Local Approximate MLE of Time-Varying Stochastic Volatility Model Based on Log-Likelihood Expansion\textsuperscript{1}

Tian-Xiu Wang
College of Mathematic and Information Science
Henan Normal University, Henan, 453007, China

Ji-Xia Wang
College of Mathematic and Information Science
Henan Normal University, Henan, 453007, China

Ya-Meng Zhang
College of Mathematic and Information Science
Henan Normal University, Henan, 453007, China

Copyright © 2017 Tian-Xiu Wang, Ji-Xia Wang and Ya-Meng Zhang. This article is distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The aim of this paper is to study the local approximate maximum likelihood estimations of time-varying stochastic volatility (SV) model. The approximate log-transition density functions of the SV model are obtained by using the Hermite and Kolmogorov methods. The performance of the approximate transition probability density is demonstrated by Choi (2015). The approximate maximum likelihood estimations (MLE) are obtained by the approximate log-transition density functions. We prove the asymptotic properties of the approximate log-transition density of the SV model.

\textsuperscript{1}This work is supported by the National Natural Science Foundation (NNSF) of China (No. U1504701); Doctoral research startup project of Henan Normal University (No. qd15184); Program for Graduate Innovative Research of Henan Normal University (No: YL201703).
Mathematics Subject Classification: 62G05; 62M09

Keywords: local approximate maximum likelihood estimation, stochastic volatility model, log-likelihood expansion; asymptotic property

1 Introduction

The SV model accurately characterize the various characteristics of the financial model. The earliest researches on this field are Clark (1973), Taylor (1982) and so on. Heston (1993) has derived a closed-form solution for the price of a European call option of SV model. The transition probability density can be used to estimate the parameters and to simulate the data. The approximate transition density function of the univariate diffusion model has been established by using the Hermite series expansion by Aït-Sahalia (2002). Aït-Sahalia (2008) has obtained closed-form log-likelihood expansions for multivariate diffusions. His method is applied to the article of Choi (2013) and Choi (2015) et al. We use a similar method to obtain the log-transition density function of the SV model and estimate the parameters of the model.

Harvey and Shephard (1996) have estimated the SV model by transforming the quasi-maximum likelihood process in the form of linear state space. The moments and the asymptotic distribution of the actual volatility error have derived by Barndorff-Nielsen and Shephard (2002), which can be used to estimate the parameters of the SV model. Another method is that use the Malliavin weight asymptotic expansion to estimate the error of the SV model, such as Takahashi and Yamada(2014). Kanaya and Kristensen(2015) study the two-step estimation method of SV models.

In this paper, we study the local approximate maximum likelihood estimations of time-varying stochastic volatility (SV) model. Choi (2015) studied the explicit form of approximate transition probability density functions of diffusion processes. This article is built on Choi (2015) to extend his results to approximate log-transition density function cases. We obtain the approximate log-transition density function of the SV model by using the Hermite and Kolmogorov methods. Meanwhile, we show the asymptotic properties of the approximate log-likelihood function of the SV model.

The structure of this paper is organized as follows. Section 2 sets out the model and assumptions. In Section 3, we describe how to find the log-transition probability density function of SV model. Estimations of the parameters of the log-transition probability density function using the approximate MLE method are discussed in section 4. Section 5 studies the asymptotic properties of the approximate log-likelihood function.
2 Model and Assumptions

We consider the following time-inhomogenous SV model represented by

\[\begin{align*}
    dX_{1t} &= \mu f(X_{1t})dt + G(X_{1t})\sqrt{X_{2t}}(\sqrt{1-\rho^2}dB_{1t}^{(1)} + \rho dB_{1t}^{(2)}) \\
    dX_{2t} &= (\alpha(t) + \beta(t)X_{2t})dt + \sigma(t)\sqrt{X_{2t}}dB_{2t}^{(2)},
\end{align*}\]

where \(\mu\) is a non-time-varying drift parameter, \(\alpha(t)\) and \(\beta(t)\) are time-varying drift parameters. \(\sigma(t)\) is a volatility function. \(f(\cdot)\) and \(G(\cdot)\) are known differentiable functions. \(\{(B_{1t}^{(1)}, B_{1t}^{(2)})\}\) is a two dimensional Brown motion with instantaneous correlation \(\rho\),

\[dB_{1t}^{(1)}dB_{1t}^{(2)} = \rho dt.\]

The model (1) can be expressed as the following form

\[dX_t = b(t, X_t)dt + v(t, X_t)dB_t,\]

where \(X_t = \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix}\), \(b(t, X_t) = \begin{pmatrix} \mu f(X_{1t}) \\ \alpha(t) + \beta(t)X_{2t} \end{pmatrix}\), \(B_t = \begin{pmatrix} B_{1t}^{(1)} \\ B_{1t}^{(2)} \end{pmatrix}\),

\[v(t, X_t) = \begin{pmatrix} G(X_{1t})\sqrt{X_{2t}}(1-\rho^2) \\ G(X_{1t})\sqrt{X_{2t}\rho} \end{pmatrix},\]

and \(S_X\) is a subset of \(R^2\). \(S_X\) is the domain of the diffusion \(X_t\). Note that \(X_{2t}\) is also the volatility function of \(\{X_{1t}\}\). For a function \(\psi(x) = (\psi_1(x), \psi_2(x))^T\), differentiable in \(x\), where \((\psi_1(x), \psi_2(x))^T\) is transposition of vector \((\psi_1(x), \psi_2(x))\). \(\nabla \psi(x)\) is the Jacobian matrix of \(\psi\), \(\nabla \psi(x) = [\partial \psi_i(x)/\partial x_j]_i = 1, 2, j = 1, 2.\)

For the existence of the solution to the SV model (1) and approximate transition probability density function, some conditions need to be exerted on model (1).

**Assumption 1.** \(S_X\) is a product of two intervals with lower limit \(X_i\) and upper limit \(\overline{X}_i\), where probably \(X_i = -\infty\) and/or \(\overline{X}_i = +\infty\), in which case, the intervals are open at infinite limits.

Characteristic description of a diffusion relies on the following variance-covariance matrix, \(\nu(t, x)\) rather than the volatility function, \(v(t, x)\). Let \(\nu(t, x) \equiv v(t, x)v^T(t, x)\).

**Assumption 2.** The diffusion matrix \(\nu(t, x)\) is positive definite for all \((t, x)\) in the interior of \([0, \infty) \times S_X\). (See Choi (2013))

In order to ensure the existence and uniqueness of the solution to model (1) and make possible the computation of likelihood expansions, we need the following assumption.

**Assumption 3.** For \(i, j = 1, 2\), \(b_i(t, x)\) and \(\nu_{ij}(t, x)\) are infinitely differentiable with respect to \(x \in S_X\) and \(t \in [0, \infty)\), respectively.
It is required that \( b(t, x) \) and \( v(t, x) \) have local Lipschitz continuity in \( x \). Then we can ensure the uniqueness of the solution to model (1). There is a constant \( C_c > 0 \) such that for every \( c > 0, t \in [0, \infty) \), \( \|x\| \leq C \) and \( \|x'\| \leq c \),
\[
\|b(t, x) - b(t, x')\| + \|v(t, x) - v(t, x')\| \leq C_c \|x - x'\|.
\]
In that way, if the solution to model (1) exists, it is unique. For \( x \in \mathbb{R}^2 \), \( \| \cdot \| \) denotes the usual Euclidean norm.

The following condition restricts the growth behavior of the coefficients near the boundaries of the domain. The drift and diffusion function need to meet this conditions.

**Assumption 4.** There is a constant \( D > 0 \) such that for all \((t, x) \in [0, \infty) \times S_X\),
\[
\|b_i(t, x)\|_2 \leq D(1 + \|x\|_2), \quad \|v_{ij}(t, x)\|_2 \leq D(1 + \|x\|_2).
\]
where \( i, j = 1, 2 \). Their derivatives exhibit at most polynomial growth. Assumption 4 ensures the existence of a solution to the model (1).

### 3 The Log-likelihood Expansion

There exists a one-to-one transformation of the diffusion \( X_t \) into the diffusion \( Y_t \) whose diffusion matrix \( v_Y \) is the identity matrix. That is, there exists an invertible function \( \psi(x) \), infinitely differentiable in \( t \) and \( x \) on \([0, \infty) \times S_X\), such that \( Y_t \equiv \psi(t, X_t) \) satisfies the stochastic differential equation
\[
dY_t = b_{Y_i}(t, Y_t) dt + dB_t \tag{2}
\]
on the domain \( S_Y \). Using It\(\hat{\text{o}}\)'s lemma and \( \nabla \psi(t, x) = v^{-1}(t, x) \), then the drift function of \( Y_t \) is
\[
b_{Y_i}(t, y) = \frac{\partial \psi_i(t, \psi_{inv}(t, y))}{\partial t} + \sum_{p=1}^{2} b_p(t, \psi_{inv}(t, y)) \frac{\partial \psi_i(t, \psi_{inv}(t, y))}{\partial x_p}
\]
\[
+ \frac{1}{2} \sum_{p=1}^{2} \sum_{q=1}^{2} \sum_{r=1}^{2} v_{pr}(t, \psi_{inv}(t, y)) v_{qr}(t, \psi_{inv}(t, y)) \frac{\partial^2 \psi_i(t, \psi_{inv}(t, y))}{\partial x_p \partial x_q},
\]
where \( \psi_{inv}(t, y) \) denotes the inverse transformation of \( \psi(t, y) \). \( v^{-1}(t, x) \) is the inverse matrix of \( v(t, x) \).

The first conversion is making the volatility term equal to 1, so that the tail behavior of \( P_Y \) which is the transition probability density of \( Y \) close to Gaussian. However, when the time difference between the two consecutive observations becomes small, \( P_Y \) gets concentrated and peaked around \( y_0 \). Therefore,
we need the second transformation from $Y$ to $Z$,

$$Z \equiv \frac{y - y_0}{\sqrt{\Delta}},$$

where $\Delta = t - t_0$. This transformation makes $P_Z$ which is the transition density of $Z_t$ close to normal. Thus, we can obtain Hermite-expansion of $P_Z$ around standard normal density function, $\phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$,

$$P_Z^{(j)}(t, z|t_0, y_0) \equiv \phi(z) \sum_{j=0}^{J} \eta_z^{(j)}(t, t_0, y_0) H_j(z)$$

can produce an exact approximate density with a few coefficients. $\phi(z)$ is the two dimensional multivariate matrix. Multivariate Hermite polynomials are defined by

$$H_h(Z) = (-1)^{|h|} \frac{\partial^{|h|} \phi(z)}{\partial Z_1^{h_1} \partial Z_2^{h_2}},$$

where $h = (h_1, h_2)^T$, and $h_1, h_2$ are non-negative integers. Then $|h|$ is defined as the sum of all elements of 2 dimensional vector $h$, i.e, $|h| = h_1 + h_2$. The Hermite polynomials are orthogonal in the following cases

$$\int_{R^2} H_h(z) H_k(z) dx = \begin{cases} h_1!h_2! & \text{if } h=k \\ 0 & \text{otherwise.} \end{cases}$$

We can obtain the coefficients $\eta_z^{(j)}(t, t_0, y_0)$ by using the orthogonality of Hermite polynomials, namely

$$\eta_z^{(j)}(t, t_0, y_0) = \frac{1}{h_1!h_2!} \int_{-\infty}^{+\infty} H_j(z) P_Z(t, z|t_0, y_0) dz$$

$$= \frac{1}{h_1!h_2!} \int_{-\infty}^{+\infty} H_j(z) \sqrt{\Delta} P_Y(t, \sqrt{\Delta} z + y_0|t_0, y_0) dz$$

$$= \frac{1}{h_1!h_2!} \int_{-\infty}^{+\infty} H_j(\frac{z-y_0}{\sqrt{\Delta}}) P_Y(t, y|t_0, y_0) dy$$

$$= \frac{1}{h_1!h_2!} E \left[ H_j(\frac{Y_t-y_0}{\sqrt{\Delta}})|Y_{t_0} = y_0 \right].$$

The conditional expectation in the above equation can be approximated up to any order by Taylor expansion using infinitesimal operator, $A_Y$ defined by

$$A_Y \circ f(t, y, t_0, y_0) = \frac{\partial f(t, y, t_0, y_0)}{\partial t} + \sum_{i=1}^{2} b_Y(t, y) \frac{\partial f(t, y, t_0, y_0)}{\partial y_i}$$

$$+ \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial^2 f(t, y, t_0, y_0)}{\partial y_i \partial y_j}.$$ 

Thus, for any infinitely differentiable function $f$,

$$E \left[ f(t, Y_t, t_0, Y_{t_0})|Y_{t_0} = y_0 \right] = \sum_{i=0}^{k} \frac{\Delta^i}{i!} A_Y^{(i)} \circ f(t, y, t_0, y_0)|_{y=y_0, t=t_0} + O(\Delta^{k+1}).$$
So, the condition expectation of order $K$ is obtained by ignoring terms of higher orders.

$$E \left[ H_j \left( \frac{y-y_0}{\sqrt{\Delta}} \right) | Y_t = y_0 \right] = \sum_{i=0}^{k} \frac{\Delta^i}{i!} A_{y}^{i} \circ H_j \left( \frac{y-y_0}{\sqrt{\Delta}} \right) |_{y=y_0}.$$ 

Therefore, the $J$-th order Hermite expansion of the two-dimensional multivariate time-inhomogeneous unit diffusion process $Y_t$ can be obtained by Egorov(2003), i.e.

$$P^{(J,K)}_Y (t, y|t_0, y_0) = \Delta^{-1} \phi \left( \frac{y-y_0}{\sqrt{\Delta}} \right) \left\{ \sum_{j \in N^2, |j| \leq J} \frac{1}{h_1!h_2!} \left[ \sum_{i=0}^{k} \frac{\Delta^i}{i!} A_{y}^{i} \circ H_j \left( \frac{y-y_0}{\sqrt{\Delta}} \right) |_{y=y_0} \right] H_j \left( \frac{y-y_0}{\sqrt{\Delta}} \right) \right\}.$$  

We rewrite $P^{(J,K)}_Y$ by setting $J$ to $\infty$ and arranging the order of $\Delta$ from small to large to get $P^{(K)}_Y = P^{(\infty,K)}_Y$ which is an alternative of $P^{(J,K)}_Y$,

$$P^{(K)}_Y (t, y|t_0, y_0) = \Delta^{-1} \phi \left( \frac{y-y_0}{\sqrt{\Delta}} \right) \exp \left[ (y_1 - y_{01}) \int_0^1 b_{Y_1} (t, y_0 + u(y - y_0)) du \right] \times \sum_{k=0}^{K} c^{(K)}_Y (t, y|t_0, y_0) \frac{\Delta^k}{k!}.$$  

Here $c^{(K)}_Y (t, y|t_0, y_0)$ can be found successively by solving Kolmogorov forward and backward partial differential equations(PDE), i.e, plugging equation (5) in the following Kolmogorov forward PDE,

$$\frac{\partial P_Y (t, y|t_0, y_0)}{\partial t} = -\sum_{i=1}^{2} \frac{\partial b_{Y_i} (t, y) P_Y (t, y|t_0, y_0)}{\partial y_i} + \frac{1}{2} \sum_{i=1}^{2} \frac{\partial^2 P_Y (t, y|t_0, y_0)}{\partial y_i^2}.$$

And then, $P_Y (t, y|t_0, y_0)$ also satisfies the backward PDE,

$$-\frac{\partial P_Y (t, y|t_0, y_0)}{\partial t_0} = \sum_{i=1}^{2} b_{Y_i} (t_0, y_0) \frac{\partial P_Y (t, y|t_0, y_0)}{\partial y_{0i}} + \frac{1}{2} \sum_{i=1}^{2} \frac{\partial^2 P_Y (t, y|t_0, y_0)}{\partial y_{0i}^2}.$$

Thus, we can get the coefficients

$$c^{(0)}_Y (t, y|t_0, y_0) = 1 \quad and \quad c^{(k)}_Y (t, y|t_0, y_0) = k \int_0^1 g^{(k)} (t, y_0 + u(y - y_0)|t_0, y_0) u^{k-1} du$$
for $k \geq 1$, where

$$
g^{(k)}(t, \omega|t_0, y_0)|_{\omega = y_0 + u(y - y_0)} = \left\{ -\sum_{i=1}^{2} b_{Y_i}(t, \omega) \frac{\partial}{\partial \omega_i} \left( \sum_{i=1}^{2} (\omega_i - y_{0i}) \int_0^1 b_{Y_i}(t, y_0 + u(\omega - y_0)) du \right) \\
- \sum_{i=1}^{2} \frac{\partial b_{Y_i}(t, \omega)}{\partial \omega_i} - \sum_{i=1}^{2} (\omega_i - y_{0i}) \int_0^1 \frac{\partial b_{Y_i}(t, y_0 + u(\omega - y_0))}{\partial t} du \\
+ \sum_{i=1}^{2} \left[ \frac{1}{2} \frac{\partial}{\partial \omega_i} \left( \sum_{i=1}^{2} (\omega_i - y_{0i}) \int_0^1 b_{Y_i}(t, y_0 + u(\omega - y_0)) du \right) \right]^2 \\
+ \sum_{i=1}^{2} \frac{1}{2} \frac{\partial^2}{\partial \omega_i^2} \left( \sum_{i=1}^{2} (\omega_i - y_{0i}) \int_0^1 b_{Y_i}(t, y_0 + u(\omega - y_0)) du \right) \right\}
$$

Taking the logarithm of (5) and Taylor-expanding it in $\Delta$, and we get the $K$-th order approximate log-transition density expansion

$$
l^{(K)}_Y(t, y|t_0, y_0) = -\ln(2\pi\Delta) + \frac{C^{(-1)}_Y(t, y|t_0, y_0)}{\Delta} + \sum_{k=0}^{K} \frac{C^{(k)}_Y(t, y|t_0, y_0) \Delta^k}{k!}
$$

with

$$
C^{(-1)}_Y(t, y|t_0, y_0) = -\frac{1}{2} \sum_{i=1}^{2} (y_i - y_{0i})^2
$$

$$
C^{(0)}_Y(t, y|t_0, y_0) = \sum_{i=1}^{2} (y_i - y_{0i}) \int_0^1 b_{Y_i}(t, y_0 + u(y - y_0)) du
$$

and for $k \geq 1$

$$
C^{(k)}_Y(t, y|t_0, y_0) = k \int_0^1 G^{(k-1)}_Y(t, y_0 + u(y - y_0)|t_0, y_0) u^{k-1} du,
$$
where

\[ G^{(1)}_Y(t, y|t_0, y_0) = -\sum_{i=1}^{2} \frac{\partial b_Y(t, y)}{\partial y_i} - \frac{\partial C^{(0)}_Y(t, y|t_0, y_0)}{\partial t} \]

\[ - \sum_{i=1}^{2} b_Y(t, y) \frac{\partial C^{(0)}_Y(t, y|t_0, y_0)}{\partial y_i} \]

\[ + \frac{1}{2} \sum_{i=1}^{2} \left\{ \frac{\partial^2 C^{(0)}_Y(t, y|t_0, y_0)}{\partial y_i^2} + \left( \frac{\partial C^{(0)}_Y(t, y|t_0, y_0)}{\partial y_i} \right)^2 \right\} \]

and for \( k \geq 2 \)

\[ G^{(k)}_Y(t, y|t_0, y_0) = \frac{1}{2} \sum_{i=1}^{2} \frac{\partial^2 C^{(k-1)}_Y(t, y|t_0, y_0)}{\partial y_i^2} - \frac{\partial C^{(k-1)}_Y(t, y|t_0, y_0)}{\partial t} \]

\[ - \sum_{i=1}^{2} b_Y(t, y) \frac{\partial C^{(k-1)}_Y(t, y|t_0, y_0)}{\partial y_i} \]

\[ + \frac{1}{2} \sum_{i=1}^{2} \sum_{h=0}^{k-1} \left( k - 1 \right) \frac{\partial C^{(h)}_Y(t, y|t_0, y_0)}{\partial y_i} \frac{\partial C^{(k-1-h)}_Y(t, y|t_0, y_0)}{\partial y_i}. \]

The coefficient of logarithmic expansion is also obtained by using the Kolmogorov-equation. i.e,

\[ \frac{\partial l_Y(t, y|t_0, y_0)}{\partial t} = -\sum_{i=1}^{2} \frac{\partial b_Y(t, y)}{\partial y_i} - \sum_{i=1}^{2} b_Y(t, y) \frac{\partial l_Y(t, y|t_0, y_0)}{\partial y_i} \]

\[ + \frac{1}{2} \sum_{i=1}^{2} \frac{\partial^2 l_Y(t, y|t_0, y_0)}{\partial y_i^2} + \frac{1}{2} \sum_{i=1}^{2} \left[ \frac{\partial l_Y(t, y|t_0, y_0)}{\partial y_i} \right]^2 \]

(6)

and

\[ - \frac{\partial l_Y(t, y|t_0, y_0)}{\partial t_0} = \sum_{i=1}^{2} b_Y(t_0, y_0) \frac{\partial l_Y(t, y|t_0, y_0)}{\partial y_{0i}} + \frac{1}{2} \sum_{i=1}^{2} \frac{\partial^2 l_Y(t, y|t_0, y_0)}{\partial y_{0i}^2} \]

(7)

\[ + \frac{1}{2} \sum_{i=1}^{2} \left[ \frac{\partial l_Y(t, y|t_0, y_0)}{\partial y_{0i}} \right]^2. \]

We are looking for the log-transition density of \( X_t \) rather than \( Y_t \), so by change of variable we get

\[ l_X(t, x|t_0, x_0) = ln\{Det[\nabla \psi(t, x)] + l_Y(t, \psi(t, x)|t_0, \psi(t_0, x_0))\}. \]

Since \( Det[\nabla \psi(t, x)] = Det[\nu^{-1}(t, x)] = Det[\nu(t, x)]^{-\frac{1}{2}} \).
We define \( D_{\nu}(t, x) \equiv \frac{1}{2}\ln(\operatorname{Det}[\nu(t, x)]) \). Then
\[
l_X(t, x|t_0, x_0) = -D_{\nu}(t, x) + l_Y(t, \psi(t, x)|t_0, \psi(t_0, x_0)).
\]
Thereby
\[
l^{(k)}_X(t, x|t_0, x_0) = -D_{\nu}(t, x) - \ln(2\pi \Delta) + C^{(k)-1}_Y(t, \psi(t, x)|t_0, \psi(t_0, x_0)) \Delta_k/k!.
\]
At this point, we obtain the logarithmic transition density of the SV model.

4 Approximate Maximum Likelihood Estimation

In this section we will discuss the approximate MLE of the parameters. Let the parameter space \( \Theta \) be a compact subset of \( R \). Parameterized \((\mu, \rho, \alpha(t), \beta(t), \sigma(t))\) with parameter vector \( \theta \in \Theta \). Because the process \( X_t \) are observed at discrete times, then \( \{X_t|t= i\Delta \ and \ i = 0, 1, \cdots, n\} \). The values of \( \Delta \) can be different, but we set each \( \Delta \) is the same. Because of the Markov property of model (1), the log-likelihood function is expressed as follows
\[
\ln(\theta) = \sum_{i=1}^{n} l_X(i\Delta, X_i\Delta|(i-1)\Delta, X_{(i-1)\Delta}; \theta),
\]
where the initial observation \( X_0 \) is neglected.

In order to discuss the MLE, we set \( \theta_1 = (\mu, \rho)^T \) as the non-time-varying parameter vector, and \( \theta_2(t) = (\alpha(t), \beta(t), \sigma(t))^T \) as the time-varying parameter vector, where \( \theta_2(t) \) and their derivatives at all orders thrice continuously differentiable. Now suppose that \( (X_{t_0}, X_{t_1}, \cdots, X_{t_n}) \) is the isometry observation sample for model (1), \( 0 \leq X_{t_0} < X_{t_1} < \cdots < X_{t_n} \leq T \). Assume that when \( \theta_i \neq \theta_j, l_i \neq l_j \), that is, \( \theta \) can be identified. Aiming at the parameter vectors \( \theta_1 \) and \( \theta_2 \), the MLE is discussed in two steps:

Step 1 First, fix the parameter \( \theta_1 = \theta_{10} = (\mu_0, \rho_0)^T \). Consider the local MLE of time-varying parameter \( \theta_2(t) \). Since \( \theta_2(t) \) is time-varying, it is approximated by a local constant approximation method, namely, for any given time point \( t_0 \), in a small neighborhood, let
\[
\theta_2(t) \approx \theta_2(t_0) = (\alpha(t_0), \beta(t_0), \sigma(t_0))^T.
\]
Then \( \alpha(t_0), \beta(t_0), \sigma(t_0) \) are the approximation of \( \alpha(t), \beta(t), \sigma(t) \), respectively.
Next, we construct a kernel function weighted log-likelihood function for \( \theta_2(t_0) \), namely

\[
\ln(\theta_2(t_0)) = \sum_{i=1}^{n} l_X(i\Delta, X_{i\Delta} | (i-1)\Delta, X_{(i-1)\Delta}; \theta_2(t_0))K_h(t_i - t_0),
\]

where \( K(\cdot) \) is weight function, called kernel function, and \( K_h(\cdot) = K(\cdot/h)/h, h \) represents the size of the neighborhood, called the bandwidth parameter. Since the log-likelihood function is difficult to be calculated in the practical problem, the \( K \)-order weighted log-likelihood function of \( \theta_2(t_0) \) is introduced by

\[
Ln^{(K)}(\theta_2(t_0)) = \sum_{i=1}^{n} l_X^{(K)}(i\Delta, X_{i\Delta} | (i-1)\Delta, X_{(i-1)\Delta}; \theta_2(t_0))K_h(t_i - t_0).
\]

Thus it can be obtained that the approximate MLE is

\[
\hat{\theta}_2^{(K)}(t_0) = (\hat{\alpha}^{(K)}(t_0), \hat{\beta}^{(K)}(t_0), \hat{\sigma}^{(K)}(t_0))^T = \arg \max_{\theta_2(t_0)} L_n^{(K)}(\theta_2(t_0)).
\]

At this point, we obtain the approximate MLE \( \hat{\theta}_2^{(K)}(t_0) \) of \( \theta_2(t_0) \), then \( \hat{\theta}_2^{(K)}(t_0) \) is the approximate MLE of \( \theta_2(t) \).

**Step 2** Fix \( \hat{\theta}_2^{(K)}(t_0) \), estimate \( \theta_1 \). First, we take the MLE of parameter \( \theta_1 \), and the Step 1 of the analogy is obtained by the \( K \)-order approximation maximum likelihood function

\[
Ln^{(K)}(\theta_1) = \sum_{i=1}^{n} l_X^{(K)}(i\Delta, X_{i\Delta} | (i-1)\Delta, X_{(i-1)\Delta}; \theta_1).
\]

By taking the partial derivative of \( \theta_1 \) with respect to \( Ln^{(K)}(\theta_1) \), the approximate MLE \( \hat{\theta}_1^{(K)} \) is

\[
\hat{\theta}_1^{(K)} = \arg \max Ln^{(K)}(\theta_1).
\]

At this point, we get the approximate \( \hat{\theta}_1^{(K)} \) of \( \hat{\theta}_1 \), and obtain the approximate MLE of \( \theta_1 \) is \( \hat{\theta}_1^{(K)} \).

### 5 Asymptotic Properties

Chio(2015) given the convergence of the transition probability density function of the time-inhomogenous diffusion process. In this section we will discuss the convergence of the approximate log-likelihood function by analogy Chio(2015).

From the previous section we can observe that \( l_x^{(k)}(\tilde{l}_X^{(k)}) \) converges to \( l_x \) and therefore \( l_n^{(k)}(\theta, \Delta)(\tilde{l}_n^{(k)}(\theta, \Delta)) \) converges to \( ln(\theta, \Delta) \). Suppose that \( ln(\theta, \Delta) \) has
a unique MLE $\hat{\theta}_{n,\Delta} \in \Theta$, then the approximate MLE $\hat{\theta}^{(k)}_{n,\Delta}$ of $\theta$ can be obtained by maximizing $l^{(k)}_n(\theta, \Delta)\bigl(\hat{\theta}^{(k)}_n(\theta, \Delta)\bigr)$.

**Theorem 1** For any $n$,

$$\sup_{\theta \in \Theta} |l^{(k)}_n(\theta, \Delta) - l^{(k)}_n(\theta, \Delta)| \to 0$$

(9)

in probability as $\Delta \to 0$. In the reducible case, the same holds for $l^{(k)}_n$. The approximate MLE sequence $\hat{\theta}^{(k)}_{n,\Delta}$ exists almost surely and satisfies $\hat{\theta}^{(k)}_{n,\Delta} - \hat{\theta}^{(k)}_{n,\Delta} \to 0$ in probability as $\Delta \to 0$.

**Proof:** Since the logarithm is continuous, we first prove the convergence of the transition probability density function (TPDF), which can push the convergence of the log-likelihood function. Taylor expands $\exp(l^{(k)}_n)$ in $(t, x)$ about $(t_0, x_0)$ produce TPDF $\tilde{P}^{(k)}_X$, which satisfies the Kolmogorov backward equation

$$-\frac{\partial P_X(t, x|t_0, x_0)}{\partial t_0} = \sum_{i=1}^2 b_i(t_0, x_0) \frac{\partial P_X(t, x|t_0, x_0)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 \nu_{ij}(t_0, x_0) \frac{\partial^2 P_X(t, x|t_0, x_0)}{\partial x_i \partial x_j}.$$  

(10)

Because (10) is supported by $P^{(k)}_X$, and when we define $\tilde{r}^{(k)}_X(t, x|t_0, x_0) = \tilde{P}^{(k)}_X(t, x|t_0, x_0) - P_X(t, x|t_0, x_0)$. $\tilde{r}^{(k)}_X$ also satisfies (10) and has the same residual term. Moreover, $\tilde{r}^{(k)}_X(t, x|t_0, x_0) \to 0$ as $\Delta \to 0$, due to the fact that both $\tilde{P}^{(k)}_X$ and $P_X$ converge to a Dirac-Mass at $(t_0, x_0)$ as $\Delta \to 0$. At the same time we can see that $\tilde{r}^{(k)}_X(t, x|t_0, x_0) = O(\Delta^k)$ is uniformly convergent for all $(t_0, x_0)$ and $(t, x)$ in $[0, \infty) \times S_X$ and for all $\theta$ in $\Theta$. Let

$$\tilde{R}^{(k)}_X(t, x|t_0, x_0) = \sup_{\theta \in \Theta} |\tilde{r}^{(k)}_X(t, x|t_0, x_0)|.$$  

(11)

Then

$$E[\tilde{R}^{(k)}_X(t, X_0|t_0, X_{t_0})|X_0 = x_0]$$  

$$= \int_{S_X} \tilde{R}^{(k)}_X(t, x|t_0, x_0) P_X(t, x|t_0, x_0) dx$$  

$$= \int_{N} \tilde{R}^{(k)}_X(t, x|t_0, x_0) P_X(t, x|t_0, x_0) dx$$  

$$+ \int_{N \setminus S_X} \tilde{R}^{(k)}_X(t, x|t_0, x_0) P_X(t, x|t_0, x_0) dx,$$

where $N = \prod_{i=1}^2 [x_0i - \sqrt{\Delta}c_\Delta, x_0i + \sqrt{\Delta}c_\Delta]$ is a neighborhood of $x_0$, $N \setminus S_X$ is its complement and $c_\Delta$ is a positive numbers sequence such that $\sqrt{\Delta}c_\Delta \to 0$. 

First, since $\tilde{r}^{(k)}(t, x|t_0, x_0) = O(\Delta^k)$, $\tilde{R}^{(k)}_X(t, x|t_0, x_0) \leq M\Delta^k$ for some constant $M$ and for all $x \in N$. Thus

$$\int_N \tilde{R}^{(k)}_X(t, x|t_0, x_0) P_X(t, x|t_0, x_0) dx \to 0$$

in probability as $\Delta \to 0$.

Next, the tail behavior of $P_X$ is driven by exponential term in $\Delta = 0$ neighborhood. Since $\tilde{R}^{(k)}_X$ grows at a polynomial rate, the expected value of $\tilde{R}^{(k)}_X$ outside of $N$ of the form

$$\sqrt{\Delta} \int_{\sqrt{\Delta}c\Delta}^{\infty} |x - x_0|^b \exp\left(-\frac{(x - x_0)^2}{\nu(t_0, x_0)}\right) dx$$

$$= \Delta^{\frac{b}{2}} \int_{c\Delta}^{\infty} |z - z_0|^b \exp\left(-\frac{(z - z_0)^2}{\nu(t_0, x_0)}\right) dz$$

on the interval $(\sqrt{\Delta}c\Delta, \infty)$, where $b \geq 0$ and similarly on the interval $(-\infty, -\sqrt{\Delta}c\Delta)$. Hence

$$\int_{N\setminus S_X} \tilde{R}^{(k)}_X(t, x|t_0, x_0) P_X(t, x|t_0, x_0) dx \to 0$$

in probability as $\Delta \to 0$.

Thereby, $E[\tilde{R}^{(k)}_X(t, X_t|t_0, X_{t_0})|X_0 = x_0] \to 0$ as $\Delta \to 0$ from (12) and (13) equation.

Following the previous procedure we can prove that $Var[\tilde{R}^{(k)}_X(t, X_t|t_0, X_{t_0})|X_{t_0} = x_0] \to 0$ as $\Delta \to 0$. Through these two results and applying Lebesgue’s dominated convergence theorem, we obtain $\tilde{R}^{(k)}_X(t, X_t|t_0, X_{t_0}) \to 0$ in probability, namely $\tilde{R}^{(k)}_X \to P_X$. Therefore, $\tilde{R}^{(k)}_X(\theta) \to l_X(\theta)$ in probability and so does $\tilde{R}^{(k)}_n(\theta) \to l_n(\theta)$ for fixed $n$ and the theorem is proved.

References


Local approximate MLE of time-varying stochastic volatility model


Received: November 17, 2017; Published: December 11, 2017