0) = 0$.

Terme (3.16)

Posons : $\Lambda_n(u) = \sum_{k=1}^n \Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) (X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])$. Montrons que

pour tout $B > 0$, $\sup_{|u| \leq B} |\Lambda_n(u)| \xrightarrow{L^1(P_0)} 0$, lorsque $nh_n \rightarrow \infty$ et $(h_n)_n$ est bornée. Soit n tel que $nh_n \geq 1$. On a :

$$\begin{aligned} & E_0 \left[(\Lambda_n(u) - \Lambda_n(v))^2 \right] \\ &= \sum_{k=1}^n E_0 \left[\left(\Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) - \Lambda \left(\frac{v}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right)^2 (X_{kh_n} - E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^2 \right] \end{aligned}$$

Appliquant l'inégalité de Schwarz et utilisant les propriétés 1 et 3-(3.21), pour tout $B > 0$, il existe C_B telle que pour tout u et tout v , $|u| \leq B$, $|v| \leq B$:

$$E_0 \left[(\Lambda_n(u) - \Lambda_n(v))^2 \right] \leq C_B \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{(nh_n)^2} \right)^{\frac{1}{2}} (v - u)^2$$

et :

$$E_0 [\Lambda_n^2(u)] = E_0 \left[(\Lambda_n(u) - \Lambda_n(0))^2 \right] \leq BC_B \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{(nh_n)^2} \right)^{\frac{1}{2}}$$

De plus, Λ_n est un processus continu car, pour tout x , $u \rightarrow \Lambda(u, x)$ est continue. On peut donc appliquer le lemme 7. On a donc pour une constante K_B finie :

$$E_0 [V(\Lambda_n, B, 2B)] \leq K_B \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{(nh_n)^2} \right)^{\frac{1}{4}}$$

Or : $\sup_{|u| \leq B} |\Lambda_n(u)| = \sup_{|u| \leq B} |\Lambda_n(u) - \Lambda_n(0)| \leq V(\Lambda_n, B, 2B)$. D'où :

$$E_0 \left[\sup_{|u| \leq B} |\Lambda_n(u)| \right] \leq K_B \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{(nh_n)^2} \right)^{\frac{1}{4}}$$

avec $Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{(nh_n)^2}$ qui tend vers 0, lorsque $nh_n \rightarrow \infty$. On a donc la convergence annoncée.

Terme (3.17)

Montrons que pour $B > 0$, $\sup_{|u| \leq B} \left| R_n \left(\frac{u}{\sqrt{nh_n}} \right) \right| \xrightarrow{L^1(P_0)} 0$, lorsque $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$. Soit $B > 0$ et n tel que $nh_n \geq 1$. Du fait que b est de Lipschitz, on a pour tout $x \in R$:

$$\sup_{|u| \leq B} \left| b(x) - b \left(x - \frac{u}{\sqrt{nh_n}} \right) \right| \leq K \frac{B}{\sqrt{nh_n}}$$

On obtient alors :

$$\begin{aligned} & E_0 \left[\sup_{|u| \leq B} \left| R_n \left(\frac{u}{\sqrt{nh_n}} \right) \right| \right] \\ & \leq \frac{KB}{\sqrt{nh_n}} \sum_{k=1}^n E_0 \left[|E_0[X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0)| \right] \end{aligned}$$

Or, d'après la propriété 2, il existe une constante finie C_B telle que pour tout n , $nh_n \geq 1$:

$$E_0 \left[\sup_{|u| \leq B} \left| R_n \left(\frac{u}{\sqrt{nh_n}} \right) \right| \right] \leq C_B \sqrt{nh_n^3}$$

Si $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$, on a donc :

$$\sup_{|u| \leq B} \left| R_n \left(\frac{u}{\sqrt{nh_n}} \right) \right| \xrightarrow{L^1(P_0)} 0$$

Terme (3.18)

Par application du lemme 3 : $\left| \frac{1}{n} \sum_{k=1}^n Db^2(X_{(k-1)h_n}) - I(0) \right| \xrightarrow{L^1(P_0)} 0$.

Terme (3.19)

D'après la propriété 3-(3.22), il existe C_B finie telle que si $nh_n \geq 1$:

$$E_0 \left[h_n \sum_{k=1}^n \sup_{|u| \leq B} \Lambda^2 \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right] \leq C_B \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{nh_n} \right)$$

avec $Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{nh_n} \rightarrow 0$ lorsque $nh_n \rightarrow \infty$. On a donc une convergence L^1 vers 0 de la variable considérée.

Terme (3.20)

Enfin, par application de l'inégalité de Schwarz et de la propriété 3-(3.22), il existe C_B telle que :

$$\begin{aligned} & E_0 \left[\sqrt{\frac{h_n}{n}} \sum_{k=1}^n |Db(X_{(k-1)h_n})| \sup_{|u| \leq B} \left| \Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right| \right] \\ & \leq \sqrt{\frac{h_n}{n}} K \sum_{k=1}^n E_0 \left[\sup_{|u| \leq B} \Lambda^2 \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right]^{\frac{1}{2}} \\ & \leq C_B K \sqrt{\frac{h_n}{n}} \frac{n}{\sqrt{nh_n}} \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{nh_n} \right)^{\frac{1}{2}} \\ & \leq C_B K \left(Q \left(\frac{B}{\sqrt{nh_n}} \right) + \frac{1}{nh_n} \right)^{\frac{1}{2}} \end{aligned}$$

On a donc :

$$\sqrt{\frac{h_n}{n}} \sum_{k=1}^n |Db(X_{(k-1)h_n})| \sup_{|u| \leq B} \left| \Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right| \xrightarrow{L^1(P_0)} 0$$

Chacun des cinq termes, (3.16) à (3.20), converge vers 0 au sens L^1 sous P_0 ; le théorème 6 est donc démontré.

Nous avons donc trouvé une décomposition de Z_n du type (3.13) telle que, sous **(A)**, les conditions du théorème 4 sont satisfaites sous la condition $nh_n \rightarrow \infty$ et $nh_n^3 \rightarrow 0$. Le théorème 2 est ainsi démontré.

3.6 Applications : Modèle CTAR(1)

Dans ce paragraphe, nous allons spécifier les résultats précédents pour un modèle CTAR(1) à raccordement continu défini par :

$$dX_t = (-a_1 X_t I_{(X_t > 0)} - a_2 X_t I_{(X_t \leq 0)}) dt + dB_t \quad (3.23)$$

Notons que les CTAR sans hypothèse de continuité sur la dérive sont étudiés dans [1], [2] et [10].

Dans le cas considéré, le paramètre de seuil est $r_0 = 0$. Commençons par examiner les conditions **(A)**. **(A 1)** est aisément vérifiée.

La fonction d'échelle associée à (3.23) est donnée par :

$$s_0(x) = \exp(a_1 x^2) I_{(x > 0)} + \exp(a_2 x^2) I_{(x \leq 0)}$$

(A 2) est ainsi satisfaite si et seulement si $a_1 > 0$ et $a_2 > 0$. La densité de la loi invariante μ_0 est :

$$q_0(x) = \left(\frac{\sqrt{a_2}}{\sqrt{a_1} + \sqrt{a_2}} \right) 2\sqrt{\frac{a_1}{\pi}} \exp(-a_1 x^2) I_{(x > 0)} + \left(\frac{\sqrt{a_1}}{\sqrt{a_1} + \sqrt{a_2}} \right) 2\sqrt{\frac{a_2}{\pi}} \exp(-a_2 x^2) I_{(x \leq 0)}$$

On vérifie facilement que $I(0) = \left(\frac{\sqrt{a_2 a_1^2 + \sqrt{a_1} a_2^2}}{\sqrt{a_1} + \sqrt{a_2}} \right)$ et que $I(0) > 0$ pour $a_1 > 0$ et $a_2 > 0$.

D'après la remarque 1 faite au § 3.3, **(A4)** est satisfaite si $a_1 > 0$ et $a_2 > 0$. On a en effet :

$$\lim_{|x| \rightarrow +\infty} q_0(x) = 0, \quad \lim_{|x| \rightarrow +\infty} \frac{1}{2b(x)} = 0$$

Etude expérimentale par simulation :

Nous prendrons comme domaine de recherche de r_0 , l'intervalle $[a, b]$ avec $a = -4$ et $b = 5$. Pour simuler le point initial $X_0(\omega)$ selon la loi μ_0 **(A 3)**, on se base sur la propriété suivante :

Propriété 4 Soient U une v.a. uniforme sur $[0, 1]$ et X une v.a. gaussienne centrée réduite. On suppose que U et X sont indépendantes. Posons :

$$Y = \frac{1}{\sqrt{2a_1}} I_{(0 \leq U \leq \frac{\sqrt{a_2}}{\sqrt{a_1} + \sqrt{a_2}})} - \frac{1}{\sqrt{2a_2}} I_{(\frac{\sqrt{a_2}}{\sqrt{a_1} + \sqrt{a_2}} \leq U \leq 1)}$$

Alors, la v.a. Z définie par $Z = Y|X|$ suit la loi μ_0 .

Démonstration de la propriété 4 :

Pour tout $z \leq 0$,

$$\begin{aligned} P(Z \leq z) &= P\left(Z \leq z, Y = \frac{1}{\sqrt{2a_1}}\right) + P\left(Z \leq z, Y = -\frac{1}{\sqrt{2a_2}}\right) \\ &= P\left(Y = \frac{1}{\sqrt{2a_1}}\right) P(|X| \leq \sqrt{2a_1}z) + P\left(Y = -\frac{1}{\sqrt{2a_2}}\right) P(|X| \geq -\sqrt{2a_2}z) \\ &= \frac{\sqrt{a_1}}{\sqrt{a_1} + \sqrt{a_2}} \left(\int_{-\sqrt{2a_2}z}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx + \int_{-\infty}^{\sqrt{2a_2}z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \right) \\ &= \frac{\sqrt{a_1}}{\sqrt{a_1} + \sqrt{a_2}} \int_{-\infty}^z 2\sqrt{\frac{a_2}{\pi}} \exp(-a_1 x^2) dx \end{aligned}$$

De même, pour tout $z \geq 0$, on montre que :

$$\begin{aligned}
P(Z \leq z) &= P\left(Y = \frac{1}{\sqrt{2a_1}}\right) P(-\sqrt{2a_1}z \leq X \leq \sqrt{2a_1}z) + P\left(Y = -\frac{1}{\sqrt{2a_2}}\right) \\
&= \frac{\sqrt{a_1}}{\sqrt{a_1} + \sqrt{a_2}} + \frac{\sqrt{a_2}}{\sqrt{a_1} + \sqrt{a_2}} \int_0^{\sqrt{2a_1}z} \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\
&= \frac{\sqrt{a_1}}{\sqrt{a_1} + \sqrt{a_2}} + \frac{\sqrt{a_2}}{\sqrt{a_1} + \sqrt{a_2}} \int_0^z 2\sqrt{\frac{a_1}{\pi}} \exp(-a_1x^2) dx
\end{aligned}$$

On a donc $P(Z \leq z) = \int_{-\infty}^z q_0(x) dx$ où q_0 est la densité de μ_0 .

Pour $a_1 = 1$, $a_2 = 2$ et $r_0 = 0$, nous simulons n observations à partir d'un schéma d'Euler à pas $\frac{h_n}{100}$. Comme la normalité asymptotique de \hat{r}_n est obtenue pour une suite telle que $nh_n^3 \rightarrow 0$, nous avons choisi $h_n = (n \ln(\ln(n)))^{-\frac{1}{3}}$. On a alors pour $n = 5000$, $h_n \simeq 0.04536$ et $T = nh_n \simeq 226.83$.

Cette étude numérique est menée en supposant dans un premier temps les paramètres a_1 et a_2 connus. Nous calculons $N = 100$ estimations indépendantes du paramètre de seuil. Nous notons $r_N = \frac{1}{N} \sum_{i=1}^N \hat{r}_n^i$ la moyenne empirique et $S_N = \frac{1}{N-1} \sum_{i=1}^N (\hat{r}_n^i - r_N)^2$ la variance empirique des estimations. r_N est à comparer à $r_0 = 0$ et TS_N à $\frac{1}{I(0)} = \frac{1+\sqrt{2}}{4+\sqrt{2}} \simeq 0.4459$.

Nous complétons cette étude numérique en testant la normalité $\mathcal{N}\left(0, \frac{1}{I(0)}\right)$ de la loi limite des $(\sqrt{T}\hat{r}_n^i)_{i=1, N}$ par un test du Khi-Deux. Pour ce faire, on choisit 20 classes $C_i =]b_{i-1}, b_i]$, avec $b_0 = -\infty$ et $b_{20} = +\infty$, sur lesquelles la loi $\mathcal{N}\left(0, \frac{1}{I(0)}\right)$ a une probabilité $p_i \simeq 0,05$. Pour tout $i = 1, 20$, on note $N_i = \sum_{j=1}^N I_{(\sqrt{nh_n}\hat{r}_n^j \in]b_{i-1}, b_i])}$ l'effectif observé pour la classe C_i , l'effectif théorique étant Np_i . On définit alors $K_N = \sum_{i=1}^N \frac{(N_i - Np_i)^2}{Np_i}$. Si la loi limite est la loi $\mathcal{N}\left(0, \frac{1}{I(0)}\right)$, la loi de K_N est un Khi-Deux à 19 degrés de liberté. La zone de rejet choisie pour un risque de première espèce de 5% est $\{K_N > 30.1\}$. Les effectifs empiriques et théoriques sont donnés dans le tableau 3.3.

Dans un deuxième temps, nous nous placerons dans le cas où a_1 et a_2 sont également des paramètres à estimer. Pour r fixé tel que $\min_{k=1, n} X_{(k-1)h_n} \leq r < \max_{k=1, n} X_{(k-1)h_n}$, les estimateurs du minimum de contraste de a_1 et a_2 associés à \mathcal{U}_n sont :

$$\begin{aligned}
\hat{a}_{1n}(r) &= \frac{1}{h_n} \frac{\sum_{k=1}^n (r - X_{(k-1)h_n}) (X_{kh_n} - X_{(k-1)h_n}) I_{(X_{(k-1)h_n} > r)}}{\sum_{k=1}^n (r - X_{(k-1)h_n})^2 I_{(X_{(k-1)h_n} > r)}} \\
\hat{a}_{2n}(r) &= \frac{1}{h_n} \frac{\sum_{k=1}^n (r - X_{(k-1)h_n}) (X_{kh_n} - X_{(k-1)h_n}) I_{(X_{(k-1)h_n} \leq r)}}{\sum_{k=1}^n (r - X_{(k-1)h_n})^2 I_{(X_{(k-1)h_n} \leq r)}}
\end{aligned}$$

On remplace alors dans le contraste a_1 et a_2 par leur r -estimation et on minimise en r la fonction ainsi construite. L'estimation \hat{r}_n étant obtenue, on estime a_1 et a_2 par $\hat{a}_{1n}(\hat{r}_n)$ et $\hat{a}_{2n}(\hat{r}_n)$. Nous indiquerons par * les résultats obtenus relativement à cette procédure d'estimation.

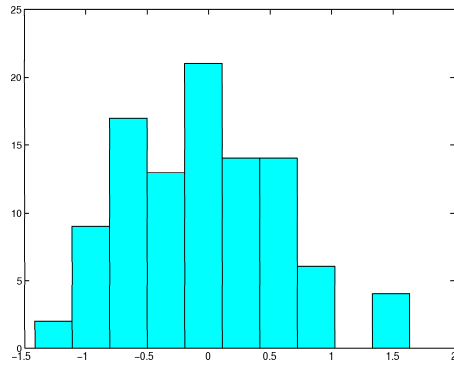


Figure 3.1: Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{r}_n^* - r_0)$ lorsque $a_1 = 1$ et $a_2 = 2$ sont supposés connus.

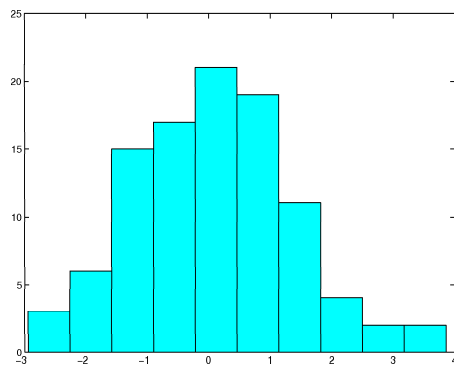


Figure 3.2: Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{r}_n^* - r_0)$ lorsque $a_1 = 1$ et $a_2 = 2$ sont estimés.

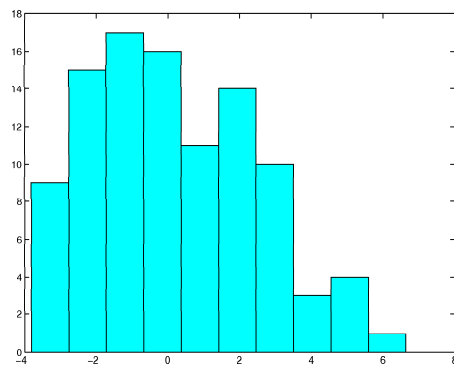


Figure 3.3: Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{a}_{1n} - a_1)$ lorsque $a_1 = 1$ et $a_2 = 2$.

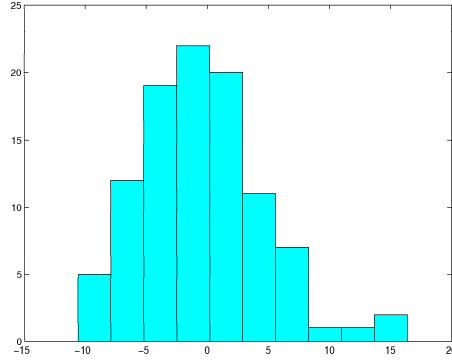


Figure 3.4: Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{a}_2 - a_2)$ lorsque $a_1 = 1$ et $a_2 = 2$.

Conclusion :

Que a_1 et a_2 soient estimés ou non, les moyennes empiriques pour le paramètre de seuil approchent r_0 avec une précision de l'ordre de 10^{-3} . Alors que la variance empirique est proche de la variance théorique si a_1 et a_2 sont connus, elle s'en éloigne naturellement lorsque a_1 et a_2 doivent être estimés, et cela bien que ces estimations soient en moyenne proches à 10^{-2} près des vraies valeurs des paramètres a_1 et a_2 . Lorsque $a_1 = 1$ et $a_2 = 2$ sont connus, le test du Khi-Deux nous permet de confirmer, au niveau 5%, que la loi limite commune des $(\sqrt{nh_n}\hat{r}_n^i)_{i=1,N}$ est la loi $\mathcal{N}\left(0, \frac{1}{I(0)}\right)$. On trouve en effet $K_N \simeq 22.59$.

Pour ce modèle à seuil particulier, lorsque a_1 et a_2 ne sont pas estimés, les résultats numériques sont conformes aux résultats théoriques énoncés aux § 3.4 et § 3.5.

Si on se réfère à [4], on peut conjecturer que l'estimateur du minimum de contraste $(\hat{a}_{1n}, \hat{a}_{2n}, \hat{r}_n)$ associé à (a_1, a_2, r_0) est tel que $(\hat{a}_{1n}, \hat{a}_{2n})$ est asymptotiquement normal avec une vitesse standard et indépendant de \hat{r}_n , la loi asymptotique de \hat{r}_n n'étant pas affectée par l'estimation de (a_1, a_2) .

3.7 Estimation du paramètre de seuil pour les chaînes de Markov

Dans [4], Chan étudie les propriétés asymptotiques de l'estimateur des moindres carrés conditionnels du paramètre de seuil dans le cas d'un modèle autorégressif à l'ordre p à temps discret ou TAR(p) (voir Tong [18] et [19]) :

$$x_k = \begin{cases} a_{10} + a_{11}x_{k-1} + \dots + a_{1p}x_{k-p} + c_1e_k, & \text{si } x_{k-d} \leq r_0 \\ a_{20} + a_{21}x_{k-1} + \dots + a_{2p}x_{k-p} + c_2e_k, & \text{si } x_{k-d} > r_0 \end{cases}$$

r_0	r_N	$I(0)^{-1}$	TS_N
0	-0.001974	0.4459	0.397615

Table 3.1: Moyenne et variance empiriques du paramètre de seuil pour $N = 100$ lorsque $a_1 = 1$ et $a_2 = 2$ sont connus.

a_N^1	a_N^2	r_N^*	TS_N^*
1.016415	1.969189	0.004132	1.689259

Table 3.2: Moyennes empiriques des estimateurs de a_1 , a_2 et r_0 pour $N = 100$ et $a_1 = 1$, $a_2 = 2$ et $r_0 = 0$.

Classe	effectif théorique : Np_i	effectif observé : N_i
$] -\infty; -1, 1352]$	4,46	2
$] -1, 1352; -0, 8746]$	5,05	6
$] -0, 8746; -0, 7078]$	4,95	5
$] -0, 7078; -0, 5742]$	5,03	11
$] -0, 5742; -0, 4607]$	5,02	4
$] -0, 4607; -0, 3605]$	4,95	3
$] -0, 3605; -0, 2671]$	5	6
$] -0, 2671; -0, 1736]$	5,28	4
$] -0, 1736; -0, 0868]$	5,09	4
$] -0, 0868; 0]$	5,17	6
$] 0; 0, 0868]$	5,17	9
$] 0, 0868; 0, 1736]$	5,09	5
$] 0, 1736; 0, 2671]$	5,28	8
$] 0, 2671; 0, 3605]$	5	1
$] 0, 3605; 0, 4607]$	4,95	4
$] 0, 4607; 0, 5742]$	5,02	4
$] 0, 5742; 0, 7078]$	5,03	7
$] 0, 7078; 0, 8746]$	4,95	6
$] 0, 8746; 1, 1352;]$	5,05	1
$] 1, 1352; +\infty[$	4,46	4

Table 3.3: Résultats des effectifs par classe relatifs au test du Khi-Deux pour le paramètre de seuil lorsque $a_1 = 1$ et $a_2 = 2$ ne sont pas estimés.

Notons que des auteurs tels que Petrucci [15] et Woolford [14] ont étudié les propriétés de cet estimateur dans le cas d'un TAR(1) (voir aussi [5]) :

$$x_k = \begin{cases} a_{10} + a_{11}x_{k-1} + c_1e_k, & \text{si } x_{k-1} \leq r_0 \\ a_{20} + a_{21}x_{k-1} + c_2e_k, & \text{si } x_{k-1} > r_0 \end{cases}$$

Dans [4], sous des conditions d'ergodicité géométrique et de stationnarité de la solution, Chan montre que l'estimateur de r_0 est consistant avec vitesse non standard de l'ordre de N et une loi limite liée à un processus de Poisson composé. Dans le cas considéré, la fonction autorégressive présente une réelle discontinuité en r_0 : on suppose qu'il existe (z_0, \dots, z_{p-1}) tel que $(a_{10} - a_{20}) + (a_{11} - a_{21})z_{p-1} + \dots + (a_{1p} - a_{2p})z_0 \neq 0$ et $z_{p-d} = r_0$. Une des questions posée par l'auteur est de déterminer la loi limite de cet estimateur lorsque cette discontinuité disparaît. Nous allons adapter les techniques mises en oeuvre précédemment pour apporter une réponse dans le cas d'un modèle autorégressif à l'ordre 1 du type :

$$x_k = b(x_{k-1} - r_0) + e_k \quad (3.24)$$

Ici, on suppose que b est une fonction connue appartenant à $C_{\{0\}}^2$. $(e_k)_{k \geq 1}$ est une suite de variables indépendantes, identiquement distribuées et centrées, telle que e_k soit indépendante de $\mathcal{F}_{k-1} = \sigma(x_0, x_1, \dots, x_{k-1})$.

On considère le groupe d'hypothèses **(B)** :

- (B 1)** On suppose que la loi des $(e_k)_k$ admet une densité g par rapport à la mesure de Lebesgue, g étant strictement positive sur R . De plus, il existe $\alpha \geq 1$ tel que $E[|e_1|^{4\alpha}] < \infty$.
On note σ^2 la variance commune des e_k .
- (B 2)** Il existe λ , $0 < \lambda < 1$, et $K \geq 0$ telles que, pour tout $x \in R$, $|b(x)| \leq \lambda|x| + K$.
Sous **(B 1)** et **(B 2)**, il existe une unique loi invariante μ_{r_0} qui admet un moment d'ordre 4α . De plus, μ_{r_0} a une densité q_{r_0} par rapport à la mesure de Lebesgue, q_{r_0} étant strictement positive sur R ([6], [3]).
- (B 3)** La loi de x_0 est μ_{r_0} .
- (B 4)** b' et b'' sont à croissance polynomiale d'ordre α sur R : il existe $C < \infty$ telle que, pour tout $x \in R$, $|b'(x)| + |b''(x)| \leq C(1 + |x|)^\alpha$.
- (B 5)** $I(r_0) = \frac{1}{\sigma^2} \int_R [(b')^2((x - r_0)^+) I_{(x > r_0)} + (b')^2((x - r_0)^-) I_{(x \leq r_0)}] q_{r_0}(x) dx > 0$

Sous **(B 1)** et **(B 2)**, (3.24) admet donc une unique solution stationnaire notée $(x_k)_k$. $(x_k)_k$ est une chaîne de Markov homogène. Notons Π_{r_0} sa probabilité de transition. On sait de plus que $(x_k)_k$ est récurrente positive ([6], théorème 1.IV.26, p.29). Si on note P_{r_0} la loi de la chaîne, on a alors :

Lemme 12 Pour toute fonction $h \in L^1(\mu_{r_0})$, $\frac{1}{n} \sum_{k=1}^n h(x_k) \xrightarrow{P_{r_0} \text{ p.s.}} \int_R h(x) q_{r_0}(x) dx$

Soient a et b deux réels tels que $r_0 \in]a, b[$. L'estimateur des moindres carrés conditionnels est donné par :

$$\tilde{r}_N = \inf_{r \in [a, b]} L_N(r), \quad L_N(r) = \frac{1}{N\sigma^2} \sum_{k=1}^N (x_k - b(x_{k-1} - r))^2$$

Théorème 7 *Sous (B), \tilde{r}_N est un estimateur consistant de r_0 . De plus :*

$$\sqrt{N} (\tilde{r}_N - r_0) \xrightarrow{\mathcal{L}(P_{r_0})} \mathcal{N} \left(0, \frac{1}{I(r_0)} \right)$$

L'estimateur des moindres carrés est asymptotiquement normal à la vitesse standard \sqrt{N} . La démonstration de ce théorème repose sur une adaptation des techniques développées dans le cas des diffusions à seuil. En annexe I, nous donnons les différentes étapes de la preuve et les lemmes techniques nécessaires. Notons que les calculs sont simplifiés par rapport au cas d'une diffusion, ceci pour deux raisons. La première est l'absence du terme d'erreur R_n dans les développements. En second lieu, les termes $X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}]$ sont remplacés par e_k supposés indépendants de \mathcal{F}_{k-1} , ce qui compense l'absence du caractère lipschitz de b .

Bibliography

- [1] Brockwell, P. J. & Hyndman, R. J. (1992) *On continuous-time threshold autoregression.*- International Journal of Forecasting 8, North-Holland, 157-173.
- [2] Brockwell, P. J., Hyndman, R. J. & G. K. Grunwald (1991) *Continuous time threshold autoregressive models.*- Statistica Sinica, Vol. 1, No. 2, 401-410.
- [3] Chan, K.S. & Tong, H. (1985) *On the use of the deterministic lyapunov function for the ergodicity of stochastic difference equations.*- Adv. Appl. Prob. 17, 666-678.
- [4] Chan, K. S. (1993) *Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model.*- The Annals of Statistics, Vol. 21, No. 1, 520-533.
- [5] Chan, K. S., Petruccielli, J. D., Tong, H., & Woolford, S. W. (1985) *A multiple threshold AR(1) model.*- J. Appl. Prob. 22, 267-279.
- [6] Duflo M. (1996) *Algorithmes stochastiques*- Mathématiques & Applications 23, Ed. Springer.
- [7] Dacunha-Castelle, D. & Duflo, M. (1993) *Probabilités et Statistiques : Problèmes à temps mobile.*- 2^{eme} Ed., Masson.
- [8] Florens-Zmirou, D. (1989) *Approximate discrete time schemes for statistics of diffusion processes.*- Statistics 20, 547-557.
- [9] Hansen, L. P. & Scheinkman, J.A. (1995) *Back to the future : generating moment implications for continuous-time markov processes.*- Econometrica, Vol. 63, No. 4, 767-804.
- [10] Hyndman, R. J. (1992) *Continuous time threshold autoregressive models.*- PhD thesis, University of Melbourne, Department of statistics.
- [11] Ibragimov, I. A. & Has'minskii, R. Z. (1981) *Statistical Estimation- Asymptotic Theory.*- Springer-Verlag.
- [12] Karatzas, I. & Shreve, S.E. (1996) *Brownian Motion and Stochastic Calculus.*- 2nd Ed., Springer.
- [13] Kutoyants, A. Y. (1984) *Parameter estimation for stochastic processes.*- Research and Exposition in math. 6, Heldermann Verlag, Berlin.
- [14] Petruccielli, J. D. & Woolford, S. W. (1984) *A threshold AR(1) model.*- J. Appl. Prob. 21, 270-286.

- [15] Petrucci, J. D. (1986) *On the consistency of least squares estimator of a threshold AR(1) model.*- J. Time Ser. Anal. 7, 269-278.
- [16] Souchet, S. (1998) *Schéma de discrétisation anticipatif et estimation du paramètre de dérive d'une diffusion.*- Note au CRAS, 327, série 1, 897-900.
- [17] Souchet, S. (1999) *Estimation des paramètres d'une diffusion ergodique observée à temps discret.*- Thèse de doctorat de l'Université Paris I.
- [18] Tong, H. (1987) *Non-linear time series models of regularly sampled data : A review.*- In *Proc. First World Congress of Bernoulli Society, Tashkent, USSR 2*, 355-367. VNU Science Press, The Netherlands.
- [19] Tong, H. (1990) *Non-linear Time Series : A Dynamical System Approach.*- Oxford statistical science series 6, Ed. Oxford science publications.

Annexes

A Démonstration du lemme 2

Dans la suite, on note : $S_h(a) = h(a^+) - h(a^-)$

1. Montrons que si $f \in C^2_{\{a\}}$, f est la différence de deux fonctions convexes.

Soit $x < a$. Pour tout $y \in \mathbb{R}$, on définit $G(y) = \int_x^y f''(u) du$.

Intégrant par parties, on obtient :

$$\forall t > 0, G(a+t) = f'(a+t) - f'(x) - S_{f'}(a)$$

De plus :

$$G(a-t) = f'(a-t) - f'(x)$$

On a donc :

$$\forall y \in \mathbb{R}, G(y) = f'(y^-) - f'(x) - S_{f'}(a) I_{(a < y)} \quad (3.25)$$

D'où :

$$\forall z \in \mathbb{R}, \int_x^z G(y) dy = f(z) - f(x) - f'(x)(z-x) - S_{f'}(a)(z-a)^+$$

Donc :

$$\begin{aligned} f(z) &= f_1(z) - f_2(z) \\ f_1(z) &= f(x) + f'(x)(z-x) + [S_{f'}(a)]^+(z-a)^+ + \int_x^z \int_x^y [f''(u)]^+ dudy \\ f_2(z) &= [S_{f'}(a)]^-(z-a)^+ - \int_x^z \int_x^y [f''(u)]^- dudy \end{aligned}$$

f_1 et f_2 étant des fonctions convexes, on a le résultat annoncé.

2. Pour achever la démonstration du lemme 2, on applique le changement de variable généralisé à f_1 et f_2 (théorème 7.1 [12], p.218). On obtient ainsi :

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \int_R L_t(x) \nu(dx)$$

où ν est la mesure définie par $\nu([u, v]) = f'(v^-) - f'(u^-)$. Or, d'après l'équation (3.25), on a pour $u < v$:

$$f'(v^-) - f'(u^-) = S_{f'}(a) \delta_a([u, v]) + \int_u^v f''(t) dt$$

On obtient donc :

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \int_R L_t(x) f''(x) dx + S_{f'}(a) L_t(a)$$

avec : $\int_R L_t(x) f''(x) dx = \frac{1}{2} \int_0^t f''(X_s) d \langle M \rangle_s$ ([12], p.218).

On a donc le développement annoncé.

B Démonstration de la propriété 1

Soient k et n quelconques. On a :

$$X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] = \Delta_{kh_n} B + \int_{(k-1)h_n}^{kh_n} (b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}]) ds$$

Or, pour tous réels x et y , $(x + y)^4 \leq 8(x^4 + y^4)$. Donc :

$$(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^4 \leq 8 \left[(\Delta_{kh_n} B)^4 + \left(\int_{(k-1)h_n}^{kh_n} (b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}]) ds \right)^4 \right]$$

Par une double application de l'inégalité de Schwarz, on obtient :

$$\begin{aligned} \left(\int_{(k-1)h_n}^{kh_n} (b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}]) ds \right)^4 &\leq h_n^3 \int_{(k-1)h_n}^{kh_n} (b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}])^4 ds \\ &\leq 8h_n^3 \int_{(k-1)h_n}^{kh_n} (b_s^4(0) + E_0 [b_s^4(0) | \mathcal{F}_{(k-1)h_n}]) ds \end{aligned}$$

En passant à l'espérance sous la loi invariante μ_0 , on a :

$$E_0 \left[\left(\int_{(k-1)h_n}^{kh_n} (b_s(0) - E_0 [b_s(0) | \mathcal{F}_{(k-1)h_n}]) ds \right)^4 \right] \leq 16h_n^4 \|b\|_{4,\mu_0}^4 \quad (3.26)$$

De plus :

$$E_0 [(\Delta_{kh_n} B)^4] = 3h_n^2$$

Si $(h_n)_n$ est bornée, il existe une constante C finie telle que :

$$E_0 \left[(X_{kh_n} - E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}])^4 \right] \leq Ch_n^2$$

On a le résultat annoncé.

C Démonstration de la propriété 2

D'après le lemme 1, on a :

$$|\eta_{(k-1)h_n}| \leq \int_{(k-1)h_n}^{kh_n} (kh_n - s) |h(X_s)| ds + h_n |b'(0^+) - b'(0^-)| (L_{kh_n}(0) - L_{(k-1)h_n}(0))$$

car $s \rightarrow L_s(0)$ est un processus croissant. D'où :

$$\begin{aligned} E_0 [|\eta_{(k-1)h_n}|] &\leq \int_{(k-1)h_n}^{kh_n} (kh_n - s) E_0 [|h(X_s)|] ds \\ &+ h_n |b'(0^+) - b'(0^-)| E_0 [(L_{kh_n}(0) - L_{(k-1)h_n}(0))] \end{aligned}$$

Compte tenu du fait que μ_0 est la loi invariante, qu'elle admet des moments de tous ordres et que h est à croissance polynomiale, on a :

$$\int_{(k-1)h_n}^{kh_n} (kh_n - s) E_0 [|h(X_s)|] ds \leq \frac{\|h\|_{1,\mu_0}^2}{2} h_n^2$$

De plus, d'après la formule de Tanaka-Meyer ([12], p.220), on a :

$$|X_{t+s}| = |X_t| + \int_t^{t+s} \text{sign}(X_u) dB_u + \int_t^{t+s} \text{sign}(X_u) b(X_u) du + L_{t+s}(0) - L_t(0)$$

On obtient donc en passant à l'espérance sous la loi invariante :

$$\begin{aligned} E_0 [(L_{kh_n}(0) - L_{(k-1)h_n}(0))] &= \int_{(k-1)h_n}^{kh_n} E_0 [\text{sign}(X_u) b(X_u)] du \\ &\leq \|b\|_{1,\mu_0} h_n \end{aligned}$$

De ce qui précède, on déduit que :

$$E_0 [|\eta_{(k-1)h_n}|] \leq Ch_n^2 \tag{3.27}$$

où C est une constante finie indépendante de n . Pour conclure, il suffit de constater que d'après le lemme 1, on a pour tout k :

$$E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0) = E_0 [\eta_{(k-1)h_n} | \mathcal{F}_{(k-1)h_n}]$$

Or :

$$E_0 [|E_0 [\eta_{(k-1)h_n} | \mathcal{F}_{(k-1)h_n}]|] \leq E_0 [|\eta_{(k-1)h_n}|]$$

Donc :

$$E_0 [|E_0 [X_{kh_n} | \mathcal{F}_{(k-1)h_n}] - X_{(k-1)h_n} - h_n b_{(k-1)h_n}(0)|] \leq Ch_n^2$$

D Démonstration du lemme 8

D'après le lemme 11, on a pour $\Delta \geq r > 0$ et tout x :

$$b(x) - b(x-r) = rb'(x^-) - (b'(0^+) - b'(0^-))(r-x)I_{(0 \leq x \leq r)} - r^2 \int_0^1 \int_0^t b''(x-rs) ds dt$$

On a donc :

$$\begin{aligned} (b(x) - b(x-r))^2 &\geq r^2 (b'(x^-))^2 - 2r^2 |b'(0^+) - b'(0^-)| |b'(x^-)| I_{(0 \leq x \leq \Delta)} \\ &\quad - 2r^3 [|b'(x^-)| + |b'(0^+) - b'(0^-)|] \int_0^1 \int_0^t |b''(x-rs)| ds dt \end{aligned}$$

Par hypothèse, b'' est à croissance polynomiale, il existe donc $\alpha > 0$ et C finie telles que, pour tout $x \in R$:

$$\forall r \in [0, \Delta], \int_0^1 \int_0^t |b''(x-rs)| ds dt \leq C(1 + |x| + |\Delta|)^\alpha$$

De plus, pour tout $x \in R$, $|b'(x)| \leq K$. D'où, pour $H > 0$ et $n \geq n_H^\Delta = \inf \left\{ n \in N^*, \frac{H}{\sqrt{nh_n}} \leq \Delta \right\}$:

$$\inf_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \left(\frac{b(x) - b(x-r)}{r} \right)^2 \geq (b'(x^-))^2 - g_1^\Delta(x) - \Delta g_2^\Delta(x)$$

avec :

$$g_1^\Delta(x) = 4K^2 I_{(0 \leq x \leq \Delta)}$$

et

$$g_2^\Delta(x) = 4KC(1 + |x| + |\Delta|)^\alpha$$

On a donc, si $H > 0$ et $n \geq n_H^\Delta$:

$$\inf_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{n} \sum_{k=1}^n \left(\frac{b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)}{r} \right)^2 \geq \frac{1}{n} \sum_{k=1}^n \left(b' \left(X_{(k-1)h_n}^- \right) \right)^2 - G_n^1(\Delta) - \Delta G_n^2(\Delta)$$

où : $G_n^i(\Delta) = \frac{1}{n} \sum_{k=1}^n g_i^\Delta(X_{(k-1)h_n})$ ($i=1,2$).

Sous (A), $(b')^2$, g_1^Δ et g_2^Δ sont dans $L^2(\mu_0)$. On peut alors définir $G_i(\Delta) = \mu_0(g_i^\Delta)$ ($i=1,2$) et appliquer le lemme 6 à chacune de ces variables. Par conséquent, pour tout $\varepsilon > 0$ et tout $\eta > 0$, il existe n_0 tel que $\forall n \geq n_0$:

$$\begin{aligned} P_0 \left(\left| \frac{1}{n} \sum_{k=1}^n \left(b' \left(X_{(k-1)h_n}^- \right) \right)^2 - \|b'\|_{2,\mu_0}^2 \right| \leq \eta \right) &\geq 1 - \frac{\varepsilon}{3} \\ i = 1, 2, \quad P_0 \left(|G_n^i(\Delta) - G_i(\Delta)| \leq \eta \right) &\geq 1 - \frac{\varepsilon}{3} \end{aligned}$$

On a donc :

$$P_0 \left(\forall i = 1, 2, |G_n^i(\Delta) - G_i(\Delta)| \leq \eta, \left| \frac{1}{n} \sum_{k=1}^n \left(b' \left(X_{(k-1)h_n}^- \right) \right)^2 - \|b'\|_{\mu_0}^2 \right| \leq \eta \right) \geq 1 - \varepsilon$$

Or :

$$\begin{aligned}
\inf_{\Delta \geq r \geq \frac{H}{\sqrt{n}h_n}} \frac{1}{n} \sum_{k=1}^n \left(\frac{b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)}{r} \right)^2 &\geq \|b'\|_{2,\mu_0}^2 - G_1(\Delta) - \Delta G_2(\Delta) \\
&- \left| \frac{1}{n} \sum_{k=1}^n \left(b'(X_{(k-1)h_n}^-) \right)^2 - \|b'\|_{2,\mu_0}^2 \right| \\
&- |G_n^1(\Delta) - G_1(\Delta)| - \Delta |G_n^2(\Delta) - G_2(\Delta)| \\
&\geq \|b'\|_{2,\mu_0}^2 - 4K^2q(\Delta) - \Delta G_2(\Delta) \\
&- \left| \frac{1}{n} \sum_{k=1}^n \left(b'(X_{(k-1)h_n}^-) \right)^2 - \|b'\|_{2,\mu_0}^2 \right| \\
&- |G_n^1(\Delta) - G_1(\Delta)| - \Delta |G_n^2(\Delta) - G_2(\Delta)|
\end{aligned}$$

avec : $q(\Delta) = \int_0^\Delta \mu_0(dx)$. Donc pour $\varepsilon > 0$, $\eta > 0$ et $H > 0$, il existe $n_1 = \max\{n_0, n\frac{\Delta}{H}\}$ tel que $\forall n \geq n_1$:

$$P_0 \left(\inf_{\Delta \geq r \geq \frac{H}{\sqrt{n}h_n}} \frac{1}{n} \sum_{k=1}^n \left(\frac{b_{(k-1)h_n}(0) - b_{(k-1)h_n}(r)}{r} \right)^2 \geq \phi(\Delta) - (2 + \Delta)\eta \right) \geq 1 - \varepsilon$$

avec $\phi(\Delta) = \|b'\|_{2,\mu_0}^2 - 4K^2q(\Delta) - \Delta G(\Delta)$, $G = G_2$.

E Démonstration du lemme 9

La démonstration de (3.10) s'inspire de la démarche développée par Chan [4].

Pour tout $d > 1$, $H > 0$ et $n \in N^*$, on définit $I(d, H, n) = \left\{ i \in N, \frac{d^{i+1}H}{\sqrt{nh_n}} \leq b \right\}$ et, pour

$\Delta \in]0, b]$, $N(\Delta, d, H) = \inf \left\{ n \in N^*, \frac{dH}{\sqrt{nh_n}} \leq \Delta, nh_n \geq 1 \text{ et } h_n \leq 1 \right\}$.

Montrons que, pour $\varepsilon > 0$, $\eta > 0$ et $d > 1$, il existe $H_0 > 0$ telle que $\forall H \geq H_0$ et $\forall n \geq N(b, d, H)$, on a :

$$P_0 \left(\sup_{i \in I(d, H, n)} \frac{1}{nh_n} \left| \frac{M_n(0) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right)}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \right| > \eta \right) \leq \varepsilon \quad (3.28)$$

On a :

$$P_0 \left(\sup_{i \in I(d, H, n)} \frac{1}{nh_n} \left| \frac{M_n(0) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right)}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \right| > \eta \right) \leq \sum_{i \in I(d, H, n)} P_0 \left(\frac{1}{nh_n} \left| \frac{M_n(0) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right)}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \right| > \eta \right)$$

Or, le lemme 4 implique qu'il existe C , une constante finie qui ne dépend que de a et b , telle que :

$$\forall n \geq N(b, d, H), \forall i \in I(d, H, n), E_0 \left[\left| \frac{1}{nh_n} \frac{M_n(0) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right)}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \right|^2 \right] \leq \frac{C}{(d^2)^i H^2}$$

D'où :

$$\forall n \geq N(b, d, H), P_0 \left(\sup_{i \in I(d, H, n)} \left| \frac{1}{nh_n} \frac{M_n(0) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right)}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \right| > \eta \right) \leq \frac{C}{(1-d^{-2})\eta^2 H^2}$$

Pour obtenir le résultat souhaité (3.28), il suffit de prendre $H \geq H_0 = \sqrt{\frac{C}{(1-d^{-2})\eta^2 \varepsilon^2}}$.

Nous allons montrer que, pour $\varepsilon > 0$, $\eta > 0$ et $d > 1$, il existe $H_0 > 0$ telle que, $\forall H \geq H_0$ et $\forall n \geq N(b, d, H)$, on a :

$$P_0 \left(\sup_{i \in I(d, H, n)} \sup_{r \in \left[\frac{d^i H}{\sqrt{nh_n}}, \frac{d^{i+1} H}{\sqrt{nh_n}} \right]} \frac{1}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \frac{1}{nh_n} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| > \eta \right) \leq \varepsilon \quad (3.29)$$

Pour tout $i \in I(d, H, n)$ et tout $r \in \left[\frac{d^i H}{\sqrt{nh_n}}, \frac{d^{i+1} H}{\sqrt{nh_n}} \right]$, on a :

$$\begin{aligned} \frac{1}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \frac{1}{nh_n} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| &= \frac{1}{(d^2)^i H^2} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| \\ &\leq \frac{1}{(d^2)^i H^2} V \left(M_n, \frac{d^{i+1} H}{\sqrt{nh_n}}, \frac{d^i (d-1) H}{\sqrt{nh_n}} \right) \end{aligned}$$

Or, par application des lemmes 4 et 7, on sait qu'il existe une constante C finie, indépendante de n , i , H et d , telle que :

$$\begin{aligned} E_0 \left[V \left(M_n, \frac{d^{i+1}H}{\sqrt{nh_n}}, \frac{d^i(d-1)H}{\sqrt{nh_n}} \right) \right] &\leq C \sqrt{nh_n} \frac{d^{i+1}H}{\sqrt{nh_n}} \sqrt{\frac{d^i(d-1)H}{\sqrt{nh_n}}} \\ &\leq C (d^i H)^{\frac{3}{2}} d \sqrt{\frac{d-1}{\sqrt{nh_n}}} \end{aligned}$$

D'où :

$$\begin{aligned} E_0 \left[\sup_{r \in \left[\frac{d^i H}{\sqrt{nh_n}}, \frac{d^{i+1}H}{\sqrt{nh_n}} \right]} \frac{1}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \frac{1}{nh_n} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| \right] &\leq \frac{C (d^i H)^{\frac{3}{2}} d \sqrt{d-1}}{(d^i H)^2 (nh_n)^{\frac{1}{4}}} \\ &\leq \frac{Cd \sqrt{d-1}}{\sqrt{d^i H} (nh_n)^{\frac{1}{4}}} \end{aligned}$$

Or, $n \geq N(b, d, H)$:

$$E_0 \left[\sup_{r \in \left[\frac{d^i H}{\sqrt{nh_n}}, \frac{d^{i+1}H}{\sqrt{nh_n}} \right]} \frac{1}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \frac{1}{nh_n} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| \right] \leq \frac{Cd \sqrt{d-1}}{\sqrt{d^i H}}$$

où $C < \infty$ est indépendante de n , i , d et H . En appliquant l'inégalité de Markov à l'ordre 1, on montre que, pour $\eta > 0$ et $n \geq N(b, d, H)$:

$$P_0 \left(\sup_{i \in I(d, H, n)} \sup_{r \in \left[\frac{d^i H}{\sqrt{nh_n}}, \frac{d^{i+1}H}{\sqrt{nh_n}} \right]} \frac{1}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \frac{1}{nh_n} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| > \eta \right) \leq \frac{Cd \sqrt{d-1}}{\sqrt{H} \eta (1 - d^{-\frac{1}{2}})} \quad (3.30)$$

Il suffit alors de prendre $H \geq \left(\frac{Cd \sqrt{d-1}}{\varepsilon \eta (1 - d^{-\frac{1}{2}})} \right)^2$ pour obtenir (3.29).

Enfin, on remarque que, pour tout $\Delta \in]0, b[$, $H > 0$ et $d > 1$, $\forall n \geq N(\Delta, d, H)$, on a :

$$\begin{aligned} \sup_{r \in \left[\frac{H}{\sqrt{nh_n}}, \Delta \right]} \frac{1}{nh_n} \left| \frac{M_n(0) - M_n(r)}{r^2} \right| &\leq \sup_{i \in I(d, H, n)} \sup_{r \in \left[\frac{d^i H}{\sqrt{nh_n}}, \frac{d^{i+1}H}{\sqrt{nh_n}} \right]} \frac{1}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \frac{1}{nh_n} \left| M_n(r) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right) \right| \\ &\quad + \sup_{i \in I(d, H, n)} \frac{1}{nh_n} \left| \frac{M_n(0) - M_n \left(\frac{d^i H}{\sqrt{nh_n}} \right)}{\left(\frac{d^i H}{\sqrt{nh_n}} \right)^2} \right| \end{aligned}$$

(3.28) et (3.29) impliquent donc (3.10).

F Démonstration du lemme 10

Soient $\Delta \in]0, b[$ et $H > 0$ quelconques. Pour tout $n \in N^*$ tel que $\frac{H}{\sqrt{nh_n}} \leq \Delta$, on a :

$$\sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left| \frac{R_n(r)}{r^2} \right| \leq \frac{1}{H} \frac{1}{\sqrt{nh_n}} \sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \left| \frac{R_n(r)}{r} \right|$$

Or, d'après le lemme 5-(3.9), on sait qu'il existe C finie qui ne dépend que de a et b telle que :

$$E_0 \left[\sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left| \frac{R_n(r)}{r} \right| \right] \leq Cnh_n^2$$

On en déduit que :

$$E_0 \left[\sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \left| \frac{R_n(r)}{r^2} \right| \right] \leq \frac{C}{H} \sqrt{nh_n^3}$$

Par hypothèse, la suite $(nh_n^3)_n$ est bornée. Si on note $M > 0$ cette borne, on a par application de l'inégalité de Markov à l'ordre 1 :

$$P_0 \left(\sup_{\Delta \geq r \geq \frac{H}{\sqrt{nh_n}}} \frac{1}{nh_n} \left| \frac{R_n(r)}{r^2} \right| > \eta \right) \leq \frac{C}{H} \sqrt{M}$$

Pour obtenir (3.11), il suffit de prendre $H \geq H_0 = \frac{C\sqrt{M}}{\varepsilon\eta}$.

G Démonstration du théorème 4

Si on montre que, pour tout $\delta > 0$, $\lim_{n \rightarrow \infty} P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right) = 0$, alors $\sqrt{nh_n} \hat{r}_n$ converge en loi vers la loi limite de $-\Delta_n$ et on aura le résultat annoncé.

Pour tout $A > 0$, on a :

$$\begin{aligned} P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right) &\leq P_0 \left(\left| \sqrt{nh_n} \hat{r}_n \right| > A \right) + P_0 \left(|\Delta_n| > A \right) \\ &+ P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A, |\Delta_n| \leq A \right) \end{aligned}$$

La condition 1 du théorème 4 implique que :

$\forall \varepsilon > 0$, $\exists A_1 = A_1(\varepsilon)$ et $\exists n_1 = n_1(\varepsilon)$ tels que, $\forall n \geq n_1$, $P_0 \left(|\Delta_n| > A_1 \right) \leq \varepsilon$.

La tension de la suite $(\sqrt{nh_n} \hat{r}_n)$ implique que :

$\forall \varepsilon > 0$, $\exists A_2 = A_2(\varepsilon)$ et $\exists n_2 = n_2(\varepsilon, A_2)$ tels que, $\forall n \geq n_2$, $P_0 \left(\left| \sqrt{nh_n} \hat{r}_n \right| > A \right) \leq \varepsilon$.

Montrons que :

$$P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A, |\Delta_n| \leq A \right) \leq P_0 \left(\sup_{|u| \leq A} |\Psi_n(u)| \geq I(0) \frac{\delta^2}{2} \right)$$

On a :

$$\forall u, |u| \leq A, \inf_{|u| \leq A} \frac{Z_n(u)}{2I(0)} \leq \frac{Z_n(u)}{2I(0)} \leq u\Delta_n + \frac{u^2}{2} + \frac{1}{2I(0)} \sup_{|u| \leq A} |\Psi_n(u)| \quad (3.31)$$

Or, sous $\left\{ \left| \sqrt{nh_n} \hat{r}_n \right| \leq A \right\}$, on a :

$$\inf_{|u| \leq A} \frac{Z_n(u)}{2I(0)} = \frac{Z_n(\sqrt{nh_n} \hat{r}_n)}{2I(0)}$$

Sous $\left\{ \left| \sqrt{nh_n} \hat{r}_n \right| \leq A, \left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right\}$,

$$\frac{Z_n(\sqrt{nh_n} \hat{r}_n)}{2I(0)} \geq \frac{\delta^2}{2} - \frac{\Delta_n^2}{2} - \frac{1}{2I(0)} \sup_{|u| \leq A} |\Psi_n(u)|$$

Compte tenu de (3.31), on obtient, sous $\left\{ |\Delta_n| \leq A, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A, \left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right\}$:

$$\frac{\delta^2}{2} - \frac{\Delta_n^2}{2} - \frac{1}{2I(0)} \sup_{|u| \leq A} |\Psi_n(u)| \leq -\Delta_n^2 + \frac{\Delta_n^2}{2} + \frac{1}{2I(0)} \sup_{|u| \leq A} |\Psi_n(u)|$$

Donc :

$$\left\{ |\Delta_n| \leq A, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A, \left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right\} \subset \left\{ \sup_{|u| \leq A} |\Psi_n(u)| \geq I(0) \frac{\delta^2}{2} \right\}$$

On a l'inégalité annoncée.

La condition 2 implique alors que $\forall \delta > 0$, $\forall \varepsilon > 0$, $\forall A > 0$, $\exists n_3 = n_3(\varepsilon, \delta, A)$ tel que, $\forall n \geq n_3$ $P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A, |\Delta_n| \leq A \right) \leq \varepsilon$.

Pour $\delta > 0$ et $\varepsilon > 0$ fixés, on pose $A_0 = \max \{A_1, A_2\}$ et $n_0 = \max \{n_1, n_2, n_3(A_0)\}$.

Pour tout $n \geq n_0$, on a :

$$\begin{aligned}
& P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right) \\
& \leq P_0 \left(\left| \sqrt{nh_n} \hat{r}_n \right| > A_0 \right) + P_0 (|\Delta_n| > A_0) + P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A_0, |\Delta_n| \leq A_0 \right) \\
& \leq P_0 \left(\left| \sqrt{nh_n} \hat{r}_n \right| > A_2 \right) + P_0 (|\Delta_n| > A_1) + P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta, \left| \sqrt{nh_n} \hat{r}_n \right| \leq A_0, |\Delta_n| \leq A_0 \right) \\
& \leq 3\varepsilon
\end{aligned}$$

Donc, pour tout $\delta > 0$ et tout $\varepsilon > 0$, il existe n_0 tel que, $\forall n \geq n_0$,

$$P_0 \left(\left| \sqrt{nh_n} \hat{r}_n + \Delta_n \right| > \delta \right) \leq \varepsilon$$

H Démonstration de la propriété 3

Soit n tel que $nh_n \geq 1$. On a pour tout x , tout u , $|u| \leq B$, et tout v , $|v| \leq B$:

$$\begin{aligned}
& \left| \Lambda \left(\frac{u}{\sqrt{nh_n}}, x \right) - \Lambda \left(\frac{v}{\sqrt{nh_n}}, x \right) \right| \\
& \leq |b'(0^+) - b'(0^-)| \left| \left(\frac{v}{\sqrt{nh_n}} - x \right)^+ - \left(\frac{u}{\sqrt{nh_n}} - x \right)^+ \right| I_{(x>0)} \\
& + |b'(0^+) - b'(0^-)| \left| \left(x - \frac{v}{\sqrt{nh_n}} \right)^+ - \left(x - \frac{u}{\sqrt{nh_n}} \right)^+ \right| I_{(x \leq 0)} \\
& + \int_{\frac{u}{\sqrt{nh_n}}}^{\frac{v}{\sqrt{nh_n}}} \int_0^y |b''(x-z)| dz dy \\
& \leq |b'(0^+) - b'(0^-)| \left| \frac{v-u}{\sqrt{nh_n}} \right| I_{(|x| \leq \frac{B}{\sqrt{nh_n}})} + \int_{\frac{u}{\sqrt{nh_n}}}^{\frac{v}{\sqrt{nh_n}}} \int_0^y |b''(x-z)| dz dy
\end{aligned}$$

Or, b'' est à croissance polynomiale, il existe donc une constante finie C et $\alpha > 0$ tels que :

$$\forall x \in \mathbf{R}, \forall z \in \left[-\frac{B}{\sqrt{nh_n}}, \frac{B}{\sqrt{nh_n}} \right], |b''(x-z)| \leq C \left(1 + |x| + \frac{B}{\sqrt{nh_n}} \right)^\alpha$$

Puisque $nh_n \geq 1$, on a : $\forall x \in \mathbf{R}, \forall z \in \left[-\frac{B}{\sqrt{nh_n}}, \frac{B}{\sqrt{nh_n}} \right], |b''(x-z)| \leq C(1 + |x| + B)^\alpha$.

Posons : $G_B(x) = C(1 + |x| + B)^\alpha$. D'après ce qui précède, on a pour tout x , si $|u| \leq B$, $|v| \leq B$ et $nh_n \geq 1$:

$$\left| \Lambda \left(\frac{u}{\sqrt{nh_n}}, x \right) - \Lambda \left(\frac{v}{\sqrt{nh_n}}, x \right) \right| \leq |b'(0^+) - b'(0^-)| \left| \frac{v-u}{\sqrt{nh_n}} \right| I_{(|x| \leq \frac{B}{\sqrt{nh_n}})} + BG_B(x) \left| \frac{v-u}{nh_n} \right|$$

D'où :

$$\begin{aligned}
& \left[\Lambda \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) - \Lambda \left(\frac{v}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) \right]^4 \\
& \leq 8 [b'(0^+) - b'(0^-)]^4 \left(\frac{v-u}{\sqrt{nh_n}} \right)^4 I_{(|X_{(k-1)h_n}| \leq \frac{B}{\sqrt{nh_n}})} + 8B^4 (G_B)^4 (X_{(k-1)h_n}) \left(\frac{v-u}{nh_n} \right)^4
\end{aligned}$$

et

$$\begin{aligned}
\sup_{|u| \leq B} \Lambda^2 \left(\frac{u}{\sqrt{nh_n}}, X_{(k-1)h_n} \right) & \leq 2 [b'(0^+) - b'(0^-)]^2 \frac{B^2}{nh_n} I_{(|X_{(k-1)h_n}| \leq \frac{B}{\sqrt{nh_n}})} \\
& + 2 (G_B)^2 (X_{(k-1)h_n}) \left(\frac{B^2}{nh_n} \right)^2
\end{aligned}$$

Pour obtenir (3.21) et (3.22), il suffit de passer à l'espérance sous μ_0 dans les deux inégalités précédentes : $(G_B)^2$ et $(G_B)^4$ sont intégrables car, par hypothèse, μ_0 admet des moments de tous ordres. On a ainsi, dans les membres de droite, deux constantes, qui ne dépendent que de B , et $Q(\varepsilon) = \int_{-\varepsilon}^\varepsilon \mu_0(dx)$. μ_0 étant absolument continue par rapport à la mesure de Lebesgue, Q est continue et vaut 0 en 0.

I Schéma de la démonstration du théorème 7

On suppose que $r_0 = 0$. Posons : $\widetilde{M}_N(r) = \sum_{k=1}^N b(x_{k-1} - r) e_k$.

I.1 Préliminaires

Nous commencerons par montrer que notre procédure d'estimation s'inscrit dans le cadre de l'estimation par minimum de contraste. Pour ce faire, nous utilisons le résultat suivant :

Lemme 13 *Sous (B), il existe une constante C finie telle que :*

$$\begin{aligned} \forall (r, s) \in [a, b]^2, \quad E_0 \left[\left(\widetilde{M}_N(r) - \widetilde{M}_N(s) \right)^2 \right] &\leq NC (r - s)^2 \\ \forall r \in [a, b], \quad E_0 \left[\widetilde{M}_N(r)^2 \right] &\leq NC \end{aligned}$$

Démonstration :

Par hypothèse, e_k est indépendant de $(x_0, e_1, \dots, e_{k-1})$ et est centrée. Donc, pour tout r et tout s dans $[a, b]$:

$$\begin{aligned} E_0 \left[\left(\widetilde{M}_N(r) - \widetilde{M}_N(s) \right)^2 \right] &\leq N\sigma^2 \int (b(x-r) - b(x-s))^2 \mu_0(dx) \\ E_0 \left[\widetilde{M}_N(r)^2 \right] &\leq N\sigma^2 \int b^2(x-r) \mu_0(dx) \end{aligned}$$

Compte tenu de l'hypothèse de sous-linéarité de b et de croissance polynomiale de b' , on a pour tout $x \in R$:

$$b^2(x-r) \leq (G_{a,b})^2(x) \text{ et } (b(x-r) - b(x-s))^2 \leq (r-s)^2 (G_{a,b})^2(x)$$

où : $G_{a,b}(x) = C(1 + |x| + M_{a,b})^{\max(1, \alpha)}$. $(G_{a,b})^2$ étant intégrable par rapport à μ_0 , on a les résultats annoncés.

Lemme 14 *Sous (B), pour tout $r \in [a, b]$:*

$$L_N(r) - L_N(0) \xrightarrow{P_0} L(r, 0) = \frac{1}{\sigma^2} \int_R (b(x-r) - b(x))^2 \mu_0(dx)$$

où $r \rightarrow L(r, 0)$ est continue, positive et admet un unique minimum en $r = 0$.

Démonstration :

Pour tout $r \in [a, b]$, on a :

$$\begin{aligned} L_N(r) - L_N(0) &= \frac{2}{N\sigma^2} \left(\widetilde{M}_N(0) - \widetilde{M}_N(r) \right) \\ &\quad + \frac{1}{N\sigma^2} \sum_{k=1}^N (b(x_{k-1}) - b(x_{k-1} - r))^2 \end{aligned}$$

Le lemme 13 implique que, si $N \rightarrow \infty$ alors $\frac{1}{N} \left(\widetilde{M}_N(0) - \widetilde{M}_N(r) \right) \xrightarrow{L^2(P_0)} 0$.

De plus, b est à croissance sous linéaire sur R et μ_0 admet un moment d'ordre 4α ($\alpha \geq 1$).

On a ainsi: $x \rightarrow (b(x) - b(x - r))^2 \in L^1(\mu_0)$. Par application du lemme 12, on obtient alors :

$$\frac{1}{N\sigma^2} \sum_{k=1}^N (b(x_{k-1}) - b(x_{k-1} - r))^2 \xrightarrow{P_0} L(r, 0)$$

Pour montrer que $L(\cdot, 0)$ est une fonction continue qui admet un unique minimum en $r = 0$, on procède comme dans la démonstration du lemme 6.

On vérifie aisément que L_N et L satisfont les conditions de la définition 3.2.7 de [7].

I.2 Consistance

Pour montrer que \tilde{r}_N est consistant, on procède comme dans la démonstration du théorème 1. Pour tout η , $0 < \eta \leq b - a$, on a :

$$V(L_N, M_{a,b}, \eta) \leq \frac{2}{N\sigma^2} V(\tilde{M}_N, M_{a,b}, \eta) + \frac{(b-a)\eta}{\sigma^2} \frac{2}{N} \sum_{k=1}^N (G_{a,b})^2(x_{k-1})$$

Utilisant les lemmes 13 et 7, on montre que si $N \rightarrow \infty$, $\frac{1}{N} V(\tilde{M}_N, M_{a,b}, \eta) \xrightarrow{L^1(P_0)} 0$. De plus, $(G_{a,b})^2 \in L^1(\mu_0)$. Par application du lemme 12, on obtient :

$$\frac{1}{N} \sum_{k=1}^N (G_{a,b})^2(x_{k-1}) \xrightarrow{P_0 \text{ P.S.}} \|G_{a,b}\|_{2, \mu_0}^2$$

Si on pose $V_N(\eta) = \frac{2}{N\sigma^2} V(\tilde{M}_N, M_{a,b}, \eta) + \frac{(b-a)\eta}{\sigma^2} \frac{2}{N} \sum_{k=1}^N (G_{a,b})^2(x_{k-1})$, on a alors :

$V_N(\eta) \geq V(L_N, M_{a,b}, \eta)$ et $V_N(\eta) \xrightarrow{P_0} \phi(\eta) = 2 \frac{(b-a)}{\sigma^2} \|G_{a,b}\|_{2, \mu_0}^2 \eta$ avec $\lim_{\eta \rightarrow 0} \phi(\eta) = 0$.

I.3 Normalité asymptotique

Nous allons procéder comme au § 3.5 : après avoir identifié la vitesse de convergence de \tilde{r}_N vers 0, nous montrons que le contraste se développe localement autour de 0 et satisfait une condition LAN.

Théorème 8 *Sous (B), pour tout $\varepsilon > 0$, il existe H et $N_H \in N^*$ tels que $\forall N \geq N_H$:*

$$P_0 \left(|\tilde{r}_N| \leq \frac{H}{\sqrt{N}} \right) \leq 1 - \varepsilon$$

La démonstration de ce résultat de tension peut être calquée sur la preuve du théorème 3 : il repose en particulier sur les deux lemmes suivants :

Lemme 15 Sous (B), pour tout $\Delta \in]0, b]$, $\varepsilon > 0$, $H > 0$ et tout $\eta > 0$, il existe N_0 tel que, $\forall N \geq N_0$:

$$P_0 \left(\inf_{\Delta \geq |r| \geq \frac{H}{\sqrt{N}}} \frac{1}{N} \sum_{k=1}^N \left(\frac{b(x_{k-1}) - b(x_{k-1} - r)}{r} \right)^2 \geq \phi(\Delta) - (2 + \Delta)\eta \right) \geq 1 - \varepsilon$$

avec $\phi(\Delta) = \|b'\|_{2, \mu_0}^2 - 2|b'(0^+) - b'(0^-)| \|b'\|_{2, \mu_0} q(\Delta) - \Delta G(\Delta)$, où q et G sont des fonctions continues, croissantes, telles que $q(0) = 0$ et $G(0) < \infty$.

et :

Lemme 16 Sous (B), pour $\Delta \in]0, b[$, $\varepsilon > 0$ et $\eta > 0$, il existe $H_0 > 0$ telle que $\forall H \geq H_0$, $\exists N_H \in N^*$, $\forall N \geq N_H$:

$$P_0 \left(\sup_{\Delta \geq |r| \geq \frac{H}{\sqrt{N}}} \frac{1}{N} \left| \frac{\widetilde{M}_N(r) - \widetilde{M}_N(0)}{r^2} \right| > \eta \right) \leq \varepsilon$$

On définit alors pour tout u :

$$\begin{aligned} \widetilde{Z}_N(u) &= N \left(L_N \left(\frac{u}{\sqrt{N}} \right) - L_N(0) \right) \\ &= \frac{2}{\sigma^2} \left(\widetilde{M}_N(0) - \widetilde{M}_N \left(\frac{u}{\sqrt{N}} \right) \right) + \frac{1}{\sigma^2} \sum_{k=1}^N \left(b \left(x_{k-1} - \frac{u}{\sqrt{N}} \right) - b(x_{k-1}) \right)^2 \end{aligned}$$

D'après le lemme 11 et avec les notations introduites en (3.14) et (3.15), on obtient :

$$\begin{aligned} \widetilde{Z}_N(u) &= 2u \frac{1}{\sigma^2} \frac{1}{\sqrt{N}} \sum_{k=1}^N Db(x_{k-1}) e_k - \frac{2}{\sigma^2} \sum_{k=1}^N \Lambda \left(\frac{u}{\sqrt{N}}, x_{k-1} \right) e_k \\ &+ \frac{u^2}{\sigma^2} \left(\frac{1}{N} \sum_{k=1}^N Db^2(x_{k-1}) - \sigma^2 I(0) \right) + u^2 I(0) \\ &- 2 \frac{u}{\sqrt{N} \sigma^2} \sum_{k=1}^N Db(x_{k-1}) \Lambda \left(\frac{u}{\sqrt{N}}, x_{k-1} \right) + \frac{1}{\sigma^2} \sum_{k=1}^N \Lambda^2 \left(\frac{u}{\sqrt{N}}, x_{k-1} \right) \end{aligned}$$

avec $I(0)$ définie en (B 5). Posons :

$$\begin{aligned} \Delta_N &= \frac{1}{\sigma^2 I(0) \sqrt{N}} \sum_{k=1}^N Db(x_{k-1}) e_k \\ \Psi_N(u) &= - \frac{2}{\sigma^2} \sum_{k=1}^N \Lambda \left(\frac{u}{\sqrt{N}}, x_{k-1} \right) e_k + \frac{u^2}{\sigma^2} \left(\frac{1}{N} \sum_{k=1}^N Db^2(x_{k-1}) - \sigma^2 I(0) \right) \\ &- 2 \frac{u}{\sqrt{N} \sigma^2} \sum_{k=1}^N Db(x_{k-1}) \Lambda \left(\frac{u}{\sqrt{N}}, x_{k-1} \right) + \frac{1}{\sigma^2} \sum_{k=1}^N \Lambda^2 \left(\frac{u}{\sqrt{N}}, x_{k-1} \right) \end{aligned}$$

Pour prouver la normalité asymptotique de \widetilde{r}_N , il nous reste à prouver que Δ_N et Ψ_N satisfont les conditions du théorème 4.

1. Sous **(B)**, $\Delta_N \xrightarrow{\mathcal{L}(P_0)} \mathcal{N}\left(0, \frac{1}{I(0)}\right)$.

Si on note, $\tilde{D}_N = \sum_{k=1}^N Db(x_{k-1}) e_k$, il suffit de montrer que $\frac{1}{\sqrt{N}} \tilde{D}_N \xrightarrow{\mathcal{L}(P_0)} \mathcal{N}(0, \sigma^4 I(0))$. Comme pour la démonstration du théorème 5, on vérifie successivement les deux conditions d'application du théorème 2.8.41. de [7] :

(a) $\frac{1}{N} \langle \tilde{D} \rangle_N \xrightarrow{P_0} \sigma^4 I(0)$.

(b) $\forall \varepsilon > 0, \frac{1}{N} \sum_{k=1}^N E_0 \left[\left| \tilde{D}_k - \tilde{D}_{k-1} \right|^2 I_{(|\tilde{D}_k - \tilde{D}_{k-1}| \geq \varepsilon \sqrt{N})} \mid \mathcal{F}_{k-1} \right] \xrightarrow{L^1(P_0)} 0$.

Sous **(B)**, \tilde{D}_N est une martingale de carré intégrable et de crochet $\langle \tilde{D} \rangle_N = \sigma^2 \sum_{k=1}^N Db^2(x_{k-1})$. La première condition est vérifiée par application du lemme 12. Il nous reste donc à établir la condition de Lindberg.

$$\begin{aligned} & \frac{1}{N} \sum_{k=1}^N E_0 \left[E_0 \left[\left| \tilde{D}_k - \tilde{D}_{k-1} \right|^2 I_{(|\tilde{D}_k - \tilde{D}_{k-1}| \geq \varepsilon \sqrt{N})} \mid \mathcal{F}_{k-1} \right] \right] \\ &= \frac{1}{N} \sum_{k=1}^N E_0 \left[I_{(|\tilde{D}_k - \tilde{D}_{k-1}| \geq \varepsilon \sqrt{N})} \right]^{\frac{1}{2}} E_0 \left[(Db)^4(x_{k-1}) e_k^4 \right]^{\frac{1}{2}} \\ &\leq \sqrt{m_4} \|Db\|_{4, \mu_0}^2 \frac{1}{N} \sum_{k=1}^N P_0 \left(\left| \tilde{D}_k - \tilde{D}_{k-1} \right| \geq \varepsilon \sqrt{N} \right)^{\frac{1}{2}} \end{aligned}$$

où m_4 est le moment d'ordre 4 des $(e_k)_k$. Or, par application de l'inégalité de Markov, on a :

$$P_0 \left(\left| \tilde{D}_k - \tilde{D}_{k-1} \right| \geq \varepsilon \sqrt{N} \right) \leq \frac{\sigma^2 \|Db\|_{2, \mu_0}^2}{\varepsilon^2 N}$$

Pour tout $\varepsilon > 0$, il existe donc une constante $C_\varepsilon > 0$ telle :

$$\frac{1}{N} \sum_{k=1}^N E_0 \left[E_0 \left[\left| \tilde{D}_k - \tilde{D}_{k-1} \right|^2 I_{(|\tilde{D}_k - \tilde{D}_{k-1}| \geq \varepsilon \sqrt{N})} \mid \mathcal{F}_{k-1} \right] \right] \leq \frac{C_\varepsilon}{\sqrt{N}}$$

La condition de Lindeberg est ainsi satisfaite et on a la convergence annoncée. Pour conclure, il suffit de remarquer que $\Delta_N = \frac{1}{\sigma^2 I(0) \sqrt{N}} \tilde{D}_N$.

2. Sous **(B)**, pour tout $B > 0$, $\sup_{|u| \leq B} |\Psi_N(u)| \xrightarrow{P_0} 0$. Pour montrer ce résultat, il suffit d'adapter la démonstration du théorème 6.

Les deux conditions du théorème 4 sont donc satisfaites pour la décomposition donnée, on obtient ainsi la normalité asymptotique de \tilde{r}_N .

Chapter 4

Estimation d'un CAR(p) incomplètement observé à partir des équations de YULE-WALKER

4.1 Introduction :

Rappelons la définition formelle d'un CAR(p) ([3], [8], [9]). On considère l'équation différentielle stochastique p -dimensionnelle suivante :

$$dY(t) = AY(t) dt + \sigma b dW(t) \quad (4.1)$$

où $(W(t))_{t \geq 0}$ est un mouvement brownien standard unidimensionnel et $Y(0)$ est indépendant de W . A et b sont respectivement une matrice de dimensions $p \times p$ et un vecteur de R^p définis par :

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & \cdots & -\alpha_{p-1} \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (4.2)$$

Notons P le polynome caractéristique de A :

$$P(z) = z^p + \alpha_{p-1}z^{p-1} + \cdots + \alpha_0 \quad (4.3)$$

Si toutes les racines de P ont une partie réelle négative, (4.1) admet une unique solution stationnaire $(Y(t))_{t \geq 0}$ définie par ([1], p. 118-119) :

$$\begin{aligned} Y(t) &= \exp(At) Y(0) + \sigma \int_0^t \exp(A(t-s)) b dW(s) \\ Y(0) &= \mathcal{N}_p \left(0, \sigma^2 \int_0^{+\infty} \exp(As) b^t b \exp({}^t As) ds \right) \end{aligned} \quad (4.4)$$

Dans toute la suite, nous supposons vérifiée cette condition qui assure l'existence d'une solution stationnaire. Si $\theta = (\alpha, \sigma)$ et $\alpha = (\alpha_0, \dots, \alpha_{p-1})$, on note alors μ_θ la loi stationnaire, P_θ

la loi du processus sous μ_θ et E_θ l'espérance sous P_θ .

Nous dirons que $(X(t))_{t \geq 0}$ est un CAR(p) de paramètre θ , si pour tout $t \geq 0$, $X(t) = eY(t)$, $e = (1, 0, \dots, 0) \in \mathbb{R}^p$. Par (4.1) et (4.2), X est donc un processus dérivable à l'ordre $(p-1)$. Si on note $X^{(j)}(t) = \frac{d^j}{dt^j} X(t)$, on a :

$$Y(t) = \begin{pmatrix} X(t) \\ \vdots \\ X^{(p-1)}(t) \end{pmatrix}$$

De plus, (4.1) s'écrit :

$$dX^{(p-1)}(t) + [\alpha_0 X(t) + \dots + \alpha_{p-1} X^{(p-1)}(t)] dt = \sigma dW(t) \quad (4.5)$$

Comme dans le cas des modèles autorégressifs à temps discret, AR(p), les paramètres d'un CAR(p) peuvent être estimés à partir des équations de Yule-Walker qui lui sont associées. Dans le cas continu, ces équations sont basées sur les fonctions de covariance dérivée (DCVF) $(D_{j,k})$ définies par (cf. [9]) :

$$\begin{aligned} \forall (j, k) \in \{0, 1, \dots, p-1\}, \quad D_{j,k}(h) &= E_\theta [X^{(j)}(h) X^{(k)}(0)] \\ D_{j,p}(h) &\stackrel{L^2(P_\theta)}{=} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T X^{(j)}(t+h) dX^{(p-1)}(t) \end{aligned}$$

Les équations de Yule-Walker sont obtenues en multipliant l'équation (4.5) par $X^{(j)}(t+h)$ et en passant à l'espérance sous P_θ . On a ainsi, pour tout $j = 0, \dots, p-1$ et tout $h \in \mathbb{R}$:

$$D_{j,p}(h) + \alpha_{p-1} D_{j,p-1}(h) + \dots + \alpha_0 D_{j,0}(h) = 0 \quad (4.6)$$

Si l'on estime correctement les $(D_{j,k})$, on estime par la même α . Dans [9], Hyndman utilise des estimateurs des $(D_{j,k}(0))$ basés sur une observation continue et complète de Y sur $[0, T]$, c'est-à-dire que les p composantes de Y sont observées à temps continu sur $[0, T]$. Il obtient alors un estimateur de α consistant, asymptotiquement normal et efficace, lorsque $T \rightarrow +\infty$.

Dans notre contexte, on ne dispose que d'une observation discrétisée, à pas $\delta > 0$ fixé, de la première composante X de Y . Nous allons cependant adopter une démarche similaire à celle de Hyndman en définissant des estimateurs naturels des $(D_{j,k})$ et en déduisant, par l'intermédiaire des équations de Yule-Walker, un estimateur de α . Sans modifier l'esprit de la méthode d'estimation, on peut cependant choisir d'estimer les équations de Yule-Walker en $h = 0$ mais également en n'importe quel point de \mathbb{R} . Or, comme le remarque Hyndman dans la cas d'un CAR(2) et comme nous le montrons plus généralement pour un CAR(p) (cf. lemme 3, p.138), l'estimateur empirique associé à $D_{p-1,p}(0)$ présente un biais systématique (multiplicatif). Afin de résoudre ce problème, nous allons estimer les équations de Yule-Walker en $h = -(p-1)\delta$ et non en $h = 0$.

Dans le deuxième paragraphe, nous commençons ainsi par définir les estimateurs empiriques des dérivées de X dont nous donnons un développement en δ . Dans la troisième partie, nous présentons les estimateurs des $(D_{j,k})$ associés. Nous exposons ensuite la méthode d'estimation de α basée sur l'estimation des équations de Yule-Walker et nous étudions les

propriétés asymptotiques de l'estimateur obtenu : nous montrons en particulier que cet estimateur converge vers α^δ , où α^δ approche α à δ près. Dans le quatrième paragraphe, nous proposons, dans le cas d'un CAR(2), de réduire le biais d'estimation associé à α à l'aide d'une méthode basée sur le schéma d'approximation du trapèze. L'estimateur associé est alors convergent avec un biais asymptotique de l'ordre de δ^2 . Le cinquième paragraphe est consacré à l'estimation de σ^2 : là encore, l'estimateur obtenu est convergent avec un biais explicite en δ . Dans la dernière partie, nous mettons en oeuvre, expérimentalement par simulations, les différentes méthodes d'estimation pour un CAR(2).

4.2 Rappels et résultats préliminaires

Dans toute la suite, la notation abrégée $k = m, n$ se lit $k \in \{m, \dots, n\}$.

Comme dans le cas des AR(p), les équations de Yule-Walker peuvent être exprimées en fonction de la covariance r de X , $r(h) = E_\theta[X(h)X(0)]$, car les $(D_{j,k})$ et r sont liés fonctionnellement. En effet, r est dérivable jusqu'à l'ordre $(2p-2)$ sur R et $r^{(2p-2)}$ est dérivable partout sauf en 0 [4] et on a ([9], théorème 3.1) :

Lemme 1 *Pour tout $h \geq 0$ et tout $j, k = 0, p-1$, on a :*

$$D_{j,k}(h) = (-1)^k r^{(j+k)}(h) \quad (4.7)$$

$$D_{j,k}(-h) = D_{k,j}(h) = (-1)^j r^{(j+k)}(h) \quad (4.8)$$

Pour $h \geq 0$ et $j = 0, p-2$:

$$D_{j,p}(h) = -D_{j+1,p-1}(h) = (-1)^p r^{(j+p)}(h) \quad (4.9)$$

et, pour $h \neq 0$:

$$D_{p-1,p}(h) = (-1)^p r^{(2p-1)}(h) \quad (4.10)$$

$$D_{p-1,p}(0) = -\frac{\sigma^2}{2} \quad (4.11)$$

Les équations de Yule-Walker s'écrivent donc en fonction de r et de ses dérivées comme suit :

$$\forall h \geq 0, \forall j = 0, p-1, \quad r^{(j+p)}(h) + \alpha_{p-1} r^{(j+p-1)}(h) + \dots + \alpha_0 r^{(j)}(h) = 0 \quad (4.12)$$

$$(-1)^p r^{(j+p)}(h) + (-1)^{p-1} \alpha_{p-1} r^{(j+p-1)}(h) + \dots + \alpha_0 r^{(j)}(h) = 0 \quad (4.13)$$

où $r^{(2p-1)}(0)$ est la dérivée à droite de $r^{(2p-2)}$ en 0.

Dans la suite, nous utiliserons selon les cas l'une ou l'autre des trois expressions (4.6), (4.13) et (4.12).

Les $(D_{j,k})$ sont directement liés aux dérivées de X qui ne sont pas observées. Pour estimer les DCVF, nous sommes donc amenés à définir des approximations des dérivés de X .

Si $\Delta X(t) = X(t) - X(t-\delta)$ et $\Delta^{(i)}$ est le $i^{\text{ème}}$ itéré de cet opérateur, il est naturel d'approximer $X^{(j)}(t+j\delta)$ par $\delta^{-j} \Delta^{(j)} X(t+j\delta)$.

La suite de ce paragraphe est consacré à l'établissement d'un développement en δ des $(\Delta^{(j)} X)_{j=1,p}$. Ce résultat est crucial pour l'explicitation du biais asymptotique de l'estimateur dérivé des

équations de Yule-Walker.

Pour tout $p \geq 1$ et tout $k \in N$, on définit :

$$D_p^k = \sum_{l=0}^{p-1} C_p^l (-1)^l (p-l)^k$$

On a le lemme combinatoire suivant :

Lemme 2 *On a :*

$$\forall k \geq 1, D_{p+1}^k = \sum_{j=0}^{k-1} C_k^j D_p^j - D_p^0 \quad (4.14)$$

$$D_p^{p+1} = \frac{p}{2} (p+1)! \quad (4.15)$$

Pour tout $k = 1, p-1$, $D_p^k = 0$. De plus, $D_p^0 = (-1)^{p+1}$ et $D_p^p = p!$.

Ce lemme est démontré en annexe A.

Notons, pour tout $i = 0, p-1$ et tout $j \geq 1$:

$$R_i^p(j, t) = \frac{1}{(p-1)!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t+j\delta-l\delta-u)^{p-1} \right] X^{(i)}(u) du \quad (4.16)$$

$$B^p(j, t) = \frac{1}{(p-1)!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t+j\delta-l\delta-u)^{p-1} \right] \sigma dW(u) \quad (4.17)$$

Le théorème qui suit donne un développement en δ des quantités $(\Delta^{(i)} X(t+i\delta))_{i=1,p}$. La démonstration en est donnée en annexe B.

Théorème 1 *Avec les notations (4.16) et (4.17), on a pour $j = 1, p-1$:*

$$\Delta^{(j)} X(t+j\delta) = \sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} X^{(k)}(t) + R^p(j, t) + B^p(j, t) \quad (4.18)$$

$$\Delta^{(p)} X(t+p\delta) = R^p(p, t) + B^p(p, t) \quad (4.19)$$

avec : $R^p(j, t) = -\sum_{i=0}^{p-1} \alpha_i R_i^p(j, t)$.

De plus, notant $\beta_0 = \alpha_{p-1} \alpha_0$ et, pour $i = 1, p-1$, $\beta_i = \alpha_{p-1} \alpha_i - \alpha_{i-1}$, on a :

$$\begin{aligned} \Delta^{(p-1)} X(t+(p-1)\delta) &= \delta^{p-1} X^{(p-1)}(t) - \delta^p \frac{D_{p-1}^p}{p!} \sum_{i=0}^{p-1} \alpha_i X^{(i)}(t) \\ &+ \sum_{i=0}^{p-1} \beta_i R_i^{p+1}(p-1, t) + B^p(p-1, t) - \alpha_{p-1} B^{p+1}(p-1, t) \end{aligned} \quad (4.20)$$

Définissant $\gamma_0 = -\beta_{p-1}\alpha_0$ et, pour $i = 1, p-1$, $\gamma_i = \beta_{i-1} - \beta_{p-1}\alpha_i$, on obtient :

$$\begin{aligned} \Delta^{(p)} X(t + p\delta) &= -\delta^p \sum_{i=0}^{p-1} \alpha_i X^{(i)}(t) + \delta^{p+1} \frac{D_p^{p+1}}{(p+1)!} \sum_{i=0}^{p-1} \beta_i X^{(i)}(t) \\ &+ \sum_{i=0}^{p-1} \gamma_i R_i^{p+2}(p, t) \\ &+ B^p(p, t) - \alpha_{p-1} B^{p+1}(p, t) + [(\alpha_{p-1})^2 - \alpha_{p-2}] B^{p+2}(p, t) \end{aligned} \quad (4.21)$$

4.3 Estimation du paramètre à partir des équations de Yule-Walker

Dans toute la suite, $[s]$ représente la partie entière de s .

Le théorème 1 montre que, pour $j = 1, p-1$, $\delta^{-j} \Delta^{(j)} X(t + j\delta)$ approxime $\frac{d}{dt} X^{(j-1)}(t)$ à δ près. Si on dispose de n observations à pas δ et si on note $n_p = \left\lfloor \frac{n}{p} \right\rfloor$, alors, pour $0 \leq j \leq p-1$ et $0 \leq k \leq p$, un estimateur empirique naturel de $D_{j,k}(0)$ est :

$$\widehat{D}_{j,k} = \frac{\delta^{-(j+k)}}{n_p} \sum_{i=0}^{n_p-1} \Delta^{(j)} X(ip\delta + j\delta) \Delta^{(k)} X(ip\delta + k\delta), \quad n_p = \left\lfloor \frac{n}{p} \right\rfloor \quad (4.22)$$

On pourrait utiliser ces estimateurs afin d'estimer les équations de Yule-Walker (4.6) en $h = 0$, comme le fait Hyndman dans [9] (§7). Il apparait cependant que $\widehat{D}_{p-1,p}$ présente un biais systématique. Précisons cela.

Lemme 3 *On a :*

$$E_\theta \left[\widehat{D}_{p-1,p} \right] = C(p) D_{p-1,p}(0) + \delta C_\theta + o(\delta), \quad C_\theta = \frac{(-1)^p}{2} \sum_{k=0}^{p-1} \alpha_k r^{(p+k)}(0) + \alpha_{p-1} D(p) \sigma^2$$

avec :

$$\begin{aligned} C(p) &= \frac{1}{[(p-1)!]^2} \sum_{k=1}^{p-1} \int_{k-1}^k \left[\sum_{l=0}^{p-k} C_p^l (-1)^l (p-l-u)^{p-1} \right]^2 du \\ D(p) &= \frac{1}{(p!)^2} \sum_{k=1}^{p-1} \left[\left(\sum_{l=0}^{p-1-k} C_{p-1}^l (-1)^l (p-1-l-u)^p \right) \left(\sum_{i=0}^{p-k} C_p^i (-1)^i (p-i-u)^p \right) \right]_{u=k-1}^{u=k} \end{aligned}$$

La démonstration de ce résultat est donnée en annexe E.

Pour contourner ce problème, nous avons choisi d'estimer les équations de Yule-Walker en $h = -(p-1)\delta$. Nous considérons ainsi les estimateurs empiriques des $D_{j,k}(-(p-1)\delta)$. Ces estimateurs sont de la forme :

$$\tilde{D}_{j,k} = \frac{\delta^{-(j+k)}}{n_{2p-1}} \sum_{i=0}^{n_{2p-1}-1} \Delta^{(j)} X(i(2p-1)\delta + j\delta) \Delta^{(k)} X(i(2p-1)\delta + (p-1)\delta + k\delta), \quad n_{2p-1} = \left\lfloor \frac{n}{2p-1} \right\rfloor$$

Notons, pour tout $h \in R$:

$$\Gamma(h) = (D_{j,k}(h))_{j,k=0,p-1}, \quad \gamma(h) = {}^t(D_{0,p}(h), \dots, D_{p-1,p}(h))$$

D'après (4.6), α satisfait pour tout h :

$$\gamma(h) + \Gamma(h) \alpha = 0$$

Comme nous utilisons les $\tilde{D}_{j,k}$ pour estimer le système précédent en $h = -(p-1)\delta$, il est naturel de choisir comme estimateur de α , une solution (α^n) du système suivant :

$$\gamma_\delta^n + \Gamma_\delta^n \alpha^n = 0, \quad \alpha^n = {}^t(\alpha_0^n, \dots, \alpha_{p-1}^n) \quad (4.23)$$

où :

$$\Gamma_\delta^n = (\tilde{D}_{j,k})_{j,k=0,p-1} \quad \text{et} \quad \gamma_\delta^n = {}^t(\tilde{D}_{0,p}, \dots, \tilde{D}_{p-1,p})$$

Toutefois, les $\tilde{D}_{j,k}$ approchent les $D_{j,k}(-(p-1)\delta)$ seulement à δ près. Comme nous le verrons, cette erreur se transmet à l'estimation de α et provoque un biais asymptotique de l'ordre de δ . Si on note $\mathcal{D}_{j,k} = E_\theta [\tilde{D}_{j,k}]$, on a :

Lemme 4

$$\forall j = 0, p-1, \forall k = 0, p, \quad \mathcal{D}_{j,k} = D_{j,k}(-(p-1)\delta) + \delta \Psi_{j,k}(\delta) + o(\delta)$$

avec :

$$\Psi_{p-1,p-1}(\delta) = \frac{(p-1)}{2} \sigma^2 \quad (4.24)$$

$$\Psi_{p-1,p}(\delta) = \frac{(-1)^p}{2} \sum_{k=0}^{p-1} \alpha_k r^{(p+k)} ((p-1)\delta) - \alpha_{p-1} \frac{(p-1)}{2} \sigma^2 \quad (4.25)$$

et, dans les autres cas :

$$\Psi_{j,k}(\delta) = \frac{(-1)^j}{2} (k-j) r^{(k+j+1)} ((p-1)\delta) \quad (4.26)$$

La démonstration de ce lemme est donnée en annexe F.

De ce fait, α^n ne va pas estimer directement α mais α^δ où α^δ est solution du système :

$$\gamma_\delta + \Gamma_\delta \alpha^\delta = 0, \quad \alpha^\delta = {}^t(\alpha_0^\delta, \dots, \alpha_{p-1}^\delta) \quad (4.27)$$

avec :

$$\Gamma_\delta = (\mathcal{D}_{j,k})_{j,k=0,p-1} \quad \text{et} \quad \gamma_\delta = {}^t(\mathcal{D}_{0,p}, \dots, \mathcal{D}_{p-1,p})$$

Cependant, compte tenu des résultats du lemme 4, α^δ approche α à δ près.

Lemme 5 *Il existe $\delta_0 > 0$ tel que pour tout δ , $0 < \delta \leq \delta_0$, le système (4.27) admet une unique solution $\alpha^\delta = -[\Gamma_\delta]^{-1} \gamma_\delta$.*

De plus :

$$\alpha^\delta = \alpha - \delta [\Gamma(0)]^{-1} B_\theta + o(\delta), \quad B_\theta = {}^t (B_\theta^0, \dots, B_\theta^{p-1})$$

avec, pour tout $j = 0, p-1$:

$$B_\theta^j = \frac{(-1)^j}{2} \left[\sum_{k=0}^{p-1} \alpha_k (k-p) r^{(j+k+1)}(0) \right]$$

Démonstration du lemme 5 :

Le lemme 4 implique que : $\Gamma_\delta = \Gamma(-(p-1)\delta) + O(\delta)$. Or, d'après l'équation (3.9) de [9], $\Gamma(h) = \exp(Ah) \Gamma(0)$ et donc $\Gamma(h) = \Gamma(0) + O(h)$. On sait de plus que $\Gamma(0)$ est inversible (proposition 3.1 de [9]). Pour h proche de 0, $\Gamma(h)$ est donc également inversible. Il existe par conséquent $\delta_0 > 0$, tel que pour δ , $0 < \delta \leq \delta_0$, Γ_δ est inversible. Le système (4.27) admet donc une unique solution $\alpha^\delta = -[\Gamma_\delta]^{-1} \gamma_\delta$.

Il nous reste à expliciter l'écart entre α^δ et α . On a :

$$\alpha^\delta - \alpha = -[\Gamma_\delta]^{-1} [\gamma_\delta + \Gamma_\delta \alpha]$$

Notons $N_\delta = {}^t (\Psi_{0,p}(\delta), \dots, \Psi_{p-1,p}(\delta))$ et $M_\delta = (\Psi_{j,k}(\delta))_{j,k=0,p-1}$, où les $(\Psi_{j,k}(\delta))$ sont définis par (4.24), (4.25) et (4.26). D'après le lemme 4, on a : $\gamma_\delta = \gamma(-(p-1)\delta) + \delta N_\delta + o(\delta)$ et $\Gamma_\delta = \Gamma(-(p-1)\delta) + \delta M_\delta + o(\delta)$. Donc :

$$\alpha^\delta - \alpha = -[\Gamma_\delta]^{-1} [\gamma(-(p-1)\delta) + \Gamma(-(p-1)\delta) \alpha + \delta N_\delta + \delta M_\delta \alpha + o(\delta)]$$

Les équations de Yule-Walker (4.6) en $h = -(p-1)\delta$ impliquent que :

$$\alpha^\delta - \alpha = -\delta [\Gamma_\delta]^{-1} [N_\delta + M_\delta \alpha + o(\delta)]$$

avec :

$$N_\delta + M_\delta \alpha = \left(\Psi_{j,p}(\delta) + \sum_{k=0}^{p-1} \alpha_k \Psi_{j,k}(\delta) \right)_{j=0,p-1}$$

Or, pour tout $j = 0, p-2$, on a :

$$\begin{aligned} \Psi_{j,p}(\delta) + \sum_{k=0}^{p-1} \alpha_k \Psi_{j,k}(\delta) &= \frac{(-1)^j}{2} \left[(p-j) r^{(j+1+p)}((p-1)\delta) + \sum_{k=0}^{p-1} \alpha_k (k-j) r^{(j+1+k)}((p-1)\delta) \right] \\ &= \frac{(-1)^j}{2} \sum_{k=0}^{p-1} \alpha_k [(k-j) - (p-j)] r^{(j+1+k)}((p-1)\delta) \\ &= \frac{(-1)^j}{2} \sum_{k=0}^{p-1} \alpha_k (k-p) r^{(j+1+k)}((p-1)\delta) \end{aligned}$$

et :

$$\begin{aligned}
\Psi_{p-1,p}(\delta) + \sum_{k=0}^{p-1} \alpha_k \Psi_{p-1,k}(\delta) &= \frac{(-1)^p}{2} \sum_{k=0}^{p-1} \alpha_k r^{(p+k)} ((p-1)\delta) - \alpha_{p-1} \frac{(p-1)}{2} \sigma^2 \\
&+ \frac{(-1)^p}{2} \sum_{k=0}^{p-2} \alpha_k (p-1-k) r^{(p+k)} ((p-1)\delta) + \alpha_{p-1} \frac{(p-1)}{2} \sigma^2 \\
&= \frac{(-1)^p}{2} \sum_{k=0}^{p-1} \alpha_k (p-k) r^{(p+k)} ((p-1)\delta)
\end{aligned}$$

Il suffit alors de remarquer que : $\lim_{\delta \rightarrow 0} [N_\delta + M_\delta \alpha] = N_0 + M_0 \alpha$ et $\lim_{\delta \rightarrow 0} \Gamma_\delta = \Gamma(0)$.

L'étude du comportement asymptotique de α^n repose sur une adaptation du résultat d'ergodicité obtenu par Florens-Zmirou [6] dans le cas de modèles de diffusion unidimensionnels.

Si on note $\pi_\theta^t(x, dy) = P_\theta(Y(t) \in dy \mid Y(0) = x)$, avec $x = (x_0, \dots, x_{p-1})$ et $dy = (dy_0, \dots, dy_{p-1})$, la probabilité de transition associée à Y et $Q_\theta^t = \mu_\theta \otimes \pi_\theta^t$ la loi de $(Y(0), Y(t))$ sous la loi invariante, on montre alors que :

Lemme 6 *Si $f : R^{m+1} \rightarrow R$ est telle que $Q_\theta^{m\delta}(f^2) < \infty$, $m \in N^*$, alors, lorsque $n \rightarrow +\infty$:*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X(km\delta), X(km\delta + \delta), \dots, X((k+1)m\delta)) \xrightarrow{L^2(P_\theta)} Q_\theta^{m\delta}(f)$$

Démonstration du lemme 6 :

Notons $L_0^2(\mu_\theta) = \{h : R^p \rightarrow R, \mu_\theta(h) = 0 \text{ et } \mu_\theta({}^t h h) < \infty\}$. D'après le lemme 3 (chapitre 1.2) et la proposition 3 (chapitre 2.5) de [5], on sait qu'il existe $\lambda > 0$ tel que :

$$\forall h \in L_0^2(\mu_\theta), \|\pi_\theta^t(h)\|_{L^2(\mu_\theta)} \leq \exp(-\lambda t) \|h\|_{L^2(\mu_\theta)}$$

Adaptant la preuve du lemme 2 de [6], on montre que pour $m \in N^*$ et $h : (R^p)^{m+1} \rightarrow R$ telle que $Q_\theta^{m\delta}(h^2) < +\infty$:

$$\frac{1}{n} \sum_{k=0}^{n-1} h(Y(km\delta), Y(km\delta + \delta) \dots, Y((k+1)m\delta)) \xrightarrow{L^2(P_\theta)} Q_\theta^{m\delta}(h)$$

Ce résultat reste vrai en particulier lorsque $h = f \circ p_1$ où \circ représente le symbole de composition, f est une fonction de $R^{m+1} \rightarrow R$ et p_1 est la projection de $(R^p)^{m+1} \rightarrow R$ définie par : $\forall (y_0, \dots, y_m), y_i = {}^t(x_0^i, x_1^i, \dots, x_{p-1}^i), p_1(y_0, \dots, y_m) = (x_0^0, x_0^1, \dots, x_0^m)$. On obtient donc la convergence annoncée.

Utilisant ce résultat, on obtient :

Théorème 2 *Il existe $\delta_0 > 0$ tel que pour tout $\delta, 0 < \delta \leq \delta_0$, et pour n suffisamment grand, le système (4.23) admet une unique solution $\alpha^n = -[\Gamma_\delta^n]^{-1} \gamma_\delta^n$.*

Lorsque $n \rightarrow +\infty$, on a :

$$\alpha^n \xrightarrow{P_\theta} \alpha^\delta$$

L'estimateur dérivé des équations de Yule-Walker estimées en $h = -(p-1)\delta$ est donc asymptotiquement biaisé avec un biais explicite de l'ordre de δ .

Dans le paragraphe suivant, nous montrons que pour un CAR(2), ce biais peut être réduit à un ordre en δ^2 en utilisant une méthode d'estimation basée sur le schéma d'approximation du trapèze bidimensionnel du processus.

Démonstration du théorème 2 :

D'après le lemme 6, Γ_δ^n converge vers Γ_δ lorsque $n \rightarrow +\infty$. Or, pour δ proche de 0, Γ_δ est inversible (cf. démonstration du lemme 5. Pour n suffisamment grand, Γ_δ^n est donc inversible et α^n existe et a la forme annoncée.

La convergence de α^n vers α^δ résulte directement du lemme 6. Lorsque $n \rightarrow +\infty$, on a en effet :

$$\Gamma_\delta^n \xrightarrow{L^2(P_\theta)} \Gamma_\delta, \quad \text{et} \quad \gamma_\delta^n \xrightarrow{L^2(P_\theta)} \gamma_\delta$$

4.4 Estimation des paramètres d'un CAR(2) à partir du schéma du trapèze.

Dans ce paragraphe, nous considérons un CAR(2) :

$$dX(t) = X^{(1)}(t) dt \tag{4.28}$$

$$dX^{(1)}(t) = - \left[\alpha_0 X(t) + \alpha_1 X^{(1)}(t) \right] dt + \sigma dW(t)$$

La méthode d'estimation de α présentée ici est basée sur le schéma du trapèze bidimensionnel associé au processus. Plus précisément, nous déduisons de celui-ci un schéma d'approximation pour X . Nous associons alors ce schéma, qui est anticipatif, à une méthode d'estimation par variable instrumentale (cf. [2],[7], [11] et [12]). Le biais de l'estimateur de α ainsi obtenu est non plus d'ordre δ mais δ^2 .

Commençons par définir le schéma du trapèze associé à un CAR(2).

Lemme 7 (Schéma du trapèze pour un CAR(2))

Pour tout $t \in R^+$ et $\delta > 0$, on a :

$$X(t+\delta) - X(t) = \frac{\delta}{2} \left[X^{(1)}(t) + X^{(1)}(t+\delta) \right] + \zeta(0,t) + \xi(0,t) \tag{4.29}$$

$$\begin{aligned} X^{(1)}(t+\delta) - X^{(1)}(t) &= -\frac{\delta}{2} \left[\alpha_0 (X(t) + X(t+\delta)) + \alpha_1 (X^{(1)}(t) + X^{(1)}(t+\delta)) \right] \\ &+ \zeta(1,t) + \xi(1,t) \end{aligned} \tag{4.30}$$

avec :

$$\begin{aligned}
\zeta(0, t) &= \frac{1}{2} \int_t^{t+\delta} (t-u)(t+\delta-u) \left[\alpha_0 \alpha_1 X(u) + \left((\alpha_1)^2 - \alpha_0 \right) X^{(1)}(u) \right] du \\
\xi(0, t) &= \int_t^{t+\delta} \left[\left(t + \frac{\delta}{2} - u \right) - \frac{1}{2} (t-u)(t+\delta-u) \alpha_1 \right] \sigma dW(u) \\
\zeta(1, t) &= \frac{1}{2} \int_t^{t+\delta} (t-u)(t+\delta-u) \left[\left((\alpha_0)^2 - \alpha_0 (\alpha_1)^2 \right) X(u) + \left(2\alpha_0 \alpha_1 - (\alpha_1)^3 \right) X^{(1)}(u) \right] du \\
\xi(1, t) &= \int_t^{t+\delta} \left[1 - \left(t + \frac{\delta}{2} - u \right) \alpha_1 + \frac{1}{2} (t-u)(t+\delta-u) \left((\alpha_1)^2 - \alpha_0 \right) \right] \sigma dW(u)
\end{aligned}$$

La démonstration de ce lemme est donnée en annexe G.

Puisque la dérivée de X n'est pas observée, le schéma précédent ne peut être utilisé en l'état pour estimer α . Nous allons ainsi éliminer dans (4.29) tous les termes associés à $X^{(1)}$. Nous obtenons alors un schéma d'approximation anticipatif de pas 2δ pour X , qui dépend cette fois exclusivement de $X(t+2\delta)$, $X(t+\delta)$ et de $X(t)$. Avant d'énoncer ce résultat, on définit les coefficients numériques suivants :

$$\begin{aligned}
a_2^\delta &= \alpha_0 \alpha_1 + \delta \frac{(\alpha_0)^2}{2}, & b_2^\delta &= (\alpha_1)^2 - \alpha_0 + \delta \frac{\alpha_1 \alpha_0}{2} \\
c_2^\delta &= -\alpha_0 \alpha_1 + \delta \frac{(\alpha_0)^2}{2}, & d_2^\delta &= -(\alpha_1)^2 - \alpha_0 + \delta \frac{\alpha_1 \alpha_0}{2}
\end{aligned}$$

et, pour $k = 3, 5$, (a_k^δ, b_k^δ) et (c_k^δ, d_k^δ) sont donnés par :

$$\begin{pmatrix} a_k^\delta \\ b_k^\delta \end{pmatrix} = {}^t A^{k-2} \begin{pmatrix} a_2^\delta \\ b_2^\delta \end{pmatrix}, \quad \begin{pmatrix} c_k^\delta \\ d_k^\delta \end{pmatrix} = {}^t A^{k-2} \begin{pmatrix} c_2^\delta \\ d_2^\delta \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -\alpha_0 & -\alpha_1 \end{pmatrix}$$

où A est la matrice définie en (4.2). On définit également :

$$b_0^\delta = 1, \quad b_1^\delta = -\alpha_1 - \delta \frac{\alpha_0}{2}, \quad d_0^\delta = -1, \quad d_1^\delta = \alpha_1 - \delta \frac{\alpha_0}{2}$$

Théorème 3 Pour tout $\delta > 0$ et $t \in R^+$, on a :

$$\begin{aligned}
\Delta^{(2)} X(t+2\delta) &= -\frac{\delta \alpha_1}{2} [X(t+2\delta) - X(t)] - \frac{\delta^2 \alpha_0}{4} [X(t+2\delta) + 2X(t+\delta) + X(t)] \\
&+ \zeta(t) + \xi(t)
\end{aligned} \quad (4.31)$$

avec :

$$\begin{aligned}
\zeta(t) &= \frac{\delta^4}{12} \left[(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2 \right] X(t+\delta) + \frac{\delta^4}{12} \left[(\alpha_1)^3 - 3\alpha_1 \alpha_0 \right] X^{(1)}(t+\delta) \\
&+ \frac{\delta^6}{80} \left[(\alpha_0)^3 - (\alpha_1 \alpha_0)^2 \right] X(t+\delta) + \frac{\delta^6}{80} \left[2(\alpha_0)^2 \alpha_1 - (\alpha_1)^2 \alpha_0 \right] X^{(1)}(t+\delta) \\
&+ \int_t^{t+\delta} \left[\frac{(t-u)^5}{5!} + \frac{\delta(t-u)^4}{24!} \right] \left[c_5^\delta X(u) + d_5^\delta X^{(1)}(u) \right] du \\
&+ \int_{t+\delta}^{t+2\delta} \left[\frac{(t+2\delta-u)^5}{5!} - \frac{\delta(t+2\delta-u)^4}{24!} \right] \left[a_5^\delta X(u) + b_5^\delta X^{(1)}(u) \right] du
\end{aligned}$$

et :

$$\begin{aligned}\xi(t) &= \int_t^{t+\delta} \left(\frac{\delta}{2} + \sum_{k=1}^5 d_{k-1}^\delta \left[\frac{(t-u)^k}{k!} + \frac{\delta}{2} \frac{(t-u)^{k-1}}{(k-1)!} \right] \right) \sigma dW(u) \\ &+ \int_{t+\delta}^{t+2\delta} \left(\frac{\delta}{2} + \sum_{k=1}^5 b_{k-1}^\delta \left[\frac{(t+2\delta-u)^k}{k!} - \frac{\delta}{2} \frac{(t+2\delta-u)^{k-1}}{(k-1)!} \right] \right) \sigma dW(u)\end{aligned}$$

La démonstration de ce résultat est donnée en annexe H.

Commentaire :

La variable d'erreur $\zeta(t)$ est donc un $O_{P_\theta}(\delta^4)$ et le bruit $\xi(t)$ est indépendant de la tribu engendrée par $(X(s))_{s \in [0,t]}$ et tel que $E_\theta[\xi^2(t)]$ est d'ordre δ^2 .

Du fait du caractère anticipatif de ce schéma, l'utilisation des moindres carrés n'est pas adaptée pour estimer α . Précisons cela.

Soit $n_2 = \lfloor \frac{n}{2} \rfloor$. Notons Z_n le vecteur de R^{n_2} et X_n la matrice de dimensions $n_2 \times 2$ définis respectivement par :

$$Z_n = {}^t \left(\frac{1}{\delta^2} \Delta^{(2)} X(2\delta), \frac{1}{\delta^2} \Delta^{(2)} X(4\delta), \dots, \frac{1}{\delta^2} \Delta^{(2)} X(2n_2\delta) \right)$$

et

$$X_n = \begin{pmatrix} \frac{1}{4} [X(2\delta) + 2X(\delta) + X(0)] & \frac{1}{2\delta} [X(2\delta) - X(0)] \\ \frac{1}{4} [X(4\delta) + 2X(3\delta) + X(2\delta)] & \frac{1}{2\delta} [X(4\delta) - X(2\delta)] \\ \vdots & \vdots \\ \frac{1}{4} [X(2n_2\delta) + 2X(2n_2\delta - \delta) + X(2(n_2 - 1)\delta)] & \frac{1}{2\delta} [X(2n_2\delta) - X(2(n_2 - 1)\delta)] \end{pmatrix}$$

L'estimateur des moindres carrés dérivé du schéma précédent est donné par :

$$\check{\alpha}^n = - ({}^t X_n X_n)^{-1} ({}^t X_n Z_n)$$

Cet estimateur présente un biais systématique lié au fait que X_n et le bruit

$\xi_n = {}^t (\xi(0), \xi(2\delta), \dots, \xi(n_2\delta))$ sont corrélés. Pour éliminer ce problème, nous utilisons une méthode d'estimation par variable instrumentale (cf. [2],[7], [11] et [12]). Selon le principe de l'estimation par variable instrumentale, il nous faut remplacer X_n par une variable qui lui est liée tout en étant décorrélée de ξ_n . Si on note $n_4 = \lfloor \frac{n}{4} \rfloor$, un bon choix est la matrice X_n^I de dimensions $n_4 \times 2$ définie par :

$$X_n^I = \begin{pmatrix} X(0) & \frac{1}{2\delta} [X(2\delta) - X(0)] \\ X(4\delta) & \frac{1}{2\delta} [X(6\delta) - X(4\delta)] \\ \vdots & \vdots \\ X(4(n_4 - 1)\delta) & \frac{1}{2\delta} [X(4(n_4 - 1)\delta + 2\delta) - X(4(n_4 - 1)\delta)] \end{pmatrix}$$

L'estimateur par variable instrumentale associé au schéma (4.31) est alors :

$$\hat{\alpha}^n = - ({}^t X_n^I \hat{X}_n)^{-1} ({}^t X_n^I \hat{Z}_n)$$

avec : $\widehat{Z}_n = {}^t \left(\frac{1}{\delta^2} \Delta^{(2)} X(4\delta), \frac{1}{\delta^2} \Delta^{(2)} X(8\delta), \dots, \frac{1}{\delta^2} \Delta^{(2)} X(4n_4\delta) \right) \in R^{n_4}$ et

$$\widehat{X}_n = \begin{pmatrix} \frac{1}{4} [X(4\delta) + 2X(3\delta) + X(2\delta)] & \frac{1}{2\delta} [X(4\delta) - X(2\delta)] \\ \frac{1}{4} [X(8\delta) + 2X(7\delta) + X(6\delta)] & \frac{1}{2\delta} [X(8\delta) - X(6\delta)] \\ \vdots & \vdots \\ \frac{1}{4} [X(4n_4\delta) + 2X(4n_4\delta - \delta) + X(4n_4\delta - 2\delta)] & \frac{1}{2\delta} [X(4n_4\delta) - X(4n_4\delta - 2\delta)] \end{pmatrix}$$

Dans ce cas, la variable d'erreur ζ provoque un biais dans la procédure d'estimation. Cependant, ce biais n'est plus systématique car il est de l'ordre de δ^2 .

Si on note $\widehat{\gamma}_\delta = E_\theta \left[\frac{1}{n_4} {}^t X_n^I \widehat{Z}_n \right]$ et $\widehat{\Gamma}_\delta = E_\theta \left[\frac{1}{n_4} {}^t X_n^I \widehat{X}_n \right]$, on montre en effet que :

Théorème 4 *Il existe $\delta_0 > 0$ tel que pour tout $\delta, 0 < \delta \leq \delta_0$, et pour n suffisamment grand, $\widehat{\alpha}^n = - \left({}^t X_n^I \widehat{X}_n \right)^{-1} \left({}^t X_n^I \widehat{Z}_n \right)$ et $\widehat{\alpha}^\delta = - \left(\widehat{\Gamma}_\delta \right)^{-1} \widehat{\gamma}_\delta$ existent. Lorsque $n \rightarrow +\infty$, on a :*

$$\widehat{\alpha}^n \xrightarrow{P_\theta} \widehat{\alpha}^\delta$$

De plus :

$$\widehat{\alpha}^\delta - \alpha = -\delta^2 [\Gamma(0)]^{-1} \widehat{B}_\theta + o(\delta^2), \quad \widehat{B}_\theta = \begin{pmatrix} \frac{1}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] r(0) \\ -\frac{1}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] r^{(2)}(0) \end{pmatrix}$$

Ce théorème est démontré en annexe I.

Pour un CAR(2), cette procédure d'estimation permet donc de réduire le biais d'estimation de α .

4.5 Estimation du coefficient de diffusion σ

La méthode d'estimation de σ repose sur le résultat de l'équation (4.11) :

$$D_{p-1,p}(0) = -\frac{1}{2}\sigma^2$$

. On peut donc espérer estimer σ^2 à partir de $\widehat{D}_{p-1,p}$ qui est l'estimateur empirique de $D_{p-1,p}(0)$. D'après (4.22) :

$$\widehat{D}_{p-1,p} = \frac{\delta^{-(2p-1)}}{n_p} \sum_{i=0}^{n_p-1} \Delta^{(p-1)} X(ip\delta + (p-1)\delta) \Delta^{(p)} X((i+1)p\delta), \quad n_p = \left\lceil \frac{n}{p} \right\rceil$$

Comme nous l'avons montré dans le lemme 3, cet estimateur présente un biais systématique multiplicatif $C(p)$. Il apparait cependant que $C(p)$ est indépendante de (α, σ) et peut être calculée. On a par exemple, $C(2) = \frac{2}{3}$, $C(3) = \frac{11}{20} = 0.55$, $C(4) = \frac{151}{315} \simeq 0.48$ et $C(5) = \frac{15619}{36288} \simeq 0.43$. Il suffit ainsi de renormaliser $\widehat{D}_{p-1,p}$ par cette constante pour obtenir une estimation de $D_{p-1,p}(0)$ (et donc de σ^2) à δ près.

On définit ainsi :

$$\sigma_n^2 = -\frac{2}{C(p)} \widehat{D}_{p-1,p} \quad \text{et} \quad \sigma_\delta^2 = E_\theta [\sigma_n^2]$$

Le lemme 3 et le lemme 6 impliquent alors que :

Théorème 5 On a pour C_θ et $C(p)$ définis dans le lemme 3 :

$$\sigma_n^2 \xrightarrow{L^2(P_\theta)} \sigma_\delta^2 \quad \text{et} \quad \sigma_\delta^2 = \sigma^2 - \frac{2C_\theta}{C(p)}\delta + o(\delta)$$

4.6 Etude Expérimentale

Dans ce paragraphe, nous testons pour un CAR(2) les procédures d'estimation exposées précédemment à l'aide de simulations numériques. Dans la suite, nous indiquons par t (respectivement par y) les quantités qui se rapportent à la méthode d'estimation de α basée sur le schéma du trapèze (resp. sur les équations de Yule-Walker) et par s , celles associées à l'estimation de σ^2 . Le processus est simulé à l'aide d'un schéma d'Euler de pas 0.0001 sur un intervalle de longueur $T = n\delta$. Pour différentes décompositions (n, δ) de T , nous calculons N ($N = 100$) estimations de α (par les deux méthodes) et de σ^2 . Nous notons alors $\tilde{\alpha}_N^t = \frac{1}{N} \sum_{i=1}^N \alpha_n^{t,i}$, $\tilde{\alpha}_N^y = \frac{1}{N} \sum_{i=1}^N \alpha_n^{y,i}$ et $\tilde{\sigma}_N^2 = \frac{1}{N} \sum_{i=1}^N \sigma_n^2$ les moyennes empiriques obtenues. Celles-ci sont comparées à leur limite théorique respectives : α_δ^t , α_δ^y et σ_δ^2 . Nous comparons également $\alpha_\delta^t - \alpha$, $\alpha_\delta^y - \alpha$ et $\sigma_\delta^2 - \sigma^2$ à la partie principale de leur développement en δ (cf. lemme 5, théorèmes 4 et 5) : $B_\delta^t(\alpha, \sigma^2)$, $B_\delta^y(\alpha, \sigma^2)$ et $B_\delta^s(\alpha, \sigma^2)$.

Pour l'estimateur basé sur le schéma du trapèze, la partie principale du biais théorique $B_\delta^t(\alpha, \sigma^2)$ est :

$$B_\delta^t(\alpha, \sigma^2) = -\delta^2 \begin{pmatrix} \frac{1}{r^{(0)}} & 0 \\ 0 & -\frac{1}{r^{(2)}(0)} \end{pmatrix} \begin{pmatrix} \frac{1}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] r^{(0)}(0) \\ -\frac{1}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] r^{(2)}(0) \end{pmatrix}$$

Pour l'estimateur basé sur les équations de Yule-Walker, on a :

$$B_\delta^y(\alpha, \sigma^2) = -\delta \begin{pmatrix} \frac{1}{r^{(0)}} & 0 \\ 0 & -\frac{1}{r^{(2)}(0)} \end{pmatrix} \begin{pmatrix} -\frac{1}{2} \alpha_1 r^{(2)}(0) \\ \frac{1}{2} [2\alpha_0 r^{(2)}(0) + \alpha_1 r^{(3)}(0)] \end{pmatrix} \quad (4.32)$$

Enfin, pour le paramètre de diffusion σ , la partie principale du biais est égale à :

$$B_\delta^s(\alpha, \sigma^2) = -\delta \frac{1}{C(2)} [\alpha_0 r^{(2)}(0) + \alpha_1 r^{(3)}(0)];$$

Dans un second temps, nous proposons de réduire le biais d'estimation de α associé aux équations de Yule-Walker en débiaisant partiellement $\tilde{\alpha}_N^y$.

D'après le lemme 1, (4.32) s'écrit aussi :

$$B_\delta^y(\alpha, \sigma^2) = -\delta \begin{pmatrix} \frac{D_{1,1}(0)}{2D_{0,0}(0)} \alpha_1 \\ \frac{1}{4D_{1,1}(0)} \alpha_1 \sigma^2 - \alpha_0 \end{pmatrix}$$

Cette quantité peut être estimée par :

$$\tilde{B}_N^y(\tilde{\alpha}_N^y, \tilde{\sigma}_N^2) = -\delta \begin{pmatrix} \frac{\hat{D}_{1,1}}{2\hat{D}_{0,0}} \tilde{\alpha}_{1,N}^y \\ \frac{1}{4\hat{D}_{1,1}} \tilde{\alpha}_{1,N}^y \tilde{\sigma}_N^2 - \tilde{\alpha}_{0,N}^y \end{pmatrix}$$

où $\hat{D}_{1,1}$ et $\hat{D}_{0,0}$ sont les estimateurs naturels de $D_{1,1}(0)$ et $D_{0,0}(0)$ donnés en (4.22). L'estimateur proposé est alors :

$$\hat{\alpha}_N^y = \tilde{\alpha}_N^y - \tilde{B}_N^y(\tilde{\alpha}_N^y, \tilde{\sigma}_N^2)$$

Sa limite théorique est donnée par :

$$\widehat{\alpha}_\delta^y = \alpha_\delta^y - B_\delta^y(\alpha_\delta^y, \sigma_\delta^2)$$

On choisit : $\alpha_0 = 2$, $\alpha_1 = 3$ et $\sigma = 1$. Avec ces valeurs, la covariance vaut (cf. Doob [4]) :

$$r(h) = \frac{1}{6} \exp(-|h|) - \frac{1}{12} \exp(-2|h|)$$

Nous prenons comme condition initiale l'espérance de la loi invariante à savoir : $X(0) = 0$ et $X^{(1)}(0) = 0$.

Avec les valeurs de α_0 et α_1 choisies, on a :

$$B_\delta^t(\alpha, \sigma^2) = -\delta^2 \begin{pmatrix} \frac{5}{6} \\ \frac{3}{4} \end{pmatrix}$$

$$B_\delta^y(\alpha, \sigma^2) = -\delta \begin{pmatrix} 3 \\ \frac{5}{2} \end{pmatrix}$$

et : $B_\delta^s(\alpha, \sigma^2) = -\frac{7}{4}\delta$.

Ces différentes quantités sont comparées respectivement à $\alpha_\delta^t - \alpha$, $\alpha_\delta^y - \alpha$ et $\sigma_\delta^2 - \sigma^2$ dans le tableau 4.3.

Le tableau 4.1 regroupe les résultats obtenus pour l'estimation de α . Les résultats de l'estimation de σ^2 sont donnés dans le tableau 4.2.

Commentaires :

Comme nous l'avons vu précédemment, les biais d'estimation pour α et σ^2 s'expriment en fonction de la covariance r et de ses dérivées jusqu'à l'ordre 3. Les développements de ces biais d'estimation en fonction de δ , $B_\delta^t(\alpha, \sigma^2)$, $B_\delta^y(\alpha, \sigma^2)$ et $B_\delta^s(\alpha, \sigma^2)$, sont obtenus en supposant $r(\delta)$, $r^{(1)}(\delta)$, \dots , $r^{(3)}(\delta)$ proches de leur valeur en 0. δ doit donc être choisi de manière à ce que cette hypothèse soit vérifiée. Pour $\delta = 1$ et $\delta = 0.5$, on constate que $r(\delta)$ est éloigné de $r(0)$ (cf. tableau 4.4). Ceci explique en particulier que l'estimation obtenue pour σ^2 et α (quelque soit la méthode d'estimation utilisée) ne soit pas significative. Pour α , dans le cas où $n = 60$ et $\delta = 1$, les résultats sont de plus détériorés par la faible valeur de T ($T = 60$).

Lorsque δ est suffisamment petit selon le critère exposé précédemment, les estimations de α par le schéma du trapèze et par le débiaisage de $\tilde{\alpha}_N^y$ améliorent les résultats obtenus par les équations de Yule-Walker, notamment lorsque $T = 600$. Il apparait de plus que ces deux méthodes (trapèze et Yule-Walker débiaisé) fournissent des estimations relativement comparables et on peut penser que le biais asymptotique de $\tilde{\alpha}_N^y$ est de l'ordre de δ^2 (cf. tableau 4.3).

	δ	n	$\tilde{\alpha}_n^t$	α_δ^t	$\tilde{\alpha}_n^y$	α_δ^y	$\hat{\alpha}_n^y$	$\hat{\alpha}_\delta^y$	α
$T = 60$	1	60	-0.122	1.407	2.154	0.546	3.799	2.043	2
			0.042	2.447	4.000	1.496	7.146	3.195	3
	0.5	120	1.133	1.810	1.044	0.994	1.686	2.020	2
			3.044	2.828	2.100	2.051	3.196	3.092	3
	0.3	200	2.194	1.927	1.336	1.299	1.845	2.009	2
			3.428	2.934	2.289	2.367	3.075	3.043	3
	0.2	300	2.158	1.967	1.550	1.494	1.988	2.004	2
			3.211	2.970	2.557	2.554	3.193	3.022	3
	0.1	600	2.160	1.991	1.912	1.725	2.179	2.001	2
			3.022	2.992	2.855	2.764	3.249	3.006	3
	0.01	6000	2.138	1.999	1.925	1.970	1.954	2.000	2
			3.019	2.999	3.031	2.975	3.076	3.000	3
$T = 600$	1	600	2.200	1.407	0.511	0.546	1.088	2.043	2
			3.686	2.447	1.436	1.496	2.508	3.195	3
	0.5	1200	1.939	1.810	1.009	0.994	1.652	2.020	2
			3.044	2.828	2.072	2.051	3.132	3.092	3
	0.3	2000	1.871	1.927	1.293	1.299	1.822	2.009	2
			2.964	2.934	2.376	2.367	3.210	3.043	3
	0.2	3000	1.979	1.967	1.521	1.494	1.947	2.004	2
			3.006	2.970	2.569	2.554	3.226	3.022	3
	0.1	6000	1.996	1.991	1.749	1.725	2.000	2.001	2
			3.007	2.992	2.779	2.764	3.154	3.006	3

Table 4.1: CAR(2) : comparaison des moyennes empiriques pour 100 estimateurs de α lorsque ${}^t\alpha = (2, 3)$.

	δ	n	σ_n^2	σ_δ^2	σ^2
$T = 60$	1	60	0.215	0.212	1
	0.5	120	0.430	0.439	1
	0.3	200	0.597	0.603	1
	0.2	300	0.705	0.710	1
	0.1	600	0.827	0.841	1
	0.01	6000	0.981	0.982	1
$T = 600$	1	600	0.215	0.212	1
	0.5	1200	0.436	0.439	1
	0.3	2000	0.604	0.603	1
	0.2	3000	0.714	0.710	1
	0.1	6000	0.833	0.841	1

Table 4.2: CAR(2) : moyennes empiriques pour 100 estimateurs de σ^2 lorsque $\sigma^2 = 1$.

	$\alpha_\delta^t - \alpha$	$B_\delta^t(\alpha, \sigma^2)$	$\alpha_\delta^y - \alpha$	$B_\delta^y(\alpha, \sigma^2)$	$\hat{\alpha}_\delta^y - \alpha$	$\sigma_\delta^2 - \sigma^2$	$B_\delta^s(\alpha, \sigma)$
$\delta = 0.01$	-0.000083	-0.000083	-0.029735	-0.03	0.000016	- 0.0173	-0.0175
	-0.000075	-0.000075	-0.024851	-0.025	0.000074		
$\delta = 0.1$	-0.008300	-0.0083	-0.274995	-0.3	0.00143	- 0.158	-0.175
	-0.007473	-0.0075	-0.235682	-0.25	0.0064		
$\delta = 0.2$	-0.0328	-0.0333	-0.505	-0.6	0.0497	-0.289	-0.35
	-0.0295	-0.03	-0.445	-0.5	0.0223		
$\delta = 0.3$	-0.0723	-0.075	-0.700	-0.9	0.009	-0.396	-0.525
	-0.0653	-0.0675	-0.632	-0.75	0.043		
$\delta = 0.5$	-0.189	-0.208	-1.005	-1.5	0.0204	-0.56	-0.875
	-0.171	-0.1875	-0.948	-1.25	0.0921		
$\delta = 1.0$	-0.592	-0.8333	-1.453	-3	0.0433	-0.787	-1.75
	-0.552	-0.75	-1.503	-2.5	0.195		

Table 4.3: CAR(2) : comparaison des biais réels et des biais théoriques pour α et σ^2 .

δ	$12 r(\delta)$	$12 r(0)$
1	0.600	
0.5	0.845	
0.3	0.932	
0.2	0.967	1
0.1	0.990	
0.01	0.999	

Table 4.4: Quelques valeurs numériques de la fonction de covariance r .

Bibliography

- [1] Arató, M (1982) *Linear Stochastic Systems with Constant Coefficients : a Statistical Approach.*- Lecture Notes in control and information sciences 45. Springer, Berlin, 118-119.
- [2] Bergstrom, A.R. (1976) *Statistical inference in Continuous Time Series.*- In *Statistical inference in Continuous Time Economic Models.*, Ed. Bergstrom, North Holland.
- [3] Brockwell, P. J. & Hyndman, R. J. (1992) *On continuous-time threshold autoregression.*- International Journal of Forecasting 8, North-Holland, 157-173.
- [4] Doob, J. L. (1953) *Stochastic Process.*- Wiley, N.Y.
- [5] Doukhan, P. (1995) *Mixing : Properties and Examples.*- Lecture Notes in Statistics 85, Springer-Verlag.
- [6] Florens-Zmirou, D. (1989) *Approximate discrete schemes for statistics of diffusion processes.*- Statistics 21, 547-557.
- [7] Gouriéroux, C. & Monfort, A. *Statistique et Modèles Econométriques. Tome 1-* Economica.
- [8] Hyndman, R. J. (1992) *Continuous time threshold autoregressive models.*- PhD thesis, University of Melbourne, Department of statistics.
- [9] Hyndman, R. J. (1993) *Yule-Walker estimates for continuous-time autoregressive models.* - Journal of Time Series Analysis, Vol. 14, No 3, 281-296.
- [10] Jones, R. H. (1981) *Fitting a continuous time autoregression to discrete data.*- Applied Time Series Analysis, Vol. II (ed. D. F. Findlay). Academic Press, N.Y., p. 651-682.
- [11] Sargan, J.D. (1976) *Some discrete approximations to continuous times stochastic models.*- In *Statistical inference in Continuous Time Economic Models.*, Ed. Bergstrom, North Holland, 27-80.
- [12] Souchet, S. (1998) *Schéma de discrétisation anticipatif et estimation du paramètre de dérive d'une diffusion.*-A paraître comme note au CRAS.

Annexes

A Démonstration du lemme 2

Par définition, on a : $D_p^0 = \sum_{l=0}^{p-1} C_p^l (-1)^l$. Par la formule du binôme de Newton, on obtient : $D_p^0 = (1-1)^p - (-1)^p = -(-1)^p$.

- Montrons (4.14). Utilisant à nouveau la formule du binôme de Newton, on a pour tout $k \geq 1$:

$$\begin{aligned}
 \sum_{j=0}^{k-1} C_k^j D_p^j &= \sum_{l=0}^{p-1} C_p^l (-1)^l \left[\sum_{j=0}^{k-1} C_k^j (p-l)^j \right] \\
 &= \sum_{l=0}^{p-1} C_p^l (-1)^l (p+1-l)^k - \sum_{l=0}^{p-1} C_p^l (-1)^l (p-l)^k \\
 &= \sum_{l=0}^{p-1} C_p^l (-1)^l (p+1-l)^k + \sum_{l=1}^p C_p^{l-1} (-1)^l (p+1-l)^k \\
 &= (p+1)^k + \sum_{l=1}^{p-1} (-1)^l (C_p^l + C_p^{l-1}) (p+1-l)^k + p(-1)^p \\
 &= (p+1)^k + \sum_{l=1}^{p-1} (-1)^l C_{p+1}^l (p+1-l)^k + p(-1)^p \\
 &= \sum_{l=0}^{p-1} (-1)^l C_{p+1}^l (p+1-l)^k + (p+1)(-1)^p + (-1)^{p+1} \\
 &= \sum_{l=0}^p (-1)^l C_{p+1}^l (p+1-l)^k + (-1)^{p+1} \\
 &= D_{p+1}^k + D_p^0
 \end{aligned}$$

On a donc le résultat annoncé.

- Nous allons montrer par récurrence que, pour tout $k = 1, p-1$, $D_p^k = 0$ et $D_p^p = p!$. Pour $p = 1$, on a $D_1^1 = 1$. Pour $p = 2$, $D_2^1 = 2-2 = 0$ et $D_2^2 = 4-2 = 2$. La propriété de récurrence est donc vérifiée jusqu'au rang $p = 2$. Supposons qu'elle soit vraie jusqu'au rang p . Utilisant (4.14), on obtient pour tout $k = 1, p$:

$$D_{p+1}^k = \sum_{j=0}^{k-1} C_k^j D_p^j - D_p^0$$

Or, par hypothèse pour tout $j = 1, p-1$, $D_p^j = 0$ et $D_p^p = p!$. On obtient donc :

$$\forall k = 1, p, D_{p+1}^k = 0 \text{ et } D_{p+1}^{p+1} = C_{p+1}^p D_p^p = (p+1)p! = (p+1)!$$

- Il nous reste à établir (4.15). Au rang $p = 1$, on a : $D_1^2 = 1^2 = 1 = \frac{1}{2}2!$. La propriété est donc vraie au rang 1. Supposons qu'elle soit vérifiée jusqu'au rang p . On a donc : $D_p^{p+1} = \frac{p}{2}(p+1)!$. Utilisant (4.14), le fait que $D_p^j = 0$ pour tout $j = 0, p-1$, et $D_p^p = p!$, on obtient :

$$D_{p+1}^{p+2} = \sum_{j=p}^{p+1} C_{p+2}^j D_p^j = C_{p+2}^p D_p^p + C_{p+2}^{p+1} D_p^{p+1}$$

On utilise alors la propriété de récurrence et on a :

$$\begin{aligned} D_{p+1}^{p+2} &= \frac{(p+2)!}{2} + (p+2) \frac{p(p+1)!}{2} \\ &= \frac{(p+2)!}{2} (1+p) \end{aligned}$$

La propriété est donc vérifiée au rang $p+1$.

B Démonstration du théorème 1

- Pour montrer (4.18), nous nous baserons sur le résultat suivant dont la démonstration est donnée en annexe C :

Lemme 8 *Si X est un $CAR(p)$, alors, pour tout $\mathbf{k} = 0, \mathbf{p} - 2$, on a :*

$$\mathbf{X}^{(\mathbf{k})}(\mathbf{t} + \delta) - \mathbf{X}^{(\mathbf{k})}(\mathbf{t}) = \sum_{l=\mathbf{k}+1}^{\mathbf{p}-1} \frac{\delta^{l-\mathbf{k}}}{(l-\mathbf{k})!} \mathbf{X}^{(l)}(\mathbf{t}) + \eta^{\mathbf{p}}(\mathbf{k}, \mathbf{t}) + \varepsilon^{\mathbf{p}}(\mathbf{k}, \mathbf{t})$$

$$\mathbf{X}^{(\mathbf{p}-1)}(\mathbf{t} + \delta) - \mathbf{X}^{(\mathbf{p}-1)}(\mathbf{t}) = \eta^{\mathbf{p}}(\mathbf{p} - 1, \mathbf{t}) + \varepsilon^{\mathbf{p}}(\mathbf{p} - 1, \mathbf{t})$$

où :

$$\eta^{\mathbf{p}}(\mathbf{k}, \mathbf{t}) = - \sum_{i=0}^{\mathbf{p}-1} \alpha_i \left[\int_{\mathbf{t}}^{\mathbf{t}+\delta} \frac{(\mathbf{t} + \delta - \mathbf{u})^{\mathbf{p}-1-\mathbf{k}}}{(\mathbf{p} - 1 - \mathbf{k})!} \mathbf{X}^{(i)}(\mathbf{u}) \, d\mathbf{u} \right]$$

$$\varepsilon^{\mathbf{p}}(\mathbf{k}, \mathbf{t}) = \int_{\mathbf{t}}^{\mathbf{t}+\delta} \frac{(\mathbf{t} + \delta - \mathbf{u})^{\mathbf{p}-1-\mathbf{k}}}{(\mathbf{p} - 1 - \mathbf{k})!} \sigma d\mathbf{W}(\mathbf{u})$$

Raisonnons par récurrence. Au rang $j = 1$, on a par définition :

$$\Delta^{(1)} X(t + \delta) = X(t + \delta) - X(t)$$

Il suffit alors d'utiliser le lemme 8 avec $k = 0$ pour obtenir le résultat annoncé puisque $\varepsilon^{\mathbf{p}}(0, t) = B^{\mathbf{p}}(1, t)$ et $\eta^{\mathbf{p}}(0, t) = R^{\mathbf{p}}(1, t)$.

Supposons que la propriété soit vraie jusqu'au rang j , $j \leq p - 2$. On a :

$$\begin{aligned} \Delta^{(j+1)} X(t + (j+1)\delta) &= \Delta^{(j)} X(t + (j+1)\delta) - \Delta^{(j)} X(t + j\delta) \\ &= \sum_{k=j}^{p-2} \frac{D_j^k \delta^k}{k!} \left[X^{(k)}(t + \delta) - X^{(k)}(t) \right] \\ &\quad + R^{\mathbf{p}}(j, t + \delta) - R^{\mathbf{p}}(j, t) + B^{\mathbf{p}}(j, t + \delta) - B^{\mathbf{p}}(j, t) \end{aligned}$$

Utilisant le lemme 8, on obtient :

$$\Delta^{(j+1)} X(t + (j+1)\delta) = \sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} \left[\sum_{l=k+1}^{p-1} \frac{\delta^{l-k}}{(l-k)!} X^{(l)}(t) \right] + R + B$$

avec :

$$\begin{aligned} R &= R^{\mathbf{p}}(j, t + \delta) - R^{\mathbf{p}}(j, t) + \sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} \eta^{\mathbf{p}}(k, t) \\ B &= B^{\mathbf{p}}(j, t + \delta) - B^{\mathbf{p}}(j, t) + \sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} \varepsilon^{\mathbf{p}}(k, t) \end{aligned}$$

Or, d'après le lemme 2-(4.14), on a :

$$\begin{aligned} \sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} \left[\sum_{l=k+1}^{p-1} \frac{\delta^{l-k}}{(l-k)!} X^{(l)}(t) \right] &= \sum_{l=j+1}^{p-1} \frac{\delta^l}{l!} X^{(l)}(t) \left[\sum_{k=j}^{l-1} C_l^k D_j^k \right] \\ &= \sum_{l=j+1}^{p-1} \frac{D_{j+1}^l \delta^l}{l!} X^{(l)}(t) \end{aligned}$$

Il nous reste à évaluer R et B . On a par définition :

$$\begin{aligned} &B^p(j, t + \delta) - B^p(j, t) \\ &= \int_{t+j\delta}^{t+(j+1)\delta} (t + (j+1)\delta - u)^{p-1} \frac{\sigma}{(p-1)!} dW(u) \\ &+ \sum_{k=2}^j \int_{t+(k-1)\delta}^{t+k\delta} (t + (j+1)\delta - u)^{p-1} \frac{\sigma}{(p-1)!} dW(u) \\ &+ \sum_{k=2}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=1}^{j+1-k} (-1)^l (C_j^l + C_j^{l-1}) (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\ &- \int_t^{t+\delta} \left[\sum_{l=0}^{j-1} C_j^l (-1)^l (t + j\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\ &= \sum_{k=2}^{j+1} \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j+1-k} (-1)^l C_{j+1}^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\ &- \int_t^{t+\delta} \left[\sum_{l=0}^{j-1} C_j^l (-1)^l (t + j\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \end{aligned}$$

De plus, toujours d'après le lemme 2, on a :

$$\begin{aligned} &\sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} \varepsilon^p(k, t) \\ &= \sum_{k=1}^{p-1} \frac{D_j^k \delta^k}{k!} \varepsilon^p(k, t) \\ &= \int_t^{t+\delta} \left[\sum_{l=0}^{j-1} C_j^l (-1)^l \left(\sum_{k=1}^{p-1} C_{p-1}^k (t + \delta - u)^{p-1-k} ((j-l)\delta)^k \right) \right] \frac{\sigma}{(p-1)!} dW(u) \end{aligned}$$

Une application directe de la formule du binôme de Newton donne :

$$\begin{aligned}
\sum_{k=j}^{p-1} \frac{D_j^k \delta^k}{k!} \varepsilon^p(k, t) &= \int_t^{t+\delta} \left[\sum_{l=0}^{j-1} C_j^l (-1)^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\
&- \int_t^{t+\delta} \left[\sum_{l=0}^{j-1} C_j^l (-1)^l (t + \delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\
&= \int_t^{t+\delta} \left[\sum_{l=0}^{j-1} C_j^l (-1)^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\
&+ \int_t^{t+\delta} (-1)^j (t + \delta - u)^{p-1} \frac{\sigma}{(p-1)!} dW(u) \\
&= \int_t^{t+\delta} \left[\sum_{l=0}^j C_j^l (-1)^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u)
\end{aligned}$$

On obtient donc :

$$\begin{aligned}
B &= \sum_{k=2}^{j+1} \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j+1-k} (-1)^l C_{j+1}^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\
&+ \int_t^{t+\delta} (t + (j+1)\delta - u)^{p-1} \frac{\sigma}{(p-1)!} dW(u) \\
&+ \int_t^{t+\delta} \left[\sum_{l=1}^j (C_j^l + C_j^{l-1}) (-1)^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\
&= \sum_{k=1}^{j+1} \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j+1-k} (-1)^l C_{j+1}^l (t + (j+1)\delta - l\delta - u)^{p-1} \right] \frac{\sigma}{(p-1)!} dW(u) \\
&= B^p(j+1, t)
\end{aligned}$$

Puisque le terme R a une structure identique à celle du terme précédemment évalué, la propriété de récurrence est donc satisfaite au rang $j+1$: (4.18) est établie.

- L'équation précédente est valable jusqu'au rang p avec :

$$\Delta^{(p)} X(t + p\delta) = R^p(p, t) + B^p(p, t)$$

Nous allons cependant préciser ce développement pour $\Delta^{(p-1)} X(t + (p-1)\delta)$ et $\Delta^{(p)} X(t + p\delta)$. On utilise pour cela le résultat suivant :

Lemme 9 Avec les notations (4.16) et (4.17), on a pour $i = 0, p-2$:

$$\mathbf{R}_i^p(\mathbf{j}, t) = \delta^p \frac{\mathbf{D}_j^p}{p!} \mathbf{X}^{(i)}(t) + \mathbf{R}_{i+1}^{p+1}(\mathbf{j}, t) \quad (4.33)$$

et

$$\mathbf{R}_{p-1}^p(\mathbf{j}, t) = \delta^p \frac{\mathbf{D}_j^p}{p!} \mathbf{X}^{(p-1)}(t) - \sum_{i=0}^{p-1} \alpha_i \mathbf{R}_i^{p+1}(\mathbf{j}, t) + \mathbf{B}^{p+1}(\mathbf{j}, t) \quad (4.34)$$

La démonstration de ce résultat est donnée en annexe D.

- Montrons (4.20). On a d'après (4.18) :

$$\begin{aligned}\Delta^{(p-1)}X(t+(p-1)\delta) &= \frac{D_{p-1}^{p-1}\delta^{p-1}}{(p-1)!}X^{(p-1)}(t) + R^p(p-1,t) + B^p(p-1,t) \\ &= \delta^{p-1}X^{(p-1)}(t) + R^p(p-1,t) + B^p(p-1,t)\end{aligned}$$

Utilisant le lemme 9, le reste $R^p(p-1,t)$ s'écrit :

$$\begin{aligned}R^p(p-1,t) &= -\sum_{i=0}^{p-2}\alpha_i\left[\delta^p\frac{D_p^p}{p!}X^{(i)}(t) + R_{i+1}^{p+1}(p-1,t)\right] \\ &\quad -\alpha_{p-1}\left[\delta^p\frac{D_p^p}{p!}X^{(p-1)}(t) - \sum_{i=0}^{p-1}\alpha_i R_i^{p+1}(p-1,t) + B^{p+1}(p-1,t)\right] \\ &= -\delta^p\frac{D_p^p}{p!}\sum_{i=0}^{p-1}\alpha_i X^{(i)}(t) + \sum_{i=0}^{p-1}\beta_i R_i^{p+1}(p-1,t) - \alpha_{p-1}B^{p+1}(p-1,t)\end{aligned}$$

On a donc le développement annoncé.

- Montrons (4.21). Utilisant (4.19), on a :

$$\Delta^{(p)}X(t+p\delta) = R^p(p,t) + B^p(p,t)$$

En utilisant à nouveau le lemme 9, on a :

$$\begin{aligned}R^p(p,t) &= -\delta^p\frac{D_p^p}{p!}\sum_{i=0}^{p-1}\alpha_i X^{(i)}(t) + \sum_{i=0}^{p-1}\beta_i R_i^{p+1}(p,t) - \alpha_{p-1}B^{p+1}(p,t) \\ &= -\delta^p\sum_{i=0}^{p-1}\alpha_i X^{(i)}(t) + \sum_{i=0}^{p-1}\beta_i R_i^{p+1}(p,t) - \alpha_{p-1}B^{p+1}(p,t)\end{aligned}$$

En itérant, on obtient :

$$\begin{aligned}R^{p+1}(p,t) &= -\delta^p\sum_{i=0}^{p-1}\alpha_i X^{(i)}(t) - \alpha_{p-1}B^{p+1}(p,t) \\ &\quad + \sum_{i=0}^{p-2}\beta_i\left[\delta^{p+1}\frac{D_p^{p+1}}{(p+1)!}X^{(i)}(t) + R_{i+1}^{p+2}(p,t)\right] \\ &\quad + \beta_{p-1}\left[\delta^{p+1}\frac{D_p^{p+1}}{(p+1)!}X^{(p-1)}(t) - \sum_{i=0}^{p-1}\alpha_i R_i^{p+2}(p,t) + B^{p+2}(p,t)\right] \\ &= -\delta^p\sum_{i=0}^{p-1}\alpha_i X^{(i)}(t) + \delta^{p+1}\frac{D_p^{p+1}}{(p+1)!}\sum_{i=0}^{p-1}\beta_i X^{(i)}(t) \\ &\quad + \sum_{i=0}^{p-1}\gamma_i R_i^{p+2}(p,t) \\ &\quad - \alpha_{p-1}B^{p+1}(p,t) + [(\alpha_{p-1})^2 - \alpha_{p-2}]B^{p+2}(p,t)\end{aligned}$$

On trouve donc le résultat annoncé.

C Démonstration du lemme 8

Pour $k = p - 1$, ce résultat est immédiat. Pour $k \leq p - 2$, $t \rightarrow X^{(k)}(t)$ est de classe C^{p-1-k} presque sûrement. On utilise alors la formule de Taylor avec reste intégral et on obtient :

$$\begin{aligned} X^{(k)}(t + \delta) - X^{(k)}(t) &= \sum_{l=k+1}^{p-1} \frac{\delta^{l-k}}{(l-k)!} X^{(l)} + \int_t^{t+\delta} \frac{(t + \delta - u)^{p-2-k}}{(p-2-k)!} \left[X^{(p-1)}(u) - X^{(p-1)}(t) \right] du \\ &= \sum_{l=k+1}^{p-1} \frac{\delta^{l-k}}{(l-k)!} X^{(l)} + \int_t^{t+\delta} \frac{(t + \delta - u)^{p-2-k}}{(p-2-k)!} \left[\int_t^u dX^{(p-1)}(v) \right] du \end{aligned}$$

Notons :

$$R = \int_t^{t+\delta} \frac{(t + \delta - u)^{p-2-k}}{(p-2-k)!} \left[\int_t^u dX^{(p-1)}(v) \right] du$$

Par définition, on a :

$$dX^{(p-1)}(v) = - \sum_{l=0}^{p-1} \alpha_l X^{(l)}(v) dv + \sigma dW(v)$$

R se décompose donc comme suit :

$$\begin{aligned} R &= R_1 + R_2 \\ R_1 &= - \sum_{l=0}^{p-1} \alpha_l \int_t^{t+\delta} \frac{(t + \delta - u)^{p-2-k}}{(p-2-k)!} \left[\int_t^u X^{(l)}(v) dv \right] du \\ R_2 &= \int_t^{t+\delta} \frac{(t + \delta - u)^{p-2-k}}{(p-2-k)!} \left[\int_t^u \sigma dW(v) \right] du \end{aligned}$$

En intégrant par parties, on montre pour tout $l = 0, p - 1$:

$$\int_t^{t+\delta} \frac{(t + \delta - u)^{p-2-k}}{(p-2-k)!} \left[\int_t^u X^{(l)}(v) dv \right] du = \int_t^{t+\delta} \frac{(t + \delta - v)^{p-1-k}}{(p-1-k)!} X^{(l)}(v) dv$$

On a donc : $R_1 = \eta^p(k, t)$. Il nous reste à montrer que $R_2 = \varepsilon^p(k, t)$. On utilise pour cela le résultat suivant qui découle de l'application de la Formule de Ito au processus $(u^l W(u))_{u \geq 0}$:

$$\forall l \in N, \quad \int_t^{t+\delta} u^l \left[\int_t^u dW(v) \right] du = \int_t^{t+\delta} \left[\frac{(t + \delta)^{l+1} - v^{l+1}}{(l+1)} \right] dW(v)$$

On a donc :

$$\begin{aligned}
R_2 &= \frac{\sigma}{(p-2-k)!} \sum_{l=0}^{p-2-k} C_{p-2-k}^l (-1)^l (t+\delta)^{p-2-k-l} \int_t^{t+\delta} u^l \left[\int_t^u dW(v) \right] du \\
&= \frac{\sigma}{(p-2-k)!} \sum_{l=0}^{p-2-k} C_{p-2-k}^l (-1)^l (t+\delta)^{p-2-k-l} \int_t^{t+\delta} \left[\frac{(t+\delta)^{l+1} - v^{l+1}}{(l+1)} \right] dW(v) \\
&= \frac{\sigma}{(p-1-k)!} \sum_{l=0}^{p-2-k} C_{p-1-k}^{l+1} (-1)^l (t+\delta)^{p-2-k-l} \int_t^{t+\delta} \left[(t+\delta)^{l+1} - v^{l+1} \right] dW(v) \\
&= \frac{\sigma}{(p-1-k)!} \int_t^{t+\delta} \left[\sum_{l=0}^{p-2-k} C_{p-1-k}^{l+1} (-1)^{l+1} (t+\delta)^{p-1-k-(l+1)} v^{l+1} \right] dW(v) \\
&\quad - \frac{\sigma}{(p-1-k)!} \int_t^{t+\delta} \left[\sum_{l=0}^{p-2-k} C_{p-1-k}^{l+1} (-1)^{l+1} (t+\delta)^{p-1-k} \right] dW(v) \\
&= \frac{\sigma}{(p-1-k)!} \int_t^{t+\delta} \left[(t+\delta-v)^{p-1-k} - (t+\delta)^{p-1-k} \right] dW(v) \\
&\quad + \frac{\sigma}{(p-1-k)!} \int_t^{t+\delta} (t+\delta)^{p-1-k} dW(v) \\
&= \varepsilon^p(k, t)
\end{aligned}$$

On obtient le développement annoncé.

D Démonstration du lemme 9

Pour tout $i = 0, p - 2$ et tout $j \geq 1$, on a en intégrant par parties :

$$\begin{aligned} R_i^p(j, t) &= \frac{1}{p!} \sum_{k=1}^j \left[- \sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p X^{(i)}(u) \right]_{u=t+(k-1)\delta}^{u=t+k\delta} \\ &+ \frac{1}{p!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p \right] X^{(i+1)}(u) du \end{aligned}$$

Or :

$$\begin{aligned} &\sum_{k=1}^j \left[- \sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p X^{(i)}(u) \right]_{u=t+(k-1)\delta}^{u=t+k\delta} \\ &= \sum_{k=1}^j \sum_{l=0}^{j-k} C_j^l (-1)^l (j\delta - l\delta - (k-1)\delta)^p X^{(i)}(t + (k-1)\delta) \\ &- \sum_{k=1}^j \sum_{l=0}^{j-k} C_j^l (-1)^l (j\delta - l\delta - k\delta)^p X^{(i)}(t + k\delta) \\ &= \sum_{k=0}^{j-1} \sum_{l=0}^{j-1-k} C_j^l (-1)^l (j\delta - l\delta - k\delta)^p X^{(i)}(t + k\delta) \\ &- \sum_{k=1}^{j-1} \sum_{l=0}^{j-1-k} C_j^l (-1)^l (t + j\delta - l\delta - k\delta)^p X^{(i)}(t + k\delta) \\ &= \sum_{l=0}^{j-1} C_j^l (-1)^l (j\delta - l\delta)^p X^{(i)}(t) \\ &= \delta^p D_j^p X^{(i)}(t) \end{aligned}$$

De plus, d'après (4.17) :

$$R_{i+1}^{p+1}(j, t) = \frac{1}{p!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p \right] X^{(i+1)}(u) du$$

On obtient donc :

$$R_i^p(j, t) = \delta^p \frac{D_j^p}{p!} X^{(i)}(t) + R_{i+1}^{p+1}(j, t)$$

On procède de même pour R_{p-1}^p (avec la formule de Ito); on obtient ainsi :

$$\begin{aligned} R_{p-1}^p(j, t) &= \delta^p \frac{D_j^p}{p!} X^{(p-1)}(t) \\ &+ \frac{1}{p!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p \right] dX^{(p-1)}(u) \end{aligned}$$

Or, par définition du CAR(p), on a :

$$dX^{(p-1)}(u) = \left[- \sum_{l=0}^{p-1} \alpha_l X^{(l)}(u) \right] du + \sigma dW(u)$$

En remplaçant dans l'équation précédente, on a :

$$\begin{aligned} R_{p-1}^p(j, t) &= \delta^p \frac{D_j^p}{p!} X^{(p-1)}(t) \\ &- \sum_{l=0}^{p-1} \alpha_l \left(\frac{1}{p!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p \right] X^{(l)}(u) du \right) \\ &+ \frac{1}{p!} \sum_{k=1}^j \int_{t+(k-1)\delta}^{t+k\delta} \left[\sum_{l=0}^{j-k} C_j^l (-1)^l (t + j\delta - l\delta - u)^p \right] \sigma dW(u) \end{aligned}$$

D'après (4.16) et (4.17), on a :

$$R_{p-1}^p(j, t) = \delta^p \frac{D_j^p}{p!} X^{(p-1)}(t) - \sum_{l=0}^{p-1} \alpha_l R_l^{p+1}(j, t) + B^{p+1}(j, t)$$

E Démonstration du lemme 3 :

Calculons $E_\theta [\Delta^{(p-1)} X ((p-1)\delta) \Delta^{(p)} X (p\delta)]$. Utilisant le théorème 1, on a :

$$\begin{aligned} & E_\theta \left[\Delta^{(p-1)} X ((p-1)\delta) \Delta^{(p)} X (p\delta) \right] \\ &= -\delta^{2p-1} \sum_{k=0}^{p-1} \alpha_k D_{p-1,k}(0) + \delta^{2p} \frac{(p-1)}{2} \sum_{j=0}^{p-1} \sum_{k=0}^{p-1} \alpha_j \alpha_k D_{j,k}(0) + \delta^{2p} \frac{p}{2} \sum_{k=0}^{p-1} \beta_k D_{p-1,k}(0) \\ &+ E_\theta [B^p(p-1,0) B^p(p,0)] - \alpha_{p-1} E_\theta [B^{p+1}(p-1,0) B^p(p,0)] \\ &- \alpha_{p-1} E_\theta [B^p(p-1,0) B^{p+1}(p,0)] + o(\delta^{2p}) \end{aligned}$$

D'après (4.7), (4.8), (4.13) et (4.12), on a :

$$-\sum_{k=0}^{p-1} \alpha_k D_{p-1,k}(0) = D_{p-1,p}(0)$$

et

$$\frac{(p-1)}{2} \sum_{j=0}^{p-1} \sum_{k=0}^{p-1} \alpha_j \alpha_k D_{j,k}(0) + \frac{p}{2} \sum_{k=0}^{p-1} \beta_k D_{p-1,k}(0) = \frac{(-1)^p}{2} \sum_{k=0}^{p-1} \alpha_k r^{(p+k)}(0)$$

On a de plus :

$$\begin{aligned} E_\theta [B^p(p-1,0) B^p(p,0)] &= C(p-1,p) \sigma^2 \delta^{2p-1} \\ C(p-1,p) &= \frac{1}{[(p-1)!]^2} \sum_{k=1}^{p-1} \int_{k-1}^k \sum_{l=0}^{p-1-k} \sum_{i=0}^{p-k} \left[C_{p-1}^l C_p^i (-1)^{l+i} (p-1-l-u)^{p-1} (p-i-u)^{p-1} \right] du \end{aligned}$$

c'est-à-dire d'après (4.11) :

$$E_\theta [B^p(p-1,0) B^p(p,0)] = -2C(p-1,p) D_{p-1,p}(0) \delta^{2p-1}$$

De même :

$$\begin{aligned} E_\theta [B^p(p-1,0) B^{p+1}(p,0)] &= D_1(p-1,p) \sigma^2 \delta^{2\delta} \\ D_1(p-1,p) &= \frac{1}{p!(p-1)!} \sum_{k=1}^{p-1} \int_{k-1}^k \sum_{l=0}^{p-1-k} \sum_{i=0}^{p-k} \left[C_{p-1}^l C_p^i (-1)^{l+i} (p-1-l-u)^{p-1} (p-i-u)^p \right] du \end{aligned}$$

et :

$$\begin{aligned} E_\theta [B^{p+1}(p-1,0) B^p(p,0)] &= D_2(p-1,p) \sigma^2 \delta^{2\delta} \\ D_2(p-1,p) &= \frac{1}{p!(p-1)!} \sum_{k=1}^{p-1} \int_{k-1}^k \sum_{l=0}^{p-1-k} \sum_{i=0}^{p-k} \left[C_{p-1}^l C_p^i (-1)^{l+i} (p-1-l-u)^p (p-i-u)^{p-1} \right] du \end{aligned}$$

On remarque que :

$$\begin{aligned} & D_1(p-1,p) + D_2(p-1,p) \\ &= -\frac{1}{(p!)^2} \sum_{k=1}^{p-1} \left[\left(\sum_{l=0}^{p-1-k} C_{p-1}^l (-1)^l (p-1-l-u)^p \right) \left(\sum_{i=0}^{p-k} C_p^i (-1)^i (p-i-u)^p \right) \right]_{u=k-1}^{u=k} \end{aligned}$$

En regroupant les différents développements, on obtient :

$$\begin{aligned}
 & E_{\theta} \left[\Delta^{(p-1)} X ((p-1)\delta) \Delta^{(p)} X (p\delta) \right] \\
 &= \delta^{2p-1} C(p) D_{p-1,p}(0) + \delta^{2p} \frac{(-1)^p}{2} \sum_{k=0}^{p-1} \alpha_k r^{(p+k)}(0) + \alpha_{p-1} D(p) \sigma^2 + o(\delta^{2p})
 \end{aligned}$$

avec $D(p) = -D_1(p-1, p) - D_2(p-1, p)$ et $C(p) = 1 - 2C(p-1, p)$. Or, on montre que l'écriture de $C(p)$ se factorise comme indiqué dans l'énoncé.

F Démonstration du lemme 4

Montrons (4.25). Utilisant le développement du théorème 1, on obtient :

$$\delta^{2p-1} \mathcal{D}_{p-1,p} = -\delta^{2p-1} E_\theta \left[X^{(p-1)}(0) \left(\sum_{k=0}^{p-1} \alpha_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.35)$$

$$+ \delta^{2p} \frac{p}{2} E_\theta \left[X^{(p-1)}(0) \left(\sum_{k=0}^{p-1} \beta_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.36)$$

$$+ \delta^{2p} \frac{p-1}{2} E_\theta \left[\left(\sum_{i=0}^{p-1} \alpha_i X^{(i)}(0) \right) \left(\sum_{k=0}^{p-1} \alpha_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.37)$$

$$- \delta^{2p+1} \frac{p(p-1)}{4} E_\theta \left[\left(\sum_{i=0}^{p-1} \alpha_i X^{(i)}(0) \right) \left(\sum_{k=0}^{p-1} \beta_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.38)$$

$$+ \delta^{p-1} E_\theta \left[\left(X^{(p-1)}(0) - \delta \frac{p-1}{2} \sum_{i=0}^{p-1} \alpha_i X^{(i)}(0) \right) \left(\sum_{k=0}^{p-1} \gamma_k R_k^{p+2}(p, (p-1)\delta) \right) \right] \quad (4.39)$$

$$- \delta^p E_\theta \left[\left(\sum_{i=0}^{p-1} \beta_i R_i^{p+1}(p-1, 0) \right) \left(\sum_{k=0}^{p-1} \alpha_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.40)$$

$$+ \delta^{p+1} \frac{p}{2} E_\theta \left[\left(\sum_{i=0}^{p-1} \beta_i R_i^{p+1}(p-1, 0) \right) \left(\sum_{k=0}^{p-1} \beta_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.41)$$

$$+ E_\theta \left[\left(\sum_{i=0}^{p-1} \beta_i R_i^{p+1}(p-1, 0) \right) \left(\sum_{k=0}^{p-1} \gamma_k R_k^{p+2}(p, (p-1)\delta) \right) \right] \quad (4.42)$$

$$- \delta^p E_\theta \left[\left(B^p(p-1, 0) - \alpha_{p-1} B^{p+1}(p-1, 0) \right) \left(\sum_{k=0}^{p-1} \alpha_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.43)$$

$$+ \delta^{p+1} \frac{p}{2} E_\theta \left[\left(B^p(p-1, 0) - \alpha_{p-1} B^{p+1}(p-1, 0) \right) \left(\sum_{k=0}^{p-1} \beta_k X^{(k)}((p-1)\delta) \right) \right] \quad (4.44)$$

$$+ E_\theta \left[\left(B^p(p-1, 0) - \alpha_{p-1} B^{p+1}(p-1, 0) \right) \left(\sum_{k=0}^{p-1} \gamma_k R_k^{p+2}(p, (p-1)\delta) \right) \right] \quad (4.45)$$

Nous allons analyser successivement chacun de ces onze termes.

Terme (4.35) :

D'après (4.6), on a :

$$(4.35) = -\delta^{2p-1} \sum_{k=0}^{p-1} \alpha_k D_{p-1,k}(-(p-1)\delta) = \delta^{2p-1} D_{p-1,p}(-(p-1)\delta)$$

Termes (4.36), (4.37) et (4.38) :

On a :

$$(4.36) = \delta^{2p} \frac{p}{2} \sum_{k=0}^{p-1} \beta_k D_{p-1,k} (-(p-1)\delta)$$

$$(4.37) = \delta^{2p} \frac{p-1}{2} \sum_{i=0}^{p-1} \sum_{k=0}^{p-1} \alpha_i \alpha_k D_{i,k} (-(p-1)\delta)$$

et :

$$(4.38) = -\delta^{2p+1} \frac{p(p-1)}{4} \sum_{i=0}^{p-1} \sum_{k=0}^{p-1} \alpha_i \beta_k D_{i,k} (-(p-1)\delta)$$

Termes (4.39), (4.40), (4.41) et (4.42):

Par application de l'inégalité de Hölder, on montre qu'il existe K_p finie telle que pour tout $i = 0, p-1$:

$$\left\| R_i^{p+1}(p-1, 0) \right\|_{L^2(\mu_\theta)} \leq K_p \delta^{p+1}, \quad \text{et} \quad \left\| R_i^{p+2}(p-1, (p-1)\delta) \right\|_{L^2(\mu_\theta)} \leq K_p \delta^{p+2}$$

Utilisant l'inégalité de Schwarz, on obtient que : (4.39) + (4.40) + (4.41) $\leq O(\delta^{2p+1})$ et (4.42) $\leq O(\delta^{2p+3})$.

Termes (4.43) et (4.44) :

Le terme dominant de (4.43)+(4.44) est $-\delta^p \alpha_{p-1} E_\theta [B^p(p-1, 0) X^{(p-1)}((p-1)\delta)]$. On a :

$$\begin{aligned} \delta^p E_\theta [B^p(p-1, 0) X^{(p-1)}((p-1)\delta)] &= \delta^p E_\theta [B^p(p-1, 0) \rho^p(p-1, 0)] \\ &+ \delta^p E_\theta \left[B^p(p-1, 0) \left(\int_0^{(p-1)\delta} \sigma dW(u) \right) \right] \end{aligned}$$

avec : $\rho^p(p-1, 0) = -\sum_{i=0}^{p-1} \alpha_i \int_0^{(p-1)\delta} X^{(i)}(u) du$.

Calculons $E_\theta [(B^p(p-1, 0))^2]$.

$$\begin{aligned} E_\theta [(B^p(p-1, 0))^2] &= \delta^{2p-1} \frac{\sigma^2}{[(p-1)!]^2} \sum_{k=1}^{p-1} \int_{k-1}^k \left[\sum_{l=0}^{p-1-k} C_{p-1}^l (-1)^l (p-1-l-u)^{p-1} \right]^2 du \\ &= \delta^{2p-1} \sigma^2 C_p \end{aligned}$$

Il existe de plus K finie telle que $\|\rho^p(p-1, 0)\|_{L^2(\mu_\theta)} \leq K\delta$. Par application de l'inégalité de Schwarz, on montre que : $\delta^p E_\theta [B^p(p-1, 0) \rho^p(p-1, 0)] \leq K_p \delta^{2p+\frac{1}{2}}$.

On a enfin :

$$\begin{aligned}
& \delta^p E_\theta \left[B^p(p-1, 0) \left(\int_0^{(p-1)\delta} \sigma dW(u) \right) \right] \\
&= \delta^{2p} \frac{\sigma^2}{(p-1)!} \sum_{k=1}^{p-1} \int_{k-1}^k \left[\sum_{l=0}^{p-1-k} C_{p-1}^l (-1)^l (p-1-l-u)^{p-1} \right] du \\
&= \delta^{2p} \frac{\sigma^2 (p-1)}{2}
\end{aligned}$$

Donc :

$$-\alpha_{p-1} \delta^p E_\theta \left[B^p(p-1, 0) X^{(p-1)}((p-1)\delta) \right] = \alpha_{p-1} \delta^{2p} \frac{\sigma^2 (p-1)}{2} + o(\delta^{2p})$$

Terme (4.45) :

Là encore, le terme dominant est $E_\theta \left[B^p(p-1, 0) \left(\sum_{k=0}^{p-1} \gamma_k R_k^{p+2}(p, (p-1)\delta) \right) \right]$. D'après les résultats obtenus lors de l'étude des termes $\{(4.39), (4.40), (4.41), (4.42)\}$ et $\{(4.43), (4.44)\}$, et par application de l'inégalité de Schwarz, on a :

$$E_\theta \left[B^p(p-1, 0) \left(\sum_{k=0}^{p-1} \gamma_k R_k^{p+2}(p, (p-1)\delta) \right) \right] \leq O\left(\delta^{2p+2-\frac{1}{2}}\right)$$

En regroupant les différents développements et en utilisant l'équation de Yule-Walker (4.6) en $j = p-1$, on obtient :

$$\mathcal{D}_{p-1,p} = D_{p-1,p}(-(p-1)\delta) + \delta \Psi_{p-1,p}(\delta) + o(\delta)$$

avec :

$$\Psi_{p-1,p}(\delta) = \frac{1}{2} \left[p \sum_{k=0}^{p-1} \beta_k D_{p-1,k}(-(p-1)\delta) + (p-1) \sum_{i=0}^{p-1} \sum_{k=0}^{p-1} \alpha_i \alpha_k D_{i,k}((p-1)\delta) - \alpha_{p-1} (p-1) \sigma^2 \right]$$

Par le même type d'arguments, on montre que, pour $j = 0, p-1$ et $k = 0, p$:

$$\mathcal{D}_{j,k} = D_{j,k}(-(p-1)\delta) + \delta \Psi_{j,k}(\delta) + o(\delta)$$

avec, pour $j = 0, p-2$:

$$\Psi_{j,p}(\delta) = \frac{1}{2} \left[p \sum_{k=0}^{p-1} \beta_k D_{j,k}(-(p-1)\delta) - j \left(\sum_{k=0}^{p-1} \alpha_k D_{j+1,k}(-(p-1)\delta) \right) \right]$$

Pour $k = 0, p-2$:

$$\Psi_{p-1,k}(\delta) = \frac{1}{2} \left[-(p-1) \sum_{i=0}^{p-1} \alpha_i D_{i,k}(-(p-1)\delta) + k D_{p-1,k+1}(-(p-1)\delta) \right]$$

$$\Psi_{p-1,p-1}(\delta) = \frac{1}{2} \left[- (p-1) \sum_{i=0}^{p-1} \alpha_i D_{p-1,i}(- (p-1) \delta) - (p-1) \left(\sum_{i=0}^{p-1} \alpha_i D_{i,p-1}(- (p-1) \delta) \right) \right] + \frac{(p-1)}{2} \sigma^2$$

et, pour $j = 0, p-2$ et $k = 0, p-1$:

$$\Psi_{j,k}(\delta) = \frac{1}{2} [j D_{j+1,k}(- (p-1) \delta) + k D_{j,k+1}(- (p-1) \delta)]$$

Il suffit alors d'utiliser les équations de Yule-Walker exprimées en fonction de r , (4.13) et (4.12), pour obtenir le résultat annoncé.

G Démonstration du lemme 7

Montrons (4.29). Par définition du CAR(2), on a pour tout $t \geq 0$ et tout $\delta > 0$:

$$X(t + \delta) - X(t) = \int_t^{t+\delta} X^{(1)}(u) du$$

Par application de la formule de Ito, on a :

$$\begin{aligned} \int_t^{t+\delta} X^{(1)}(u) du &= - \left[(t + \delta - u) X^{(1)}(u) \right]_{u=t}^{u=t+\delta} \\ &+ \int_t^{t+\delta} (t + \delta - u) \left[-\alpha_0 X(u) - \alpha_1 X^{(1)}(u) \right] du \\ &+ \int_t^{t+\delta} (t + \delta - u) \sigma dW(u) \end{aligned}$$

et :

$$\begin{aligned} \int_t^{t+\delta} X^{(1)}(u) du &= - \left[(t - u) X^{(1)}(u) \right]_{u=t}^{u=t+\delta} \\ &+ \int_t^{t+\delta} (t - u) \left[-\alpha_0 X(u) - \alpha_1 X^{(1)}(u) \right] du \\ &+ \int_t^{t+\delta} (t - u) \sigma dW(u) \end{aligned}$$

Utilisant les deux développements précédents, on obtient :

$$\begin{aligned} X(t + \delta) - X(t) &= \frac{\delta}{2} \left[X^{(1)}(t) + X^{(1)}(t + \delta) \right] \\ &+ \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) \left[-\alpha_0 X(u) - \alpha_1 X^{(1)}(u) \right] du \\ &+ \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) \sigma dW(u) \end{aligned}$$

Appliquant à nouveau la formule de Ito, on obtient :

$$\begin{aligned} \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) X(u) &= -\frac{1}{2} \left[(t - u) (t + \delta - u) X(u) \right]_{u=t}^{u=t+\delta} \\ &+ \int_t^{t+\delta} \frac{1}{2} (t - u) (t + \delta - u) X^{(1)}(u) du \end{aligned}$$

D'où :

$$\int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) X(u) = \int_t^{t+\delta} \frac{1}{2} (t - u) (t + \delta - u) X^{(1)}(u) du \quad (4.46)$$

et :

$$\begin{aligned} \int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) X^{(1)}(u) du &= -\frac{1}{2} \left[(t - u) (t + \delta - u) X^{(1)}(u) \right]_{u=t}^{u=t+\delta} \\ &+ \frac{1}{2} \int_t^{t+\delta} (t - u) (t + \delta - u) \left[-\alpha_0 X(u) - \alpha_1 X^{(1)}(u) \right] du \\ &+ \int_t^{t+\delta} \frac{1}{2} (t - u) (t + \delta - u) \sigma dW(u) \end{aligned}$$

Donc :

$$\int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) X^{(1)}(u) du = \frac{1}{2} \int_t^{t+\delta} (t-u)(t+\delta-u) \left[-\alpha_0 X(u) - \alpha_1 X^{(1)}(u) \right] du \quad (4.47)$$

$$+ \int_t^{t+\delta} \frac{1}{2} (t-u)(t+\delta-u) \sigma dW(u)$$

Il suffit alors de remplacer $\int_t^{t+\delta} \left(t + \frac{\delta}{2} - u \right) \left[-\alpha_0 X(u) - \alpha_1 X^{(1)}(u) \right] du$ par les deux expressions précédentes, (4.46) et (4.47). On obtient alors :

$$X(t+\delta) - X(t) = \frac{\delta}{2} \left[X^{(1)}(t) + X^{(1)}(t+\delta) \right] + \zeta(0,t) + \xi(0,t)$$

où $\zeta(0,t)$ et $\xi(0,t)$ sont données dans le lemme.

La démonstration de (4.30) repose sur des arguments identiques.

H Démonstration du théorème 3

Utilisant (4.29) et (4.30), nous commencerons par exprimer $X^{(1)}(t + \delta)$ en fonction de $X(t + \delta)$ et $X(t)$. D'après (4.29), on a :

$$\frac{\delta}{2} [X^{(1)}(t + \delta) + X^{(1)}(t)] = X(t + \delta) - X(t) - \zeta(0, t) - \xi(0, t)$$

En remplaçant dans (4.30), on obtient :

$$\begin{aligned} X^{(1)}(t) &= X^{(1)}(t + \delta) + \frac{\delta\alpha_0}{2} [X(t) + X(t + \delta)] + \alpha_1 [X(t + \delta) - X(t)] \\ &\quad - \alpha_1 \zeta(0, t) - \alpha_1 \xi(0, t) - \zeta(1, t) - \xi(1, t) \end{aligned}$$

On réintègre alors cette expression dans (4.29) :

$$\begin{aligned} X(t + \delta) &= X(t) + \frac{\delta}{2} \left(2X^{(1)}(t + \delta) + \frac{\delta\alpha_0}{2} [X(t) + X(t + \delta)] + \alpha_1 [X(t + \delta) - X(t)] \right) \\ &\quad + \left(1 - \frac{\delta\alpha_1}{2} \right) \zeta(0, t) \left(1 - \frac{\delta\alpha_1}{2} \right) \xi(0, t) - \frac{\delta}{2} \zeta(1, t) - \frac{\delta}{2} \xi(1, t) \end{aligned}$$

D'où :

$$\begin{aligned} \frac{\delta}{2} X^{(1)}(t + \delta) &= \frac{1}{2} [X(t + \delta) - X(t)] - \frac{\delta^2\alpha_0}{8} [X(t) + X(t + \delta)] - \frac{\delta\alpha_1}{4} [X(t + \delta) - X(t)] \quad (4.48) \\ &\quad - \frac{1}{2} \left(1 - \frac{\delta\alpha_1}{2} \right) \zeta(0, t) - \frac{1}{2} \left(1 - \frac{\delta\alpha_1}{2} \right) \xi(0, t) + \frac{\delta}{4} \zeta(1, t) + \frac{\delta}{4} \xi(1, t) \end{aligned}$$

Ecrivant (4.29) en $t + \delta$, on a :

$$X(t + 2\delta) - X(t + \delta) = \frac{\delta}{2} [X^{(1)}(t + 2\delta) + X^{(1)}(t + \delta)] + \zeta(0, t + \delta) + \xi(0, t + \delta)$$

On remplace alors $X^{(1)}(t + 2\delta)$ et $X^{(1)}(t + \delta)$ par leurs expressions en fonction de X (4.48) :

$$\begin{aligned} 2[X(t + 2\delta) - X(t + \delta)] &= \left(1 - \frac{\delta\alpha_1}{2} \right) [X(t + 2\delta) - X(t)] \\ &\quad - \frac{\delta^2\alpha_0}{4} [X(t) + 2X(t + \delta) + X(t + 2\delta)] \\ &\quad + \left(1 + \frac{\delta\alpha_1}{2} \right) \zeta(0, t + \delta) - \left(1 - \frac{\delta\alpha_1}{2} \right) \zeta(0, t) + \frac{\delta}{2} \zeta(1, t + \delta) + \frac{\delta}{2} \zeta(1, t) \\ &\quad + \left(1 + \frac{\delta\alpha_1}{2} \right) \xi(0, t + \delta) - \left(1 - \frac{\delta\alpha_1}{2} \right) \xi(0, t) + \frac{\delta}{2} \xi(1, t + \delta) + \frac{\delta}{2} \xi(1, t) \end{aligned}$$

Et donc :

$$\begin{aligned} \Delta^{(2)} X(t + 2\delta) &= -\frac{\delta\alpha_1}{2} [X(t + 2\delta) - X(t)] - \frac{\delta^2\alpha_0}{4} [X(t) + 2X(t + \delta) + X(t + 2\delta)] \\ &\quad + \left(1 + \frac{\delta\alpha_1}{2} \right) \zeta(0, t + \delta) - \left(1 - \frac{\delta\alpha_1}{2} \right) \zeta(0, t) + \frac{\delta}{2} \zeta(1, t + \delta) + \frac{\delta}{2} \zeta(1, t) \\ &\quad + \left(1 + \frac{\delta\alpha_1}{2} \right) \xi(0, t + \delta) - \left(1 - \frac{\delta\alpha_1}{2} \right) \xi(0, t) + \frac{\delta}{2} \xi(1, t + \delta) + \frac{\delta}{2} \xi(1, t) \end{aligned}$$

Pour tout $k \geq 1$, on définit :

$$\begin{aligned}
A_k^\delta &= \int_{t+\delta}^{t+2\delta} \left[\frac{(t+2\delta-u)^k}{k!} - \frac{\delta(t+2\delta-u)^{k-1}}{2(k-1)!} \right] X(u) du \\
B_k^\delta &= \int_{t+\delta}^{t+2\delta} \left[\frac{(t+2\delta-u)^k}{k!} - \frac{\delta(t+2\delta-u)^{k-1}}{2(k-1)!} \right] X^{(1)}(u) du \\
e_k^\delta &= \int_{t+\delta}^{t+2\delta} \left[\frac{(t+2\delta-u)^k}{k!} - \frac{\delta(t+2\delta-u)^{k-1}}{2(k-1)!} \right] dW(u) \\
C_k^\delta &= \int_t^{t+\delta} \left[\frac{(t-u)^k}{k!} + \frac{\delta(t-u)^{k-1}}{2(k-1)!} \right] X(u) du \\
D_k^\delta &= \int_t^{t+\delta} \left[\frac{(t-u)^k}{k!} + \frac{\delta(t-u)^{k-1}}{2(k-1)!} \right] X^{(1)}(u) du \\
f_k^\delta &= \int_t^{t+\delta} \left[\frac{(t-u)^k}{k!} + \frac{\delta(t-u)^{k-1}}{2(k-1)!} \right] dW(u)
\end{aligned}$$

On note aussi $a_2^\delta = \alpha_0\alpha_1 + \delta\frac{(\alpha_0)^2}{2}$, $b_2^\delta = ((\alpha_1)^2 - \alpha_0) + \delta\frac{\alpha_1\alpha_0}{2}$, $c_2^\delta = -\alpha_0\alpha_1 + \delta\frac{(\alpha_0)^2}{2}$ et $d_2^\delta = -((\alpha_1)^2 - \alpha_0) + \delta\frac{\alpha_1\alpha_0}{2}$.

Pour montrer le résultat annoncé, nous allons développer la quantité :

$$R(t) = \left(1 + \frac{\delta\alpha_1}{2}\right) \zeta(0, t+\delta) - \left(1 - \frac{\delta\alpha_1}{2}\right) \zeta(0, t) + \frac{\delta}{2} \zeta(1, t+\delta) + \frac{\delta}{2} \zeta(1, t)$$

Or, en reprenant les expressions de $\zeta(0, t)$ et $\zeta(1, t)$, on montre que :

$$R(t) = a_2^\delta A_2^\delta + b_2^\delta B_2^\delta + c_2^\delta C_2^\delta + d_2^\delta D_2^\delta$$

On utilise alors le lemme suivant :

Lemme 10 *Pout tout $k \geq 1$, on a :*

$$\begin{aligned}
\mathbf{A}_k^\delta &= \delta^{k+1} \left(\frac{\mathbf{1}}{(k+1)!} - \frac{\mathbf{1}}{2k!} \right) \mathbf{X}(\mathbf{t} + \delta) + \mathbf{B}_{k+1}^\delta \\
\mathbf{B}_k^\delta &= \delta^{k+1} \left(\frac{\mathbf{1}}{(k+1)!} - \frac{\mathbf{1}}{2k!} \right) \mathbf{X}^{(1)}(\mathbf{t} + \delta) - \alpha_0 \mathbf{A}_{k+1}^\delta - \alpha_1 \mathbf{B}_{k+1}^\delta + \mathbf{e}_{k+1}^\delta \\
\mathbf{C}_k^\delta &= (-1)^k \delta^{k+1} \left(\frac{\mathbf{1}}{(k+1)!} - \frac{\mathbf{1}}{2k!} \right) \mathbf{X}(\mathbf{t} + \delta) + \mathbf{D}_{k+1}^\delta \\
\mathbf{D}_k^\delta &= (-1)^k \delta^{k+1} \left(\frac{\mathbf{1}}{(k+1)!} - \frac{\mathbf{1}}{2k!} \right) \mathbf{X}^{(1)}(\mathbf{t} + \delta) - \alpha_0 \mathbf{C}_{k+1}^\delta - \alpha_1 \mathbf{D}_{k+1}^\delta + \mathbf{f}_{k+1}^\delta
\end{aligned}$$

Ce résultat repose sur une application de la formule de Ito.

- Nous commencerons par traiter les intégrales en $(t + \delta, t + 2\delta)$.
Posons : $R_1^1(t) = a_2^\delta A_2^\delta + b_2^\delta B_2^\delta$. D'après le lemme 10, on a :

$$\begin{aligned} R_1^1(t) &= - \left[\frac{a_2^\delta}{12} X(t + \delta) + \frac{b_2^\delta}{12} X^{(1)}(t + \delta) \right] \delta^3 + b_2^\delta e_3^\delta + R_2^1(t) \\ R_2^1(t) &= a_3^\delta A_3^\delta + b_3^\delta B_3^\delta \end{aligned}$$

avec : $a_3^\delta = -\alpha_0 b_2^\delta$ et $b_3^\delta = a_2^\delta - \alpha_1 b_2^\delta$.

On itère le procédé :

$$\begin{aligned} R_2^1(t) &= - \left[\frac{a_3^\delta}{24} X(t + \delta) + \frac{b_3^\delta}{24} X^{(1)}(t + \delta) \right] \delta^4 + b_3^\delta e_4^\delta + R_3^1(t) \\ R_3^1(t) &= a_4^\delta A_4^\delta + b_4^\delta B_4^\delta \end{aligned}$$

avec : $a_4^\delta = -\alpha_0 b_3^\delta$ et $b_4^\delta = a_3^\delta - \alpha_1 b_3^\delta$.

On applique alors une dernière fois le lemme 10 et on obtient :

$$\begin{aligned} R_1^1(t) &= - \left[\frac{a_2^\delta}{12} X(t + \delta) + \frac{b_2^\delta}{12} X^{(1)}(t + \delta) \right] \delta^3 \\ &\quad - \left[\frac{a_3^\delta}{24} X(t + \delta) + \frac{b_3^\delta}{24} X^{(1)}(t + \delta) \right] \delta^4 \\ &\quad - \left[\frac{a_4^\delta}{80} X(t + \delta) + \frac{b_4^\delta}{80} X^{(1)}(t + \delta) \right] \delta^5 \\ &\quad + a_5^\delta A_5^\delta + b_4^\delta B_5^\delta + b_2^\delta e_3^\delta + b_3^\delta e_4^\delta + b_4^\delta e_5^\delta \end{aligned}$$

avec : $a_5^\delta = -\alpha_0 b_4^\delta$ et $b_5^\delta = a_4^\delta - \alpha_1 b_4^\delta$.

- On procède de même pour les intégrales entre $(t, t + \delta)$ et on obtient :

$$\begin{aligned} R_1^2(t) &= c_2^\delta C_2^\delta + d_2^\delta D_2^\delta \\ &= - \left[\frac{c_2^\delta}{12} X(t + \delta) + \frac{d_2^\delta}{12} X^{(1)}(t + \delta) \right] \delta^3 \\ &\quad + \left[\frac{c_3^\delta}{24} X(t + \delta) + \frac{d_3^\delta}{24} X^{(1)}(t + \delta) \right] \delta^4 \\ &\quad - \left[\frac{c_4^\delta}{80} X(t + \delta) + \frac{d_4^\delta}{80} X^{(1)}(t + \delta) \right] \delta^5 \\ &\quad + c_4^\delta C_5^\delta + d_4^\delta D_5^\delta + d_2^\delta f_3^\delta + d_3^\delta f_4^\delta + d_4^\delta f_5^\delta \end{aligned}$$

Les coefficients a^δ , b^δ , c^δ et d^δ peuvent être calculés comme suit :

$$\begin{pmatrix} a_{k+2}^\delta \\ b_{k+2}^\delta \end{pmatrix} = {}^t A^k \begin{pmatrix} a_2^\delta \\ b_2^\delta \end{pmatrix} \quad \begin{pmatrix} c_{k+2}^\delta \\ d_{k+2}^\delta \end{pmatrix} = {}^t A^k \begin{pmatrix} c_2^\delta \\ d_2^\delta \end{pmatrix}$$

A étant la matrice définie en (4.2). En additionnant $R_1^1(t)$ et $R_1^2(t)$ et après simplification, on obtient le résultat annoncé.

I Démonstration du théorème 4 :

Du fait de la stationnarité du processus, on a :

$$\widehat{\gamma}_\delta = {}^t \left(\frac{1}{\delta^2} E_\theta [X(0) \Delta^{(2)} X(4\delta)], \frac{1}{2\delta^3} E_\theta [(X(2\delta) - X(0)) \Delta^{(2)} X(4\delta)] \right)$$

et

$$\widehat{\Gamma}_\delta = \begin{pmatrix} \frac{E_\theta[X(0)(X(4\delta)+2X(3\delta)+X(2\delta))]}{E_\theta[(X(2\delta)-X(0))^4(X(4\delta)+2X(3\delta)+X(2\delta))]} & \frac{E_\theta[X(0)(X(4\delta)-X(2\delta))]}{E_\theta[(X(2\delta)-X(0))^2(X(3\delta)-X(\delta))]} \\ \frac{E_\theta[(X(2\delta)-X(0))^4(X(4\delta)+2X(3\delta)+X(2\delta))]}{8\delta} & \frac{E_\theta[(X(2\delta)-X(0))^2(X(3\delta)-X(\delta))]}{4\delta^2} \end{pmatrix}$$

Nous commencerons par exprimer $\widehat{\Gamma}_\delta$ en fonction de δ en utilisant le développement limité de r à l'ordre 3 en 0^+ .

$$r(h) = r(0) + hr^{(1)}(0) + \frac{h^2}{2}r^{(2)}(0) + \frac{h^3}{6}r^{(3)}(0) + o(h^3)$$

Or, d'après le lemme 3.1 de [9], $r^{(1)}(0) = 0$. On obtient donc :

$$r(h) = r(0) + \frac{h^2}{2}r^{(2)}(0) + \frac{h^3}{6}r^{(3)}(0) + o(h^3)$$

D'où :

$$\begin{aligned} E_\theta [X(0)(X(4\delta) + 2X(3\delta) + X(2\delta))] &= r(4\delta) + 2r(3\delta) + r(2\delta) \\ &= 4r(0) + O(\delta^2) \end{aligned}$$

$$\begin{aligned} E_\theta [X(0)(X(4\delta) - X(2\delta))] &= r(4\delta) - r(2\delta) \\ &= O(h^2) \end{aligned}$$

$$\begin{aligned} E_\theta [(X(2\delta) - X(0))(X(4\delta) + 2X(3\delta) + X(2\delta))] &= -r(4\delta) - 2r(3\delta) + 2r(\delta) + r(0) \\ &= O(\delta^2) \end{aligned}$$

et :

$$\begin{aligned} E_\theta [(X(2\delta) - X(0))(X(3\delta) - X(\delta))] &= -r(4\delta) + 2r(2\delta) - r(0) \\ &= -4\delta^2 r^{(2)}(0) + O(\delta^3) \end{aligned}$$

On en déduit que :

$$\widehat{\Gamma}_\delta = \begin{pmatrix} r(0) & 0 \\ 0 & -r^{(2)}(0) \end{pmatrix} + O(\delta) = \Gamma(0) + O(\delta)$$

$\Gamma(0)$ étant inversible (proposition 3.1 de [9]), il existe $\delta_0 > 0$ tel que pour tout δ , $0 < \delta \leq \delta_0$, $\widehat{\Gamma}_\delta$ est également inversible et donc $\widehat{\alpha}^\delta = -(\widehat{\Gamma}_\delta)^{-1} \widehat{\gamma}_\delta$ existe.

Comme dans la démonstration du théorème 2, l'existence et la convergence de $\widehat{\alpha}^n$ vers $\widehat{\alpha}^\delta$ reposent sur une simple application du lemme 6.

Il nous reste à évaluer l'écart entre $\widehat{\alpha}^\delta$ et α . Utilisant (4.31), on a :

$$\begin{aligned}
& E_\theta \left[X(0) \Delta^{(2)} X(4\delta) \right] \\
&= -\frac{\delta^2}{4} E_\theta [X(0) (X(4\delta) + 2X(3\delta) + X(2\delta))] \alpha_0 - \frac{\delta}{2} E_\theta [X(0) (X(4\delta) - X(2\delta))] \alpha_1 \\
&+ \frac{\delta^4}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] E_\theta [X(0) X(3\delta)] + \frac{\delta^4}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] E_\theta [X(0) X^{(1)}(3\delta)] + O(\delta^5) \\
&= \frac{\delta^4}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] r(0) + \frac{\delta^4}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] r^{(1)}(0) + o(\delta^4) \\
&= \frac{\delta^4}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] r(0) + o(\delta^4)
\end{aligned}$$

et :

$$\begin{aligned}
E_\theta \left[(X(2\delta) - X(0)) \Delta^{(2)} X(4\delta) \right] &= -\frac{\delta^2}{4} E_\theta [(X(2\delta) - X(0)) (X(4\delta) + 2X(3\delta) + X(2\delta))] \alpha_0 \\
&- \frac{\delta}{2} E_\theta [(X(2\delta) - X(0)) (X(4\delta) - X(2\delta))] \alpha_1 \\
&+ \frac{\delta^4}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] E_\theta [(X(2\delta) - X(0)) X(3\delta)] \\
&+ \frac{\delta^4}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] E_\theta [(X(2\delta) - X(0)) X^{(1)}(3\delta)] \\
&+ \frac{\delta^6}{80} [(\alpha_0)^3 - (\alpha_1 \alpha_0)^2] E_\theta [(X(2\delta) - X(0)) X(3\delta)] \\
&+ \frac{\delta^6}{80} [2(\alpha_0)^2 \alpha_1 - (\alpha_1)^2 \alpha_0] E_\theta [(X(2\delta) - X(0)) X^{(1)}(3\delta)] \\
&+ o(\delta^6) \\
&= -\frac{\delta^2}{4} E_\theta [(X(2\delta) - X(0)) (X(4\delta) + 2X(3\delta) + X(2\delta))] \alpha_0 \\
&- \frac{\delta}{2} E_\theta [(X(2\delta) - X(0)) (X(4\delta) - X(2\delta))] \alpha_1 \\
&- \frac{2\delta^5}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] r^{(2)}(0) + o(\delta^5)
\end{aligned}$$

On en déduit que :

$$\widehat{\alpha}^\delta = \alpha - \delta^2 \left(\widehat{\Gamma}_\delta \right)^{-1} \widehat{B}_\theta + o(\delta^2), \quad \widehat{B}_\theta = \begin{pmatrix} \frac{1}{12} [(\alpha_1)^2 \alpha_0 - 2(\alpha_0)^2] r(0) \\ -\frac{1}{12} [(\alpha_1)^3 - 3\alpha_1 \alpha_0] r^{(2)}(0) \end{pmatrix}$$

Utilisant le fait que $\lim_{\delta \rightarrow 0} \widehat{\Gamma}_\delta = \Gamma(0)$, on vérifie que le biais a bien la forme annoncée.

**Estimation de la précision associée
à un échantillonnage systématique
et méthodes transitives**

Chapter 5

Precision of systematic sampling and transitive methods

En collaboration avec
Kiên Kiêu¹ & Jacques Istas²

abstract

The use of the transitive methods for assessing the precision of systematic sampling is discussed. A key point of the transitive methods is the choice of a local model for the covariogram near the origin. The relationship between the regularity of the measurements and the regularity of their covariogram is given. This result is useful for choosing the appropriate covariogram model. A method for estimating the measurement regularity from discrete data is proposed for cases where it cannot be assessed a priori. Stereological applications where sampling is based on geometric probes such as serial sections, point or line grids are also discussed.

AMS Classification: 62G05, 62G20, 62M30, 60D05

Keywords: Spatial statistics; Transitive methods; Stereology; Geometric sampling; Precision assessment; Serial sections; Cavalieri volume estimator;

5.1 Introduction

Systematic sampling is widely used when investigating spatial structures. In this paper we consider the sampling of the structure by means of some probes, e.g. points, transects, quadrats. The parameter of interest can be expressed as an integral of some function, called the measurement function, over the probe space (or sampling space). The sampling design is systematic such that the probes form a regular pattern, e.g. a point lattice. Standard empirical estimators based on systematic samples are unbiased under uniform sampling conditions.

¹Unité de biométrie, INRA, F78026 St-Cyr.

²Unité de biométrie, INRA, F78352 Jouy-en-Josas.

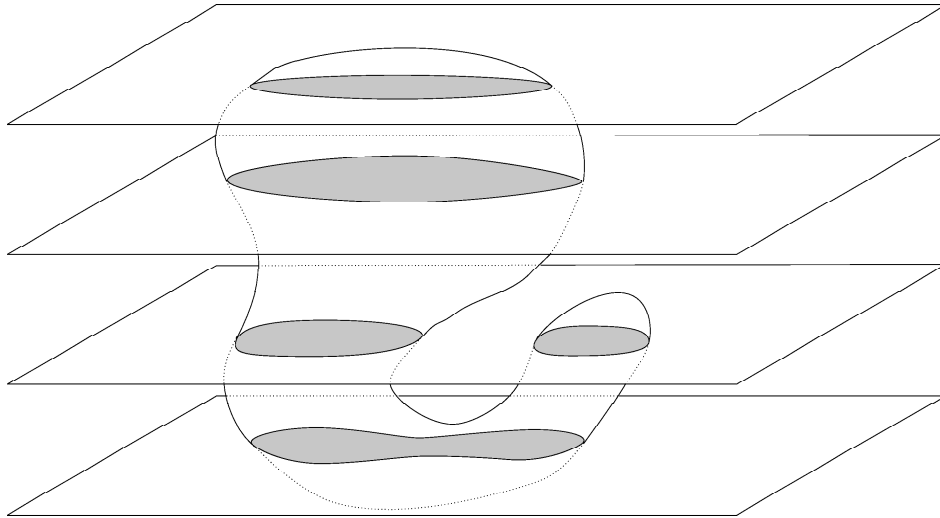


Figure 5.1: Cavalieri volume estimation. Areas of the structure profile on sections are to be measured. The sum of areas on the series of sections times the distance between 2 consecutive sections is the Cavalieri estimate of the volume of the structure.

However precision assessment is usually not an obvious task because of correlations between data. A common approach is to model the correlation structure of the data. The so-called *transitive methods* due to Matheron (1965) are alternative non-parametric and asymptotic methods. They are based on mean squared error (MSE) approximations, involving the behaviour of the covariogram of the measurement function near the origin. In his founding work, Matheron considered mainly applications to mining problems such as the estimation of the total ore quantity from density measurements at a regular grid of points.

More recently the use of transitive methods in stereology has been discussed in a series of papers, see e.g. Thioulouse et al. (1985), Gundersen and Jensen (1987), Cruz-Orive (1989) or Cruz-Orive (1993). Stereology is concerned with statistical inference about quantitative parameters of spatial structures, based on geometric samples such as plane or line sections through the structure. In stereology, systematic sampling is often preferred to simple random sampling (uniformly distributed and independent probes) both for practical and statistical reasons. A simple example is when the parameter Q to be estimated is the volume of a bounded region B of R^3 . The volume of B can be expressed as

$$Q = \int_R \text{area}(B \cap \Lambda(x)) dx,$$

where $\Lambda(x)$ is a plane with fixed orientation and position x , e.g. a horizontal plane with height x . Using a systematic set of parallel planes with constant distance T between consecutive planes, see Figure 5.1, the volume can be estimated by T times the sum of areas measured on the planar sections. This estimator is unbiased if the “first” section to hit the structure B is uniformly distributed in a slice of height T , without any assumption concerning the shape or the orientation of the structure. The volume estimator is known in stereology as the *Cavalieri estimator*, see e.g. Sterio (1984) or Cruz-Orive (1987).

In Gundersen and Jensen (1987), Cruz-Orive (1989) and Cruz-Orive (1993), it is emphasized that systematic sampling is rather efficient in stereology compared to simple random

sampling. In particular, a number of cases is discussed where the MSE rate of convergence is of the order of N^{-2} where N is the sample size. This rate is to be compared to the rate usually obtained under simple random sampling which is N^{-1} .

The standard method of assessing the precision now used in stereology is based on the assumption that the covariogram has non null slope at the origin. In Cruz-Orive (1993), an example is given where the standard assumption does not hold and an alternative method is discussed. In the present paper the relationship between the smoothness of the measurement function and the behaviour of its covariogram at the origin is further investigated. A new method for MSE estimation is presented. In a first step a smoothness parameter of the measurement function is estimated from the data. In a second step a local model for the covariogram near the origin is chosen according to the estimated smoothness parameter. The final MSE estimator is constructed by fitting the model to the data.

As an illustration we use numerical examples of the Cavalieri procedure.

5.2 Unbiased estimator of integral characteristics

We consider the general case where the parameter of interest Q can be written as the integral of some *measurement function* f over a *one-dimensional* Euclidean space, i.e.

$$Q = \int_R f(x) dx. \quad (5.1)$$

The observation is supposed to be of the form $\{(x, f(x)) : x \in \mathbf{S}\}$, \mathbf{S} being a random countable subset³ of R . We shall concentrate on the case where the sample \mathbf{S} is *systematic* and *uniform random*, i.e.

$$\mathbf{S} = \{(\mathbf{U} + k)T : k \in Z\}, \quad T > 0,$$

where \mathbf{U} is uniformly distributed in $[0, 1[$. Below, T is referred to as the *sampling period*, $F = T^{-1}$ as the *sampling frequency*. If the measurement function f has a bounded support, then the support length times F is equal to the *mean sample size* (mean number of non-null measurements) which will be denoted by N .

The standard *unbiased* estimator for Q is

$$\widehat{Q}_T = T \sum_{x \in \mathbf{S}} f(x). \quad (5.2)$$

Note that \widehat{Q}_T can be seen as the standard integral approximation from discrete data. The unbiasedness results from the fact that \mathbf{S} samples R uniformly.

Examples of measurement functions associated with the Cavalieri method are shown in Figure 5.2. The structures behind the shown measurement functions are the unions of 8 tetrahedra with a face parallel to the section planes (Figure 5.2a), 8 ellipsoids (Figure 5.2b), and 8 tetrahedra with neither face nor edge parallel to the section planes (Figure 5.2c), respectively. The sizes and locations of the 8 primary objects are the same through the 3 structures. The jumps observed in Figure 5.2a occur for sections containing a face of the structure.

The standard Euler-MacLaurin formula, see e.g. Abramowitz and Stegun (1964), provides expansions of the error involved in integral discrete approximations. However the formula

³We use bold symbols for random variables.

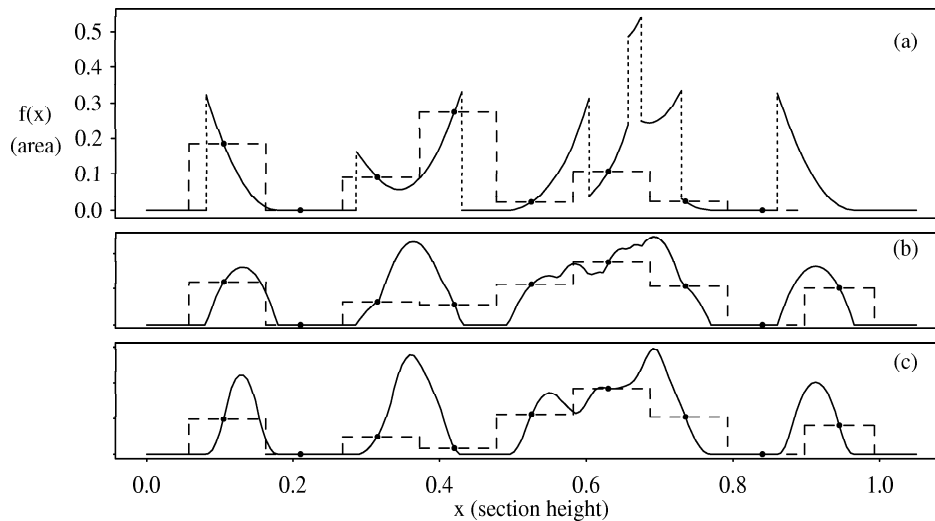


Figure 5.2: Examples of measurement functions of the Cavalieri method. Curves in solid lines are measurement (area measured on a section) functions. Dots are areas measured on a series of sections. Curves in dashed lines are estimated measurement functions based on the serial sections.

cannot be applied when the measurement function is not "smooth" between the sampling points. Such a case may happen for the measurement function shown in Figure 5.2a which is not continuous. As a matter of fact, the randomness of \mathbf{S} implies that almost surely the discontinuities occur between the sampling points. In Ki u (1997) and Souchet (1995), an alternative version of the Euler-MacLaurin formula has been derived. This latter formula can be applied when the measurement function is not "smooth" between the sampling points. The functions considered by the authors are called *piecewise smooth functions* and are defined as follows. Let m and p be non-negative integers. A function $h : R \rightarrow R$ is said to be (m, p) -piecewise smooth if the support of h is bounded and h is $(m + p)$ -times piecewise differentiable such that

- all h derivatives of order less than m are continuous on R
- the h derivatives of order from m to $m + p$ may not be continuous on a finite set, but their jumps are required to be finite.

Below the set of points where a function h is not continuous is denoted by D_h and the jump of h at a point $x \in D_h$ is denoted by $s_h(x)$. Following the terminology used in Ki u (1997), a jump of a l -th derivative is called a *transition of order l* . The primary transitions are the jumps of the first non-continuous derivative. Therefore the smoothness parameter m is called *the primary transition order*.

Observe that the class of piecewise smooth functions as defined above does not encompass unbounded functions or functions with unbounded derivatives. Such measurement functions appear in connection with the Cavalieri method e.g. if the structure is a cylinder with a revolution axis parallel to the section planes. Also very irregular functions such as Brownian

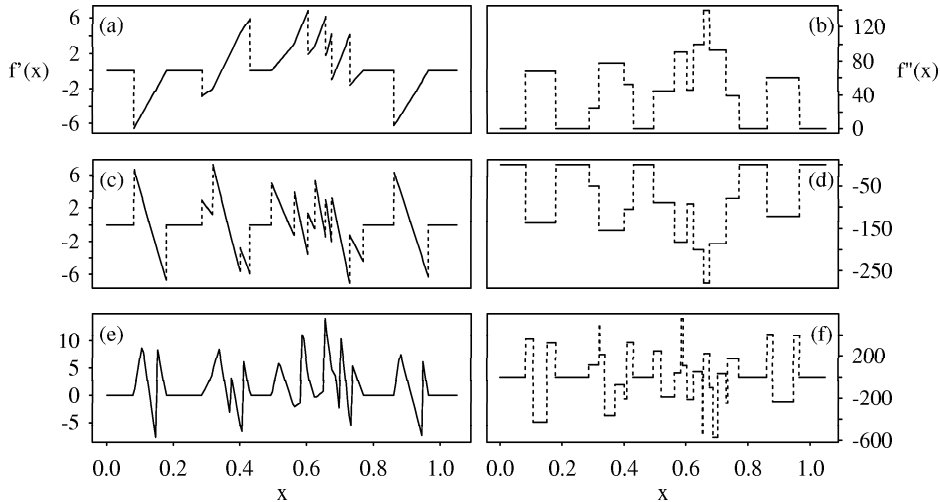


Figure 5.3: Derivatives of order 1 and 2 for the measurement functions shown in Figure 5.2. Derivatives of order larger than 2 are identically equal to 0.

motion are not piecewise smooth. Measurement functions of this type will be the object of future research.

Note that the measurement function shown in

- Figure 5.2a is $(0, \infty)$ -piecewise smooth with transitions of order 0, 1 and 2
- Figure 5.2b is $(1, \infty)$ -piecewise smooth with transitions of order 1 and 2
- Figure 5.2c is $(2, \infty)$ -piecewise smooth with transitions of order 2,

see Figure 5.2. The transitions of the measurement function involved in the Cavalieri method are related to geometrical features of the structure under study. In the examples considered in the figures, transitions of order 0 occur for section planes containing a planar face of the structure. Transitions of order 1 occur for section planes tangent to the structure. Transitions of order 2 occur for sections containing a vertex of the structure.

The alternative Euler-MacLaurin formula for piecewise smooth functions provides expansions for the error involved in integral approximations. These expansions are based on transitions as defined above. They involve the so-called Bernoulli polynomials. The Bernoulli polynomials form a sequence of polynomials indexed by non-negative integers. The polynomial with index l will be denoted by P_l and is of degree l . Explicit expressions and formal definitions can be found in e.g. Abramowitz and Stegun (1964). For any integer $l = 0, 1, 2, \dots$ and $T > 0$, define the periodic function

$$P_{l,T} : x \mapsto P_l \left(\frac{x}{T} - \left[\frac{x}{T} \right] \right)$$

where $[\cdot]$ denotes the integer part. The alternative Euler-MacLaurin formula can now be stated as follows. Let h be a $(m, 1)$ -piecewise smooth function with $m \geq 1$. Then for any

$T > 0$,

$$\begin{aligned} T \sum_{k \in Z} h(kT) - \int_R h(x) dx \\ = (-1)^m T^{m+1} \sum_{a \in D_{h^{(m)}}} s_{h^{(m)}}(a) P_{m+1,T}(a) + o(T^{m+1}). \end{aligned} \quad (5.3)$$

This result holds also for $m = 0$ under the extra condition that $D_h \cap ZT = \emptyset$. Detailed proofs are given in Souchet (1995) and Kiêu (1997). The derivation of the alternative Euler-MacLaurin formula is based on the same arguments as the standard Euler-MacLaurin formula. The integral expression of the remainder is

$$(-1)^m T^{m+1} \int_R h^{(m+1)}(x) P_{m+1,T}(x) dx.$$

Now suppose that the measurement function f is $(m, 1)$ -piecewise smooth. The alternative Euler-MacLaurin formula then provides an expansion of $\widehat{\mathbf{Q}}_T - Q$. Note that if $m = 0$, the extra condition holds almost surely because of the randomness of \mathbf{S} . One gets the following expansion which holds almost surely

$$\widehat{\mathbf{Q}}_T - Q = (-1)^m T^{m+1} \sum_{a \in D_{f^{(m)}}} s_{f^{(m)}}(a) P_{m+1,T}(a - \mathbf{U}T) + o(T^{m+1}). \quad (5.4)$$

Note that the remainder is stochastic.

Hence the worst rate of convergence is of the order of T and is reached when the measurement function f is not continuous on R . If f is continuous on R but its first derivative is not then the estimation error converges as T^2 etc... The convergence rate can also be expressed in terms of the mean sample size N . The worst rate of convergence is of the order of N^{-1} and is reached when the measurement function f is not continuous on R . If f is continuous on R but its first derivative is not then the estimation error converges as N^{-2} etc... Compare with simple random sampling where the convergence rate does not depend on the regularity of the measurement function and is always of the order of $N^{-1/2}$. The efficiency of systematic sampling compared to simple random sampling was already noticed in Gundersen and Jensen (1987) and Cruz-Orive (1989). These authors have considered mean squared error convergence rather than estimation error convergence.

5.3 Mean squared error expansion

The formula (5.4) can be used for approximating the MSE of $\widehat{\mathbf{Q}}_T$. Repeated integration by part yields

$$\mathbb{E}[P_{m+1,T}(a - \mathbf{U}T) P_{m+1,T}(b - \mathbf{U}T)] = (-1)^m P_{2m+2,T}(b - a).$$

Hence we get the following mean squared error expansion for an $(m, 1)$ -piecewise smooth function

$$\begin{aligned} \text{MSE}[\widehat{\mathbf{Q}}_T] \\ = (-1)^m T^{2m+2} \sum_{a, b \in D_{f^{(m)}}} s_{f^{(m)}}(a) s_{f^{(m)}}(b) P_{2m+2,T}(b - a) + o(T^{2m+2}). \end{aligned} \quad (5.5)$$

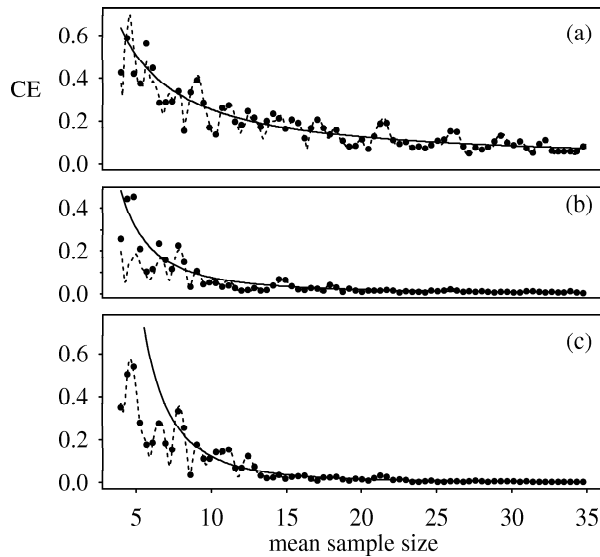


Figure 5.4: Coefficients of error for Cavalieri volume estimators. Exact coefficients of error (CE) calculated from simulations are represented by dots. Approximated CE based on primary transitions are represented by dashed line curves, cf. (5.5). Approximated CE based on primary transition amplitudes (extension term) are represented by solid line curves, cf. (5.11). Underlying structures are described in Section 5.2.

This expansion has been computed for the examples of measurement functions described in the previous section, see Figure 5.3 (dashed curves). Note that the convergence rate of the mean squared error is of the order of $T^{2m+2} \propto N^{-2m-2}$.

In Matheron (1965), Gundersen and Jensen (1987) and Cruz-Orive (1989), expansions of the mean squared error involve the so-called covariogram of the measurement function. The covariogram is defined by Matheron (1965)

$$g(y) = \int_R f(x+y) f(x) dx, \quad y \in R. \quad (5.6)$$

The covariogram describes the spatial variation of f . Note that the covariogram is an even function. Covariograms of the measurement functions shown in Figure 5.2 are presented in Figure 5.5.

It is easy to see that

$$\text{MSE} \left[\widehat{Q}_T \right] = T \sum_{k \in \mathbb{Z}} g(kT) - \int_R g(y) dy. \quad (5.7)$$

In order to calculate expansions for (5.7), Matheron assumed that the measurement function f has a bounded support $[a, b]$ and that its covariogram is non-smooth only at 0, $c = b - a$ and $-c$ (c and $-c$ are the endpoints of the covariogram's support). Expansions of (5.7) are obtained by applying the standard Euler-MacLaurin formula inside the covariogram's support and by modelling the covariogram at the neighbourhood of 0, c and $-c$ by polynomials. A more direct approach consists in applying the alternative Euler-MacLaurin formula. The

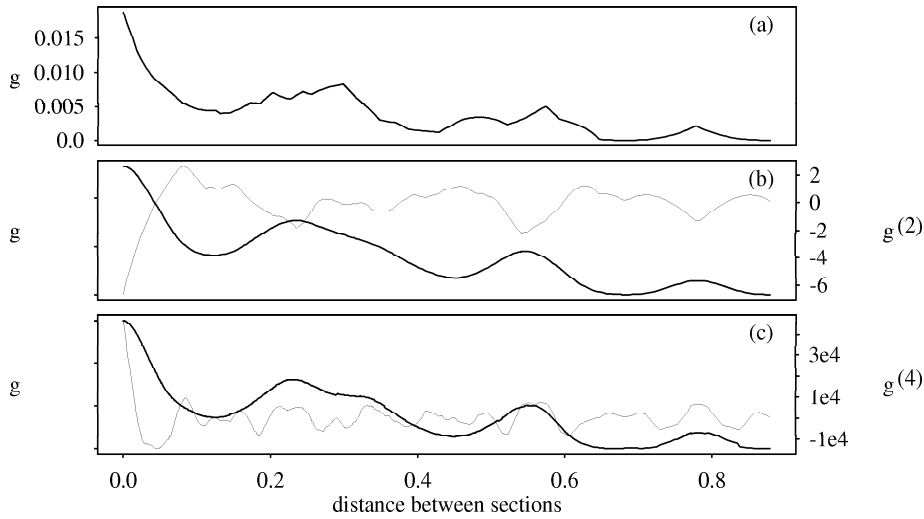


Figure 5.5: Covariograms (thick line curves) and their last continuous derivatives (thin line curves in (b) and (c)). The corresponding measurement functions are shown in Figure 5.2.

latter method provides expansions even if the covariogram is non-smooth inside its support. In view of (5.3) if the covariogram is $(q, 1)$ -piecewise smooth, (5.7) can be expanded as

$$\text{MSE} \left[\widehat{\mathbf{Q}}_T \right] = (-1)^q T^{q+1} \sum_{c \in \mathcal{D}_{g^{(q)}}} s_{g^{(q)}}(c) P_{q+1, T}(c) + o(T^{q+1}). \quad (5.8)$$

Compare the result above with (5.5). It is tempting to search for a relationship between the smoothness properties of a measurement function and the smoothness properties of its covariogram. In his paper, Matheron (1965) stated that if the measurement function is differentiable then its covariogram is twice differentiable. Furthermore, Matheron gives an expression of the covariogram second derivative in terms of the first derivative of the measurement function. In Souchet (1995) and Kiêu (1997), it is shown that if f is a (m, p) -piecewise smooth function then its covariogram g is $(2m + 1, 2p - 1)$ -piecewise smooth, see examples in Figure 5.5. The authors also derived general expressions of the covariogram derivative in terms of the derivatives of f . In particular it is shown that

$$D_{g^{(2m+1)}} = D_{f^{(m)}} - D_{f^{(m)}} \quad (5.9)$$

and that

$$s_{g^{(2m+1)}}(c) = -(-1)^m \sum_{\substack{a, b \in \mathcal{D}_{f^{(m)}} \\ b-a=c}} s_{f^{(m)}}(a) s_{f^{(m)}}(b). \quad (5.10)$$

In Matheron (1965), the quantity

$$\begin{aligned} E(T) &= -T^{2m+2} P_{2m+2}(0) s_{g^{(2m+1)}}(0) \\ &= -2T^{2m+2} P_{2m+2}(0) g^{(2m+1)}(0^+) \end{aligned} \quad (5.11)$$

is called the *extension term*. Note that the extension term depends only on the amplitudes of the transitions of the measurement function. The sum of the other terms involved in the mean squared error expansions is equal to

$$Z(T) = -T^{2m+2} \sum_{\substack{c \in D_{g^{(2m+1)}} \\ c \neq 0}} P_{2m+2,T}(c) s_{g^{(2m+1)}}(c)$$

and is called the *Zitterbewegung*. The extension term represents the central tendency of the mean squared error while the Zitterbewegung is an oscillating function of T . A detailed discussion can be found in Kiêu (1997), see also the coefficient of error approximations based only on the extension term in Figure 5.3 (solid line curves). In Matheron (1965), Gundersen and Jensen (1987) and Cruz-Orive (1989), the extension term is expressed as

$$E(T) = -T^{2m+2} \frac{B_{2m+2}}{m+1} \beta_{2m+1}$$

where $B_{2m+2} = (2m+2)!P_{2m+2}(0)$ is the $(2m+2)$ -th Bernoulli number and

$$\beta_{2m+1} = \frac{g^{(2m+1)}(0^+)}{(2m+1)!}.$$

In particular

$$E(T) = \begin{cases} -T^2 \beta_1 / 6 & \text{if } m = 0 \\ T^4 \beta_3 / 60 & \text{if } m = 1 \\ -T^6 \beta_5 / 126 & \text{if } m = 2 \\ T^8 \beta_7 / 120 & \text{if } m = 3. \end{cases}$$

Usually the mean squared error estimation focus on the extension term, see Matheron (1965), Gundersen and Jensen (1987) or Cruz-Orive (1989). These authors argue that when the measurement function is stochastic, the Zitterbewegung is expected to be small in the mean compared to the extension term. In Kiêu (1997) the particular case where the measurement function is the indicator function of an interval with random length is considered in details. It is shown that if the interval length density function is Riemann integrable then the mean Zitterbewegung can be neglected when T is sufficiently small.

5.4 Discrete approximations of the extension term

The extension term as expressed in (5.11) depends on the covariogram derivatives at the origin. However in general estimates of the covariogram are available only on a discrete set of values. By approximating the covariogram by a polynomial near the origin, the covariogram derivatives at the origin can be interpolated from discrete covariogram data. This yields approximations of the MSE extension term based on discrete covariogram data.

Below we review some of the methods discussed in Matheron (1965), Gundersen and Jensen (1987), Cruz-Orive (1989) and Cruz-Orive (1993). A more general presentation can be found in Kiêu (1997). An important feature of these methods is that the primary transition order of the measurement function (or equivalently of its covariogram) is assessed a priori.

Let us start with the case where the covariogram g of the measurement function f is assessed to be $(1, 2)$ -piecewise smooth. We have seen above that such a covariogram can be

obtained from a $(0, 2)$ -piecewise smooth measurement function⁴. Then by a Taylor expansion in a right-hand side neighbourhood of 0, we get

$$g(y) = \beta_0 + y\beta_1 + y^2\beta_2 + O(y^3), \quad (5.12)$$

where $\beta_1 = g'(0^+)$. This expansion holds for y "small enough". Now if T is "small enough", we can apply the above expansion for $y = 0, T, 2T$ in order to get an approximation of β_1 based on $g(0), g(T)$ and $g(2T)$. Finally one finds

$$E(T) \simeq E_0(T) = \frac{T}{12} (3g(0) - 4g(T) + g(2T)). \quad (5.13)$$

This approximation is suggested in Gundersen and Jensen (1987) and Cruz-Orive (1989) for use in geometric sampling including precision assessment of the Cavalieri method. The approximation error is of order T^4 for T "small enough". As a matter of fact this order holds if T is so small that g is 3 times continuously differentiable in $]0, 2T[$. In particular, this latter condition does not hold if

$$D_{g'} \cap]0, 2T[\neq \emptyset,$$

i.e. there are any transitions of the measurement function with order 0 (jumps) separated by a distance less than $2T$. In Ki u (1997) an alternative Taylor formula is derived for piecewise smooth functions. In this case it can be written as

$$g(y) = \beta_0 + y[\beta_1 + \sum_{\substack{c \in D_{g'} \\ 0 < c < y}} (1 - \frac{c}{y})s_{g'}(c)] + y^2\beta_2 + O(y^3). \quad (5.14)$$

Using the latter formula rather than (5.12) when evaluating the error approximation involved in (5.13), one finds

$$\begin{aligned} E_0(T) - E(T) &= \frac{T^2}{3} \left[\sum_{\substack{c \in D_{g'} \\ 0 < c < T}} (1 - \frac{c}{T})s_{g'}(c) - \sum_{\substack{c \in D_{g'} \\ T \leq c < 2T}} (1 - \frac{c}{T})s_{g'}(c) \right] + O(T^3). \end{aligned}$$

According to this result, the approximation error in (5.13) is of the order of T^2 if there are any transitions of the measurement function with order 0 separated by a distance less than $2T$. Note that then the approximation error is of the same order as the extension term.

The discrete approximation of the extension term (5.13) can be extended to piecewise smooth functions with arbitrary primary transition order. For instance if the measurement function is $(1, 1)$ -piecewise smooth, then for positive y "small enough", the following expansion of its covariogram holds

$$g(y) = \beta_0 + y^2\beta_2 + y^3\beta_3 + O(y^4),$$

where $\beta_3 = g^{(3)}(0^+)/6$. There is no linear term in the expansion above because the continuity of the first derivative of the covariogram implies that $g'(0) = 0$. If T is "small enough", we

⁴As stated in the previous section, if a function is $(0, 2)$ -piecewise smooth then its covariogram is $(1, 3)$ -piecewise smooth. Observe that $(1, 3)$ -piecewise smoothness is stronger than $(1, 2)$ -piecewise smoothness.

can apply the above expansion for $y = 0, T, 2T$ in order to get an approximation of β_3 based on $g(0), g(T)$ and $g(2T)$. Finally one finds

$$E(T) \simeq E_1(T) = \frac{T}{240} (3g(0) - 4g(T) + g(2T)). \quad (5.15)$$

This approximation was first proposed by Cruz-Orive (1993) when assessing the precision of the Cavalieri method for "very regular, quasi-ellipsoidal" structures. As a matter of fact it may be used for any (1,1)-piecewise smooth measurement function. For a (2,1)-piecewise smooth measurement function, one gets the approximation

$$E(T) \simeq E_2(T) = \frac{T}{8316} (10g(0) - 15g(T) + 6g(2T) - g(3T)).$$

For a (3,1)-piecewise smooth measurement function, one gets the approximation

$$E(T) \simeq E_3(T) = \frac{T}{289920} (35g(0) - 56g(T) + 28g(2T) - 8g(3T) + g(4T)).$$

A general formula for the discrete approximation of $E(T)$ for a (m, p) -piecewise smooth function is given in Kiêu (1997) together with a general discussion of the approximation error.

5.5 Extension term estimation

When the primary transition order of the measurement function is assessed a priori, the methods for estimating the extension term are straightforward. They consist in replacing the covariogram values involved in the discrete approximations of the extension term by estimates. The *empirical covariogram* is the sequence indexed by Z defined by

$$\hat{\mathbf{g}}_l = T \sum_{k \in Z} f((\mathbf{U} + k)T) f((\mathbf{U} + k + l)T), \quad l \in Z.$$

The extension term estimators are as follows

$$\text{if } m = 0, \quad \hat{\mathbf{E}}_0(T) = \frac{T}{12} (3\hat{\mathbf{g}}_0 - 4\hat{\mathbf{g}}_1 + \hat{\mathbf{g}}_2) \quad (5.16)$$

$$\text{if } m = 1, \quad \hat{\mathbf{E}}_1(T) = \frac{T}{240} (3\hat{\mathbf{g}}_0 - 4\hat{\mathbf{g}}_1 + \hat{\mathbf{g}}_2) \quad (5.17)$$

$$\text{if } m = 2, \quad \hat{\mathbf{E}}_2(T) = \frac{T}{8316} (10\hat{\mathbf{g}}_0 - 15\hat{\mathbf{g}}_1 + 6\hat{\mathbf{g}}_2 - \hat{\mathbf{g}}_3) \quad (5.18)$$

$$\text{if } m = 3, \quad \hat{\mathbf{E}}_3(T) = \frac{T}{289920} (35\hat{\mathbf{g}}_0 - 56\hat{\mathbf{g}}_1 + 28\hat{\mathbf{g}}_2 - 8\hat{\mathbf{g}}_3 + \hat{\mathbf{g}}_4) \quad (5.19)$$

These estimators are asymptotically unbiased. For instance consider the estimator $\hat{\mathbf{E}}_0(T)$. Since the empirical covariogram is unbiased, we have

$$\mathbb{E} \left[\hat{\mathbf{E}}_0(T) \right] = E_0(T).$$

We have seen in the previous sections that $E_0(T)$ differs from $E(T)$ only by an order of T^4 for T small enough and that $E(T)$ is of the order of T^2 . Accordingly we have

$$\lim_{T \rightarrow 0} \frac{\mathbb{E} \left[\hat{\mathbf{E}}_0(T) - E(T) \right]}{E(T)} = 0.$$

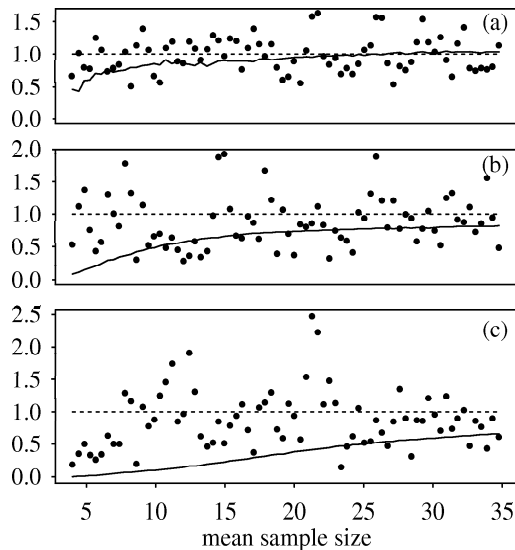


Figure 5.6: Bias of the extension term estimators. For consistency with Figure 5.3 showing coefficients of errors, the bias is measured by $\overline{\mathbf{E}}(T)^{\frac{1}{2}}/E(T)^{\frac{1}{2}}$ (solid line curve) where $\overline{\mathbf{E}}(T)$ is the average extension term estimator computed from simulations. The departures (due to the Zitterbewegung) of the exact coefficient of error from the extension term are measured by $\text{CE}[\hat{\mathbf{Q}}_T]/E(T)^{\frac{1}{2}}$ (dots). Corresponding measurement functions are shown in Figure 5.2.

This result extends to the other estimators above. Figure 5.5 shows biases for the examples of measurement functions presented in Section 5.2. The estimator $\hat{\mathbf{E}}_0(T)$ has been used for Figure 5.5a, estimator $\hat{\mathbf{E}}_1(T)$ for Figure 5.5b, estimator $\hat{\mathbf{E}}_2(T)$ for Figure 5.5c. Note that the bias is rather important for small mean sample sizes.

The critical point of the methods above concerns the assessment of the primary transition order m . As noticed by Cruz-Orive (1993) when using the estimator $\hat{\mathbf{E}}_0(T)$ while $m = 1$, one gets in the mean an extension term estimator 20 times greater than the true extension term! Conversely if $m = 0$ and estimator $\hat{\mathbf{E}}_1(T)$ is used, one gets in the mean an extension term estimator 20 times less than the true extension term. Assessing the primary transition order of a measurement function is subject to uncertainty when measurements are "complicated". Even for the Cavalieri method, there are no general results relating the smoothness properties of the measurement function to geometrical characteristics of the underlying structure.

5.6 Robust extension term estimation

In many stereological contexts, *a priori* assessment of the measurement function smoothness is not possible. Either because very little is known *a priori* about the geometry of the structures under study or because the measurement function is so complicated that predicting its smoothness is out of reach. In Istas and Lang (1995) a method for estimating the smoothness of a function from discrete data is presented. The functions considered by the authors are Gaussian processes. However the method described in Istas and Lang (1995) can be used without modification to piecewise smooth functions. The estimation of the extension term

is a two-step procedure: first the primary transition order of the measurement function is estimated from the data, then the extension term is estimated as described in the previous section assuming that the estimated primary transition order is correct.

Let us start with the case where m is not known exactly but can be assumed to be either 0 or 1. Let us consider the estimator $\widehat{\mathbf{E}}_1(T)$. We have

$$\mathbb{E} \left[\widehat{\mathbf{E}}_1(T) \right] \simeq \begin{cases} E(T)/20 = -T^2\beta_1/120 & \text{if } m = 0 \\ E(T) = T^4\beta_3/60 & \text{if } m = 1. \end{cases}$$

Now suppose that 2 measurement series are available. The sampling period for the first series is T , for the second series αT , $\alpha \neq 1$. In view of the formula given just above, we have

$$\mathbb{E} \left[\widehat{\mathbf{E}}_1(\alpha T) \right] \simeq \alpha^{2m+2} \mathbb{E} \left[\widehat{\mathbf{E}}_1(T) \right]$$

in both cases $m = 0$ or $m = 1$. This relation suggests the following procedure for estimating m from 2 measurement series. Consider the extension term estimators $\widehat{\mathbf{E}}_1(T)$ and $\widehat{\mathbf{E}}_1(\alpha T)$ associated with the 2 measurement series and estimate the primary transition order m by $\widehat{\mathbf{m}}$ such that

$$\widehat{\mathbf{E}}_1(\alpha T) \simeq \alpha^{2\widehat{\mathbf{m}}+2} \widehat{\mathbf{E}}_1(T). \quad (5.20)$$

In case where only one measurement series is available, measures with another sampling period can be obtained by subsampling. In particular for $\alpha = 2$, 2 series with sampling period $2T$ can be extracted from the initial series:

$$\mathbf{S}_o = \{(\mathbf{U} + k)T : k \text{ odd}\}, \quad \mathbf{S}_e = \{(\mathbf{U} + k)T : k \text{ even}\}.$$

Let $\widehat{\mathbf{E}}_{1,o}(2T)$, $\widehat{\mathbf{E}}_{1,e}(2T)$ be the extension term estimators based on the series \mathbf{S}_o , \mathbf{S}_e respectively. The mean of $\widehat{\mathbf{E}}_1(2T)$ can be estimated by the average of $\widehat{\mathbf{E}}_{1,o}(2T)$ and $\widehat{\mathbf{E}}_{1,e}(2T)$. Then the estimation equation (5.20) writes

$$\frac{\widehat{\mathbf{E}}_{1,o}(2T) + \widehat{\mathbf{E}}_{1,e}(2T)}{2} \simeq 2^{2\widehat{\mathbf{m}}+2} \widehat{\mathbf{E}}_1(T).$$

It is easy to see that this latter equation can be rewritten as

$$3\widehat{\mathbf{g}}_0 - 4\widehat{\mathbf{g}}_2 + \widehat{\mathbf{g}}_4 \simeq 2^{2\widehat{\mathbf{m}}+1} (3\widehat{\mathbf{g}}_0 - 4\widehat{\mathbf{g}}_1 + \widehat{\mathbf{g}}_2),$$

i.e.

$$\widehat{\mathbf{m}} = \text{round} \left(\frac{\log \frac{3\widehat{\mathbf{g}}_0 - 4\widehat{\mathbf{g}}_2 + \widehat{\mathbf{g}}_4}{3\widehat{\mathbf{g}}_0 - 4\widehat{\mathbf{g}}_1 + \widehat{\mathbf{g}}_2}}{2 \log 2} - \frac{1}{2} \right).$$

This principle can be extended to the general case where the primary transition order m is not known exactly but can be assumed to be less than or equal to some value M assessed a priori. The key point is that the extension term estimator obtained when assuming that $m = M$ has a mean equal to

$$c_m E(T) \simeq c_m T^{2m+2} \beta_{2m+1},$$

where m is the true primary transition order and c_m is a constant depending only on m . The relation above holds only if $m \leq M$. The calculation of the constants c_m is straightforward.

For instance let us consider the extension term estimator $\widehat{\mathbf{E}}_3(T)$ obtained by assuming that $m = 3$. If the true value of m is 2, the expansion of the covariogram writes

$$g(iT) \simeq \beta_0 + i^2 T^2 \beta_2 + i^4 T^4 \beta_4 + i^5 T^5 \beta_5 + i^6 T^6 \beta_6 + i^7 T^7 \beta_7,$$

for $i = 0, \dots, 4$. Calculations yield

$$\begin{aligned} \mathbb{E} \left[\widehat{\mathbf{E}}_3(T) \right] = E_3(T) &\simeq -\frac{2}{723} T^6 \beta_5 \\ &\simeq \frac{91100}{723} E(T), \end{aligned}$$

i.e. $c_2 = 91100/723 \simeq 126.0$. Extending the approach described above for the case $m = 0$ or 1, one constructs general estimators of m when it is assumed that $m \leq M$, M being fixed a priori. In particular, if m is assumed to be less than or equal to 3, then m is estimated by

$$\widehat{\mathbf{m}} = \text{round} \left(\frac{\log \frac{35\widehat{\mathbf{g}}_0 - 56\widehat{\mathbf{g}}_2 + 28\widehat{\mathbf{g}}_4 - 8\widehat{\mathbf{g}}_6 + \widehat{\mathbf{g}}_8}{35\widehat{\mathbf{g}}_0 - 56\widehat{\mathbf{g}}_1 + 28\widehat{\mathbf{g}}_2 - 8\widehat{\mathbf{g}}_3 + \widehat{\mathbf{g}}_4}}{2 \log 2} - \frac{1}{2} \right). \quad (5.21)$$

Once the transition primary order m is estimated, the extension term is estimated as described in the previous section assuming that $m = \widehat{\mathbf{m}}$. Note that this two-step method is quite unstable. For instance if $\widehat{\mathbf{m}} = 1$, one finds an extension term estimate 20 times less than if $\widehat{\mathbf{m}} = 0$. Therefore a sensible way of measuring the accuracy of the estimator $\widehat{\mathbf{m}}$ is to consider the probability that $\widehat{\mathbf{m}}$ is equal to the true primary transition order. Simulations have been carried out with the measurement functions shown in Figure 5.2. The estimator (5.21) has been used: we have assumed that $m \leq 3$ while the true primary transition order is 0 (Figure 5.2a), 1 (Figure 5.2b), 2 (Figure 5.2c). For mean sample sizes larger than 32.6 (measurement function of Figure 5.2a), 19.6 (measurement function of Figure 5.2b), 47.0 (measurement function of Figure 5.2c), the probability that $\widehat{\mathbf{m}} = m$ is close to 1.

5.7 Discussion

In this paper the use of the transitive methods for evaluating the precision of estimators based on systematic sampling designs is considered. The problem of choosing a local model of the covariogram at the origin is discussed in details. The described relationship between the smoothness of the measurement function and the smoothness of its covariogram seems to be a useful tool for choosing the appropriate model. For applications in stereology, an important open problem is now to relate geometrical features of the structure of interest to smoothness characteristics of the considered measurement function.

A method for estimating the smoothness of the measurement function from discrete data is proposed. The numerical studies of this method show that it requires rather large samples compared to the usual size of samples in practical stereology. Note that in Ki eu (1997), an alternative method is proposed which is expected to be less demanding in terms of sample size. The main idea is to use grids of points with 2 sampling periods, a large one and a small one. Pairs of points close to each other are used to estimate the primary transition order. Numerical investigations of this alternative methods are planned.

In this paper, we have not considered errors either on measurements or on the location of the sampling points. In Kiêu (1997), measurement errors are also considered. The error scheme is as follows. It is assumed that the errors are centered, that their variance only depends on the measurement location and that they are uncorrelated given the location of the sampling points. Such a model is appropriate when the measurement at a given location involve some local subsampling. A common example related to the Cavalieri method is when areas on sections are estimated by countings at point grids superimposed on the sections.

Finally note that although we have concentrated in this paper on the case where the sampling space is one-dimensional, the transitive methods can also be used for sampling spaces with higher dimensions, see e.g. Matheron (1965), Gundersen and Jensen (1987), Cruz-Orive (1989) or Chadœuf et al. (1997). Extending the approach described in this paper to higher-dimensional sampling spaces will be the object of future research. As a matter of fact, in stereology sampling spaces may also be special kinds of manifolds. For instance some stereological methods, see e.g. Weibel (1979) or Weibel (1980), involve sampling by lines with various locations and orientations. The space of lines can be seen as a cylinder, see e.g. Stoyan et al. (1987). Therefore using a grid of lines with systematically distributed locations and orientations may be seen as systematic sampling on a cylinder.

Acknowledgements

The interest that Eva B. Vedel Jensen and Hans Jørgen G. Gundersen took in this work has been an encouraging support. The authors are also indebted to Eva B. Vedel Jensen for reading and improving the manuscript.

References

- Abramowitz, M. & Stegun, I. A. (1964) *Handbook of Mathematical Functions*. Dover Publications, New-York.
- Chadœuf, J., Moran, C. & Goulard, M. (1997) A note on extension variances in R^2 . To appear in *Math. geol.*
- Cruz-Orive, L. M. (1987) Stereology: historical notes and recent evolution. In: *The commemorative-Memorial Volume. Twenty-Five Years of Stereology* (ed. by R.E. Miles). *Acta Stereol.*, **6**, Suppl. II, 43-56.
- Cruz-Orive, L. M. (1989) On the precision of systematic sampling: a review of Matheron's transitive methods. *J. Microsc.*, **153**, 3, 315-33.
- Cruz-Orive, L. M. (1993) Systematic sampling in stereology. *Bulletin de l'institut international de statistique*, Actes de la 47ième session, livraison **2**, 451-468.
- Gundersen, H. J. G. & Jensen, E. B. (1987) The efficiency of systematic sampling in stereology and its prediction. *J. Microsc.*, **147**, 3, 229-263.
- Istas, J. & Lang, G. (1997) Quadratic variations and estimation of the local Hölder index of a Gaussian process. To appear in *Annales de l'institut Henri Poincaré*.

- Kiêu, K. (1997) Three lectures on systematic geometric sampling. Memoirs **13**, Department of Theoretical Statistics, University of Aarhus.
- Matheron, G. (1965) *Les variables régionalisées et leur estimation*. Masson, Paris.
- Sandau, K. (1987) How to estimate the area of a surface using the spatial grid. *Acta Stereol.*, **6/III**, 31-36.
- Santaló, L. A. (1976) *Integral Geometry and Geometric Probability*. Encyclopedia of Mathematics and Its Applications. Addison-Wesley, Reading.
- Souchet, S. (1995) Précision de l'estimateur de Cavalieri. Rapport de stage, DEA de statistiques et modèles aléatoires appliqués à la finance, Université Paris-VII. Laboratoire de biométrie, INRA-Versailles.
- Sterio, D.C. (1984) The unbiased estimation of number and sizes of arbitrary particles using the disector. *J. Microsc.*, **134**, 127-136.
- Stoyan, D., Kendall, W.S. & Mecke, J. (1987) *Stochastic Geometry and Its Applications*. Wiley & Sons, Chichester.
- Thioulouse, J., Mathy, B. & Ploye, H. (1985) Estimation d'un volume à partir de coupes sériées : sous-échantillonnage, covariogramme transitif et calcul de précision. *Mikroskopie*, **42**, 215-224.
- Weibel, E. R. (1979) *Stereological Methods, volume 1, Practical Methods for Biological Morphometry*. Academic Press, London.
- Weibel, E. R. (1980) *Stereological Methods, volume 2, Theoretical Foundations*. Academic Press, London.

List of Tables

1.1	Modèle de O.U. : comparaison des valeurs θ_δ pour les quatre schémas.	40
1.2	Modèle de O.U. : comparaison des moyennes empiriques pour 100 réalisations.	40
1.3	Modèle de O.U. : comparaison des écarts quadratiques moyens pour 100 réalisations.	40
1.4	Modèle de C.I.R. : comparaison des valeurs (a_δ, b_δ) pour les quatre schémas.	42
1.5	Modèle de C.I.R. : comparaison des moyennes empiriques pour 100 réalisations.	42
1.6	Modèle de C.I.R. : comparaison des écarts quadratiques moyens pour le paramètre a_0 pour 100 réalisations.	42
1.7	Modèle de C.I.R. : comparaison des écarts quadratiques moyens pour le paramètre b_0 pour 100 réalisations.	42
1.8	Diffusion Log-Normale : comparaison des valeurs θ_δ pour les quatre schémas.	44
1.9	Diffusion Log-Normale : comparaison des moyennes empiriques pour 100 réalisations.	44
1.10	Diffusion Log-Normale : comparaison des écarts quadratiques moyens pour 100 réalisations.	44
2.1	Processus de O.U. et approximation à l'ordre 1 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.	78
2.2	Processus de O.U. et approximation à l'ordre 2 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.. Pour $\delta = 0.5$, nous avons obtenu 987 réalisations de $\widehat{\theta}_{p,n}^\delta$; dans les autres cas, les réalisations de S_n sont en dehors de $[-\frac{1}{2}, 0]$	78
2.3	Processus de O.U. et approximation à l'ordre 3 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.	78
2.4	Processus de O.U. et approximation à l'ordre ∞ : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.	78
2.5	Processus de C.I.R. et approximation à l'ordre 1 : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.	80
2.6	Processus de C.I.R. et approximation à l'ordre 2 : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.	80
2.7	Processus de C.I.R. et approximation à l'ordre 3 : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.	80
2.8	Processus de C.I.R. et approximation à l'ordre ∞ : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.	80
3.1	Moyenne et variance empiriques du paramètre de seuil pour $N = 100$ lorsque $a_1 = 1$ et $a_2 = 2$ sont connus.	112
3.2	Moyennes empiriques des estimateurs de a_1, a_2 et r_0 pour $N = 100$ et $a_1 = 1, a_2 = 2$ et $r_0 = 0$	113

3.3	Résultats des effectifs par classe relatifs au test du Khi-Deux pour le paramètre de seuil lorsque $a_1 = 1$ et $a_2 = 2$ ne sont pas estimés.	113
4.1	CAR(2) : comparaison des moyennes empiriques pour 100 estimateurs de α lorsque ${}^t\alpha = (2, 3)$	148
4.2	CAR(2) : moyennes empiriques pour 100 estimateurs de σ^2 lorsque $\sigma^2 = 1$. . .	148
4.3	CAR(2) : comparaison des biais réels et des biais théoriques pour α et σ^2 . . .	149
4.4	Quelques valeurs numériques de la fonction de covariance r	149

List of Figures

3.1	Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{r}_n^* - r_0)$ lorsque $a_1 = 1$ et $a_2 = 2$ sont supposés connus.	111
3.2	Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{r}_n^* - r_0)$ lorsque $a_1 = 1$ et $a_2 = 2$ sont estimés.	111
3.3	Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{a}_{1n} - a_1)$ lorsque $a_1 = 1$ et $a_2 = 2$. . .	111
3.4	Histogramme des estimateurs $(nh_n)^{\frac{1}{2}} (\hat{a}_{2n} - a_2)$ lorsque $a_1 = 1$ et $a_2 = 2$. . .	112
5.1	Cavalieri volume estimation. Areas of the structure profile on sections are to be measured. The sum of areas on the series of sections times the distance between 2 consecutive sections is the Cavalieri estimate of the volume of the structure.	176
5.2	Examples of measurement functions of the Cavalieri method. Curves in solid lines are measurement (area measured on a section) functions. Dots are areas measured on a series of sections. Curves in dashed lines are estimated measurement functions based on the serial sections.	178
5.3	Derivatives of order 1 and 2 for the measurement functions shown in Figure 5.2. Derivatives of order larger than 2 are identically equal to 0.	179
5.4	Coefficients of error for Cavalieri volume estimators. Exact coefficients of error (CE) calculated from simulations are represented by dots. Approximated CE based on primary transitions are represented by dashed line curves, cf. (5.5). Approximated CE based on primary transition amplitudes (extension term) are represented by solid line curves, cf. (5.11). Underlying structures are described in Section 5.2.	181
5.5	Covariograms (thick line curves) and their last continuous derivatives (thin line curves in (b) and (c)). The corresponding measurement functions are shown in Figure 5.2.	182
5.6	Bias of the extension term estimators. For consistency with Figure 5.3 showing coefficients of errors, the bias is measured by $\overline{\mathbf{E}}(T)^{\frac{1}{2}} / E(T)^{\frac{1}{2}}$ (solid line curve) where $\overline{\mathbf{E}}(T)$ is the average extension term estimator computed from simulations. The departures (due to the Zitterbewegung) of the exact coefficient of error from the extension term are measured by $\text{CE}[\hat{\mathbf{Q}}_T] / E(T)^{\frac{1}{2}}$ (dots). Corresponding measurement functions are shown in Figure 5.2.	186