Theory of Stochastic Processes
Vol. 17 (33), no. 1, 2011, pp. 12–15

JAYA P. N. BISHWAL

SUFFICIENCY AND RAO–BLACKWELLIZATION OF VASICEK MODEL

We use sufficiency and Rao–Blackwell theorem to obtain efficient estimators and discretize the continuous time Vasicek process optimally.

1. Introduction

Parameter estimation in a continuous time process was first studied by Arató (1962) under the directions of Kolmogorov. Consider the mean reverting Vasicek model of short term interest rate

$$d\eta_t = (\beta - \theta\eta_t)dt + dW_t$$

where W is a Brownian motion, the parameter β is the level of mean reversion and the parameter θ is the speed of mean reversion which needs to be estimated from data on $\{\eta_t\}$. We use sufficiency and Rao-Blackwell theorem to obtain efficient estimators and discretize the continuous process optimally. Sampling of diffusion processes can be done at random and nonrandom time points, see Bishwal (2008). We will consider nonrandom (deterministic) sampling. Arató and Fefyverneki (2002) devised method for the solving maximum likelihood estimator which are solutions of complicated system of equations. They used recursive methods like Newton's method and another recursive method called iteration with roots and compared the results with simulation studies. This question often aries in financial econometrics and statistics: how often should one sample a continuous time stochastic process at discrete time points? In general, it is believed that one should sample as often as possible to get the best estimate. But this is not always the case. Discretization problem of stationary processes was first studied by Vilenkin (1959) who considered estimation of mean and correlation function of a stationary process, see also Yaglom (1987). Vilenkin (1959) showed that there exists an optimal discretization based on a certain number n^* of equally spaced observation points on $[0, T]$ of the continuous stationary process for which the estimator of mean is efficient (has minimum variance). One should restrict oneself to some value of n^* which gives the optimal discretization and one should not strive to get the maximum possible number of observation points since accuracy of results obtained becomes worse for increased sample size.

2. Sufficiency and Rao-Blackwellization

Consider the stationary Ornstein-Uhlenbeck process

$$dX_t = -\theta X_t dt + dW_t, \quad 0 \le t \le T, \quad \theta > 0,\ X_0 \sim \mathcal{N}\left(0, \frac{1}{2\theta}\right).$$

2000 *Mathematics Subject Classification.* Primary 62F12, 62F15, 62M05; Secondary 60F05, 60F10, 60H10.

Key words and phrases. Itô stochastic differential equation, diffusion, discrete observations, optimal discretization, experimental design, maximum likelihood estimators, least squares estimator, sufficiency, efficiency.

Let P_θ^T be the measure generated by the process X, P_W^T be the standard Wiener measure and $\mathcal{L}$ be the Lebesgue measure. Then the Radon-Nikodym derivatives are given by

$$\frac{dP_\theta^T}{d\mathcal{L}}(X_0^T) = \sqrt{\frac{\theta}{\pi}}\exp\{-\theta X_0^2\}$$

and

$$\frac{dP_\theta^T}{d(P_W \times \mathcal{L})}(X_0^T) = \sqrt{\frac{\theta}{\pi}}\exp\left\{-\frac{\theta^2}{2}\int_0^T X_t^2 dt - \frac{\theta}{2}[X_T^2 + X_0^2] + \frac{\theta T}{2}\right\}.$$

Using

$$\int_0^T X_t dX_t = \int_0^T X_t(-\theta X_t dt + dW_t) = -\theta\int_0^T X_t^2 dt + \int_0^T X_t dW_t = \frac{1}{2}[X_T^2 - X_0^2 - T]$$

one gets

$$\frac{dP_0^T}{d(P_W \times \mathcal{L})}(X_0^T) = \sqrt{\frac{\theta}{\pi}}\exp\left\{-\frac{\theta^2}{2}\int_0^T X_t^2 dt - \theta\left[\int_0^T X_t dX_t + X_0^2\right] + \frac{\theta T}{2}\right\}.$$

Consider the Vasicek model

$$d\eta_t = (\beta - \theta X_t)dt + dW_t$$

and if θ is known then

$$\frac{dP_\beta^T}{dP_\theta^T}(\eta_0^T) = \exp\left\{\theta\beta\left[\eta_T + \eta_0 + \beta\int_0^T \eta_t dt\right] + \beta^2\left(1 + \frac{\beta T}{2}\right)\right\}.$$

From this we see that $(\eta_0 + \eta_T,\ \int_0^T \eta_t dt)$ is a system of sufficient statistics and

$$\hat{\beta}_T := \frac{\eta_0 + \eta_T + \theta\int_0^T \eta_t dt}{2 + \theta T}$$

is an efficient estimator in the sense of having minimum variance. One can use the Rao-Blackwell theorem.
When $\beta = 0$ and θ is unknown, $(X_0^2 + X_T^2,\ \int_0^T X_t^2 dt)$ is a system of sufficient statistics, but the maximum likelihood estimator is nonlinear function of sufficient statistics

$$\hat{\theta}_T := \frac{-\frac{1}{2}[X_0^2 + X_T^2 - T] + \sqrt{(\frac{1}{2}[X_0^2 + X_T^2 - T])^2 + \int_0^T X_t^2 dt}}{2\int_0^T X_t^2 dt}.$$

Using the transformation

$$s = \frac{t}{T}, \quad \xi_s = \frac{X_s}{\sqrt{T}}, \quad W_s = \frac{W_t}{\sqrt{T}}$$

we have

$$d\xi_s = -\kappa\ \xi_s\ ds\ +\ dW_s, \quad 0 \le s \le 1, \quad \xi_0 \sim \mathcal{N}\left(0, \frac{1}{2\theta T}\right).$$

Thus, without loss of generality, it is enough to consider the normalized case $T = 1$ with the unknown parameter $\kappa = \theta T$. Then $\kappa \to \infty$ gives the stationary regime and $\kappa \to -\infty$ gives the nonstationary regime.
The MLE of κ is given by

$$\hat{\kappa} := -\frac{\xi_1^2 - 1}{2\int_0^1 \xi_s^2 ds}.$$

Assuming κ known and β unknown, (if $\kappa \to 0$) consider the estimator

$$\tilde{\beta} := \frac{\eta_0 + \eta_T}{2}.$$

Proposition 1 The variance of the estimator $\tilde{\beta}$ satisfies $\mathrm{var}(\tilde{\beta}) < \mathrm{var}(\int_0^1 \eta_t dt)$ when $\kappa < 2$.

This means that one should restrict oneself to some value of N and should not strive to get the maximum possible number of observations since the accuracy of the result becomes worse with increased sample size.

Proposition 2 Given κ, there is an optimal division (minimal number of grid) points N^* such that $(0 = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = 1)$ $\tilde{\beta}_{N^*}$ is the best, i.e.,

$$\mathrm{var}(\tilde{\beta}_2) \geq \mathrm{var}(\tilde{\beta}_3) \geq \cdots \geq \mathrm{var}(\tilde{\beta}_{N^*}) < \mathrm{var}(\tilde{\beta}_{N^*+1}) < \mathrm{var}(\tilde{\beta}_{N^*+2}) < \cdots < \mathrm{var}(\tilde{\beta}_N)$$

where

$$\tilde{\beta}_m := \frac{1}{m+1}\sum_{i=0}^{m} \eta_{t_i}.$$

This estimator was studied by Vilenkin (1959) who showed that there exists a number of points $N^* < \infty$ such that

$$\mathrm{var}(\tilde{\beta}_{N^*}) \leq \mathrm{var}(I_T)$$

where

$$I_T = \frac{1}{T}\int_0^T \eta_t dt.$$

Ryzhov (1971, 1972) considered a different estimator

$$\check{\beta}_m := \frac{1}{m}\sum_{i=1}^{m} \eta_{t_i}.$$

He showed that

$$\mathrm{var}(\check{\beta}_m) \geq \mathrm{var}(I_T)$$

for any m. The problem with Ryzhov's estimator is that it not sufficient as excludes one end point observation.

It is known that the Laplace transform $\psi(p,\kappa) := E_\kappa \exp\{-p\int_0^1 \xi_s^2 ds\}$ is given by

$$\psi(p,\kappa) = (\cosh\nu + \frac{\kappa}{\nu}\sinh\nu)^{-1/2}$$

where $\nu := (\kappa^2 + 2p)^{-1}$. See Novikov (1972) and Arató (1982).

At the same time, the two dimensional Laplace transform

$$\phi(p,r;\kappa) \quad := \quad E_\kappa \exp\{\frac{-r}{2}(\xi_1^2 - 1) - p\int_0^1 \xi_s^2 ds\}$$

is given by

$$\phi(p,r;\kappa) \quad = \quad \psi\left(\frac{1}{2}(\varrho^2 - (\kappa-r)^2),\ \kappa - r\right)$$

where $\varrho := (\kappa^2 + 2p)^{1/2}$.

The density function of $\hat{\kappa}$ can be computed by Cramér's theorem. The distribution function of $\hat{\kappa}$ is given by

$$F_{\hat{\kappa}}(x) = \frac{1}{2\pi j}\int_{\mathbb{R}} \frac{\psi(jt,\kappa) - \phi(jt,-jtx;\kappa)}{t} dt$$

where $j = \sqrt{-1}$. See Rao (1978).

If $\kappa \to \infty$, then

$$\frac{\hat{\kappa}-\kappa}{\sqrt{\kappa}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,2).$$

See Arató (1982).

If $\kappa \to -\infty$, then

$$2\kappa e^{\kappa}(\hat{\kappa}-\kappa) \xrightarrow{\mathcal{D}} \mathcal{C}$$

where $\mathcal{C}$ denotes standard Cauchy distribution.

Note that a discretely observed Ornstein-Uhlenbeck process is a Gaussian autoregressive process of order 1.

Let the Ornstein-Uhlenbeck process be observed at discrete time points $\frac{i}{n}$, $i = 0, 1, \ldots, n$.
Denote $\zeta_i := \xi_{\frac{i}{n}}$, $i = 0, 1, \ldots, n$.
Then ζ_i satisfies the equation

$$\zeta_{i+1} = \gamma \zeta_i + \epsilon_i$$

where $\gamma = \exp\{-\frac{\kappa}{n}\}$ and ϵ_i is a zero mean white noise process.
The MLE of γ

$$\hat{\gamma}_n = \frac{\sum_{i=1}^n \zeta_i \zeta_{i-1}}{\sum_{i=1}^n \zeta_{i-1}^2}.$$

Proposition 3 If $(1-\gamma^2)n \to \infty$, then

$$\frac{\sqrt{n}(\hat{\gamma}_n - \gamma)}{\sqrt{(1-\gamma^2)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Due to strong Markov property of ξ_t, for $t > s$, we have

$$E(\xi_t | \mathcal{F}_s) = \xi_s e^{-\theta(t-s)}.$$

The mean integrated squared error is given by, for $t > s$

$$\begin{aligned}
& E\left[\int_0^T (\xi_t - E(\xi_t|\mathcal{F}_s))^2 dt\right] \\
= & E\left[\int_0^{t_1} (\xi_t - E(\xi_t|\mathcal{F}_s))^2 dt + \int_{t_1}^{t_n} (\xi_t - E(\xi_t|\mathcal{F}_s))^2 dt + \int_{t_n}^T (\xi_t - E(\xi_t|\mathcal{F}_s)^2 dt\right] \\
= & E\left[\int_0^{t_1} (\xi_t)^2 dt + \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} (\xi_t - E(\xi_t|\mathcal{F}_s))^2 dt + \int_{t_n}^T (\xi_t - E(\xi_t|\mathcal{F}_s))^2 dt\right]
\end{aligned}$$

Given the observations at $0 \le t_1 \le t_2 \le \cdots \le t_n \le T$, the least squares estimate is $E(\xi_t|\mathcal{F}_s) = 0$ if $t \in [0, t_1)$ and $\xi_{t_i} e^{-\theta(t-t_i)}$ if $t \in [t_i, t_{i+1})$.

The minimizer of the mean integrated squared error gives the optimum sampling time.

Proposition 4 The optimum sampling time satisfies

$$n^* = \frac{1}{2} + \frac{\log(1-2\theta)}{4\theta}.$$

References

1. M. Arató, *Estimation of the parameters of a stationary Gaussian Markov process*, Soviet Math. Dokladi **162** (1962), 905-909.
2. M. Arató, *Linear Stochastic Systems with Constant Coefficients : A Statistical Approach*, Lecture Notes in Control and Information Sciences **45** (1982), Springer, New York.
3. M. Arató and S. Fefyverneki, *New statistical investigations of the Ornstein-Uhlenbeck process*, Comput. Math. Appl. **44** (2002), 677-692.
4. J.P.N. Bishwal, *Parameter Estimation in Stochastic Differential Equations*, Springer-Verlag, 2008.
5. Novikov, A., *Sequential estimation of the parameters of diffusion type processes*, Mathematical Notes **12** (1972), 812-818.
6. M. M. Rao, *Asymptotic distribution of an estimator of the boundary parameter of an unstable process*, Ann. Statist. **6** (1978), 185-190.
7. Y. M. Ryzhov, *Estimation of the mean and the correlation function of a stationary process from discrete data*, Theory. Probab. Appl. **16** (1971), 367-369.
8. Y. M. Ryzov, *Letter to the editor*, Theory. Probab. Appl. **17** (1972), no. 1, 193.
9. S. Y. Vilenkin, *On the estimation of the mean in stationary processes*, Theory. Probab. Appl. **4** (1959) , 415-416.
10. A.M. Yaglom, *Correlation Theory of Stationary and Related Random Functions I, II*, New York: Springer, 1987.

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg, 9201 University City Blvd., Charlotte, NC 28223-0001

E-mail address: J.Bishwal@uncc.edu