UNIVERSITE PARIS 1 PANTHEON SORBONNE

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

S. SOUCHET

Prépublication du SAMOS n° 78
Juillet 1997

90, rue de Tolbiac - 75634 PARIS CEDEX 13

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

Sandie SOUCHET
SAMOS, Université Paris 1*

Juillet 97

Résumé
On considère Y une diffusion unidimensionnelle ergodique, dont la fonction de dérive dépend d'un paramètre θ_0. Une solution pour estimer θ_0, à partir d'une observation du processus à n instants équidistants $k\delta$, lorsque δ est proche de 0 mais fixé et $n \to \infty$, est de construire une approximation de la vraisemblance discrète. On remplace chaque densité conditionnelle par une densité gaussienne, dont l'espérance approche l'espérance conditionnelle à un $o(\delta^p)$ près. Dans ce but, nous proposons un schéma d'approximation discret et adapté de la diffusion à l'ordre p. Sous certaines conditions générales, on obtient ainsi un estimateur asymptotiquement normal et efficace à un biais en δ^p près.

Abstract
We consider Y a one-dimensional ergodic diffusion process, which drift depends on an unknown parameter θ_0. A way to estime θ_0 from the observation of the process at n equidistant times $k\delta$, as δ is near 0 but fixed and $n \to \infty$ is to construct an approximation to the likelihood function. Each transition density can be replaced by a Gaussian density, the mean of which is an approximation in $o(\delta^p)$ to the conditional expectation. For that purpose, we propose a p-order adapted and approximative discrete-time scheme of diffusion process. Under some general assumptions, we obtain an asymptotically normal and efficient estimator with a bias in δ^p.

Mots clefs
Processus de diffusion, estimation par minimum de contraste, espérance conditionnelle, schéma d'approximation adapté à l'ordre p, biais d'estimation, efficacité asymptotique.
Code A.M.S.
62 M 05 - 62 F 12.

1 Introduction

Considérons l'équation différentielle stochastique suivante :

$$dY_t = f(\theta_0, Y_t)\, dt + \sigma(Y_t)\, dW_t,\ Y_0 = x \tag{1}$$

où f et σ sont connues et à valeurs dans $\mathbb{R}$ et θ_0 est un paramètre inconnu de $\mathbb{R}^p$.
L'estimation du paramètre de dérive à partir d'une observation discrète à pas $\delta > 0$ fixé

*Centre Pierre Mendès France, 90, rue de Tolbiac 78634 PARIS cedex 13

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 $n \to \infty$, 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 alternative consiste à approximer les densités de transition par des densités gausiennes d'espérance $\mathbb{E}_\theta\left[Y_{k\delta} \mid Y_{(k-1)\delta}\right]$ et de variance $\mathbb{V}ar_\theta\left[Y_{k\delta} \mid Y_{(k-1)\delta}\right]$, 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 $h \in C^{2p+1}$. 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.

Bien que présentée de manière différente, 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 de σ_0, le paramètre de diffusion), lorsque $\delta_n \to 0$, $n\delta_n \to \infty$ et $n\delta_n^p \to 0$. Dans notre démarche, le pas d'observation $\delta > 0$ est fixé et proche de 0. 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 en 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 est d'ordre 1. Nous montrerons alors que l'estimateur du minimum de contraste associé est asymptotiquement efficace en variance à 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 Schéma d'approximation discret et adapté à l'ordre p de $h(Y_t)_{t\geq 0}$

Soit l'e.d.s.

$$dY_t = f(Y_t)\,dt + \sigma(Y_t)\,dW_t \tag{2}$$

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

Propriété 1
Soient h une fonction de $\mathbb{R}$ dans $\mathbb{R}$ et $p \in \mathbb{N}*$. Supposons que :

1. $h \in C^{2(p+1)}$, *f et σ* $\in C^{2p}$.

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

$$h\left(Y_{t+\delta}\right) = h\left(Y_t\right) + \sum_{l=1}^{p} \frac{\delta^l}{l} \mathcal{A}^l h\left(Y_t\right) + \eta_t^p + \tilde{W}_t^p \tag{3}$$

avec :

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

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

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 $\mathbb{R}$. 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 :

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

- Montrons la propriété pour $p = 1$:
  On pose :

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

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

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

  D'où :

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

  où η_t^1 et $\tilde{W}_t^1$ sont donnés par (3).

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

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

avec :

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

on a :

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

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

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

D'où :

$$\begin{aligned}
Y_{t+\delta} - Y_t &= \sum_{l=1}^{p} \frac{\delta^l}{l!} \mathcal{A}^l h(Y_t) + R_t + \tilde{W}_t^{p-1} \\
&= \sum_{l=1}^{p} \frac{\delta^l}{l!} \mathcal{A}^l h(Y_t) + \eta_t^p + \tilde{W}_t^p
\end{aligned}$$

où η_t^p et $\tilde{W}_t^p$ sont donnés par (3). $\square$.

Dans le paragraphe qui suit, nous utilisons ce schéma discret et adapté afin d'obtenir des développements en δ de $\mathbb{E}[Y_{t+\delta} \setminus Y_t]$ et $\mathbb{V}ar[Y_{t+\delta} \setminus Y_t]$.

3 Résultats asymptotiques

L'approximation des quantités conditionnelles, ainsi que les propriétés asymptotiques de l'estimateur associé au contraste gaussien construit à partir de celle-ci, repose sur un jeu d'hypothèses général noté **H** que nous allons préciser ci-après. Notons $f(\theta,.) = f_\theta(.)$ et $s(x,\theta) = \exp\left[-2\int^x \frac{f_\theta(u)}{\sigma^2(u)} du\right]$ la fonction d'echelle associée à (1). Soit Θ un sous-ensemble de $\mathbb{R}^p$ tel que $\theta_0 \in \overset{\circ}{\Theta}$. On désigne par g' la dérivée de g par rapport à x. Enfin, si $M = (m_{i,j})_i^j$ est une matrice de dimension $d \times d$, i représente l'indice des lignes et j celui des colonnes.

H 1 f_{θ_0} et σ sont de classe C^{2p} sur $\mathbb{R}$. De plus, il existe $K > 0$ et $\alpha \in [\frac{1}{2}, 1]$ telles que,

$$\forall (x,y) \in \mathbb{R}^2,\ |f_{\theta_0}(x) - f_{\theta_0}(y)| \leq K\,|x-y|\,,\ |\sigma(x) - \sigma(y)| \leq K\,|x-y|^\alpha$$

Sous cette hypothèse, (1) admet une unique solution $(Y_t)_{t\geq 0}$ adaptée.

H 2

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

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

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

H 3 Il existe deux constantes $C > 0$ et $M > 0$ telles que :

$$\forall |x| > M,\ \left(\frac{f_{\theta_0}\left(x\right)}{\sigma(x)} - \frac{1}{2}\sigma'\left(x\right)\right) sign(x) \leq -C$$

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

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

De plus, pour $\pi^t_{\theta_0}\left(x,dy\right) = P\left(Y_t \in dy \mid Y_0 = x\right)$ et $Q^t_{\theta_0} = \mu_{\theta_0} \otimes \pi^t_{\theta_0}$, il existe $\delta_0 > 0$ tel que:

$$\sup_{t\in[0,\delta_0]} \int \frac{|y|^p}{\sigma^4(x)} dQ^t_{\theta_0}(x,y) < \infty$$

H 5
- Pour tout $l = 0, p-1$ et tout $x \in \mathbb{R}$, $\theta \to \mathcal{A}^l_\theta f_\theta\left(x\right)$ est de classe C^3 sur Θ et à croissance pôlynomiale en x uniformément en θ, ainsi que ses dérivées partielles en θ jusqu'à l'ordre 3.
- Pour tout $l = 0, p-1$, $\left(\mathcal{A}^l_{\theta_0} f_{\theta_0}\right)'$ et $\mathcal{A}^p_{\theta_0} f_{\theta_0}$ sont à croissance polynômiale en x.

H 6 Pour tout i, j, les fonctions suivantes sont continues au voisinage de 0:

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

On note $N^p_{\theta_0} = \left(N^p_i(0)\right)_i$

H 7 $H_{\theta_0} = \left(E_{\theta_0}\left[\frac{\partial}{\partial\theta^i} f_{\theta_0}\left(Y_0\right) \frac{\partial}{\partial\theta^j} f_{\theta_0}\left(Y_0\right) / \sigma^2\left(Y_0\right)\right]\right)^j_i$ est inversible.

H 8 Θ est un compact de $\mathbb{R}^d$.

H 9 Il existe un unique $\theta_\delta \in \overset{\circ}{\Theta}$ tel que pour tout $i = 1, d$, $\frac{\partial}{\partial\theta^i} U\left(\theta_\delta, \theta_0\right) = 0$.

On note : $M\left(\theta, x\right) = \sum_{l=0}^{p-1} \frac{\delta^{l+1}}{(l+1)!} \mathcal{A}^l_\theta f_\theta\left(x\right)$.

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

$$\begin{aligned} E_{\theta_0}[Y_{t+\delta} \mid Y_t] &= Y_t + M(\theta_0, Y_t) + o(\delta^p) \\ \mathbb{V}ar_{\theta_0}[Y_{t+\delta} \mid Y_t] &= \delta\sigma^2(Y_t) + o(\delta) \end{aligned}$$

<u>Démonstration</u> :

- L'approximation à l'ordre p de $E_{\theta_0}[Y_{t+\delta} \mid Y_t]$ est obtenue en utilisant le schéma (3) avec $h(x) = x$. Compte tenu des hypothèses d'intégrablité et de croissante polynômiale, on obtient par passage à l'espérance conditionnelle :

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

  avec $E_{\theta_0}[|\eta_t^p|] \leq K \frac{\delta^{p+1}}{(p+1)!}$, où K est une constante indépendante de δ.

- On procède de même pour l'approximation de la variance avec $h(x) = x^2$, on a :

$$E_{\theta_0}\left[Y_{t+\delta}^2 \mid Y_t\right] = Y_t^2 + 2\delta Y_t f_{\theta_0}(Y_t) + \delta\sigma^2(Y_t) + o(\delta)$$

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

  $\square$

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. Le contraste gaussien associé au schéma à l'ordre p s'écrit alors :

$$U_n(\theta) = \frac{1}{n}\sum_{k=1}^{n} \Phi\left(Y_{k\delta}, Y_{(k-1)\delta}, \theta\right), \ \Phi(y, x, \theta) = \frac{(y - x - M(\theta, x))^2}{\delta\sigma^2(x)}$$

L'ensemble des développements qui suivent s'appuient sur l'application du lemme 3.2 de [1] qui généralise un résultat d'ergodicité et de normalité asymptotique établit par D. Florenz-Zmirou [4]. Rappelons ce résultat :

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

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

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

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n} g\left(Y_{k\delta}, Y_{(k-1)\delta}\right) \overset{\mathcal{D}(P_{\theta_0})}{\longrightarrow} N\left(0, Q_{\theta_0}^{\delta}\left(g^2\right)\right)$$

Notre cadre d'étude s'inscrit bien dans celui de l'estimation par minimum de contraste (c.f. [2]. chapitre 3). Notons en effet

$$U(\theta, \theta_0) = E_{\theta_0}[\Phi(Y_\delta, Y_0, \theta)]$$

Lemme 3 *Sous (H), U est la fonction de contraste associée à U_n.*

Démonstration :
(H 4) et (H 5) impliquent que $\forall\theta \in \Theta$, $(y,x) \rightarrow \Phi(y,x,\theta) \in L^2\left(Q^{\delta}_{\theta_0}\right)$. Donc, par application du lemme 2, on a :

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

U est définie sur Θ. De plus, elle positive ou nulle et admet un minimum en θ_δ (H 9). U_n et U remplissent donc conditions de la définition 3.2.7 de [2].
□
On note :

$$\widehat{\theta}_n = \arg\min_{\theta\in\Theta} U_n(\theta)$$

3.1 Calcul de l'écart entre θ_δ et θ_0

Théorème 1 *Sous (H), pour H_{θ_0} définie en (H 7) et $N^p_{\theta_0}$ défini en (H 6), on a :*

$$\theta_\delta = \theta_0 + \frac{\delta^p}{(p+1)!} H^{-1}_{\theta_0} N^p_{\theta_0} + o(\delta^p)$$

Démonstration : Nous utiliserons le résultat suivant démontré en annexe A.

Lemme 4 *Sous (H), $\theta \rightarrow U(\theta,\theta_0)$ est de classe C^3 sur Θ. De plus pour tout $i,j = 1,d$:*

$$\begin{aligned}\frac{\partial^2}{\partial\theta^j\partial\theta^i} U(\theta,\theta_0) &= 2\delta E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta^i} f_{\theta_0}(Y_0)\frac{\partial}{\partial\theta^j} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)}\right] + R(\delta,\theta)\\ |R(\delta,\theta)| &\leq K\delta\,||\theta-\theta_0|| + o(\delta)\end{aligned}$$

K étant une constante indépendante de δ et θ.

- Compte tenu de (H 9) et par application du théorème des accroissements finis, pour tout $i = 1,d$, il existe $u_i \in]0,1[$ tel que pour $\tilde{\theta}_i = u_i\theta_0 + (1-u_i)\theta_\delta$, on ait :

$$\frac{\partial}{\partial\theta^i} U(\theta_\delta,\theta_0) = 0 = \frac{\partial}{\partial\theta^i} U(\theta_0,\theta_0) + \sum_{j=1}^{d} \frac{\partial^2}{\partial\theta^j\partial\theta^i} U\left(\tilde{\theta}_i,\theta_0\right)\left(\theta^j_\delta - \theta^j_0\right)$$

- Par le lemme 4, on a pour δ proche de 0 :

$$\frac{\partial^2}{\partial\theta^j\partial\theta^i} U\left(\tilde{\theta}_i\right) = 2\delta E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta^i} f_{\theta_0}(Y_0)\frac{\partial}{\partial\theta^j} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)}\right] + o(\delta)$$

- La propriété 1 et (H 6) permettent d'écrire :

$$\begin{aligned}\frac{\partial}{\partial\theta^i} U(\theta_0,\theta_0) &= -2E_{\theta_0}\left[\frac{\partial}{\partial\theta^i} M(\theta_0,Y_0)\frac{\eta^p_0}{\delta\sigma^2(Y_0)}\right]\\ &= -2\int_0^\delta \frac{(\delta-v)^p}{p!} N^p_i(v)\,dv - 2\sum_{l=1}^{p-1}\frac{\delta^l}{(l+1)!} E_{\theta_0}\left[\frac{\partial}{\partial\theta^i} A^l_{\theta_0} f_{\theta_0}(Y_0)\frac{\eta^p_0}{\sigma^2(Y_0)}\right]\end{aligned}$$

D'après (H 4) et (H 5), la somme du développement précédent est un $o\left(\delta^{p+1}\right)$. (H 6) implique de plus que :

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

On a donc :

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

Utilisant (H 7), on obtient le résultat annoncé.

□

3.2 Convergence de $\widehat{\theta}_n$ vers θ_δ

Théorème 2 *Sous (H),* $\widehat{\theta}_n\xrightarrow{P_{\theta_0}}\theta_\delta$

Démonstration : Θ étant compact, $\theta\rightarrow U_n\left(\theta\right)$ et $\theta\rightarrow U\left(\theta,\theta_0\right)$ étant continue sur Θ (lemme 4), la convergence de $\widehat{\theta}_n$ vers θ_δ reprend exactement la démonstration du théorème 3.2.8. de [2], la condition sur le module de continuité de $U_n\left(\theta\right)$ étant remplacée par :

$$\forall\varepsilon>0,\ \lim_{\eta\rightarrow0}\lim_{n\rightarrow\infty}P_{\theta_0}\left(\sup_{\left|\left|\theta-\tilde{\theta}\right|\right|_\infty\leq\eta}\left|U_n\left(\theta\right)-U_n\left(\tilde{\theta}\right)\right|>\varepsilon\right)=0 \tag{4}$$

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

$$\sup_{\left|\left|\theta-\tilde{\theta}\right|\right|_\infty\leq\eta}\left|U_n\left(\theta\right)-U_n\left(\tilde{\theta}\right)\right|\leq\sum_{i=1}^d\left(\frac{1}{n}\sum_{k=1}^n\phi^i\left(Y_{k\delta},Y_{(k-1)\delta}\right)\right)\eta$$

Par application du lemme 2, on obtient :

$$D_n=\sum_{i=1}^d\frac{1}{n}\sum_{k=1}^n\phi^i\left(Y_{k\delta},Y_{(k-1)\delta}\right)\overset{L^1\left(Q_{\theta_0}^\delta\right)}{\rightarrow}D=\sum_{i=1}^dE_{\theta_0}\left[\phi^i\left(Y_\delta,Y_0\right)\right]<\infty$$

D'où :

$$\begin{aligned}\forall\varepsilon>0,\ P_{\theta_0}\left(\sup_{\left|\left|\theta-\tilde{\theta}\right|\right|_\infty\leq\eta}\left|U_n\left(\theta\right)-U_n\left(\tilde{\theta}\right)\right|>\varepsilon\right)&\leq P_{\theta_0}\left(D_n>\varepsilon\right)\\&\leq\frac{\eta}{\varepsilon}E_{\theta_0}\left[\left|D_n-D\right|\right]+\frac{\eta}{\varepsilon}D\end{aligned}$$

En passant successivement à la limite en n, puis en η, on obtient (4).

□

3.3 Normalité asymptotique de $\widehat{\theta}_n$

Théorème 3 *Sous (H), pour H_{θ_0} définie en (H 7), on a :*

$$\sqrt{T}\left(\widehat{\theta}_n - \theta_0\right) \overset{D(P_{\theta_0})}{\longrightarrow} N_d\left(0, V_\delta\left(\theta_0\right)\right), \quad V_\delta\left(\theta_0\right) = H_{\theta_0}^{-1}\left(I_d + o(1)\right)$$

<u>Démonstration :</u> $\theta \to U_n(\theta)$ est de classe C^3 sur Θ; on peut donc poser :

$$DU_n(\theta) = \left(\frac{\partial}{\partial\theta^i}U_n(\theta)\right)_i \text{ et } D^2U_n(\theta) = \left(\frac{\partial^2}{\partial\theta^j\partial\theta^j}U_n(\theta)\right)_i^j$$

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

$$\begin{aligned}\sqrt{n\delta}DU_n\left(\widehat{\theta}_n\right) &= 0 = \sqrt{n\delta}DU_n\left(\theta_\delta\right) + \left(D^2U_n\left(\theta_\delta\right) + R_n\right)\sqrt{n\delta}\left(\widehat{\theta}_n - \theta_\delta\right) \\ R_n &= \int_0^1 \left[D^2U_n\left(\theta_\delta + s\left(\widehat{\theta}_n - \theta_\delta\right)\right) - U_n\left(\theta_\delta\right)\right] ds\end{aligned}$$

Si l'on prouve que :

1. $\sqrt{n\delta}DU_n\left(\theta_\delta\right) \overset{D(P_{\theta_0})}{\to} N\left(0, K\left(\theta_\delta\right)\right), \; K\left(\theta_\delta\right) = 4H_{\theta_0}\delta^2 + o\left(\delta^2\right)$
2. $D^2U_n\left(\theta_\delta\right) \overset{P_{\theta_0}}{\to} 2H_{\theta_0}\delta + o(\delta)$
3. $R_n \overset{P_{\theta_0}}{\to} 0$

on aura la normalité asymptotique annoncée pour $\widehat{\theta}_n$.

1. Sous (H 4) et (H 5), $\frac{\partial}{\partial\theta^i}\Phi\left(y, x, \theta_\delta\right) \in L^2\left(Q_{\theta_0}^\delta\right)$. Compte tenu de (H 9) et par application du lemme 2, on a :

   $$\sqrt{n\delta}DU_n\left(\theta_\delta\right) \overset{D(P_{\theta_0})}{\longrightarrow} N_d\left(0, K\left(\theta_\delta\right)\right), \quad K\left(\theta_\delta\right) = \left(\delta E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}\Phi\left(Y_\delta, Y_0, \theta_\delta\right)\frac{\partial}{\partial\theta^j}\Phi\left(Y_\delta, Y_0, \theta_\delta\right)\right]\right)_i^j$$

   Lemme 5 *Sous (H), pour H_{θ_0} définie en (H 6), on a :*

   $$K\left(\theta_\delta\right) = 4H_{\theta_0}\delta^2 + o\left(\delta^2\right)$$

   La démonstration de ce résultat est donnée en annexe B. Le premier résultat de convergence faible est donc vérifié.

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

   $$D^2U_n\left(\theta_\delta\right) \overset{L^1(P_{\theta_0})}{\longrightarrow} D^2U\left(\theta_\delta, \theta_0\right) = \left(E_{\theta_0}\left[\frac{\partial^2}{\partial\theta^j\partial\theta^i}\Phi\left(Y_\delta, Y_0, \theta_\delta\right)\right]\right)_i^j$$

   Compte tenu du lemme 4 et de l'écart obtenu entre θ_δ et θ_0, on a :

   $$D^2U\left(\theta_\delta, \theta_0\right) = 2H_{\theta_0}\delta + o(\delta)$$

   Il nous reste donc à montrer la convergence de R_n vers 0.

3. Pour tout $i,j=1,d$, on a d'après la formule de Rolles :
$$\left|\int_0^1\left[\frac{\partial^2}{\partial\theta^j\partial\theta^i}U_n\left(\theta_\delta+s\left(\widehat{\theta}_n-\theta_\delta\right)\right)-\frac{\partial^2}{\partial\theta^j\partial\theta^i}U_n\left(\theta_\delta\right)\right]ds\right|\leq\sum_{l=1}^{d}D_n^{i,j,l}\left|\widehat{\theta}_n^l-\theta_\delta^l\right|$$
avec $D_n^{i,j,l}=\frac{1}{n}\sum_{k=1}^n\sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\theta^l\partial\theta^j\partial\theta^i}\Phi\left(Y_{k\delta},Y_{(k-1)\delta},\theta\right)\right|$.
Or, sous (H 4) et (H 5), $\sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\theta^l\partial\theta^j\partial\theta^i}\Phi\left(y,x,\theta\right)\right|\in L^1\left(Q_{\theta_0}^\delta\right)$. Donc, $D_n^{i,j,k}$ converge en probabilité vers une quantité finie. Compte tenu de la convergence de $\widehat{\theta}_n$ vers θ_δ, on obtient $R_n\overset{P_{\theta_0}}{\to}0$.

□

4 Etude expérimentale

Dans ce paragraphe, nous vérifierons expérimentalement quelques unes des propriétés annoncées pour $\widehat{\theta}_n$ (convergence vers θ_δ, écart entre $\widehat{\theta}_n$ 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 $\delta=0.05,0.1,0.5$ et $T=200$, le paramètre est alors estimé par les moindres carrés pour différentes p, l'ordre du schéma d'approximation. On indicera par p l'estimation obtenue. On repète 1000 fois cette expérience afin de calculer la moyenne et la variance empiriques, notées respectivement $m\left(\widehat{\theta}_n\right)$ et $S^2\left(\widehat{\theta}_n\right)$.

4.1 Processus de Ornstein-Uhlenbeck

On considère l'e.d.s. suivante:
$$dY_t=-\theta_0Y_tdt+dW_t,\ Y_0=x\tag{5}$$
(5) admet une unique solution forte, à trajectoires continues,
$$Y_t=\exp\left(-\theta_0t\right)\left(x+\int_0^t\exp\left(\theta_0s\right)dW_s\right)$$
$\pi_{\theta_0}^t(x,.)$ est donc une loi normale $N\left(x\exp\left(-\theta_0t\right),\frac{1-\exp(-2\theta_0t)}{2\theta_0}\right)$.
Si $\theta_0>0$, $(Y_t)_{t\geq0}$ est récurrente positive sur $\mathbb{R}$, de loi invariante $\mu_{\theta_0}=N\left(0,\frac{1}{2\theta_0}\right)$. En outre, pour $M>0$ et $C=\theta_0M$, (H 1.3) est vérifiée. Enfin, on a :
$$\mathcal{A}_{\theta_0}^lf_{\theta_0}(x)=\left(-\theta_0\right)^{l+1}x$$
$$P(v)=\frac{1}{2\theta_0},\ N^l(v)=\frac{\left(-\theta_0\right)^l}{2}\exp\left(-\theta_0v\right),\ H_{\theta_0}=\frac{1}{2\theta_0}$$
Notons :
$$Q_p\left(x\right)=\sum_{l=0}^{p-1}\frac{\left(-x\right)^{l+1}}{(l+1)!}\tag{6}$$

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

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

θ_δ est solution de l'équation :

$$Q_p\left(\delta\theta\right) = \exp\left(-\delta\theta_0\right) - 1$$

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, l'hypothèse (H 9) est vérifiée. On a en particulier pour $p = 1$, $p = 2$ et $p = \infty$:

$$\begin{aligned}
\theta_\delta^1 &= \frac{1}{\delta}\left(1 - \exp\left(-\delta\theta_0\right)\right) = \theta_0 - \frac{\theta_0^2}{2}\delta + o\left(\delta\right) \\
\theta_\delta^2 &= \frac{1}{\delta}\left(1 - \sqrt{2\exp\left(-\delta\theta_0\right) - 1}\right) = \theta_0 + \frac{\theta_0^3}{6}\delta^2 + o\left(\delta^2\right) \\
\theta_\delta^\infty &= \theta_0
\end{aligned}$$

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 $p = \infty$, $\widehat{\theta}_n$ correspond à l'estimateur associé au contraste gaussien basé sur les espérances conditionnelles exactes.

			$\widehat{\theta}_n \ (\theta_0 = 1)$			
	δ	n	$m\left(\widehat{\theta}_n\right)$	$S^2\left(\widehat{\theta}_n\right)$	θ_δ	$\frac{H_{\theta_0}^{-1}}{T}$
	0.05	4000	0.9878	0.0098	0.9754	
$p = 1$	0.1	2000	0.964525	0.008976	0.951626	0.01
	0.5	400	0.7999	0.0068	0.7869	

Table 1: Processus de O.U. et approximation à l'ordre 1 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.

			$\widehat{\theta}_n \ (\theta_0 = 1)$			
	δ	n	$m\left(\widehat{\theta}_n\right)$	$S^2\left(\widehat{\theta}_n\right)$	θ_δ	$\frac{H_{\theta_0}^{-1}}{T}$
	0.05	4000	1.0138	0.0109	1.0004	
$p = 2$	0.1	2000	1.016774	0.011170	1.001807	0.01
	0.5	400	1.1172	0.0344	1.076	

Table 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 $\widehat{\theta}_n$; dans les autres cas, les réalisations de R_n sont en dehors de $[-\frac{1}{2}, 0]$.

			$\widehat{\theta}_n \ (\theta_0 = 1)$			
	δ	n	$m\left(\widehat{\theta}_n\right)$	$S^2\left(\widehat{\theta}_n\right)$	θ_δ	$\frac{H_{\theta_0}^{-1}}{T}$
	0.05	4000	1.01335	0.01089	0.999756	
$p = 3$	0.1	2000	1.014716	0.011030	1.000366	0.01
	0.5	400	1.0167	0.0179	0.9925	

Table 3: Processus de O.U. et approximation à l'ordre 3 : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.

			$\widehat{\theta}_n \ (\theta_0 = 1)$			
	δ	n	$m\left(\widehat{\theta}_n\right)$	$S^2\left(\widehat{\theta}_n\right)$	θ_δ	$\frac{H_{\theta_0}^{-1}}{T}$
	0.05	4000	1.013373	0.010912	1	
$p = \infty$	0.1	2000	1.014816	0.011036	1	0.01
	0.5	400	1.0264	0.0196	1	

Table 4: Processus de O.U. et approximation à l'ordre ∞ : moyenne et variance empiriques pour 1000 réalisations de l'estimateur E.M.C.

4.2 Processus de Cox-Ingersoll-Ross

On considère l'e.d.s. suivante:

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

On a donc :

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

f est C^∞ sur $\mathbb{R}$ et lipschitzienne, σ est C^∞ sur $\mathbb{R}_+^*$ et hölderienne 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 récurrente positive sur $\mathbb{R}_+^*$, de loi invariante $\mu_\theta = \Gamma(2b, 2a)$. Cette loi admet des moments de tous ordres.
De plus, si $b > 2$, $\int \frac{1}{\sigma^8(x)} d\mu_{\theta_0}(x) < \infty$ et donc (H 1.4) est vérifiée. (H 1.3) est vérifiée pour $M > \max\{1, \frac{4b_0-1}{4a}\}$ et $C = \frac{1}{4} + a_0 M - b_0$. Enfin, on a pour tout $l \leq 0$, $\mathcal{A}_\theta^l f_\theta(x) = (-a)^l f_\theta(x)$. On montre que :

$$\begin{aligned}
N_1^p(v) &= 0, \ N_2^p(v) = -\frac{(-a_0)^{p+1}}{2b_0 - 1} \exp(-a_0 v) \\
P_{1,1}(v) &= \frac{b_0}{a_0}, \ P_{1,2}(v) = P_{2,1}(v) = \frac{\exp(-a_0 v) - 2b_0}{2b_0 - 1}, \ P_{2,2}(v) = 2\frac{a_0 b_0 (b_0 - \exp(-a_0 v))}{(2b_0 - 1)(b_0 - 1)}
\end{aligned}$$

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

Q_p étant donné par (6), l'estimateur du minimum de contraste (a_n, b_n) vérifie le système suivant :

$$\begin{aligned}
Q_p(\delta a_n) &= R_n, \ R_n = \frac{\left(\frac{1}{n}\sum_{k=1}^n \frac{1}{Y_{(k-1)\delta}}\right)\left(\frac{1}{n}\sum_{k=1}^n \left(Y_{k\delta} - Y_{(k-1)\delta}\right)\right) - \frac{1}{n}\sum_{k=1}^n \frac{\left(Y_{k\delta} - Y_{(k-1)\delta}\right)}{Y_{(k-1)\delta}}}{\left(\frac{1}{n}\sum_{k=1}^n Y_{(k-1)\delta}\right)\left(\frac{1}{n}\sum_{k=1}^n \frac{1}{Y_{(k-1)\delta}}\right) - 1} \\
b_n &= \frac{a_n}{R_n}\left[R_n \frac{1}{n}\sum_{k=1}^n Y_{(k-1)\delta} - \frac{1}{n}\sum_{k=1}^n \frac{\left(Y_{k\delta} - Y_{(k-1)\delta}\right)}{Y_{(k-1)\delta}}\right]
\end{aligned}$$

(a_δ, b_δ), est solution de :

$$\begin{aligned}
Q^p(\delta a_\delta) &= \exp(-\delta a_0) - 1 \\
b_\delta &= \frac{b_0}{a_0} a_\delta
\end{aligned}$$

a_δ vérifie donc la même équation que θ_δ pour le modèle O.U.
Pour $p = 1$, $p = 2$ et $p = \infty$, lorsque δ est proche de 0 :

$$\begin{aligned}
b_\delta^1 &= b_0 - \frac{b_0 a_0}{2}\delta + o(\delta) \\
b_\delta^2 &= b_0 + \frac{b_0 a_0^2}{6}\delta^2 + o\left(\delta^2\right) \\
b_\delta^\infty &= b_0
\end{aligned}$$

On prend $Y_0 = 1$ et $(a_0, b_0) = (1, 3)$.

	δ	n	a_n $(a_0 = 1)$ $m(a_n)$	$S^2(a_n)$	a_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1^1$	b_n $(b_0 = 3)$ $m(b_n)$	$S^2(b_n)$	b_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2^2$
$p = 1$	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 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.

	δ	n	a_n $(a_0 = 1)$ $m(a_n)$	$S^2(a_n)$	a_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1^1$	b_n $(b_0 = 3)$ $m(b_n)$	$S^2(b_n)$	b_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2^2$
$p = 2$	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 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.

	δ	n	a_n $(a_0 = 1)$ $m(a_n)$	$S^2(a_n)$	a_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1^1$	b_n $(b_0 = 3)$ $m(b_n)$	$S^2(b_n)$	b_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2^2$
$p = 3$	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 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.

	δ	n	a_n $(a_0 = 1)$ $m(a_n)$	$S^2(a_n)$	a_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_1^1$	b_n $(b_0 = 3)$ $m(b_n)$	$S^2(b_n)$	b_δ	$\left(\frac{H_{\theta_0}^{-1}}{T}\right)_2^2$
$p = \infty$	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 8: Processus de C.I.R. et approximation à l'ordre ∞ : moyenne et variance empirique pour 1000 réalisations de l'estimateur E.M.C.

References

[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.*- 2^{nd} Ed. Springer.

[7] Kessler, M. (1995) *Estimation of an ergodic diffusion from discrete observations.*- Prépublication 306. Laboratoire de Probabilité de l'Université Paris 6.

[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 approch 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. (1997) *Schéma de discrétisation anticipatif et estimation du paramètre de dérive d'une diffusion.*-Prépublicatio du SAMOS 72.

[15] Yoshida, N. (1992) *Estimation for diffusion processes from discrete observations.*- J. Multivariate Anal. 41, 220-242.

A Démonstration du lemme 4 : régularité de $\theta \to U(\theta, \theta_0)$

1. Montrons que $\theta \to U(\theta, \theta_0)$ est de classe C^3 sur Θ : (H 4) implique que pour tout couple $(y, x) \in \mathbb{R}^2$, $\theta \to \Phi(y, x, \theta)$ est de classe C^3 sur Θ. De plus, sous (H 4) et (H 5), on montre que pour tout i et tout j :

$$\sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta^i} \Phi(y, x, \theta) \right| \in L^1\left(Q^{\delta}_{\theta_0}\right)$$
$$\sup_{\theta \in \Theta} \left| \frac{\partial^2}{\partial \theta^j \theta^i} \Phi(y, x, \theta) \right| \in L^1\left(Q^{\delta}_{\theta_0}\right)$$

Par application du théorème de Lebesgue, on obtient que U est deux fois continuement différentiable par dérivation sous le signe somme.

2. Montrons que $\frac{\partial^2}{\partial \theta^j \theta^i} U(\theta, \theta_0) = M_{i,j}(0) + R(\delta; \theta)$, avec $||R(\delta, \theta)|| \leq \delta K ||\theta - \theta_0|| + o(\delta)$, K étant une constante indépendante de θ et δ.
Posons pour tout $l = 0, p-1$ et pour tout couple (i, j), $i, j = 1, d$

$$\begin{aligned}
C^l_i &= \left(E_{\theta_0} \left[\frac{\left(\sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta^i} \mathcal{A}^l_\theta f_\theta(Y_0) \right| \right)^2}{\sigma^2(Y_0)} \right] \right)^{\frac{1}{2}} \\
D^l_{i,j} &= \left(E_{\theta_0} \left[\frac{\left(\sup_{\theta \in \Theta} \left| \frac{\partial^2}{\partial \theta^j \partial \theta^i} \mathcal{A}^l_\theta f_\theta(Y_0) \right| \right)^2}{\sigma^2(Y_0)} \right] \right)^{\frac{1}{2}}
\end{aligned}$$

Sous (H 4) et (H 5), les quantités précédentes sont finies et indépendantes de θ et δ. D'après la décomposition donnée en (3), on a :

$$\begin{aligned}
\frac{\partial^2}{\partial \theta^j \partial \theta^i} U(\theta, \theta_0) &= -2E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \partial \theta^i} M(\theta, Y_0) \frac{M(\theta_0, Y_0) - M(\theta, Y_0)}{\delta \sigma^2(Y_0)} \right] \\
&- 2E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^j \partial \theta^i} M(\theta, Y_0) \frac{\eta^p_0}{\delta \sigma^2(Y_0)} \right] \\
&+ 2E_{\theta_0} \left[\frac{\frac{\partial}{\partial \theta^i} M(\theta, Y_0) \frac{\partial}{\partial \theta^j} M(\theta, Y_0)}{\delta \sigma^2(Y_0)} \right] \\
&= 2\delta E_{\theta_0} \left[\frac{\frac{\partial}{\partial \theta^i} f_{\theta_0}(Y_0) \frac{\partial}{\partial \theta^j} f_{\theta_0}(Y_0)}{\sigma^2(Y_0)} \right] \\
&+ R_1(\delta, \theta) + R_2(\delta, \theta) + R_3(\delta, \theta) + R_4(\delta, \theta)
\end{aligned}$$

avec :

$$
\begin{aligned}
R_1(\delta,\theta) &= -2\sum_{l,k=0}^{p-1}\frac{\delta^{l+k+1}}{(l+1)!(k+1)!}E_{\theta_0}\left[\frac{\partial^2}{\partial\theta^j\theta^i}\mathcal{A}_\theta^l f_\theta(Y_0)\left(\frac{\mathcal{A}_\theta^l f_{\theta_0}(Y_0)-\mathcal{A}_\theta^k f_\theta(Y_0)}{\sigma^2(Y_0)}\right)\right]\\
R_2(\delta,\theta) &= -2\sum_{l=0}^{p-1}\frac{\delta^l}{(l+1)!}E_{\theta_0}\left[\frac{\partial^2}{\partial\theta^j\theta^i}\mathcal{A}_\theta^l f_\theta(Y_0)\frac{\eta_0^p}{\delta\sigma^2(Y_0)}\right]\\
R_3(\delta,\theta) &= 2\sum_{l,k=1}^{p-1}\frac{\delta^{l+k+1}}{(l+1)!(k+1)!}E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta^i}\mathcal{A}_\theta^l f_\theta(Y_0)\frac{\partial}{\partial\theta^j}\mathcal{A}_\theta^k f_\theta(Y_0)}{\sigma^2(Y_0)}\right]\\
R_4(\delta,\theta) &= 2\delta\left(E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta^i}f_\theta(Y_0)\frac{\partial}{\partial\theta^j}f_\theta(Y_0)}{\sigma^2(Y_0)}\right]-E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\partial}{\partial\theta^j}f_{\theta_0}(Y_0)}{\sigma^2(Y_0)}\right]\right)
\end{aligned}
$$

L'application de la formule de Rolles et de l'inégalité de Schwarz donnent :

$$|R_1(\delta,\theta)|\leq 2\sum_{l,k=0}^{p-1}\frac{\delta^{l+k+1}}{(l+1)!\,(k+1)!}D_{j,i}^l\sum_{m=1}^{d}C_m^k\,|\theta^m-\theta_0^m|$$

$$|R_2(\delta,\theta)|\leq 2\sum_{l=0}^{p-1}\frac{\delta^{l+p+1}}{p!(l+1)!}D_{j,i}^l\sup_{v\in[0,\delta_0]}\left(E_{\theta_0}\left[\frac{\left(\sup_{\theta\in\Theta}|\mathcal{A}_\theta^p f_\theta(Y_v)|\right)^2}{\sigma^2(Y_0)}\right]\right)^{\frac{1}{2}}$$

De plus :

$$|R_3(\delta,\theta)|\leq 2\sum_{l,k=1}^{p-1}\frac{\delta^{l+k+1}}{(l+1)!(k+1)!}C_i^lC_j^k$$

Enfin :

$$|R_4(\delta,\theta)|\leq 2\delta\sum_{m=1}^{d}\left(D_{m,i}^0C_j^0+C_i^0D_{m,j}^0\right)|\theta^m-\theta_0^m|$$

D'où :

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta^j\theta^i}U(\theta,\theta_0) &= 2\delta E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\partial}{\partial\theta^j}f_{\theta_0}(Y_0)}{\sigma^2(Y_0)}\right]+R(\delta,\theta)\\
|R(\delta,\theta)| &\leq K\delta\,||\theta-\theta_0||+o(\delta)
\end{aligned}
$$

K étant une constante indépendante de δ et θ.

□

B Démonstration du lemme 5 : développement de $K(\theta_\delta)$ en fonction de δ

Démonstration : Pour simplifier, nous établierons ce résultat pour $p = 1$. Le cas général s'obtient de façon analogue.

$$K(\theta_\delta) = (K_{i,j}(\theta_\delta))_i^j\,,\;\; K_{i,j}(\theta_\delta) = E_{\theta_0}\left[\delta\frac{\partial}{\partial\theta^i}\Phi(Y_\delta, Y_0, \theta_\delta)\frac{\partial}{\partial\theta^j}\Phi(Y_\delta, Y_0, \theta_\delta)\right]$$

On montre que sous (H) $\theta \to K(\theta)$ est de classe C^2 sur Θ par dérivation sous le signe somme. Pour tout $i, j = 1, d$, la formule de Taylor-Young donne :

$$K_{i,j}(\theta_\delta) = K_{i,j}(\theta_0) + \sum_{k=1}^{d}\frac{\partial}{\partial\theta^k}K_{i,j}(\theta_0)(\theta_\delta - \theta_0) + o(\theta_\delta - \theta_0) \tag{7}$$

Développons $K_{i,j}(\theta_0)$ en fonction de δ : du fait de (3), on a,

$$K_{i,j}(\theta_\delta) = 4\delta E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0, \theta_0)\frac{\partial}{\partial\theta^j}f_{\theta_0}(Y_0, \theta_0)\frac{\left(\eta_0^1 + \tilde{W}_0^1\right)^2}{\sigma^4(Y_0)}\right]$$

- Etudions $A_0 = E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0, \theta_0)\frac{\partial}{\partial\theta^j}f_{\theta_0}(Y_0, \theta_0)\frac{(\tilde{W}_0^1)^2}{\delta^2\sigma^4(Y_0)}\right]$:

$$\begin{aligned} A_0 &= \int_0^\delta P_{i,j}(v)\,dv + 2\int_0^\delta(\delta - v)E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\sigma^2(Y_v)f'_{\theta_0}(Y_v)}{\sigma^4(Y_0)}\right] \\ &+ \int_0^\delta(\delta - v)^2E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\sigma^2(Y_v)\left(f'_{\theta_0}\right)^2(Y_v)}{\sigma^4(Y_0)}\right]\end{aligned}$$

  Deux applications de l'inégalité de Schwarz, (H 4) et (H 5) montrent qu'il existe une constante $K > 0$, indépendante de δ, telle que :

$$\begin{aligned}\left|\int_0^\delta(\delta - v)E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\sigma^2(Y_v)f'_{\theta_0}(Y_v)}{\sigma^4(Y_0)}\right]\right| &\le K\delta^2 \\ \left|\int_0^\delta(\delta - v)^2E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\sigma^2(Y_v)\left(f'_{\theta_0}\right)^2(Y_v)}{\sigma^4(Y_0)}\right]\right| &\le K\delta^3\end{aligned}$$

  Compte tenu de (H 6), on a :

$$A_0 = \delta P_{i,j}(0) + o(\delta)$$

- Appliquant à nouveau à deux reprises l'inégalité de Schwarz, on obtient :

$$\left|E_{\theta_0}\left[\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)\frac{\partial}{\partial\theta^j}f_{\theta_0}(Y_0)\frac{(\eta_0^1)^2}{\sigma^4(Y_0)}\right]\right| \le C_iC_jD\delta^4$$

  avec :

$$C_i = E_{\theta_0}\left[\left(\frac{\frac{\partial}{\partial\theta^i}f_{\theta_0}(Y_0)}{\sigma(Y_0)}\right)^4\right],\;\; D = \sup_{v\in[0,\delta_0]}\left(E_{\theta_0}\left[\left(\frac{A_{\theta_0}f_{\theta_0}(Y_v)}{\sigma(Y_0)}\right)^4\right]\right)^{\frac{1}{2}}$$

- Majorant le double produit restant grâce à l'inégalité de Schwarz, on montre qu' il est de l'ordre d'un $o\left(\delta^2\right)$.

On a donc : $K_{i,j}\left(\theta_0\right) = 4\delta^2 P_{i,j}(0) + o\left(\delta^2\right)$
Par le même type d'arguments, on a pour tout $k = 1, d$:

$$\left|\frac{\partial}{\partial\theta^k} K_{i,j}\left(\theta_0\right)\right| \leq O\left(\delta^2\right)$$

L'écart entre θ_δ et θ_0 étant en δ, (7) s'écrit : $K_{i,j}\left(\theta_\delta\right) = H_{\theta_0}\delta^2 + o\left(\delta^2\right)$.
□

Prépublications du SAMOS (depuis 1995)

1995

40 - Claude BOUZITAT, Gilles PAGES
Pour quelques images de plus..., 12 p.

41 - Jean-Claude FORT, Gilles PAGES
About the Kohonen algorithm: Strong or Weak Self-organisation? 15 p.

42 - Marie COTTRELL, Patrick LETREMY
Classification et analyse des correspondances au moyen de l'algorithme de Kohonen: application à l'étude de données socio-économiques. 10 p.

43 - Joël CHADOEUF, Xavier GUYON, Jian-Feng YAO
Sur l'ergodicité de l'estimation par Restauration-Estimation de modèles incomplètement observés. 11 p.

44 - Jean-Gabriel ATTALI, Gilles PAGES
Approximation of functions by perceptrons, a new approach. 11 p.

45 - Marie COTTRELL, Bernard GIRARD, Yvonne GIRARD, Corinne MULLER, Patrick ROUSSET
Daily electrical power curves : classification and forecasting using a Kohonen map. 8 p.

46 - Fabienne COMTE, Cécile HARDOUIN
Regression on log-regularized periodogram for fractional models at low frequencies. 19 p.

47 - Fabienne COMTE, Cécile HARDOUIN
Regression on log-regularized periodogram under assumption of bounded spectral densities: the non fractional and the fractional cases. 14 p.

48 - Patrick ROUSSET
Prévision des courbes demi-horaires au moyen d'une classification de Kohonen. 25 p.

49 - Smail IBBOU et Marie COTTRELL
Multiple Correspondence analysis of a crosstabulations matrix using the Kohonen algorithm. 6 p.

50 - Philippe JOLIVALDT
Schémas de discrétisation pour la simulation et l'estimation d'un CAR(2): une étude expérimentale. 22 p.

51 - Philippe JOLIVALDT
Utilisation de méthodes implicites pour la simulation et l'estimation de modèles CAR(2) . 14 p.

1996

52 - Samuel BAYOMOG
Estimation of a Markov field dynamic. 14 p.

53 - Morgan MANGEAS et Jian-feng YAO
Sur l'estimateur des moindres carrés des modèles auto-régressifs fonctionnels. 19 p.

54 - Marie COTTRELL, Florence PIAT, Jean-Pierre ROSPARS
A Stochastic Model for Interconnected Neurons. 17 p.

55 - Marie COTTRELL, Jean-Claude FORT, Gilles PAGES
Two or three mathematical things about the Kohonen algorithm. 31 p.

56 - Marie COTTRELL, Bernard GIRARD, Patrick ROUSSET

Forecasting of curves using a Kohonen classification. 14 p.

57 - Jean-Claude FORT, Gilles PAGES
Quantization vs Organization in the Kohonen S.O.M. 5 p.

58 - Eric de BODT, Marie COTTRELL, Michel LEVASSEUR
Réseaux de neurones en finance.33 p.

59 - Marie COTTRELL, Eric de BODT, Emmanuel HENRION, Ismaïl IBBOU, Annick WOLFS, Charles Van WYMEERSCH
Comprendre la décision à l'aide d'une carte de Kohonen. Une étude empirique. 16 p.

60 - Marie COTTRELL, Eric de BODT
Understanding the leasing decision with the help of a Kohonen map. An empirical study of the Belgian market. 5p.

61 - Marie COTTRELL, Eric de BODT, Philippe GREGOIRE
The relation between interest rate shocks and the initial rate structure: an empirical study using a Kohonen map. 16p.

62 - Marie COTTRELL, Eric de BODT, Philippe GREGOIRE
A kohonen map representation to avoid misleading interpretation. 8p.

63 - Marie COTTRELL, Eric de BODT
Analyzing shocks on the interest rate structure with Kohonen map. 6p.

64 - Marie COTTRELL, Eric de BODT, Philippe GREGOIRE
Simulating interest rate structure evolution on a long term horizon. A kohonen map application. 5p.

65 - Fabienne COMTE, Cécile HARDOUIN
Log-regularized periodogram regression. 22p.

66 - Jian Feng YAO
Simulation et optimisation sous contrainte par une dynamique de Metropolis. 6p.

67 - Morgan MANGEAS, Jian-feng YAO
On least squares estimation for nonlinear autoregressive processes. 16p.

68 - Carlo GAETAN, Xavier GUYON
Simulation des modèles de Gibbs spatiaux par chaine de Markov. 28p.

69 - François GARDES, Patrice GAUBERT, Patrick ROUSSET
Cellulage de données d'enquêtes de consommation par une méthode neuronale. 41p.

70 - Serge IOVLEFF, José R. León
High-Frequency approximation for the Helmholtz equation :a probabilistic approach. 14p.

1997

71 - Xavier Guyon et Jian-feng Yao
On description of wrong parametrisation sets of a model . 23p.

72 - Sandie Souchet
Schéma de discrétisation anticipatif et estimation du paramètre d'une diffusion. 36p.

73 - M. Cottrell, J.C. Fort, G. Pagès
Theoretic aspects of the SOM algorithm 22p.

74 - M. Benaïm, J.C. Fort & G. Pagès
Convergence of the one-dimensional Kohonen algorithm. 23p.

75 - C. Bouton & G. Pagès
About the multidimensional competitive learning vector quantization algorithm with constant gain. 36p.

76 - M. Cottrell, P. Rousset
The Kohonen algorithm: a powerful tool for analyzing and reprenting multidimensional quantitative and qualitative data 12p.

77 - M. Cottrell, B. Girard et P. Rousset
Long term forecasting by combining Kohonen algorithm and standard prevision . 6p.

78 - S. Souchet
Schéma d'approximation adapté à l'ordre *p* et estimation du paramètre de dérive d'une diffusion