Journal of Econometrics 172 (2013) 106–126

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

A quasi-maximum likelihood method for estimating the parameters of multivariate diffusions

A.S. Hurn[a,*], K.A. Lindsay[a,b], A.J. McClelland[c]

[a] Queensland University of Technology, 2 George Street, Brisbane, 4001, Australia

[b] University of Glasgow, 16 University Gardens, Glasgow, Scotland, G12 8QW, United Kingdom

[c] Sydney Numerix, Level 34, 50 Bridge Street, Sydney, 2047, Australia

ARTICLE INFO

Article history:
Received 4 October 2010
Received in revised form
1 May 2012
Accepted 5 September 2012
Available online 19 September 2012

JEL classification:
C22
C52

Keywords:
Stochastic differential equations
Parameter estimation
Quasi-maximum likelihood
Moments

ABSTRACT

A quasi-maximum likelihood procedure for estimating the parameters of multi-dimensional diffusions is developed in which the transitional density is a multivariate Gaussian density with first and second moments approximating the true moments of the unknown density. For affine drift and diffusion functions, the moments are exactly those of the true transitional density and for nonlinear drift and diffusion functions the approximation is extremely good and is as effective as alternative methods based on likelihood approximations. The estimation procedure generalises to models with latent factors. A conditioning procedure is developed that allows parameter estimation in the absence of proxies.

1. Introduction

Estimation of the parameters of systems of stochastic differential equations by maximum likelihood poses a number of considerable challenges. First and foremost among these is that the likelihood function is seldom known in closed-form. Maximum likelihood estimation, particularly in multiple dimensions, is therefore infeasible for all practical purposes apart from a few trivial cases. The most straightforward way to overcome this problem is to approximate the unknown transitional density function with the Gaussian density function, which is known to be an excellent approximation to the true density for intervals of short duration. Consequently, the transitional probability density function from which the log-likelihood function is constructed is not the true transitional density function and the maximum likelihood estimator based on this misspecified distribution is referred to as the quasi-maximum likelihood estimator.

Quasi-maximum likelihood estimation of the parameters of stochastic differential equations based on the Gaussian approximation is not new (Fisher and Gilles, 1996; Duffee, 2002). The simplest implementation, known as discrete maximum likelihood, requires that the stochastic differential equations be discretised and the discrete drift and diffusion functions so obtained be used to approximate the mean and variance of the true transitional distribution. As these moments are determined by the initial point of each transition and do not change as the process evolves, discrete maximum likelihood is generally not a consistent estimation method. Elerian (1998), Shoji and Ozaki (1998), Kessler (1997) and Huang (2011) all develop ways of improving the Gaussian approximation.

In the univariate case, the most refined form of a quasi-maximum likelihood approach to parameter estimation is the method of Aït-Sahalia (2002), in which the approximating density is expanded in terms of a series of orthogonal polynomials in such a way that the approximating density approaches the true density in the limit.[1] This polynomial expansion approach works particularly well in the cases in which it has been applied (see Jensen and Poulsen, 2002, Bakshi et al., 2006 and Hurn et al., 2007 for the time-homogeneous diffusions and Egorov et al., 2003 for the time-inhomogeneous diffusions) and is widely regarded as the method

* Corresponding author. Tel.: +61 7 31385066; fax: +61 7 31381500.
E-mail addresses: s.hurn@qut.edu.au (A.S. Hurn), k.lindsay@qut.edu.au (K.A. Lindsay), clelland@numerix.com (A.J. McClelland).

[1] If the approximating density is a Gaussian density, then the appropriate polynomials are Hermite polynomials. Other approximating densities will use their associated polynomial basis.

0304-4076/$ – see front matter
doi:10.1016/j.jeconom.2012.09.002

of choice for estimating the parameters of univariate diffusions. The polynomial expansion method translates into the multivariate sphere provided a distinction is drawn between reducible and irreducible models. Aït-Sahalia (2008) develops a closed-form approximation to the multivariate log-likelihood function that is similar in spirit to the polynomial expansion and Aït-Sahalia and Kimmel (2007, 2010) apply this likelihood approximation to stochastic volatility and interest rate models respectively.

Rather than seek yet more complex ways in which to tackle the issue of parameter estimation for multivariate diffusions, this paper returns to the most simple implementation of quasi-maximum likelihood, namely to approximate the unknown transitional distribution of the process with the multivariate Gaussian density. The focus is on developing accurate estimates of the first two moments of the true transitional probability distribution for use with the Gaussian approximation. In the process of doing this a number of new results are derived and the it is shown, both under simulation and in terms of an empirical application, that in the multivariate sphere the vanilla version of quasi-maximum likelihood is very effective.

The main contributions made by this paper may be succinctly summarised as follows. *First*, for systems of stochastic differential equations with affine drift and diffusion functions, it is shown that it is always possible to provide closed-form expressions for the first and second moments of the true but unknown transitional probability distribution. Use of these expressions with the Gaussian approximation will result in consistent parameter estimates of the model parameters irrespective of the fact that a misspecified transitional probability distribution has been used (see, for example Bollerslev and Wooldridge, 1992). Furthermore, increasing the dimension of the system to be estimated does not pose any particular difficulties. *Second*, for non-affine drift and diffusion functions it remains possible to solve for the first and second moments of the transitional distribution numerically. The use of the Gaussian approximating density has the important by-product of allowing crucial integrals in the expressions for the moments to be computed to very high accuracy using Gaussian quadratures. *Third*, new analysis is presented to demonstrate the limiting behaviour of the root mean square error in the Gaussian approximation to the true transitional distribution in multiple dimensions for intervals of short duration. A general theoretical result is obtained which is verified analytically and numerically by means of a novel application of Parseval's theorem. *Fourth*, a new approach to dealing with unobserved state variables within the quasi-maximum likelihood framework is presented, which is based on conditioning the multivariate Gaussian density function on the unobserved states. Unlike the algorithm developed by Bates (2006), this conditioning approach based on the Gaussian density is applicable to non-affine stochastic differential equations. The resultant filtering algorithm is illustrated in the context of Heston's stochastic volatility model (Heston, 1993) and a non-affine extension of this model. *Fifth*, simulation evidence is presented to show that in a multivariate setting, the performance under simulation of the quasi-maximum likelihood based on a Gaussian approximation is comparable with that of the closed-form likelihood approximations reported by Aït-Sahalia and Kimmel (2007).

2. Specification and estimation

Suppose the N-dimensional process $\mathbf{X}(t) = (X_1, \ldots, X_N)$ with sample space $\mathscr{S}$ satisfies the stochastic differential equation

$$dX_k = \mu_k(\mathbf{X};\theta)\,dt + \sum_{j=1}^{M} \sigma_{kj}(\mathbf{X};\theta)\,dW_j, \quad k = 1, \ldots, N, \tag{1}$$

where θ is a vector of model parameters, $\sigma(\mathbf{X};\theta)$ is an array of dimension $N \times M$ with $M \le N$, and dW_j is the increment in the j-th component of the M dimensional vector Wiener process $W(t) = (W_1, \ldots, W_M)$ with $M \times M$ covariance matrix $Q = [Q_{ij}]$ defined by $Q_{ij}\,dt = \mathbb{E}\left[dW_i\,dW_j\right]$.

Let $f_0(\mathbf{X}, t \mid \mathbf{X}_0, \theta)$ denote the true transitional probability density function of the process $\mathbf{X}$ at time $t > 0$ starting initially at $\mathbf{X}_0$, then the k-th component of the probability flux vector associated with Eq. (1) is

$$J_k = \mu_k(\mathbf{X};\theta) f_0(\mathbf{X}, t \mid X_0, \theta) - \frac{1}{2}\sum_{j=1}^{N} \frac{\partial\left(g_{jk}(\mathbf{X};\theta) f_0(\mathbf{X}, t \mid \mathbf{X}_0, \theta)\right)}{\partial X_j}, \tag{2}$$

where $G = [g_{jk}(\mathbf{X};\theta)]$ is the $N \times N$ diffusion matrix given by $\sigma\, Q\, \sigma^{\mathrm{T}}$. Conservation of probability density requires that $f_0(\mathbf{X}, t \mid \mathbf{X}_0, \theta)$ satisfies the Fokker–Planck equation

$$\frac{\partial f_0}{\partial t} + \sum_{k=1}^{N} \frac{\partial}{\partial X_k}\left(\mu_k f_0 - \frac{1}{2}\sum_{j=1}^{N}\frac{\partial\left(g_{jk} f_0\right)}{\partial X_j}\right) = 0, \quad (\mathbf{X}, t) \in \mathscr{S} \times (0, \infty), \tag{3}$$

with boundary conditions

$$\sum_{k=1}^{N} n_k\left(\mu_k f_0 - \frac{1}{2}\sum_{j=1}^{N}\frac{\partial\left(g_{jk} f_0\right)}{\partial X_j}\right) = 0, \quad (\mathbf{X}, t) \in \partial\mathscr{S} \times (0, \infty), \tag{4}$$

where $\mathbf{n}$ is the unit outward normal to $\partial\mathscr{S}$. When the state $\mathbf{X}_0$ is fully observed, the initial density will be a product of delta functions of the observed state variables, that is,

$$f_0(\mathbf{X}, 0 \mid \mathbf{X}_0, \theta) = \prod_{k=1}^{N} \delta(X_k - X_{0,k}), \quad \mathbf{X} \in \mathscr{S}, \tag{5}$$

but otherwise it will be a product of delta functions of the observed state variables and the conditional density of the unobserved state variables.

The parameters of the model are to be estimated using observed data consisting of a sequence of observations, $X_0, \ldots, X_T$, of the system at discrete times $t_0, \ldots, t_T$. The maximum-likelihood estimator of θ, which maximises the conditional log-likelihood function of the observed sample with respect to the parameters θ is

$$\widetilde{\theta} = \arg\max_{\theta} \frac{1}{T}\sum_{p=1}^{T} \log f_0(\mathbf{X}_p, \Delta_p \mid \mathbf{X}_{p-1}, \theta), \tag{6}$$

where Δ_p is the duration of the interval between observations X_{p-1} and X_p. The primary difficulty with this approach, however, is that the log-likelihood function in Eq. (6) is seldom known in closed-form, with the vast majority of known cases relating to univariate models. Consequently maximum-likelihood estimation is infeasible for most practical applications of interest.

A simple alternative to using the true (but unknown) transitional density function in the construction of the log-likelihood function is to replace $f_0(\mathbf{X}, t \mid \mathbf{X}_p, \theta)$ in Eq. (6) with the multivariate Gaussian density

$$f(\mathbf{X}, t \mid \mathbf{X}_p, \theta) = \frac{1}{(2\pi)^{N/2}}\frac{1}{|\Sigma|^{1/2}} \times \exp\left[-\frac{1}{2}\sum_{j,k=1}^{N}(X_j - m_j)\Sigma_{jk}^{-1}(X_k - m_k)\right], \tag{7}$$