Computational Statistics and Data Analysis 103 (2016) 426–437

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Efficient computation of the quasi likelihood function for discretely observed diffusion processes

Lars Josef Höök [a], Erik Lindström [b,*]

[a] Division of Scientific Computing, Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden
[b] Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Box 118, SE-221 00 Lund, Sweden

ARTICLE INFO

Article history:
Received 22 September 2015
Received in revised form 23 May 2016
Accepted 24 May 2016
Available online 9 June 2016

Keywords:
Quasi likelihood
Diffusion process
Conditional moment
Maximum likelihood
Stochastic differential equation

ABSTRACT

An efficient numerical method for nearly simultaneous computation of all conditional moments needed for quasi maximum likelihood estimation of parameters in discretely observed stochastic differential equations is presented. The method is not restricted to any particular dynamics of the stochastic differential equation and is virtually insensitive to the sampling interval. The key contribution is that computational complexity is sublinear in terms of expensive operations in the number of observations as all moments can be computed offline in a single operation. Simulations show that the bias of the method is small compared to the random error in the estimates, and to the bias of comparable methods. Furthermore the computational cost is comparable (actually faster for moderate and large data sets) to the simple, but in some applications badly biased, the Euler–Maruyama approximation.

1. Introduction

Most applications such as simulation or estimation involving Itō stochastic differential equations (SDEs) are in one way or another linked to the transition probabilities of the process. For example, it would be straightforward to estimate parameters using the maximum likelihood method if the transition probability density was known, but this is rarely the case in practice.

However, it is often possible to approximate the transition probability density. The probability density was obtained by brute force numerical computation of the solution to the Fokker–Planck equation (a partial differential equation) in Lo (1988); Lindström (2007) while Monte Carlo based approaches were proposed in Pedersen (1995b), Durham and Gallant (2002), Beskos et al. (2009), Pastorello and Rossi (2010) and Lindström (2012b) and references therein. Those methods are computationally expensive, making them unsuitable for large data sets. A Gauss–Hermite series expansion of the transition probability density was proposed by Aït-Sahalia (2002), although that approach is limited to models with a specific structure.

The recent advances in collecting and storing large amounts of data are shifting the focus away from computationally slow but statistically efficient maximum likelihood methods towards computationally faster, yet not quite as statistically efficient quasi-maximum likelihood methods as the abundance of data often more than makes up for the loss of efficiency.

A simple approach based on the quasi maximum likelihood technique was introduced in Florens-Zmirou (1989) where the conditional mean and variance were obtained from an Euler–Maruyama discretization of the model, see Kloeden and Platen (1992). This is very efficient from a computational point of view and it was shown in Florens-Zmirou (1989) that their method is equivalent to the maximum likelihood estimator as the sampling interval goes to zero, as the bias vanishes.

* Corresponding author.
E-mail address: erikl@maths.lth.se (E. Lindström).

http://dx.doi.org/10.1016/j.csda.2016.05.014
0167-9473/

A higher order version of this approach is proposed in Kessler (1997). Quasi maximum likelihood methods for diffusion processes are consistent provided that the mean and variance are correctly specified, see Bibby and Sørensen (1995) and Sørensen (2012). The bias in the Florens-Zmirou (1989) and Kessler (1997) methods therefore explicitly depends on the quality of the approximation of conditional moments, see Höök and Lindström (2014).

The purpose of this paper is to develop a computationally fast quasi maximum likelihood estimator for discretely observed diffusion processes that is suitable for moderate to large data sets. We show that the computational cost is sublinear rather than superlinear due to how the conditional moments are computed. Our simulations show that the computational complexity is comparable to that of the Euler–Maruyama scheme, and hence magnitudes faster than any approximate maximum likelihood method. This will be achieved without the bias problems associated with the Euler–Maruyama method, a property that is virtually independent of the sampling interval.

The outline of the paper is as follows. In Section 2 we formulate the statistical problem and discuss some alternative techniques for calculating conditional moments. This is followed by Section 3 where we present a numerical implementation that results in sublinear complexity. The resulting parameter estimation algorithm is demonstrated in Section 4 on two qualitatively different diffusion processes as well a randomly sampled data followed by conclusions being drawn in Section 5.

2. Diffusion processes and conditional moments

Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ be a filtered probability space and let $X_t(\theta)$ be a stochastic process defined on that space that solves the following one dimensional stochastic differential equation (SDE)

$$dX_t = a_\theta(X_t)dt + b_\theta(X_t)dW_t, \qquad X_{t_0} = x. \tag{1}$$

We assume throughout the paper that the drift and diffusion terms are regular enough (e.g. bounded growth and local Lipschitz, see Karatzas and Shreve (2012) for alternative conditions) to ensure existence and uniqueness of the solution. The optimal method for estimating the parameters, θ, is the maximum likelihood estimator. Let $x_k = x(t_k),\ k = 1, \ldots, K$ be observations generated from Eq. (1). The maximum likelihood estimator is defined as

$$\hat{\theta}_{MLE} = \underset{\theta\in\Theta}{\operatorname{argmax}}\, \ell(\theta), \tag{2}$$

where the log-likelihood function is given by

$$\ell(\theta) = \log p_\theta(x_0) + \sum_{k=1}^{K} \log p_\theta(x_k|x_{k-1}). \tag{3}$$

The transition probability densities, $p_\theta(x_k|x_{k-1})$ are obtained as the solution to the Fokker–Planck equation (also known as the Kolmogorov forward equation). That partial differential equation (PDE) is given by equation,

$$\frac{\partial}{\partial t} p_\theta(x, t) = \mathcal{L}^* p_\theta(x, t), \tag{4}$$

where the differential operator $\mathcal{L}^*$ is given by

$$\mathcal{L}^* = -\frac{\partial}{\partial x} a_\theta(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2} b_\theta^2(x). \tag{5}$$

The initial condition for the Fokker–Planck equation when starting from a known location is a Dirac delta function, $p_\theta(x|x_k) = \delta(x - x_k)$. This initial condition is likely to cause problems for numerical implementations of Eq. (4) due to discontinuity, see the implementation in Lo (1988) and the remedy proposed in Lindström (2007). A multivariate version of the Fokker–Planck equation can for example be found in Lindström et al. (2015, p. 286), but we focus on the univariate version in this paper for simplicity.

Another method for computing the transition probability is to use the Markov property and law of total probability, adding and integrating out an intermediate state variable, see Pedersen (1995a,b). Define s such that $t_{k-1} < s < t_k$. It then holds that

$$\begin{aligned} p_\theta(x_k|x_{k-1}) &= \int p_\theta(x_k, x_s|x_{k-1})dx_s \\ &= \mathbf{E}_\theta\left[p_\theta(x_k|x_s)|x_{k-1}\right]. \end{aligned} \tag{6}$$

Monte Carlo methods can easily approximate that expected value, but the use of variance reduction techniques is needed for most applications, see Durham and Gallant (2002) and Lindström (2012a). However, we cannot expect to be able to solve either the Fokker–Planck equation (4) or the conditional expectation in Eq. (6) in closed form for more complex models. That means that the computational complexity of any of these approximate maximum likelihood method will be linear (in terms of expensive operations) in the number of observations.

A possible remedy is the Gauss–Hermite or saddle-point expansion, see Aït-Sahalia and Yu (2006) and Varughese (2013). These can be very accurate for frequently sampled data but there are also important limitations. The accuracy often relies