Journal of Econometrics 180 (2014) 73–80

Contents lists available at ScienceDirect

Journal of Econometrics

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

Maximum likelihood estimation of partially observed diffusion models ☆

Tore Selland Kleppe [a], Jun Yu [b,*], Hans J. Skaug [a]

[a] University of Bergen, Norway

[b] Singapore Management University, Singapore

ARTICLE INFO

Article history:
Received 13 December 2012
Received in revised form
21 November 2013
Accepted 6 February 2014
Available online 26 February 2014

JEL classification:
C11
C15
G12

Keywords:
Closed-form approximation
Diffusion model
Efficient importance sampler

ABSTRACT

This paper develops a maximum likelihood (ML) method to estimate partially observed diffusion models based on data sampled at discrete times. The method combines two techniques recently proposed in the literature in two separate steps. In the first step, the closed form approach of Aït-Sahalia (2008) is used to obtain a highly accurate approximation to the joint transition probability density of the latent and the observed states. In the second step, the efficient importance sampling technique of Richard and Zhang (2007) is used to integrate out the latent states, thereby yielding the likelihood function. Using both simulated and real data, we show that the proposed ML method works better than alternative methods. The new method does not require the underlying diffusion to have an affine structure and does not involve infill simulations. Therefore, the method has a wide range of applicability and its computational cost is moderate.

1. Introduction

Continuous time diffusion models have long proven useful in economics and finance. For example, they provide a convenient mathematical framework for the development of financial economics and option pricing theory (Black and Scholes, 1973; Heston, 1993; Duffie and Kan, 1996). The separate treatment of stock and flow variables in macroeconomics (Bergstrom, 1984) and the formulation of continuous time inter-temporal optimization models (Turnovsky, 2000) represent additional usage of diffusion models in economics. Not surprisingly, fitting diffusion models based on data sampled at discrete times has received a great deal of attention in econometrics.

The case when some of the state variables are latent is often encountered in practical applications. One example of such partially observed diffusion models is the continuous time stochastic volatility models with the volatility being the latent state; see Hull and White (1987) and Heston (1993). Another example is the continuous time stochastic mean model of Balduzzi et al. (1998), in which the mean is a latent state. Obviously, the combination of a latent volatility and a latent mean also makes a partially observed diffusion model. This class of models has found a wide range of applications in the term structure literature Duffie and Kan (1996), Dai and Singleton (2000) and in the option pricing literature Duffie et al. (2000).

It has been argued, on the basis of asymptotic properties, that the preferred choice of estimation method for diffusion models should be maximum likelihood (ML) (Aït-Sahalia, 2002; Durham and Gallant, 2002). The ML estimation of partially observed diffusions necessitates the computation of the joint transition probability density (TPD) of the observed and the latent state variables as well as the marginalization of the latent variable from the joint density.

When the transition density of the state variables does not have a closed-form expression, it has to be approximated. Many approximation methods have been proposed in the literature which is reviewed in Jensen and Poulsen (2002), Hurn et al. (2007) and Phillips and Yu (2009). Broadly speaking, the methods can be classified

☆ Kleppe gratefully acknowledges the hospitality during his research visit to School of Economics and Sim Kee Boon Institute for Financial Economics at Singapore Management University. Yu thanks the Singapore Ministry of Education for Academic Research Fund under grant number MOE2011-T2-2-096. We are grateful to the co-Editor Yacine Aït-Sahalia and two anonymous referees for constructive comments that have greatly improved the paper. Moreover, we are indebted to Kenneth Lindsay for advise on implementing the method of Hurn et al. (2013) and to Roman Liesenfeld and Atle Øglend for invaluable comments.

* Correspondence to: School of Economics, 90 Stamford Road, Singapore 178903, Singapore. Tel.: +65 68280858; fax: +65 68280833.

E-mail address: yujun@smu.edu.sg (J. Yu).

http://dx.doi.org/10.1016/j.jeconom.2014.02.002
0304-4076/

into two classes. In the first class, the diffusion model is approximated by a discrete time model whose TPD is available in closed form. The well-known Euler–Maruyama (EM) approximation and the approximations of Bergstrom (1966), Nowman (1997), Milstein (1979) and Shoji and Ozaki (1998) all belong to this class. These methods in general lead to a bias in the calculation of the likelihood function that does not vanish asymptotically (see Aït-Sahalia, 2002 and references therein). We shall refer to such a bias as the discretization bias. In the second class, the TPD of the diffusion model is approximated directly, including the infill approximations and the closed-form approximations. To the best of our knowledge, the first contribution on the topic of infill approximations is in Pedersen (1995).[1] The closed-form approximation techniques include the Hermite approximation of Aït-Sahalia (2002), the polynomial approximation of Aït-Sahalia (2008), and the saddlepoint approximation of Aït-Sahalia and Yu (2006). Both the infill techniques and the closed-form techniques can reduce the discretization bias. Compared to the infill techniques, the closed-form techniques are computationally much cheaper and the approximation errors are smaller (Aït-Sahalia, 2002).

When the diffusion specifying the observed and latent states is not a linear Gaussian process, the marginalization of the latent state variable cannot be achieved analytically. Consequently, various importance sampling techniques have been proposed to integrate out the latent state variable from the joint density via simulations. The importance sampler of Shephard and Pitt (1997) and Durbin and Koopman (1997) (see also Sandmann and Koopman, 1998, Durham, 2006, and Skaug and Yu, forthcoming) used a global, multivariate Gaussian approximation to the joint density as the importance density, whereas Richard and Zhang (2007) (see also Liesenfeld and Richard, 2003, 2006) used a product of univariate Gaussian approximations to the conditional TPD as the importance density. While both methods work well for estimating discrete time stochastic volatility models, as shown in Lee and Koopman (2004), the importance sampler of Richard and Zhang (2007) tends to be more well-behaved than that of Shephard and Pitt (1997) and Durbin and Koopman (1997).

For most partially observed diffusion models, almost all of the afore-mentioned ML methods are not directly applicable. Arguably, the most widely used ML approach to estimating partially observed diffusion models is to use the EM approximation to discretize the diffusion models and then use an importance sampler to marginalize out the unobserved latent state variables. This approach naturally leads to the discretization bias when a sampling interval is fixed.

Several approaches have been proposed in the literature to provide ML estimation of partially observed diffusion models with the discretization bias being controlled. Bates (2006) proposed a frequency domain filtering method to compute the likelihood function of latent affine diffusion models via the conditional characteristic functions. This technique is not feasible for non-affine models for which the conditional characteristic functions are not available in closed form.

Aït-Sahalia and Kimmel (2007) proposed to approximate the volatility (latent state) using the implied volatility computed from the underlying options. Consequently, no state variable is latent in the continuous time stochastic volatility model and the closed form approximation of Aït-Sahalia (2008) is directly applicable. However, it is well-known that option prices are derived from the risk neutral measure (Heston, 1993). As a result, using data from both the spot market and the options market jointly, one can simultaneously learn about both the physical measure and the risk-neutral measure. Naturally, this benefit is achieved at a cost. To connect the physical measure to the risk-neutral measure, the functional form of the market price of risk has to be specified. If one's interest is to learn about the physical measure only, the implied volatility is less useful. Moreover, in some cases, such as for models with a stochastic mean, it is not clear how to extract latent variables from derivative prices. Perhaps most importantly, when the latent volatility is approximated by the implied volatility, approximation errors are introduced. How these errors influence the estimated price dynamics remains to be answered.

More recently, a quasi-ML (QML) approach was proposed by Hurn et al. (2013) to estimate partially observed diffusion models. However, the discretization bias cannot be completely removed by this method. While the infill technique combined with the importance sampler via the global approximation has been introduced to provide the ML estimation (see, for example, Durham and Gallant, 2002), it is computationally very expensive.

In this paper, we introduce a new ML method to estimate partially observed diffusion models. Our ML method combines the closed-form approach of Aït-Sahalia (2008) for approximating the joint TPD of the observed and the latent state variables and the efficient importance sampler (EIS) of Richard and Zhang (2007) for integrating out latent states from the joint density. Our method inherits two nice features of the closed-form approximation techniques of Aït-Sahalia (2002) and Aït-Sahalia (2008). First, it can practically remove the discretization bias and, hence, leads to more accurate likelihood values than the QML and the EM methods. Second, it is computationally inexpensive, especially relative to the infill methods. Moreover, our method is very general in the sense that only weak assumptions regarding the structure of the underlying diffusion must be made. Most notably, an affine structure does not need to be assumed.

The paper is organized as follows. Section 2 proposes the new estimation method. Section 3 illustrates our method using the GARCH diffusion model of Nelson (1990) and investigates the performance of the proposed method relative to alternative methods, including the EM method and the QML of Hurn et al. (2013), using simulated data. In Section 4, we fit the GARCH diffusion to real data. Finally, Section 5 concludes and outlines some further applications and implications of the approach.

2. Methodology

2.1. Model specifications

Let the time-homogeneous diffusion be denoted by

$$dX_\tau = a(X_\tau;\theta)d\tau + b(X_\tau;\theta)dB_\tau, \tag{1}$$

where X_τ and $a(X_\tau;\theta)$ are q-vectors, and $b(X_\tau;\theta)$ is a $q \times q$ matrix, with B_τ being a q-dimensional uncorrelated Brownian motion. θ is the vector of parameters to be estimated. We assume that (1) admits a unique solution and that $b(X_\tau;\theta)b(X_\tau;\theta)'$ is positive definite for all admissible values of X_τ and θ. Moreover, we assume that $a(\cdot;\theta)$ and $b(\cdot;\theta)$ are infinitely differentiable. Let $x_t = X_{t\Delta}$ $(t = 1,\ldots,T)$ be the value of X_τ sampled at frequency $1/\Delta$ and $\mathbf{x} = (x_1,\ldots,x_T)$ be the collection of such values.

Due to the Markovian property of X_τ, the joint probability density function (PDF) of $\mathbf{x}$ may be written as

$$p(\mathbf{x};\theta) = p(x_1;\theta)\prod_{t=2}^{T} p_t(x_t|x_{t-1};\theta). \tag{2}$$

Here $p_t(x_t|x_{t-1};\theta)$ is the TPD associated with (1) and $p(x_1;\theta)$ is the, possibly degenerate, density of the initial state x_1. We assume that the first q_y elements of x_t, denoted by y_t, are observed at frequency $1/\Delta$ and $t = 1,\ldots,T$. The remaining $q_z = q - q_y$ elements of

[1] Elerian et al. (2001) and Eraker (2001) use the infill methods to conduct Bayesian inference of continuous time models.