



Recursive estimation for continuous time stochastic volatility models

H. Gong^a, A. Thavaneswaran^{b,*}

^a Department of Statistics, Temple University, Philadelphia, PA 19122, USA

^b Department of Statistics, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

ARTICLE INFO

Article history:

Received 4 March 2009

Received in revised form 20 June 2009

Accepted 22 June 2009

Keywords:

Recursive estimation

Stochastic volatility

Ito's formula

ABSTRACT

Volatility plays an important role in portfolio management and option pricing. Recently, there has been a growing interest in modeling volatility of the observed process by nonlinear stochastic process [S.J. Taylor, *Asset Price Dynamics, Volatility, and Prediction*, Princeton University Press, 2005; H. Kawakatsu, Specification and estimation of discrete time quadratic stochastic volatility models, *Journal of Empirical Finance* 14 (2007) 424–442]. In [H. Gong, A. Thavaneswaran, J. Singh, Filtering for some time series models by using transformation, *Math Scientist* 33 (2008) 141–147], we have studied the recursive estimates for discrete time stochastic volatility models driven by normal errors. In this paper, we study the recursive estimates for various classes of continuous time nonlinear non-Gaussian stochastic volatility models used for option pricing in finance.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

In the last two decades, volatility models have received considerable attention with the emphasis being placed on state space models. From an econometric standpoint time-varying volatility models have been widely developed, recognizing the essence that the volatility and the correlation of assets change over time (see for example Heston and Nandi [1]). Although state space models in which the conditional mean of the observed process is modeled as stochastic process are useful in parameter estimation, it is widely recognized that *stochastic volatility models*, which model the volatility as a stochastic process [13,14], should be employed to estimate the volatility parameters.

A filtering procedure has been suggested for discrete time stochastic volatility models, for instance, see Gong et al. [2], and Kirby [3] with normal errors. For stochastic volatility models with time-varying parameters, when a new observation is coming in, a new volatility parameter is added, and hence it is almost impossible to estimate the time-varying volatility parameters. In order to construct an optimal recursive estimate for *non-normal stochastic volatility models*, we start with the following discrete time example.

Consider a nonlinear state space model given in Shiryaev [4]

$$\begin{aligned}\theta_{t+1} &= a\theta_t + (1 + \theta_t)\eta_{t+1} \\ y_{t+1} &= A\theta_t + z_{t+1}\end{aligned}\quad (1.1)$$

where $z_t \stackrel{iid}{\sim} N(0, \sigma_z^2)$, $\eta_t \stackrel{iid}{\sim} N(0, \sigma_\eta^2)$, and the sequences $\{z_t\}$ and $\{\eta_t\}$ are independent, where $\{y_t\}$ process is observed and all θ_t is the parameter process. Then in Gong et al. [2], we have the estimate $\hat{\theta}_{t+1}$ and $\gamma_{t+1} = E[(\theta_{t+1} - \hat{\theta}_{t+1})^2 | F_{t+1}^y]$ as

$$\hat{\theta}_{t+1} = a\hat{\theta}_t + \frac{Aa\gamma_t}{A^2\gamma_t + \sigma_z^2}(y_{t+1} - A\hat{\theta}_t)\quad (1.2)$$

* Corresponding author. Tel.: +1 204 474 8984; fax: +1 204 474 7624.

E-mail address: thavane@cc.umanitoba.ca (A. Thavaneswaran).

$$\gamma_{t+1} = a^2\gamma_t + b_1^2\sigma_\eta^2 - \frac{(aA\gamma_t)^2}{A^2\gamma_t + \sigma_z^2} = \frac{a^2\sigma_z^2\gamma_t}{A^2\gamma_t + \sigma_z^2} + b_1^2\sigma_\eta^2. \tag{1.3}$$

where $b_1 = \sqrt{E(1 + \theta_t)^2}$. However, the recursive system above is based on two very important assumptions: $z_t \stackrel{iid}{\sim} N(0, \sigma_z^2)$, $\eta_t \stackrel{iid}{\sim} N(0, \sigma_\eta^2)$. If the error terms in (1.1) are not normally distributed then we cannot interpret the recursive estimates as a conditional mean. Here, as an alternative we propose an *optimal (minimum mean square error) estimate* without the Gaussian assumptions.

Lemma 1.1. *In the class of all estimates of the form $\hat{\theta}_{t+1} = a\hat{\theta}_t + G_t(y_{t+1} - A\hat{\theta}_t)$, the G_t which minimizes the mean square error $\gamma_t = E[(\theta_t - \hat{\theta}_t)^2 | \mathcal{F}_t^y]$ is given by $\hat{G}_t = \frac{aA\gamma_t}{A^2\gamma_t + \sigma_z^2}$ and the recursive estimates are given by*

$$\hat{\theta}_{t+1} = a\hat{\theta}_t + \frac{Aa\gamma_t}{A^2\gamma_t + \sigma_z^2}(y_{t+1} - A\hat{\theta}_t) \tag{1.4}$$

$$\gamma_{t+1} = \frac{a^2\sigma_z^2\gamma_t}{A^2\gamma_t + \sigma_z^2} + \left(\frac{1 - a^2}{1 - a^2 - \sigma_\eta^2}\right)\sigma_\eta^2. \tag{1.5}$$

Proof. The difference of $\theta_{t+1} - \hat{\theta}_{t+1} = a(\theta_t - \hat{\theta}_t) - G_t(y_{t+1} - A\hat{\theta}_t) + (1 + \theta_t)\eta_{t+1}$. Then

$$\gamma_{t+1} = E[(\theta_{t+1} - \hat{\theta}_{t+1})^2 | \mathcal{F}_{t+1}^y] = (a - AG_t)^2\gamma_t + \left(\frac{1 - a^2}{1 - a^2 - \sigma_\eta^2}\right)\sigma_\eta^2 + \sigma_z^2G_t^2, \tag{1.6}$$

by differentiating γ_{t+1} and set the first derivative function to zero,

$$\frac{\partial \gamma_{t+1}}{\partial G_t} = -2A(a - AG_t)\gamma_t + 2\sigma_z^2G_t \stackrel{\text{set}}{=} 0.$$

We have $\hat{G}_t = \frac{aA\gamma_t}{A^2\gamma_t + \sigma_z^2}$. The second derivative of γ_t with respect to G_t is positive, hence γ_t attains its minimum value at \hat{G}_t . \square

The optimal linear estimate of θ_t and the MSE γ_t turn out to be the same as the one given in Gong et al. [2]. However, we do not make any distributional assumptions to obtain the optimal estimates. Moreover, it is of interest to note that if we let $b_1 = \sqrt{E(1 + \theta_t)^2}$, then $b_1^2 = \frac{1 - a^2}{1 - a^2 - \sigma_\eta^2}$, and γ_{t+1} in (1.5) turns out to be $\frac{a^2\sigma_z^2\gamma_t}{A^2\gamma_t + \sigma_z^2} + b_1^2\sigma_\eta^2$. In the remainder of this paper, Section 2 discusses the application of the optimal MSE approach to continuous time SV models, with several examples for illustration.

2. Continuous time SV models

Let (Ω, \mathcal{F}, p) be a complete probability space, and let $(\mathcal{F}_t), 0 \leq t \leq T$, be a non-decreasing family of right continuous σ -algebra of \mathcal{F} , satisfying the usual conditions.

$$d\theta_t = f(\theta_t)dt + dm_t \tag{2.1}$$

$$dy_t = h(\theta_t)dt + dM_t \tag{2.2}$$

where y_t is the observed process and θ_t the unobserved volatility process, m_t, M_t are uncorrelated square integrable martingales with variance process $\langle m, m \rangle_t = \int_0^t \sigma_1^2(s, y_s)ds, \langle M, M \rangle_t = \int_0^t \sigma_2^2(s, y_s)ds, f(\theta_t), h(\theta_t) \in C^1$ (continuously differentiable), $\sup_{t \leq T} E[f^2(\theta_t)] < \infty, \sup_{t \leq T} E[h^2(\theta_t)] < \infty$ and then apply to recently proposed stochastic volatility models.

Theorem 2.1.

$$d\hat{\theta}_t = f(\hat{\theta}_t)dt + \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)}(dy_t - h(\hat{\theta}_t)dt), \tag{2.3}$$

$$\frac{d\gamma_t}{dt} = \dot{\gamma}_t = 2f^{(1)}(\hat{\theta}_t)\gamma_t + \sigma_1^2(t, y_t) - \frac{[h^{(1)}(\hat{\theta}_t)]^2}{\sigma_2^2(t, y_t)}\gamma_t^2, \tag{2.4}$$

where $\gamma_t = E[(\theta_t - \hat{\theta}_t)^2 | \mathcal{F}_t^y]$.

Proof. Eq. (2.3) is the continuous time analog of Eq. (1.2) and the detailed proof of (2.3) using estimating functions follows from Thompson and Thavaneswaran [5] or Thavaneswaran and Thompson [6].

For the proof of (2.4), let $\delta_t = \theta_t - \hat{\theta}_t$, and then $\gamma_t = E[(\theta_t - \hat{\theta}_t)^2 | \mathcal{F}_t^y] = E(\delta_t^2)$. The stochastic differential equation of δ_t is obtained by

$$\begin{aligned} d\delta_t &= d\theta_t - d\hat{\theta}_t \\ &= f(\theta_t)dt + dm_t - f(\hat{\theta}_t)dt - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} [dy_t - h(\hat{\theta}_t)dt] \\ &= (f(\theta_t) - f(\hat{\theta}_t))dt + dm_t - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} [(h(\theta_t) - h(\hat{\theta}_t))dt + dM_t]. \end{aligned}$$

Applying Itô's formula to δ_t^2 , and using first order approximation for $f(\theta_t)$ and $h(\theta_t)$,

$$\begin{aligned} d\delta_t^2 &= 2\delta_t \left[dm_t - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} dM_t \right] + \left[2\delta_t \left((f(\theta_t) - f(\hat{\theta}_t)) - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} (h(\theta_t) - h(\hat{\theta}_t)) \right) \right. \\ &\quad \left. + \frac{1}{2}(2) \left(\sigma_1^2(t, y_t) + \frac{\sigma_t^2 (h^{(1)}(\hat{\theta}_t))^2}{\sigma_2^4(t, y_t)} \sigma_2^2(t, y_t) \right) \right] dt \\ &= 2\delta_t \left[dm_t - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} dM_t \right] + \left[2\delta_t (\theta_t - \hat{\theta}_t) \left(f^{(1)}(\hat{\theta}_t) - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} \cdot h^{(1)}(\hat{\theta}_t) \right) \right. \\ &\quad \left. + \frac{1}{2}(2) \left(\sigma_1^2(t, y_t) + \frac{\sigma_t^2 (h^{(1)}(\hat{\theta}_t))^2}{\sigma_2^4(t, y_t)} \sigma_2^2(t, y_t) \right) \right] dt \\ &= 2\delta_t \left[dm_t - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} dM_t \right] + \left[2\delta_t^2 \left(f^{(1)}(\hat{\theta}_t) - \frac{\gamma_t h^{(1)}(\hat{\theta}_t)}{\sigma_2^2(t, y_t)} \cdot h^{(1)}(\hat{\theta}_t) \right) \right. \\ &\quad \left. + \left(\sigma_1^2(t, y_t) + \frac{\sigma_t^2 (h^{(1)}(\hat{\theta}_t))^2}{\sigma_2^4(t, y_t)} \sigma_2^2(t, y_t) \right) \right] dt \tag{2.5} \end{aligned}$$

and here

$$\begin{aligned} \gamma_t &= \gamma_0 + \int_0^t \left(2\gamma_s \left(f^{(1)}(\hat{\theta}_s) - \frac{\gamma_s (h^{(1)}(\hat{\theta}_s))^2}{\sigma_2^2(s, y_s)} \right) + \sigma_1^2(s, y_s) + \frac{\gamma_s^2 (h^{(1)}(\hat{\theta}_s))^2}{\sigma_2^2(s, y_s)} \right) ds \\ &= \gamma_0 + \int_0^t \left(2\gamma_s f^{(1)}(\hat{\theta}_s) + \sigma_1^2(s, y_s) - \frac{\gamma_s^2 (h^{(1)}(\hat{\theta}_s))^2}{\sigma_2^2(s, y_s)} \right) ds. \quad \square \end{aligned}$$

Corollary. For the following Gaussian state space model:

$$\begin{aligned} d\theta_t &= a(t, y)\theta_t dt + b(t, y)dW_1(t) \\ dy_t &= A(t, y)\theta_t dt + B(t, y)dW_2(t) \end{aligned}$$

where $W_1(t)$ and $W_2(t)$ are two Wiener processes, and $\int_0^t b^2(s, y)ds < \infty$ and $\int_0^t B^2(s, y)ds < \infty$.

Setting $\sigma_1^2(t, y_t) = b^2(t, y)$, $\sigma_2^2(t, y_t) = B^2(t, y)$, $f(\theta_t) = a(t, y)\theta_t$, and $h(\theta_t) = A(t, y)\theta_t$ in (2.1) and (2.2), the recursive estimates turn out to the one given by Liptser and Shiryaev [15]:

$$\begin{aligned} d\hat{\theta}_t &= a(t, y)\hat{\theta}_t dt + \frac{\gamma_t A(t, y)}{B^2(t, y)} [dy_t - A(t, y)\hat{\theta}_t dt], \\ \dot{\gamma}_t &= 2a(t, y)\gamma_t + b^2(t, y) - \frac{A^2(t, y)\gamma_t^2}{B^2(t, y)}. \end{aligned}$$

2.1. Klebaner's model

Klebaner [7] considered the following state space model

$$\begin{aligned} d\theta_t &= \left(\mu + \frac{1}{2}\sigma^2 \right) \theta_t dt + \sigma\theta_t dW_{2t} \\ dy_t &= \theta_t dt + dW_{1t}, \end{aligned}$$

where $\{W_{1t}\}$ and $\{W_{2t}\}$ are two independent wiener processes, μ and σ are constants. Let $b_1 = \sqrt{E(\sigma\theta_t)^2}$ and $W_{2t}^* = \frac{\theta_t}{b_1} W_{2t}$, and then $d\theta_t = (\mu + \frac{1}{2}\sigma^2)\theta_t dt + \sigma b_1 dW_{2t}^*$. Then $\sigma_1^2(t, y_t) = \sigma^2 b_1^2$, $\sigma_2^2(t, y_t) = 1$, $f(\theta_t) = (\mu + \frac{1}{2}\sigma^2)\theta_t$, and $h(\theta_t) = \theta_t$.

Hence the optimal linear recursive estimates are given by

$$\begin{aligned} d\hat{\theta}_t &= \left(\mu + \frac{1}{2}\sigma^2\right)\hat{\theta}_t dt + \gamma_t[dy_t - dt], \\ \dot{\gamma}_t &= 2\left(\mu + \frac{1}{2}\sigma^2\right)\gamma_t + \sigma^2 b_1^2 - \gamma_t^2. \end{aligned}$$

2.2. Hull & White model

For the Hull and White [8] SV model of the form

$$d\theta_t^2 = a\theta_t^2 dt + b\theta_t^2 dW_{2t} \tag{2.6}$$

$$dy_t = \alpha y_t dt + \theta_t y_t dW_{1t} \tag{2.7}$$

where, α , a and b are constants, and $\{W_{1t}\}$ and $\{W_{2t}\}$ are two independent wiener processes. Since in (2.7) the drift term does not include the parameter process θ_t , in order to use Theorem 2.1, we can apply Itô formula to y_t^2 .

$$dy_t^2 = (2\alpha y_t^2 + \theta_t^2 y_t^2)dt + 2\theta_t y_t^2 dW_{1t}. \tag{2.8}$$

The diffusion term in (2.6) and (2.8) contain the parameter process θ_t , and by setting $b_1^2 = E(\theta_t^2)$ and $b_2^2 = E(\theta_t^4)$, the model becomes

$$d\theta_t^2 = a\theta_t^2 dt + b b_2 dW_{2t}^*$$

$$dy_t^2 = (2\alpha y_t^2 + \theta_t^2 y_t^2)dt + 2b_1 y_t^2 dW_{1t}^*$$

Hence by Theorem 2.1, the recursive estimates are given by

$$d\hat{\theta}_t^2 = a\hat{\theta}_t^2 + \frac{\gamma_t}{4b_1^2 y_t^2} [dy_t^2 - (2\alpha y_t^2 + \hat{\theta}_t^2 y_t^2)dt]$$

$$\dot{\gamma}_t = 2\alpha \gamma_t + b^2 b_2^2 - \frac{1}{4b_1^2} \gamma_t^2.$$

2.3. Heston's SV model

Stein and Stein [9] and Heston [10] considered the SV model of the form

$$d\theta_t = k(\beta - \theta_t)dt + cdW_{2t} \tag{2.9}$$

$$dy_t = \alpha y_t dt + \theta_t y_t dW_{1t} \tag{2.10}$$

where, α , β , k , and c are constants, and $\{W_{1t}\}$ and $\{W_{2t}\}$ are two independent wiener processes. Since in (2.10) the drift term does not include the parameter process θ_t , in order to use Theorem 2.1, we first apply Itô formula to $\log y_t$ and then setting $b_1^2 = E(\theta_t^2)$,

$$d \log y_t = \left(\alpha - \frac{1}{2}\theta_t^2\right)dt + b_1 y_t^2 dW_{1t}^*.$$

Then by Theorem 2.1, the recursive estimates are given by

$$d\hat{\theta}_t^2 = k(\beta - \hat{\theta}_t)dt - \frac{\gamma_t \hat{\theta}_t}{b_1^2} \left[d \log y_t - \left(\alpha - \frac{1}{2}\hat{\theta}_t^2\right) dt \right]$$

$$\dot{\gamma}_t = -2k\gamma_t + c^2 - \frac{\hat{\theta}_t^2}{b_1^2} \gamma_t^2.$$

2.4. Aihara & Bagchi model

Aihara and Bagchi [11] considered the following model

$$d\theta_t = \alpha(m - \theta_t)dt + k\sqrt{\theta_t}dW_{2t} \tag{2.11}$$

$$dy_t = Tr(Q)\theta_t dt + dW_{1t} \tag{2.12}$$

where $Tr(Q)$ (see Aihara and Bagchi [11] for details), α , m , and k are constants, and $\{W_{1t}\}$ and $\{W_{2t}\}$ are two independent

Wiener processes. By setting $b_1^2 = E(\sqrt{\theta_t})^2$ in (2.11) and applying Theorem 2.1, the recursive estimates are given by

$$\begin{aligned} d\hat{\theta}_t &= a(m - \hat{\theta}_t)dt + \frac{\gamma_t \text{Tr}(Q)}{k^2 b_1^2} (dy_t - \text{Tr}(Q)\hat{\theta}_t dt) \\ \dot{\gamma}_t &= -2\alpha\gamma_t + 1 - \frac{\text{Tr}^2(Q)}{k^2 b_1^2} \gamma_t^2. \end{aligned}$$

2.5. Christoffersen–Heston–Jacobs [12] model

For the SV model studied in Christoffersen, Heston and Jacobs [12]

$$\begin{aligned} d\theta_t &= (a - c\theta_t)dt + \xi\sqrt{\theta_t}dZ_{2t} \\ dy_t &= (R + u\theta_t)dt + \sqrt{\theta_t}dZ_{1t} \end{aligned}$$

where the Wiener processes $\{Z_{1t}\}$ and $\{Z_{2t}\}$ are correlated with correlation coefficient ρ . R , u , a , c , and ξ are constants. This correlated model can be transformed to a model with uncorrelated noises. Let $\{W_{1t}\}$ and $\{W_{2t}\}$ be two standard uncorrelated Wiener processes. Define $dZ_{1t} = dW_{1t}$ and $dZ_{2t} = \rho dW_{1t} + \sqrt{1 - \rho^2}dW_{2t}$, then the above model may be written by using only dW_{1t} and dW_{2t} .

$$\begin{aligned} d\theta_t &= (a - c\theta_t)dt + \xi\sqrt{\theta_t}\rho dW_{1t} + \xi\sqrt{\theta_t}\sqrt{1 - \rho^2}dW_{2t} \\ dy_t &= (R + u\theta_t)dt + \sqrt{\theta_t}dW_{1t} \end{aligned}$$

By setting $b_1^2 = E(\sqrt{\theta_t})^2$ and applying Theorem 2.1, the recursive estimates are given by

$$\begin{aligned} d\hat{\theta}_t &= (a - c\hat{\theta}_t)dt + \frac{\gamma_t u}{b_1^2} [dy_t - (R + u\hat{\theta}_t)dt] \\ \dot{\gamma}_t &= -2c\gamma_t + \xi^2 b_1^2 - \frac{u^2}{b_1^2} \gamma_t^2, \end{aligned}$$

Note: $\sigma_1^2 = \xi^2 b_1^2 \rho^2 + \xi^2 b_1^2 (\sqrt{1 - \rho^2})^2 = \xi^2 b_1^2$.

3. Conclusions

We have obtained the optimal recursive estimates for various classes of nonlinear non-normal stochastic volatility models. Recursive estimates for recently proposed stochastic volatility models are also considered in some detail.

References

- [1] S.L. Heston, S. Nandi, A close-from GARCH option valuation model, *The Review of Financial Studies* 13 (2000) 585–625.
- [2] H. Gong, A. Thavaneswaran, J. Singh, Filtering for some time series models by using transformation, *The Mathematical Scientist* 33 (2008) 141–147.
- [3] C. Kirby, Linear filtering for asymmetric stochastic volatility models, *Journal of Economics Letters* 92 (2006) 284–292.
- [4] A.N. Shiryaev, *Probability*, 2nd ed., Springer-Verlag New York Inc., 2005.
- [5] M.E. Thompson, A. Thavaneswaran, Filtering via estimating functions, *Applied Mathematics Letters* 12 (1999) 61–67.
- [6] A. Thavaneswaran, M.E. Thompson, A criterion for filtering in semimartingale models, *Stochastic Processes and Their Applications* 28 (1988) 259–265.
- [7] F. Klebaner, *Introduction to Stochastic Calculus with Applications*, Imperial College Press, 2007.
- [8] J. Hull, A. White, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42 (2) (1987) 281–300.
- [9] E.M. Stein, J.C. Stein, Stock price distributions with stochastic volatility – An analytic approach, *Review of Financial Studies* 4 (1991) 727–752.
- [10] S. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6 (1993) 327–343.
- [11] S. Aihara, A. Bagchi, Filtering and identification of stochastic volatility for parabolic type factor models, *International Journal of Innovative Computing, Information and Control* 2 (5) (2006) 917–925.
- [12] P. Christoffersen, S. Heston, K. Jacobs, Option valuation with conditional skewness, *Journal of Econometrics* 131 (2006) 253–284.
- [13] H. Kawakatsu, Specification and estimation of discrete time quadratic stochastic volatility models, *Journal of Empirical Finance* 14 (2007) 424–442.
- [14] S.J. Taylor, *Asset Price Dynamics, Volatility, and Prediction*, Princeton University Press, 2005.
- [15] R. Liptser, A.N. Shiryaev, *Statistics of Random Processes I and II*, 2nd ed., Springer-Verlag, New York, Inc., 2001.