Numerical simulation on Maximum likelihood estimation of Diffusion Processes

Jingyu Wang [1, 2]

[1] *Science College, Inner Mongolia University of Technology,*
Hohhot 010051, Inner Mongolia , China
[2] *Science College, Qiqihaer University,*
Qiqihaer 161006, Heilongjiang, China

Junfeng Lai

Science College, Inner Mongolia University of Technology,
Hohhot 010051 , Inner Mongolia, China

Zaizai Yan*

Science College, Inner Mongolia University of Technology,
Hohhot 010051, Inner Mongolia, China
**Corresponding Author (zz.yan@163.com)*

Abstract
In this paper, we consider parametric estimation problem of a continuous type stochastic mathematical model (stochastic differential equation) in a wide engineering field. On analyzing the probability characteristics of process, the density function is determined by using $It\hat{o}$ differential law. The maximum-likelihood estimating (MLE) algorithm of unknown parameter is obtained. The approximation is calculated by using numerical solution techniques for diffusion process. Finally, we consider three methods for solving the Cox-Ingersoll-Ross process as a numerical example.
Key words: STOCHASTIC DIFFERENTIAL EQUATIONS, DIFFUSION PROCESSES, COX-INGERSOLL-ROSS PROCESS

1. Introduction

Stochastic differential equations (SDEs)

$$dX = \mu(X;\theta)dt + \sigma(X;\theta)dW \qquad (1)$$

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 Guass density by chi-squared density which is derived from the Milstein scheme [8]. Shoji and Ozaki [12] obtained an approximating for an Ornstein-Uhlenbeck process. Chan, Karolyi, Longstaff, and Sanders [4] use moments based one equation：

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

Gallant and Tauchen, Bibby [2] 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 the parametric estimation problem for $It\hat{o}$ process using the method of maximum likelihood (ML) with discretely sampled data. Section 2 we consider the main result of the paper, a characterization of the exact likelihood function of the discrete sample as the solution to a particular functional partial differential equation(p.d.e).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 review three important models. In section 4 we consider maximum likelihood estimation and show some approaches. In section 5, we present a numerical example.

2 Results

We consider the stochastic differential equation (1) where W is an r-dimensional Wiener process. The $\theta \in \Theta \subseteq \mathbf{R}^p$ is an unknown parameter.

If the transition densities $p(\Delta t, X_i \mid X_{i-1};\theta)$ of X are known, we can use the log-likelihood function

$$l_n(\theta) = \sum_{i=1}^{\infty} \log(p(\Delta t, X_i \mid X_{i-1};\theta)) \quad (3)$$

For θ. The maximum likelihood estimator $\hat{\theta}_n$ is known to have the usual good properties(see Billingsley, Florens-Zmirou [9]). In the case of time-equidistant observations $(t_i = i\Delta t, i = 0,1,...,n)$ for some fixed $\Delta t > 0$ Dacunha-Castelle [3] prove consistency and asymptotic normality of $\hat{\theta}_n$ as $n \to \infty$ irrespective of the value of Δt.

Unfortunately the transition densities of X are usually unknown. Suppose the process $X(t)$ is sampled at $n+1$ discrete points in time $t_0, t_1, ..., t_n$, not necessarily equally spaced apart and let $X \equiv (X_0, X_1, ..., X_n)$. Denote this random sample where $X_k \equiv X(t_k)$. Given the discretely sampled data X and the stochastic specification of the process X(t), denote by $P(X_0, X_1, ..., X_n;\theta)$ the finite-dimensional distribution of X and let $\rho(X;\theta)$ denote the density representation of P. Since $X(t)$ is a Markov process, the joint density ρ may be rewritten as the following product of conditional densities:

$$\rho(X) = \rho_0(X_0)\prod \rho_k(\Delta t, X_i \mid X_{i-1};\theta) \quad (4)$$

Theorem 1:

The likelihood function ρ_k solves the following functional partial equation:

$$\frac{\partial}{\partial t}(\rho_k) = -\frac{\partial}{\partial X}(a\rho_k) + \frac{1}{2}\frac{\partial^2}{\partial X^2}(b^2\rho_k) \quad (5)$$

Because the equation (5) is a partial differential equation, it may be solved by standard methods to yield the likelihood function. Also, additional restrictions upon the coefficient functions may simplify these calculations. The a and b satisfy the following condition

$$\frac{\partial}{\partial x}[\frac{1}{b^2}\frac{\partial b}{\partial t} - \frac{\partial}{\partial x}[\frac{a}{b}] + \frac{1}{2}\frac{\partial^2 b}{\partial x^2}] = 0 \quad (6)$$

It may be shown that there exists a transformed process Z(t)of X(t) for which the coefficient functions are independent of Z(t). That is, for some suitable change of variables $F[X(t)] \equiv Z(t)$, an application of $It\hat{o}$'s lemma will yield:

$$dZ = p(t;\theta)dt + q(t;\theta)dW$$

In this case the transition density function for the transformed data is readily derived as

$$\rho_k(Z,t) = [2\pi\int_{t_{k-1}}^{t} q^2 d\tau]^{-\frac{1}{2}} exp[-\frac{(Z - Z_{k-1} - \int_{t_{k-1}}^{t} pd\tau)^2}{2\int_{t_{k-1}}^{t} q^2 d\tau}]$$

3 Important process

3.1 Orenstein-Uhlenbeck or Vasicek model

The Ourenstein-Uhlenbeck processes is the unique solution to the following stochastic differential equation

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t \qquad (7)$$

For $\theta_2 > 0,$ the invariant law is the Gaussian density with mean $\frac{\theta_1}{\theta_2}$ and variance $\frac{\theta_3^2}{2\theta_2}$, $X_t \sim N(\frac{\theta_1}{\theta_2}, \frac{\theta_3^2}{2\theta_2})$.

For any $t \ge 0,$ the Ornstein-Uhlenbeck processes has a Gaussian transition density $p_\theta(t, X_t | X_0 = x_0).$ The of the distribution of X_t given $X_0 = x_0$ with mean and variance respectively

$$m(t,x) = E_\theta(X_t | X_0 = x_0) = \frac{\theta_1}{\theta_2} + (x_0 - \frac{\theta_1}{\theta_2})e^{-\theta_2 t}$$

If $\theta_1 = 0,$ the stochastic differential equation becomes

$$dX_t = -\theta_2 X_t dt + \theta_3 dW_t \qquad (8)$$

and the parameters are θ_2 and θ_3. In this case, if the sampling rate Δt is fixed, the maximum likelihood estimator of θ_2 is available in explicit form and takes the form

$$\hat{\theta}_{2,n} = \frac{-1}{\Delta t}\log(\frac{\sum_{i=1}^{n} X_{i-1}X_i}{\sum_{i=1}^{n} X_{i-1}^2}) \qquad (9)$$

which is defined only if $\sum_{i=1}^{n} X_{i-1}X_i > 0.$ (see Pedersen [11]) .The maximum likelihood estimator of θ_3^2 is given by

$$\hat{\theta}_{3,n}^{\,2} = \frac{2\hat{\theta}_{2,n}}{n(1-e^{-2\Delta\hat{\theta}_{2,n}})}\sum_{i=1}(X_i - X_{i-1}e^{-\Delta\hat{\theta}^2})^2 \qquad (10)$$

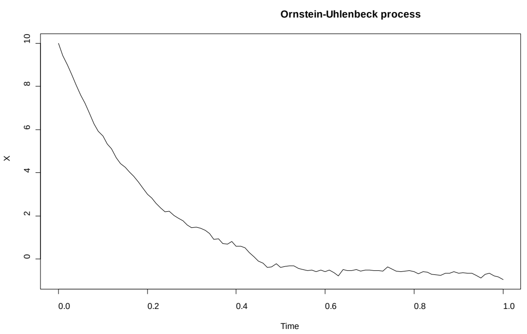

Figure 1. Ornstein-Uhlenbeck process

3.2 The Black and Scholes or geometric Brownian motion model

The Black and Scholes, or geometric Brownian motion model solves the stochastic differential equation

$$dX_t = \theta_1 X_t dt + \theta_2 X_2 dW_t,\ X_0 = x_0, \theta_1 \in R, \theta_2 \in R_+$$

The conditional density function $p_\theta(t,/x)$ is log-normal with mean and variance

$$m(t,x) = xe^{\theta_1 t}, v(t,x) = x^2 e^{2\theta_1 t}(e^{\theta_2^2 t} - 1) \quad (11)$$

Therefore

$$p_\theta(t, y \mid x) = \frac{1}{\theta_2 y\sqrt{2\pi t}}\exp(\frac{\log y - \log x + (\theta 1 - \frac{1}{2}\theta_2^2 t)^2}{-2\theta_2^2 t})$$

3.3 The Cox-Ingersoll-Ross model

Another interesting process is the Cox-Ingersoll-Ross model [5] solution to the stochastic differential equation

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3\sqrt{X_t}dW_t \qquad (12)$$

where

$\theta_1, \theta_2, \theta_3 \in R_+$.

Under this configuration of the parameters, the conditional density $p_\theta(t,.| x)$ follows a non-central χ^2 distribution

$$p_\theta(t, y/x) = ce^{-u-v}(\frac{u}{v})^{\frac{q}{2}} I_q(2\sqrt{uv}), ...(13)$$

where

$$c = \frac{2\theta_2}{\theta_3^2(1 - e^{-\theta_2 t})}, q = \frac{2\theta_1}{\theta_3^2} - 1$$

$$u = cxe^{-\theta_2 t}, v = cy$$

Here $I_q(.)$ is the modified Basel function of the first kind of order

$$I_q(x) = \sum_{k=0}^{\infty}(\frac{x}{2})^{2k+q}\frac{1}{k!\Gamma(k+q+1)}, x \in R,$$

Where $\Gamma(.)$ is the Gamma function,

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx, z \in R_+ \qquad (14)$$

The Cox-Ingersoll-Ross conditional density can be approximated using the mixing-Gamma density [7].

4 Maximum Likelihood Estimation

Because the usual methods for solving partial differential equations are in some cases quite cumbersome, solutions are often obtained by "educated guess." In this section we consider parametric estimation problems for diffusion processes sampled at discrete times. We observe at discrete assume that the process is observed at discrete times

$t_i = i\Delta t, i = 0,1,2,3..., n, and, T = n\Delta t.$

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 are $L_n(\theta)$ and all the derived quantities. There are different ideas to deal with this problem and we will show some options in what follows.

4.1 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 $[t, t+\Delta t)$:

$$X_{t+\Delta t} - X_t = b(X_t,\theta)\Delta t + \sigma(X_t,\theta)(W_{t+\Delta t} - W_t) \quad (15)$$

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

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

This approximation is good if Δ is very small.

Moreover, if the parameters in the vector θ are different for the drift and the diffusion parts, some reasonable results can be obtained. So we assume that $\sigma(x,\theta) = \theta > 0$ is constant and that all the other parameters are in the drift coefficient $b(x,\theta); \sigma$ is not one of the parameters in $\theta.$

4.2 Elerian method

Elerian [9] proposed to use the transition density derived from the Milstein scheme

$$X_{t+\Delta t} = X_t + b(t,X_t)dt + \sigma(t,X_t)(W_{t+dt} - W_t) + \frac{1}{2}\sigma(t,X_t)\sigma_X(t,W_t)((W_{t+dt} - W_t)^2 - dt) \quad (16)$$

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

4.3 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

$$dX_t = b(X_t)dt + \theta dW_t \quad (17)$$

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

$$\frac{d^2x_t}{dt^2} = b_x(x_t)\frac{dx_t}{dt}$$

suppose now that $b_x(x)$ 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

$$x_{t+\Delta t} = x_t + \frac{b(x_t)}{b_x(x_t)}(e^{b_x(x_t\Delta t)} - 1) \quad (18)$$

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

$$X_{t+\Delta t} = X_t e^{K_t\Delta t} + \sigma\int_t^{t+\Delta t} e^{K_t(t+\Delta t-u)}dW_u \quad (19)$$

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

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

$$K_t = \frac{1}{\Delta t}\log(1 + \frac{b(X_t)}{X_t b_X(X_t)}(e^{b_x(X_t)} - 1) \quad (20)$$

In particular, we have that

$$X_{t+\Delta t} / X_t = x \sim N(E_x, V_x) \quad (21)$$

where

$$E_x = x + \frac{b(x)}{b_x(x)}(e^{b_x(x)\Delta t} - 1) \quad (22)$$

$$V_x = \sigma^2 \frac{e^{2K_x \Delta T} - 1}{2K_x}. \quad (23)$$

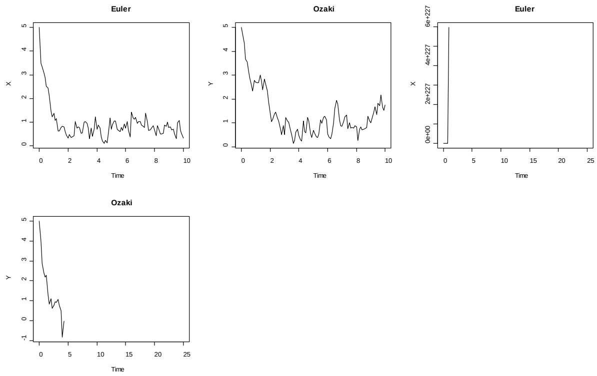

Figure 2. Euler and Ozkai graph

5 Simulation and analysis

We simulate one long trajectory of the Cox-Ingersoll-Ross process with parameters: $(\theta_1, \theta_2, \theta_3) = (0.4, 0.2, \sqrt{0.03})$.

We set N equal to 500000 and $\Delta = 0.001$ and then sample the trajectory two different value of Δt to show the convergence of approximations to the likelihood.

We compare the true against the Euler, Elerian, and Ozaki methods for fixed values of $\theta_1 = 0.4$ and $\theta_3 = \sqrt{0.03}$ as a function of θ_2 in tables.

Table 1. Different estimates of the parameter ($\Delta = 4$)

N	True	Euler	Elerian	Ozkai
126	0.2072	0.1951	0.1940	0.1963
251	0.1956	0.1935	0.1928	0.1942

Table 2. Different estimates of the parameter ($\Delta = 2$)

N	True	Euler	Elerian	Ozkai
126	0.1937	0.1937	0.1934	0.1941
251	0.1941	0.1941	0.1940	0.1943

6 Conclusion

In this paper, we construct three kinds of methods for solving equation (1). the Euler methods ,the Elerian method, and Ozkai methods .The numerical results show that these methods are meaningful methods .Finally, an extension of these result to multivariate diffusion will be investigated.

Acknowledgments

This work was supported by National Natural Foundation of china (11161031,11461051) and National Science and Technology Support Plan of china (2013BAK12B0803).

References

1. Arnold, L. (1974) *Stochastic Differential Equations: Theory and Applications*. John Wiley and Sons, New York.
2. Bibby, M. (1995) Martingale estimating functions for discretely observed diffusion processes. *Bernoulli*, 1(2), p.p.17-39.
3. Cacunha-Castelle (1986) Estimation of the coefficients of a diffusion from discrete observations. *Stochastics*, 19(2),*263-284*.
4. Chan, K.C., Sanders, A.B.(1992) An empirical investigation of alternative models of the short-term interest rate. *J. Finance*, 47(3), p.p.1209-1227.
5. Cox, J.C., Ingersoll, E., Ross, S.A. (1985) A theory of the term structure of interest of rates. *Econometrica*, 53(2), p.p.385-408.
6. Dalgaard, P. (2002) *Introductory Statistics with R.* Springer, New York.
7. Devroye, L.(1986) Pseudo random number generation by nonlinear method. *Int. Stat. Rev*, 63(2), p.p.247-255.
8. Elerian, O. (1998) A note on the existence of a closed form conditional density for the Milstein scheme, Working Paper, Nuffield College, Oxford University. Availablle at http://www.nuff.ox.ac.uk/economics/papers/.
9. Florens-Zmirou, D. (1989) Approximate discrete time schemes for statistics of diffusion processes. *Statistics*, 20(4), p.p.547-557.
10. Kutoyants, Y. (1994) *Identification of Dynamical Systems with Small Noise*. Kluwer, Dordrecht.
11. Pedersen, A, R. (1995) A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observation. *Scandinavian Journal of Statistics*, 22(1), p.p.55-71.

12. Shoji, O. (1998) Estimation for nonlinear stochastic differential equations by a local linearization method. *Stochastic Anal.Appl*, 14(4), p.p.733-752.
13. Yoshida, N. (1992) Estimation for diffusion processes from discrete observation. *J. Multivar. Anal*, 41(2), p.p.220-242.
14. 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.