Bayesian inference for Diffusion Processes

Junfeng Lai [1], Zaizai Yan [1], Jingyu Wang [12]
1. Science College, Inner Mongolia University of Technology, Hohhot 010051 , Inner Mongolia, China
2. Science College, Qiqihaer University, Qiqihaer 161006, Heilongjiang, China
E-mail: 2531365039@qq.com

Abstract:.In this paper,a numerical scheme for parametric estimation is proposed. we consider parametric estimation problem of a continuous type stochastic mathematical model(stochastic differential equation) in a wide engineering field.We deals with a Bayesian method is proposed for the analysis of discretely sampled diffusion processes .An MCMC algorithm is used to sample from the posterior distribution of the parameter of diffusion processes. The approximation is calculated by using numerical solution techniques for diffusion process. Finally, we apply the method for solving the CEV process as a numerical example.The numerical examples are given to verify the efficiency and accuracy of the proposed computational methods.
Keywords: stochastic differential equation; numerical example; CEV process

1. Introduction

Stochastic differential equations (SDEs)

$$dY = \mu(Y;\theta)dt + \sigma(Y;\theta)dW \tag{1}$$

The transition density

$$p_t(y\,|\,x) = \frac{d}{dy}\,p(Y_t \le y\,|\,Y_0 = x) \tag{2}$$

Of the diffusion process is the unique solution of the partial differential equation(PDE)

$$\frac{\partial}{\partial t}\,p_t(y\,|\,x) = \mu(x\,|\,\theta)\frac{\partial}{\partial x}\,p_t(y\,|\,x) + \frac{1}{2}\sigma^2(x\,|\,\theta)\frac{\partial^2}{\partial x^2}\,p_t(y\,|\,x) \tag{3}$$

A diffusion process has been instrumental to the modeling of wide variety of phenomena, with applications in Chemistry.However,inference for this model has challenged statisticians for years ,as the likelihood function

$$L(\theta\,|\,y) \propto \prod_{i=0}^{N-2} p_\Delta(y_{i+1}\,|\,y_i,\theta) \tag{4}$$

does not exist in closed form.

Provide a convenient way to describe the dynamics process .This has led to growing interest in methods for estimating SDEs (1). Elerian suggests replacing the Gauss density by chi-squared density which is derived from the Milstein scheme [1]. Shoji and Ozaki [2] obtained an approximating for an Ornstein-Uhlenbeck process. Chan, Karolyi, Longstaff, and Sanders [3] use moments based one equation：

$$X_{i+1} = X_i + \mu(X_i;\theta)\Delta_i + \sigma(X_i;\theta)\Delta_i^{\frac{1}{2}}\varepsilon_i \tag{5}$$

Gallant and Tauchen, Bibby [4] compute expectations by using simulation-based methods. The simulation-based methods can be computationally costly, but have the advantage of being easily adapted to diffusion with unobserved state variables. Stochastic volatility models are important applications and these techniques have been found useful.

In this paper, we consider a simulation based approach that relies on a Bayesian formulation of the problem. Our approach utilizes the recent and yet rapidly growing research on condition simulation strategies in Bayesian statistics, known as Markov Chain Monte Carlo(MCMC) .

The remainder of this paper is organised as follows. In section 2 ,we recall some main result of the paper.When the solution does exist it may often be obtained by solving this equation via standard methods to yield the likelihood function. In section 3 we consider maximum likelihood estimation and show some approaches. In Section 4 we consider the model and detail how MCMC methods can be used to analyse

diffusions, focusing on the Hasting-Metropolis sampler.In section 5, we present a numerical example and conclusions are drawn.

2.Maximum Likelihood Estimation

We consider the stochastic differential equation (1) where W is an r-dimensional Wiener process. The $\theta \in \Theta$ is an unknown parameter.

If the transition densities (2) of Y are known, we can use the log-likelihood function for θ. The maximum likelihood estimator $\hat{\theta}_n$ is known to have the usual good properties(see Billingsley, Florens-Zmirou [5]). In the case of time-equidistant observations $(t_i = i\Delta t, i = 0,1,...,n)$ for some fixed $\Delta t > 0$ Dacunha-Castelle [6] prove consistency and asymptotic normality of $\hat{\theta}_n$ as $n \to \infty$ irrespective of the value of Δt .We observe at discrete assume that the process is observed at discrete times

$$t_i = i\Delta t, i = 0,1,2,3...,n, T = n\Delta t. \tag{6}$$

Δ_n various and it is assumed that $n\Delta_n^k \to 0$ for some power $k \le 2$. The asymptotic is considered as $n \to \infty$, which is equivalent to $T \to \infty$ .Since $p_\theta(t,.|x)$ is usually not known explicitly, so likelihood function (4) can be computed .

2.1 Euler method

Consider a process solution of the general stochastic differential equation (1). If the coefficients of the stochastic differential equation above are constant over small intervals $[t, t+\Delta t)$:

$$Y_{t+\Delta t} - Y_t = b(Y_t,\theta)\Delta t + \sigma(Y_t,\theta)(W_{t+\Delta t} - W_t) \tag{7}$$

and the increments $Y_{t+\Delta t} - Y_t$ are then independent Gaussian random variable with mean $b(Y_t,\theta)\Delta t$ and variance $\sigma^2(Y_t,\theta)\Delta t.$ Therefore the transition density of the process can be written as

$$p_\theta(t,y\,|\,x) = \frac{1}{\sqrt{2\pi t\sigma^2(y,\theta)}} exp\{\frac{-1}{2}\frac{(y-x-b(y,\theta)t)^2}{t\sigma^2(y,\theta)}\}$$

This approximation is good if Δ is very small.

2.2 Elerian method

Elerian [7] proposed to use the transition density derived from the Milstein scheme

$$Y_{t+\Delta t} = Y_t + b(t,Y_t)dt + \sigma(t,Y_t)(W_{t+dt} - W_t) + \frac{1}{2}\sigma(t,Y_t)\sigma_X(t,W_t)((W_{t+dt} - W_t)^2 - dt) \tag{8}$$

When the process has constant volatility or at least $\sigma_Y \approx 0$, the transition density proposed by Elerian reduces to the Euler scheme.

2.3 Local methods

The local method consists in approximating locally the drift of the stochastic differential equation with a linear function. The main idea is that a linear approximation is better than simple constant approximation made by the Euler method. The first approach we present is the Ozaki method, and it works for homogeneous stochastic differential equations. Consider the stochastic differential equation

$$dY_t = b(Y_t)dt + \theta dW_t \tag{9}$$

where σ is supposed to be constant. The construction of the method starts from the corresponding deterministic dynamical system $\frac{dy_t}{dt} = b(y_t)$ where y_t has to be a smooth function of t in the sense that it is two times differentiable with respect to t. we have $\frac{d^2y_t}{dt^2} = b_y(y_t)\frac{dy_t}{dt}$

suppose now that $b_x(y)$ is constant in the interval $[t, t+\Delta t)$, and hence by iterated integration of both sides of the equation above, first from to $u \in [t, t+\Delta t)$ and then from t to $t+\delta$, we obtain the difference equation

$$y_{t+\Delta t} = y_t + \frac{b(y_t)}{b_y(y_t)}(e^{b_y(y_t\Delta t)} - 1) \tag{10}$$

Now we translate the result above back to the stochastic dynamical system in equation (1). So, suppose $b(x)$ is approximated by the linear function $K_t x$, where K_t is constant in the interval $[t, t+\Delta t)$. The solution to the stochastic differential equation is

$$Y_{t+\Delta t} = Y_t e^{K_t \Delta t} + \sigma \int_t^{t+\Delta t} e^{K_t(t+\Delta t-u)} dW_u \tag{11}$$

Now what remains to be done is to determine the constant K_t. The main assumption is that the conditional expectation of $Y_{t+\Delta}$ given Y_t. $E(Y_{t+\Delta t} \mid Y_t) = Y_t e^{K_t \Delta t}$,

From the above, we obtain the constant K_t very easily:

$$K_t = \frac{1}{\Delta t}\log(1 + \frac{b(Y_t)}{Y_t b_Y(Y_t)}(e^{b_x(Y_t)} - 1) \tag{12}$$

In particular, we have that

$$Y_{t+\Delta t} / Y_t = y \sim N(E_y, V_y) \tag{13}$$

where

$$E_y = y + \frac{b(y)}{b_y(y)}(e^{b_y(y)\Delta t} - 1) \tag{14}$$

$$V_y = \sigma^2 \frac{e^{2K_y \Delta T} - 1}{2K_y}. \tag{15}$$

3. Bayesian inference for parameter estimation methods

We begin with a brief review of Bayesian inference as it applies to our problem. We assume that parameter vector θ is a random variable with a prior distribution $\pi(\theta)$. The posterior distribution of θ ,denoted by $\pi(\theta \mid y)$,represents the updated distribution of θ after taking the data y into account. The prior and posterior are related via Bayes'rule

$$\pi(\theta \mid y) = \frac{\pi(y \mid \theta)\pi(\theta)}{\pi(y)} \propto \pi(y \mid \theta)\pi(\theta) \tag{16}$$

We intend to obtain the posterior $\pi(\Theta|y)$ using the gamma prior and the likelihood. The likelihood $\pi(y|\Theta)$ may be obtained as follows

$$\begin{aligned} \pi(y \mid \theta) &= \left(Y(s_0) = y_0, Y(s_1) = y_1, \cdots, Y(s_m) = y_m \mid \theta\right) \\ &= \prod_{i=1}^{m} p(y_i \mid y_{i-1}, \theta) p(y_0 \mid \theta) \\ &= \prod_{i=1}^{m} p(y_i \mid y_{i-1}, \theta) \end{aligned} \tag{17}$$

We assume the process (1).The essential idea is to substitute the missing data, $x^*_{t_i}$ with simulation.We refer to the collection of simulated data and observation as the "augument data". Let

$$\widehat{Y} = \left(x(t_0) \quad x^*(t_1) \quad \cdots \quad x^*(t_M) \quad \cdots \quad x^*(t_{n-1}) \quad x(t_n)\right)$$

$\widehat{Y}_i$ denote the i'th column of $\widehat{Y}$.

We concentrate of the case where the Gibbs sampler updates one column vector of $\widehat{Y}$.This density is defined by the proportionality relationship:

$$\pi(\widehat{Y}_i \mid \widehat{Y}_{i-1}, \widehat{Y}_{i+1}, \theta) \propto \exp\left\{-\frac{1}{2}\left\|\left(\Delta\widehat{Y}_i - \mu_{i-1}\Delta t\right)\sigma_{i-1}^{-1}\left(\Delta t\right)^{-\frac{1}{2}}\right\| - \frac{1}{2}\left\|\left(\Delta\widehat{Y}_{i+1} - \mu_i\Delta t\right)\sigma_i^{-1}\left(\Delta t\right)^{-\frac{1}{2}}\right\|\right\}$$

We have defined $\Delta\widehat{Y}_i = \widehat{Y}_i - \widehat{Y}_{i-1}, \mu_i = \mu(\widehat{Y}_i, \theta), \sigma_i = \sigma(\widehat{Y}_i, \theta)$, $\|\cdot\|$ denotes the usual Euclidean norm.

4. Simulation and analysis

We follow the approach [Gianni Gilioli,2008] and simulate one long trajectory of the Constant elasticity of variance (CEV) models(Bjorn Eraker 2001)

$$dY_t = \left(\theta_1 + \theta_2 Y_t\right)dt + \theta_3 Y_t^{\beta} dW_t$$

We get the likelihood of date

$$\pi(y \mid \theta) \propto \prod_{i=1}^{n} \frac{1}{\rho_3 \widehat{Y}_{i-1}^{\ \beta}} \exp\left\{\frac{-1\left(\widehat{Y}_i - \widehat{Y}_{i-1} - \left(\rho_1 + \rho_2 \widehat{Y}_{i-1}\right)\Delta t\right)^2}{2\rho_3^{\ 2}\widehat{Y}_{i-1}^{\ 2\beta}\Delta t}\right\}$$

Let $X = \left[\sqrt{\Delta t}\widehat{Y}_{i-1}^{\ -\beta}, \sqrt{\Delta t}\widehat{Y}_{i-1}^{\ 1-\beta}\right]$

Consequently

$$\left(\theta_1, \theta_2 \mid \theta_3\right) \sim N\left(\left(X^T X\right)^{-1}(X^T y), \theta_3\left(X^T X\right)^{-1}\right)$$

$$\theta_3^{-2} \mid \widehat{Y} \sim IG(n-2, s^2)$$

We now following sampling steps for the CEV model:

Step1 Initialize all unknowns

Step 2 Using Metropolis-Hasting with proposal density $N((\frac{\widehat{Y}_{i-1}^{\ (h)} + \widehat{Y}_{i+1}^{\ (h-1)}}{2}), [\frac{\sigma_{i-1}^{\ (h)}}{2}]^2 \Delta t)$

Step 3 Draw $(\theta_1, \theta_2)^{(h)}$ $(\theta_3)^{(h)}$

Step 4 Increase h and return step 2

In figure2 we can see the estimation of $(\theta_1, \theta_2)^{(h)}$.

5. Conclusion

This paper presented a Bayesian approach to estimate the parameter of CEV .We first discretize the equation (1) by Euler method, and use MCMC method to estimate unknown theta of equation (1). The numerical results show that the proposed numerical method is efficient.

Acknowledgments

This work was supported by National Natural Foundation of Inner Mongolia (2016MS0117, 2016MS0110).

6. References

[1] Reinker S, Altman R M, Timmer J. Parameter estimation in stochastic biochemical reactions[J]. IEE Proceedings-Systems Biology, 2006, 153(4): 168-178.

[2] Andreychenko A, Mikeev L, Spieler D, et al. Parameter identification for Markov models of biochemical reactions[C]//Computer Aided Verification. Springer Berlin Heidelberg, 2011: 83-98.

[3] Golightly, A. and Wilkinson,D.J.(2006) .Bayesian inference for nonlinear multivariate diffusions. Statistics and Computing 16,p.p 323-338.

[4] Zhou, H. (2001)Finite sample properties of EMM, GMM, QMLE, and MLE for a square-root interest rate diffusion model. Journal of Computational Finance, 2(5), p.p.89-122.
[5] Tian T, Xu S, Gao J, et al. Simulated maximum likelihood method for estimating kinetic rates in gene expression[J]. Bioinformatics, 2007, 23(1): 84-91.
[6] Eraker,B.(2001).MCMC analysis of diffusion models with application to finance. Journal of Business and Economic Statistics 19, p.p177-191.
[7] Dalgaard, P. (2002) Introductory Statistics with R. Springer, New York.
[8] Timm A ,Yin J.Kinetics of viruses production from single cells.Virology.2012;p.p 424:11-17.
[9] Gupta A.Parameter estimation in deterministic and stochastic models of biological systems,Ph.D.thesis,University of Wisconsin-Madison,Madison,WI,2013.
[10] Wilkinson DJ.Stochastic Modelling for Systems Biology.Vol.44.Boca Raton.FL:CRC press,2012.
[11] Poovathingal S K, Gunawan R. Global parameter estimation methods for stochastic biochemical systems[J]. BMC bioinformatics, 2010, 11(1): 414.
[12] Wang Y, Christley S, Mjolsness E, et al. Parameter inference for discretely observed stochastic kinetic models using stochastic gradient descent[J]. BMC systems biology, 2010, 4(1): 99.
[13] Daigle B J, Roh M K, Petzold L R, et al. Accelerated maximum likelihood parameter estimation for stochastic biochemical systems[J]. BMC bioinformatics, 2012, 13(1): 68.
[14] Munsky B, Khammash M. Identification from stochastic cell-to-cell variation: a genetic switch case study[J]. IET systems biology, 2010, 4(6): 356-366.
[15] wang, jingyu, yan, zaizi, lai, junfeng. Numerical simulation on maximum likelihood estimation of diffusion processes[J]. Metallurgical and Mining Industry, 2015, 7(3): 244-249