

# Estimation du coefficient de diffusion de la volatilité d'un modèle à volatilité stochastique

Arnaud GLOTER

Équipe d'analyse et de mathématiques appliquées, Université de Marne-la-Vallée, bâtiment Copernic, cité Descartes, 5, boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 02, France  
Courriel : gloter@math.univ-mlv.fr

(Reçu le 2 novembre 1999, accepté le 24 novembre 1999)

---

## Résumé.

Nous nous intéressons, dans le cadre des modèles à volatilité stochastique introduits par Hull et White [6], à l'estimation du coefficient de diffusion de la volatilité. Notre observation, de type discrète, est faite sur un intervalle de temps fini et aucune hypothèse d'ergodicité n'est nécessaire pour la volatilité. Nous construisons un estimateur, montrons sa convergence et établissons que sa vitesse de convergence est en  $N^{\frac{1}{4}}$  (où  $N$  est le nombre d'observations). © 2000 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

## *Estimation of the volatility diffusion coefficient for a stochastic volatility model*

## Abstract.

We study, in the stochastic volatility model introduced by Hull and White [6], the estimation of the diffusion coefficient for the volatility process. The model is discretely observed on a fixed length time interval and no ergodicity assumption is needed for the volatility process. We construct an estimator, show its consistency and establish that its rate of convergence is  $N^{\frac{1}{4}}$  ( $N$  is the number of observations). © 2000 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

---

## *Abridged English version*

Let  $(Y, X)$  be the strong solution of the two-dimensional stochastic differential equation

$$\begin{aligned} dY_t &= \sigma_t dW_t, & Y_0 &= \eta, \\ X_t &= \sigma_t^2, & dX_t &= \theta a(X_t) dB_t + b(X_t) dt, & X_0 &= x_0 \in (0, \infty), \end{aligned}$$

where  $(B, W)$  is a standard Brownian motion on  $\mathbb{R}^2$  independent of  $\eta$  and  $\theta$  is a parameter to be estimated. The function  $a$  is known, whereas  $b$  may be known or unknown. Let  $T$  be a fixed real number, and  $n, m$  be two integers  $\geq 1$ ; we observe the sample path  $(Y_t)_{0 \leq t \leq T}$ , according to a double discretization

---

Note présentée par Paul DEHEUVELS.

## A. Gloter

at times  $(\frac{i}{n} + \frac{j}{nm})T$ ,  $i = 0, \dots, n-1$ ,  $j = 0, \dots, m$ . From these  $N = nm + 1$  observations, we define, for  $i = 0, \dots, n-1$ , the following variables, upon which we shall construct our estimator,

$$\widehat{X}_{i,n}^m = nT^{-1} \sum_{j=0}^{m-1} \left( Y_{(\frac{i}{n} + \frac{j+1}{nm})T} - Y_{(\frac{i}{n} + \frac{j}{nm})T} \right)^2.$$

Our set of assumptions for  $X$  is the following:

(A1)  $P(\forall t, X_t > 0) = 1$ .

(A2)  $a, b \in \mathcal{C}^2(0, \infty)$  and there exists  $c \geq 0$  such that for all  $x \in (0, \infty)$ :

$$|a^{-1}(x)| + |a(x)| + |a'(x)| + |a''(x)| + |b(x)| + |b'(x)| + |b''(x)| \leq c(x^{-c} + x^c).$$

Introduce the sigma field,

$$\mathcal{G}_t = \sigma((B_s, W_s), s \leq t, \eta).$$

(A3)  $\exists K_0 > 0, \forall k \in [0, K_0), \exists c(k), \forall t \geq 0, \mathbb{E}(\sup_{s \in [t, t+1]} X_s^{-k} | \mathcal{G}_t) \leq c(k)(1 + X_t^{-k})$ ;

$\exists K_\infty > 0, \forall k \in [0, K_\infty), \exists c(k), \forall t \geq 0, \mathbb{E}(\sup_{s \in [t, t+1]} X_s^k | \mathcal{G}_t) \leq c(k)(1 + X_t^k)$ .

This last assumption is rather technical, but may be checked for many usual models of diffusion for  $X$ . Let us stress that no assumption about ergodicity of  $X$  is needed in our work.

First, we study the behaviour of the quadratic variation of  $(\widehat{X}_{i,n}^{m_n})$  as  $n$  and  $m$  tend to  $\infty$ .

PROPOSITION. – Assume (A1)–(A3) with  $K_0 = K_\infty = \infty$  and let  $m_n$  be a sequence of integers such that  $\frac{\sqrt{n}}{m_n} \xrightarrow{n \rightarrow \infty} 0$ , then

$$\sum_{i=0}^{n-2} (\widehat{X}_{i+1,n}^{m_n} - \widehat{X}_{i,n}^{m_n})^2 = \frac{2}{3} \int_0^T \theta^2 a^2(X_s) ds + \frac{4n}{m_n T} \int_0^T X_s^2 ds + o_{\mathbf{P}}(1).$$

This proposition suggests to introduce the following estimator,

$$\widehat{\theta}_{n,m}^2 = \frac{3}{2} \sum_{i=0}^{n-2} \left\{ \frac{(\widehat{X}_{i+1,n}^m - \widehat{X}_{i,n}^m)^2}{a^2(\widehat{X}_{i,n}^m)} - \frac{4}{m} \frac{(\widehat{X}_{i,n}^m)^2}{a^2(\widehat{X}_{i,n}^m)} \right\}.$$

THEOREM. – Assume (A1)–(A3) with  $K_0 = K_\infty = \infty$  and  $n^{\frac{2}{3}} m_n^{-1} \xrightarrow{n \rightarrow \infty} 0$ . Then,  $\widehat{\theta}_{n,m_n}^2 \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \theta^2$ .

Furthermore, we have the following expansion:

$$\sqrt{n}(\widehat{\theta}_{n,m_n}^2 - \theta^2) = Z_n + ((m_n^{-1}n)^{\frac{1}{2}} + (m_n^{-1}n)^{\frac{3}{2}})B_n,$$

where  $Z_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \frac{9}{4}\theta^4)$  and  $(B_n)$  is a tight sequence.

In the special case  $m_n = n$ , we deduce that  $\sqrt{n}(\widehat{\theta}_{n,m_n}^2 - \theta^2)$  is tight. Hence, our estimator, based on  $N = n^2 + 1$  observations of  $Y$ , has rate  $n^{\frac{1}{2}}$ , which means a rate  $N^{\frac{1}{4}}$ .

If  $nm_n^{-1} \rightarrow 0$ , the estimator is asymptotically normal, but with a rate slower than  $N^{\frac{1}{4}}$ .

## 1. Modèle, hypothèses et observations

Considérons le couple  $(Y, X)$  solution de l'équation différentielle stochastique :

## Estimation du coefficient de diffusion de la volatilité d'un modèle à volatilité stochastique

$$dY_t = \sigma_t dW_t, \quad Y_0 = \eta, \quad (1)$$

$$X_t = \sigma_t^2, \quad dX_t = \theta a(X_t) dB_t + b(X_t) dt, \quad X_0 = x_0 \in (0, \infty), \quad (2)$$

où  $(B, W)$  est un mouvement brownien standard de  $\mathbb{R}^2$ , indépendant de la variable aléatoire  $\eta$  et  $\theta \in (0, \infty)$  est un paramètre que l'on cherche à estimer. La fonction  $a$  est supposée connue, mais  $b$  et  $x_0$  ne le sont pas nécessairement (en particulier,  $b$  peut dépendre de  $\theta$ ).

Nous observons une discrétisation de  $(Y_t)_{t \in [0, T]}$ , de pas  $\frac{T}{nm}$ , où  $n$  et  $m$  sont des entiers supérieurs ou égaux à 1 et  $T > 0$  est un réel fixé. Notre échantillon est donc  $(Y_{(\frac{i}{n} + \frac{j}{nm})T})_{0 \leq i \leq n-1, 0 \leq j \leq m}$ , constitué de  $N = nm + 1$  valeurs. Notre asymptotique sera de faire tendre  $n$  et  $m = m_n$  vers l'infini.

Cette double discrétisation permet d'introduire, pour  $i = 0, \dots, n-1$ , les variables observées suivantes :

$$\widehat{X}_{i,n}^m = nT^{-1} \sum_{j=0}^{m-1} \left( Y_{(\frac{i}{n} + \frac{j+1}{nm})T} - Y_{(\frac{i}{n} + \frac{j}{nm})T} \right)^2 \quad (3)$$

qui sont les approximations des variables inobservées,

$$\overline{X}_{i,n} = nT^{-1} \int_{\frac{i}{n}T}^{\frac{i+1}{n}T} X_s ds. \quad (4)$$

Introduisons l'erreur d'approximation entre ces deux quantités :

$$E_{i,n,m} = \widehat{X}_{i,n}^m - \overline{X}_{i,n}. \quad (5)$$

Rappelons que de nombreux auteurs (par exemple [3,7]) ont étudié l'estimation des paramètres de  $X$  lorsque seule  $Y$  est observée. Dans ces travaux,  $X$  est supposée être ergodique et le temps d'observation tend vers l'infini. Aucune de ces deux hypothèses n'est faite dans notre cas.

Nos hypothèses sur la diffusion  $X$  sont les suivantes :

(A1) l'équation (2) admet une unique solution forte  $(X_t)$  qui vérifie  $P(\forall t, X_t > 0) = 1$  ;

(A2)  $a, b \in \mathcal{C}^2(0, \infty)$  et il existe  $c \geq 0$  tel que pour tout  $x \in (0, \infty)$  :

$$|a^{-1}(x)| + |a(x)| + |a'(x)| + |a''(x)| + |b(x)| + |b'(x)| + |b''(x)| \leq c(x^{-c} + x^c).$$

Dans ce qui suit, la notation  $c$  représentera une constante générique, qui sera toujours indépendante des entiers  $i, n, m$ .

Introduisons la tribu

$$\mathcal{G}_t = \sigma((B_s, W_s), s \leq t, \eta). \quad (6)$$

(A3)  $\exists K_0 > 0, \forall k \in [0, K_0), \exists c, \forall t \geq 0, \quad E(\sup_{s \in [t, t+1]} X_s^{-k} | \mathcal{G}_t) \leq c(1 + X_t^{-k})$  ;

$\exists K_\infty > 0, \forall k \in [0, K_\infty), \exists c, \forall t \geq 0, \quad E(\sup_{s \in [t, t+1]} X_s^k | \mathcal{G}_t) \leq c(1 + X_t^k)$ .

Cette dernière hypothèse exprime que la diffusion ne s'approche pas trop des bords 0 et  $\infty$ . Elle est un outil important dans nos preuves. On peut par exemple vérifier (A1)–(A3) lorsque  $X$  est l'une des diffusions suivantes :

–  $dX_t = \mu(X_t - \mu') dt + \theta X_t^\psi dB_t$ , pour  $\mu < 0, \mu' > 0$  et, soit,  $\psi \in (1/2, 1]$ , soit,  $\psi = 1/2$  et  $2|\mu|\mu'/\theta^2 > 1$ .

Si  $\psi > 1/2$ , alors les valeurs de  $K_0, K_\infty$  sont  $K_0 = K_\infty = \infty$ . Si  $\psi = 1/2$ , alors  $K_0 = 2|\mu|\mu'/\theta^2 - 1, K_\infty = \infty$ .

–  $dX_t = \exp(Z_t)$ , où  $dZ_t = \mu Z_t dt + \theta dB_t$ , pour  $\mu \leq 0$ . Ici,  $K_0 = K_\infty = \infty$ . (En particulier, si  $\mu = 0, X_t = \exp(\theta B_t)$  est une diffusion transiente.)

## 2. Résultats

Nous étudions les variables  $\overline{X}_{i,n}$  et donnons des majorations de l'erreur  $E_{i,n,m}$ . Puis, nous déduisons un estimateur de  $\theta^2$ .

### 2.1. Développements asymptotiques

PROPOSITION 2.1. – *Introduisons  $\mathcal{G}_i^n = \mathcal{G}_{\frac{i}{n}T}$  (voir(6)) et supposons (A1)–(A3) avec  $K_0 = K_\infty = \infty$ , alors on a le développement suivant :*

$$\overline{X}_{i+1,n} - \overline{X}_{i,n} = \theta a\left(X_{\frac{i}{n}T}\right) U_{i,n} \left(\frac{T}{n}\right)^{\frac{1}{2}} + \varepsilon_{i,n}, \quad (7)$$

où  $\exists c, \forall i, n \geq 0, E(\varepsilon_{i,n}^2 | \mathcal{G}_i^n) \leq cn^{-2}(X_{\frac{i}{n}T}^c + X_{\frac{i}{n}T}^{-c})$  et

$$U_{i,n} = \left(\frac{T}{n}\right)^{-\frac{3}{2}} \int_{\frac{i}{n}T}^{\frac{i+1}{n}T} \left(s - \frac{i}{n}T\right) dB_s + \left(\frac{T}{n}\right)^{-\frac{3}{2}} \int_{\frac{i+1}{n}T}^{\frac{i+2}{n}T} \left(\frac{i+2}{n}T - s\right) dB_s.$$

Par un simple calcul, on vérifie que  $(U_{i,n})_{i=0,\dots,n-1}$  est un vecteur gaussien centré avec la structure de covariance d'un processus MA(1) :

$$\text{cov}(U_{i,n}, U_{j,n}) = \begin{cases} 2/3 & \text{si } i = j, \\ 1/6 & \text{si } |i - j| = 1, \\ 0 & \text{si } |i - j| \geq 2. \end{cases} \quad (8)$$

Grâce à ce développement, on peut établir le résultat suivant sur la variation quadratique du processus  $(\overline{X}_{i,n})$  (voir [5]).

COROLLAIRE 2.2. – *Supposons (A1)–(A3) avec  $K_0 = K_\infty = \infty$ , alors*

$$\sum_{i=0}^{n-2} (\overline{X}_{i+1,n} - \overline{X}_{i,n})^2 = \frac{2}{3} \int_0^T \theta^2 a^2(X_s) ds + o_P(1).$$

Remarque 2.3. – 1) Dans le cas où  $X$  est un processus d'Ornstein–Uhlenbeck, solution de  $dX_t = \mu X_t dt + \theta dB_t$ , il est établi dans [4], que l'on a le développement exact suivant, analogue à (7),

$$\overline{X}_{i+1,n} - e^{\mu \frac{T}{n}} \overline{X}_{i,n} = \frac{\theta}{\mu} \int_{\frac{i}{n}T}^{\frac{i+1}{n}T} \left(e^{\mu \frac{T}{n}} - e^{\mu(\frac{i+1}{n}T-s)}\right) dB_s + \frac{\theta}{\mu} \int_{\frac{i+1}{n}T}^{\frac{i+2}{n}T} \left(e^{\mu(\frac{i+2}{n}T-s)} - 1\right) dB_s.$$

2) La variation quadratique des  $(\overline{X}_{i,n}, i \leq n-2)$  a donc un comportement très différent de celle des  $(X_{\frac{i}{n}T}, i \leq n-1)$  puisqu'il apparaît le facteur  $\frac{2}{3}$  provenant de la variance des  $U_{i,n}$ .

Nous pouvons calculer des moments conditionnels, pour l'erreur  $E_{i,n,m}$  (voir (5)).

PROPOSITION 2.4. – *Introduisons la tribu*

$$\mathcal{G}_{i,B}^n = \sigma\left(B_u, u \geq 0, W_s, s \leq \frac{i}{n}T, \eta\right). \quad (9)$$

Supposons (A1)–(A3), avec  $K_0 = K_\infty = \infty$ . L'erreur vérifie les relations suivantes :

$$\begin{aligned} \mathbb{E}(E_{i,n,m} | \mathcal{G}_{i,B}^n) &= 0, \\ \mathbb{E}(E_{i,n,m}^2 | \mathcal{G}_{i,B}^n) &= 2m^{-1}X_{\frac{i}{n}T}^2 + r_{i,n,m}, \\ \mathbb{E}(E_{i+1,n,m}^2 | \mathcal{G}_{i,B}^n) &= 2m^{-1}X_{\frac{i}{n}T}^2 + s_{i,n,m}. \end{aligned}$$

Les deux termes  $r_{i,n,m}$  et  $s_{i,n,m}$  vérifient :

$$\exists c, \forall n, m \geq 0, \forall i \leq n, \quad \mathbb{E}(r_{i,n,m}^2 + s_{i,n,m}^2 | \mathcal{G}_{i,B}^n) \leq cn^{-1}m^{-2}(1 + X_{\frac{i}{n}T}^c).$$

Par cette proposition, on voit que l'erreur  $E_{i,n,m}$  est d'ordre  $m^{-\frac{1}{2}}$  et que sa variance conditionnelle ne dépend, au premier ordre, ni de la fonction  $b$ , ni de la fonction  $\theta a$ , et donc ne dépend pas de  $\theta$ .

## 2.2. Estimation du paramètre

Grâce aux développements du paragraphe précédent, on peut établir le comportement asymptotique de la variation quadratique de  $(\hat{X}_{i,n}^m, i \leq n-2)$ .

PROPOSITION 2.5. – Supposons (A1)–(A3) avec  $K_0 = K_\infty = \infty$  et soit  $m_n$  une suite d'entiers tels que  $\frac{\sqrt{n}}{m_n} \xrightarrow{n \rightarrow \infty} 0$ , alors

$$\sum_{i=0}^{n-2} (\hat{X}_{i+1,n}^{m_n} - \hat{X}_{i,n}^{m_n})^2 = \frac{2}{3} \int_0^T \theta^2 a^2(X_s) ds + \frac{4n}{m_n T} \int_0^T X_s^2 ds + o_{\mathbf{P}}(1).$$

*Éléments de démonstration.* – On écrit le développement suivant pour les  $\hat{X}_{i,n}^m$ , basé sur la proposition 2.1 et (5).  $\hat{X}_{i+1,n}^m - \hat{X}_{i,n}^m = \theta a(X_{\frac{i}{n}T})U_{i,n}(\frac{T}{n})^{\frac{1}{2}} + E_{i+1,n,m} - E_{i,n,m} + \varepsilon_{i,n}$ . On voit donc, par (8) et la proposition 2.4 que  $\mathbb{E}((\hat{X}_{i+1,n}^m - \hat{X}_{i,n}^m)^2 | \mathcal{G}_i^n) = \frac{2}{3} a^2(X_{\frac{i}{n}T})\frac{T}{n} + \frac{4}{m} X_{\frac{i}{n}T}^2 + a_{i,n}$ , où  $\sup_{i,n} \mathbb{E}(|a_{i,n}|) \leq cn^{-\frac{3}{2}}$ .

Grâce à un théorème limite de martingales, nous déduisons le résultat souhaité.  $\square$

On voit, en particulier, que cette variation quadratique peut tendre vers l'infini (à cause de la variance des  $E_{i,n,m}$ ), si  $nm_n^{-1}$  n'est pas bornée.

Ce résultat nous conduit à introduire l'estimateur de  $\theta^2$  suivant :

$$\hat{\theta}_{n,m}^2 = \frac{3}{2} \sum_{i=0}^{n-2} \left\{ \frac{(\hat{X}_{i+1,n}^m - \hat{X}_{i,n}^m)^2}{a^2(\hat{X}_{i,n}^m)} - \frac{4}{m} \frac{(\hat{X}_{i,n}^m)^2}{a^2(\hat{X}_{i,n}^m)} \right\}.$$

Remarquons que cet estimateur est très différent de celui basé sur une observation discrète de  $X$  (voir [1,2])

$$\theta_n^{2*} = \sum_{i=0}^{n-1} \frac{(X_{\frac{i+1}{n}T} - X_{\frac{i}{n}T})^2}{a^2(X_{\frac{i}{n}T})}.$$

Une difficulté apparaît lors de l'étude de  $\hat{\theta}_{n,m}^2$ , car  $a^{-1}$  peut ne pas être bornée en 0 et même si  $X$  reste à distance finie positive de 0, son approximation  $\hat{X}_{i,n}^m$  peut prendre des valeurs proches de 0. Par exemple, si  $m = 1$ , conditionnellement à  $\mathcal{G}_{0,B}^n$  (voir (9)),  $\hat{X}_{i,n}^1$  est le carré d'une variable gaussienne. Donc  $\mathbb{E}((\hat{X}_{i,n}^1)^{-k} | \mathcal{G}_{0,B}^n) = \infty$  presque sûrement dès que  $k > \frac{1}{2}$ .

Pour étudier notre estimateur, nous avons établi le lemme suivant.

## A. Gloter

LEMME 2.6. – Supposons (A1)–(A3), alors on a l'inégalité suivante :

$$\forall k \in [0, K_0), \exists c, \forall m \geq 2k + 3, \forall i, n \geq 0, \quad E((\widehat{X}_{i,n}^m)^{-k} | \mathcal{G}_i^n) \leq c (1 + X_{\frac{i}{n}T}^{-k}). \quad (10)$$

Nous montrons, alors, la convergence de  $\widehat{\theta}_{n,m}^2$  et étudions sa loi asymptotique.

THÉORÈME 2.7. – Supposons (A1)–(A3) avec  $K_0 = K_\infty = \infty$  et  $n^{\frac{2}{3}} m_n^{-1} \xrightarrow{n \rightarrow \infty} 0$ . Alors,

$$\widehat{\theta}_{n,m_n}^2 \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \theta^2.$$

De plus, la loi de l'estimateur se décompose en :

$$\sqrt{n} (\widehat{\theta}_{n,m_n}^2 - \theta^2) = Z_n + ((m_n^{-1}n)^{\frac{1}{2}} + (m_n^{-1}n)^{\frac{3}{2}}) B_n,$$

où  $Z_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \frac{9}{4}\theta^4)$  et  $(B_n)$  est une suite tendue.

*Éléments de démonstration.* – On pose  $Z_n = \sqrt{n}(\widetilde{\theta}_n^2 - \theta^2)$ , où  $\widetilde{\theta}_n^2 = \frac{3}{2} \sum_{i=0}^{n-2} \frac{(\overline{X}_{i+1,n} - \overline{X}_{i,n})}{a^2(\overline{X}_{i,n})}$ . En utilisant la proposition 2.1, on démontre que  $Z_n$  est asymptotiquement gaussien. Ensuite, grâce à la proposition 2.4 et au lemme 2.6, on étudie la différence entre  $\sqrt{n}(\widehat{\theta}_{n,m_n}^2 - \theta^2)$  et  $Z_n$ , il apparaît le terme  $B_n$  qui contient les erreurs d'approximation  $E_{i,n,m}$ .  $\square$

*Remarque 2.8.* – 1) Si l'on choisit, en particulier,  $m_n = n$ , alors la suite  $n^{\frac{1}{2}}(\widehat{\theta}_{n,n}^2 - \theta^2)$  est tendue. Notre estimateur est donc à vitesse  $n^{\frac{1}{2}}$ , mais par rapport au nombre total d'observations,  $N = n^2 + 1$ , sa vitesse est  $N^{\frac{1}{4}}$ . Nous ne savons pas quelle est la vitesse exacte d'estimation du paramètre  $\theta$  dans notre modèle.

2) Si  $nm_n^{-1} \rightarrow 0$ , l'estimateur est asymptotiquement normal, mais sa vitesse est plus lente que  $N^{\frac{1}{4}}$ .

## Références bibliographiques

- [1] Dohnal G., On estimating the diffusion coefficient, J. Appl. Probab. 24 (1987) 105–114.
- [2] Genon-Catalot V., Jacod J., On the estimation of the diffusion coefficient for multidimensional diffusion processes, Ann. Inst. H.-Poincaré, Probab.-Statist. 29 (1993) 119–151.
- [3] Genon-Catalot V., Jeantheau T., Laredo C., Parameter estimation for discretely observed stochastic volatility models, Bernoulli 5 (5) (1999) 855–872.
- [4] Gloter A., Parameter estimation for a discrete sampling on an integrated Ornstein–Uhlenbeck process, Prépublication de l'université de Marne-la-Vallée, n° 14/98, 1998.
- [5] Gloter A., Discrete sampling of an integrated diffusion process and a comparison with the Euler scheme, Prépublication de l'université de Marne-la-Vallée, n° 23/98, 1998.
- [6] Hull J., White A., The pricing of options on assets with stochastic volatilities, J. Finance 42 (1987) 281–300.
- [7] Sørensen M., Prediction-based estimating functions, Preprint, 1999.