Estimating the Parameters of Stochastic Differential Equations

A S Hurn
Institute of Economics and Statistics, Oxford,

K A Lindsay
Department of Mathematics, University of Glasgow.

November 6, 1997

Abstract

Two maximum likelihood methods for estimating the parameters of stochastic differential equations (SDE) from time-series data are proposed. The first is that of simulated maximum likelihood in which a nonparametric kernel is used to construct the transitional density of an SDE from a series of simulated trials. The second approach uses a spectral technique to solve the Kolmogorov equation satisfied by the transitional probability density. The exact likelihood function for a geometric random walk is used as a benchmark against which the performance of each method is measured. Both methods perform well, with the spectral method returning results which are practically identical to those derived from the exact likelihood. The technique is illustrated by modelling interest rates in the UK gilts market using a fundamental one-factor term-structure equation for the instantaneous rate of interest.

Keywords

maximum likelihood,
transitional density,
kernel,
Kolmogorov equation,
spectral integration,
interest rates.

1 Introduction

The notion of a stochastic system of differential equations (SDEs), defined as a deterministic system of differential equations perturbed by random disturbances that are not necessarily small, has been used profitably in a variety of disciplines including *inter alia* engineering, environmetrics, physics, population dynamics and medicine. Whilst the analysis of stochastic processes has received extensive treatment over a long period, the estimation of the parameters of such processes has until recently received less attention. Indeed it has recently been argued with some force that the difficulty of obtaining consistent estimates of the parameters of nonlinear SDEs is one of the most pressing difficulties in statistical inference of continuous-time processes with discretely sampled data [3].

This article investigates the feasibility of estimating the parameters of a nonlinear stochastic differential equation by two possible strategies, both of which are based on the maximum likelihood principle. The first approach simulates the likelihood by a nonparametric method whereas the second obtains the likelihood by numerical solution of the fundamental partial differential equation underlying its propagation, often called the Kolmogorov or Fokker-Plank equation.

2 Overview of the parameter estimation problem

Suppose that $x_0, x_1, \cdots, x_n$ is a sequence of $n+1$ historical observations of the random variable $X(t)$ sampled at non-stochastic dates $t_0 < t_1 < \ldots < t_n$. The joint transitional density of this sample is

$$f(x_0 \cdots x_n;\theta) = f_0(x_0|\theta) \prod_{k=1}^{n} f(x_k, t_k|x_{k-1}, t_{k-1};\theta) \tag{1}$$

where f_0 is the density of the initial state and θ is a vector of parameters. The optimal values of θ, denoted here by $\hat{\theta}$, may be estimated by maximising the joint transitional density function (1) with respect to the parameters θ. In fact, it is normally more convenient to minimise the negative log-likelihood function

$$\mathcal{L}(\theta) = -\log f_0(\theta|x_0) - \sum_{k=1}^{n} \log f(\theta; x_k, t_k|x_{k-1}, t_{k-1}).$$

As usual, $\hat{\theta}$ is consistent and satisfies the limiting distributional property (see [3])

$$\sqrt{n}(\theta - \hat{\theta}) \sim N\left(0, \mathcal{I}^{-1}(\theta)\right), \qquad \mathcal{I}(\theta) = \lim_{n\to\infty} \frac{1}{n} \mathrm{E}\left[\frac{\partial^2 \mathcal{L}(\theta)}{\partial\theta\partial\theta'}\right]_{\hat{\theta}}.$$

The general SDE in one dimension has format

$$dX(t) = a(t, X;\theta)\,dt + \sigma(t, X;\theta)\,dW(t) \tag{2}$$

where $dW(t)$ is the differential of the Wiener process $W(t)$ and θ is a vector of system parameters. The function $a(t, X;\theta)$ is often called the instantaneous drift of $X(t)$ while $\sigma^2(t, X;\theta)$ is its instantaneous variance. The probability density function $f(t, x;\theta)$ corresponding to (2) satisfies the forward Kolmogorov equation

$$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(t, x;\theta) f\right) - \frac{\partial}{\partial x}\left(a(t, x;\theta) f\right). \tag{3}$$

If a closed form expression for f is available then computation of $\hat{\theta}$ is straightforward. However, this is rarely the case and so it is reasonable to explore other ways to estimate the transitional density f.

2.1 Simulated maximum likelihood

The simulation approach to the construction of the likelihood function is based on the estimation of probability density using a kernel. Given m observations $x_1 \cdots x_m$ of the random variable X, the kernel has generic form

$$\hat{f}(x) = \frac{1}{mh} \sum_{j=1}^{m} K\left(\frac{x - x_j}{h}\right) \tag{4}$$

where h is the kernel bandwidth and $K(.)$ is a suitable non-negative function enclosing unit mass. By construction, $\hat{f}$ satisfies automatically the properties of a probability density function. The operating efficacy of the kernel is dominated by the efficiency with which the bandwidth h can be determined. In this application, the normal kernel was selected since it is known to be only marginally inferior to the optimal Epinechnikov kernel but has a "plug in" bandwidth given by $h = 0.9\sigma m^{-1/5}$ where σ is the standard deviation of the data ([10, 11]).

The implementation of a nonparametric kernel is well documented and understood and so merits no further attention. The simulation procedure for computing the transitional density of the stochastic process described by (2) is as follows:

1. Generate S, an $m \times n$ matrix of vector increments of a Wiener process. The i-th row of S is used in the i-th simulation of the SDE and consists of vector entries S_{ij}, $(j = 1 \cdots n)$ each of which contributes a Gaussian distributed random variable with variance $(t_j - t_{j-1})$. For small intervals (t_{j-1}, t_j), vectors of dimension 1 may be adequate but for larger time intervals, more stochastic realisation may be necessary to achieve satisfactory accuracy.

2. Use a standard algorithm, for example the Milstein scheme ([8]), to integrate the SDE numerically between observations using the stochastic increments contained in S. Store the simulations in T, an $m \times n$ matrix of trials.

3. At each time step t_j, $j = 1 \cdots n$, use the j-th column of the trial matrix T to construct a kernel estimate of the transitional density from t_{j-1} to t_j.

4. Use the kernel to estimate the density of the field data point x_j. The product of these estimates for all n field points is the likelihood function to be maximised. In practice, it is convenient to minimise the negative log-likelihood function.

2.2 "Exact" maximum likelihood

It has already been indicated that every system of SDE's has an associated partial differential equation (Kolmogorov equation) that describes the density of the solutions of the SDE's in time and state space — in effect, the solution of the Kolmogorov equation is the limiting density that would be achieved by an infinite number of simulations of the SDE's. Traditional econometrics has focused on the need to provide closed-form solutions to Kolmogorov's equation. However, this equation is now amenable to modern numerical methods since it is linear and of second order in state space. Therefore it is not necessary to confine attention to the class of SDE's for which a solution is known.

Here the Kolmogorov equation is integrated by a spectral method. The procedure in one dimension is now summarised but generalises to many dimensions. Briefly, the original partial differential equation is first mapped into the finite interval $[-1,1]$. The solution in this interval is then approximated by a spectral series (often based on Chebyshev polynomials) whose coefficients are functions of time only. The partial differential equation is then used to construct a set of ordinary differential equations for the time evolution of these coefficients. Spectral approximations exhibit exponential accuracy as the number of polynomials is increased. In practice machine accuracy is often reached with comparatively small numbers of polynomials.

The procedure is now illustrated for a solution procedure based on Chebychev polynomials. Some preliminary results are established.

2.2.1 Chebychev polynomials

The family of polynomials $T_n(x)$ is defined by the relation

$$T_n(\cos\theta) = \cos(n\theta)\,, \qquad n = 0, 1, \cdots\,. \tag{5}$$

From their definition (5), it can be shown easily that Chebyshev polynomials satisfy the important relations

$$\begin{aligned} 2T_n(z)T_m(z) &= T_{m+n}(z) + T_{|n-m|}(z)\,, \\ 2T_n(z) &= \frac{T'_{n+1}(z)}{n+1} - \frac{T'_{n-1}(z)}{n-1}\,, \qquad n > 1\,, \\ \int_{-1}^{1} \frac{T_n(z)T_m(z)}{\sqrt{1-z^2}}\,dz &= \frac{\pi}{2}c_n\delta_{nm}\,, \qquad c_0 = 2\,, c_n = 1 \quad n \geq 1\,. \end{aligned} \tag{6}$$

Suppose now that g is a function defined in $[-1,1]$ with Chebyshev expansion

$$g(z) = \sum_{k=0}^{\infty} g_k T_k(z)\,, \tag{7}$$

then by multiplying both sides of this equation by $T_n(z)/\sqrt{1-z^2}$ and integrating over $[-1,1]$, it follows immediately that

$$g_n = \frac{2}{\pi c_n}\int_{-1}^{1} \frac{g(z)T_n(z)}{\sqrt{1-z^2}}\,dz = \frac{2}{\pi c_n}\int_0^{\pi} g(\cos\theta)\cos(n\theta)\,d\theta\,. \tag{8}$$

Furthermore, if g is differentiable with respect to z in $(-1,1)$ then property (6_2) can be used to verify that

$$\frac{dg}{dz} = \sum_{k=0}^{\infty} g_k^{(1)} T_k(z)\,, \qquad c_k g_k^{(1)} = g_{k+2}^{(1)} + 2(k+1)g_{k+1} \quad k \geq 0\,. \tag{9}$$

Suppose that the representation (7) involves only $(N+1)$ polynomials so that $g_{N+1} = g_{N+2} = \cdots = 0$, then under these circumstances, $g_N^{(1)} = g_{N+1}^{(1)} = \cdots = 0$, and the iterative relationship (9) now becomes a practical algorithm for the calculation of $g_0^{(1)}, \cdots, g_{N-1}^{(1)}$, the spectral coefficients of $g'(z)$.

2.2.2 Mappings

To take adavntage of these properties of Chebyshev polynomials, it is first necessary to map the sample space of the SDE into $[-1,1]$. If the sample space is $[0,\infty)$, a situation that occurs often in practice, then for any $L>0$, the mapping

$$z=\frac{x-L}{x+L}\,,\qquad x\in[0,\infty)\,,\quad z\in[-1,1)\,,\tag{10}$$

is one popular way to connect x in the sample space with z in $[-1,1)$. If $g(z)$ is the functional form of $f(x)$ when x is replaced by z from formula (10), it can be shown that differentiation with respect to x and z are related by the useful formula

$$x\frac{\partial f}{\partial x}=\frac{1-z^2}{2}\frac{\partial g}{\partial z}\,.\tag{11}$$

Hence if $g(z)$ and $g'(z)$ have Chebyshev spectral representations (7) and (9) then $x\,df/dx$ has spectral expansion

$$\begin{aligned} x\frac{df}{dx}=\sum_{k=1}^{\infty}\hat{g}_k^{(1)}T_k(z) &= \frac{(1-z^2)}{2}\frac{dg}{dz}=\frac{1}{4}\Big(T_0(z)-T_2(z)\Big)\sum_{k=0}^{\infty}g_k^{(1)}T_k(z)\\ &= \frac{1}{2}\left(2g_0^{(1)}-g_2^{(1)}\right)T_0(z)+\frac{1}{2}\left(g_1^{(1)}-g_3^{(1)}\right)T_1(z)\\ &\quad -\frac{1}{2}g_0^{(1)}T_2(z)+\frac{1}{2}\sum_{k=2}^{\infty}\left(2g_k^{(1)}-g_{k-2}^{(1)}-g_{k+2}^{(1)}\right)T_k(z)\,. \end{aligned}$$

In conclusion,

$$\begin{aligned} \hat{g}_0^{(1)}&=\frac{1}{4}\left(2g_0^{(1)}-g_2^{(1)}\right)\,, &\qquad \hat{g}_1^{(1)}&=\frac{1}{4}\left(g_1^{(1)}-g_3^{(1)}\right)\,,\\ \hat{g}_2^{(1)}&=\frac{1}{4}\left(2g_2^{(1)}-2g_0^{(1)}-g_4^{(1)}\right)\,, &\qquad \hat{g}_k^{(1)}&=\frac{1}{4}\left(2g_k^{(1)}-g_{k-2}^{(1)}-g_{k+2}^{(1)}\right)\,. \end{aligned}\tag{12}$$

2.2.3 Kolmogorov's equation

Many popular models in the literature are special cases of the stochastic differential equation

$$dx=\mu(x-\bar{x})dt+\sigma x^{\gamma}dW\tag{13}$$

where $W(t)$ is a standard Wiener process and μ, $\bar{x}$, σ and γ are model parameters. It is therefore convenient and sensible to demonstrate the spectral technique with reference to equation (13). The Kolmogorov equation corresponding to (13) has sample space $[0,\infty)$ and can be expressed in the form

$$\begin{aligned} \frac{\partial f}{\partial t} &= \frac{x^{2(\gamma-1)}\sigma^2}{2}\left[x\frac{\partial}{\partial x}\left(x\frac{\partial f}{\partial x}\right)+(4\gamma-1)x\frac{\partial f}{\partial x}+2\gamma(2\gamma-1)f\right]\\ &\quad -\mu\left[f+x\frac{\partial f}{\partial x}-\bar{x}\frac{\partial f}{\partial x}\right]\,. \end{aligned}\tag{14}$$

The mapping (10) is now used to change variables from x to z in the differential equation (14). When $\bar{x}=0$ and $\gamma=1$, equation (13) describes a geometric random walk with drift μx and variance σ^2x^2. Let $f(t,x)=g(t,z)$ then in this case g satisfies

$$\begin{aligned} \frac{\partial g}{\partial t}&=\frac{\sigma^2}{8}(1-z^2)\frac{\partial}{\partial z}\left((1-z^2)\frac{\partial g}{\partial z}\right)+\frac{(3\sigma^2-2\mu)}{4}(1-z^2)\frac{\partial g}{\partial z}+(\sigma^2-\mu)g\,,\\ g(0,z)&=\delta(z-z_0)\,,\qquad z_0=(1-L)(1+L)\,. \end{aligned}\tag{15}$$

Now suppose that $g(t,z)$ has Chebyshev spectral expansion

$$g(t,z) = \sum_{k=0}^{N} g_k(t) T_k(z) \; . \tag{16}$$

By multiplying (15) by $T_j(z)/\sqrt{1-z^2}$ and integrating over $[-1,1]$, it can be shown that $g_0(t)\,, g_1(t)\,\cdots, g_{N-2}(t)$ satisfy the ordinary differential equations

$$\frac{dg_j}{dt} = \frac{\sigma^2}{8}\hat{g}_j^{(2)}(t) - \frac{(3\sigma^2-2\mu)}{4}\hat{g}_j^{(1)}(t) + (\sigma^2-\mu) g_j \; , \qquad j = 0,1\cdots N-2 \; , \tag{17}$$

where $\hat{g}_j^{(2)}$ and $g_j^{(2)}$ are related through an equation akin to (12) while $g_j^{(2)}$ and $g_j^{(1)}$ are related through an equation similar to (9). In this instance, the coefficients $g_{N-1}(t)$ and $g_N(t)$ are determined algebraically from the relations

$$g(t,-1) = \sum_{k=0}^{N} (-1)^k g_k(t) = 0 \; , \qquad\qquad g(t,1) = \sum_{k=0}^{N} g_k(t) = 0 \tag{18}$$

and correspond to the boundary conditions $f(t,x) \to 0$ as $x \to 0$ and $x \to \infty$ respectively. Self evidently,

$$g_{N-1} = -\frac{1}{2}\sum_{k=0}^{N-2}(1-(-1)^{N-k})g_k \; , \qquad g_N = -\frac{1}{2}\sum_{k=0}^{N-2}(1+(-1)^{N-k})g_k \; . \tag{19}$$

Hence equations (17) are regarded as a system of ordinary differential equations for $g_0 \cdots g_{N-2}$ where it is understood that occurrences of g_{N-1} and g_N on the right hand side of (17) are replaced by expressions (19). Finally, the initial conditions for $g_0\,, g_1\,, \cdots, g_{N-2}$ are derived from the initial condition in (15_2). If $z_0 = \cos\theta_0$ then $\theta_0 = 2\tan^{-1}(\sqrt{L})$ and the values of the coefficients of g corresponding to a delta-like initial condition are

$$g_k(0) = \frac{2}{\pi c_k}\cos(2k\tan^{-1}\sqrt{L}) \; , \qquad k = 0,1\cdots N-2 \; . \tag{20}$$

3 Results

The SDE describing a geometric random walk is an ideal candidate on which to test these ideas as it is not only of practical relevance in stochastic modelling but also has an exact solution and probability density function that are known. If $X(0) = x_0$ then the exact solution and conditional probability density function are respectively

$$\begin{aligned} x(t) &= x_0 \exp[(\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t] \; , \\ f(t,x) &= \frac{1}{2\sigma x\sqrt{\pi t}} \exp\left(- \frac{\left(\log(x/x_0) - (\mu - \frac{\sigma^2}{2})t\right)^2}{2\sigma^2 t} \right) . \end{aligned} \tag{21}$$

Table 1 demonstrates that spectral integration of the Kolmogorov equation can be performed with high accuracy, despite the fact that the initial condition is delta-like. As expected, larger numbers of polynomials are required to resolve the structure of the delta

N	Error	$\Delta t = 0.01$	$\Delta t = 0.05$	$\Delta t = 0.1$	$\Delta t = 0.25$	$\Delta t = 0.5$
20	Absolute Relative	4.8576 ∞	0.8550 ∞	0.2307 ∞	0.0369 ∞	0.0031 0.2685
40	Absolute Relative	2.4720 ∞	0.0665 0.4682	0.0027 0.0582	0.0005 0.0289	0.0001 0.0076
60	Absolute Relative	1.0150 ∞	0.0020 0.0496	0.0011 0.0294	0.0002 0.0088	0.0001 0.0026
80	Absolute Relative	0.3273 0.1046	0.0003 0.0072	0.0004 0.0050	0.0001 0.0035	0.0000 0.0026
100	Absolute Relative	0.0786 0.1271	0.0005 0.0130	0.0000 0.0021	0.0000 0.0006	0.0000 0.0003
150	Absolute Relative	0.0025 0.0089	0.0002 0.0033	0.0001 0.0021	0.0000 0.0009	0.0000 0.0001
200	Absolute Relative	0.0008 0.0027	0.0000 0.0005	0.0000 0.0011	0.0000 0.0003	0.0000 0.0002

Table 1: Comparison of the exact and spectral solutions to Kolmogorov's equation for the geometric random walk with parameters $\mu = 1.0$ and $\sigma = 0.5$. The comparison is made over all nodes whose density exceeds 1% of the peak density.

function for small time intervals. As the time interval increases, significantly less polynomials are required to achieve the same accuracy. In fact, when the spectral solution is used to estimate model parameters in parallel with the exact solution, no discernable differences are encountered for more than 60 polynomials. This is a remarkable result given that each integration step starts from a delta function.

The results of a simulation exercise to compare the kernel and exact/spectral estimates for the parameters of the geometric random walk are reported in Table 2. There are a number of interesting points to note.

1. As expected, the parameter values and their standard errors are uniformly better for the exact/spectral solution.

2. The kernel method performs with credit. The size of the time interval does not seem to influence the estimates unduly provided the time intervals over which the simulations are done is relatively small (the case here).

3. The results indicate that increasing the time span of the integration (changing Δt with a fixed number of observations) improves the precision with which $\hat{\mu}$ is estimated. The best estimator of $\hat{\sigma}^2$, however, would be based on as many observations as possible, regardless of the sampling interval. The consistency of $\hat{\sigma}^2$ and the inconsistency of $\hat{\mu}$ under continuous-record asymptotics was first observed by Merton ([9]) for geometric Brownian motion, and this is confirmed here, at least for $\hat{\mu}$.

Number of trials (m)	Time interval ($\Delta t = 0.10$)	Time interval ($\Delta t = 0.05$)	Time interval ($\Delta t = 0.01$)
Kernel 100 trials	$\mu = 1.044 \pm 0.179$ $\sigma = 0.489 \pm 0.041$	$\mu = 1.045 \pm 0.258$ $\sigma = 0.486 \pm 0.049$	$\mu = 1.042 \pm 0.517$ $\sigma = 0.491 \pm 0.040$
Kernel 50 trials	$\mu = 1.066 \pm 0.184$ $\sigma = 0.493 \pm 0.042$	$\mu = 1.057 \pm 0.271$ $\sigma = 0.492 \pm 0.053$	$\mu = 1.096 \pm 0.536$ $\sigma = 0.497 \pm 0.044$
Kernel 25 trials	$\mu = 1.090 \pm 0.203$ $\sigma = 0.504 \pm 0.048$	$\mu = 1.115 \pm 0.289$ $\sigma = 0.506 \pm 0.057$	$\mu = 1.116 \pm 0.557$ $\sigma = 0.508 \pm 0.046$
Exact/Spectral solution	$\mu = 0.997 \pm 0.141$ $\sigma = 0.498 \pm 0.033$	$\mu = 0.996 \pm 0.209$ $\sigma = 0.496 \pm 0.034$	$\mu = 1.007 \pm 0.412$ $\sigma = 0.498 \pm 0.033$

Table 2: Kernel estimates for the geometric random walk with parameters $\mu = 1.0$, $\sigma = 0.5$ calculated from n sequential observations ($n = 25, 50, 100$) and 2000 replications.

4 Application to the instantaneous interest rate

There is a large literature documenting the link between the instantaneous short-term interest rate and the pricing of zero-coupon bonds for various terms to maturity. The classic one-factor model of the term structure is based on the assumption that the instantaneous interest rate $r(t)$ evolves in time according to the SDE

$$dr = -\alpha(r - \bar{r})dt + \sigma r^{\gamma} dW \tag{22}$$

where $W(t)$ is a standard Wiener process. Many models in the literature are special cases of this equation. For example, the case $\bar{r} = 0$, $\gamma = 1$ corresponds to geometric Brownian motion; the case $\gamma = 0$ has been treated by Vasicek [12]; the case $\gamma = 0.5$ is well known as the Cox, Ingersoll and Ross model [4]) and the case $\gamma = 1.0$ has been treated by Brennan and Schwartz [2]. Of course, it is not possible to observe the instantaneous interest rate and hence any data used to estimate these parameters will be subject to measurement error[1]. Clearly the resolution of this problem is a project for future research. However, notwithstanding the known shortcomings of (22) as a practical model of the short-term interest rate (see [1]), given data on UK 2-year gilts yields,[2] it is possible to illustrate the methods outlined above and reach some tentative conclusions on the magnitudes of the parameters of interest.

In order to implement the simulation approach a scheme for the numerical integration of (22) is required. The Milstein scheme is used in this application. For the general SDE (2) the scheme is given by

$$X_{j+1} = X_j + \alpha_j \Delta_j + \sigma_j \Delta W_j + \frac{\sigma_j}{2} \frac{\partial \sigma_j}{\partial X_j} \left((\Delta W_j)^2 - \Delta_j \right) \tag{23}$$

[1]Interest rates are both quoted and observed discretely.

[2]The dataset consists of 167 monthly observations of nominal UK gilts yields from 1982 to mid-1996 (available on request). Whilst data at the shorter end of the maturity spectrum might be considered more desirable, such as overnight rates in the London interbank market, institutional features relating to Bank of England operations in the money markets make these rates particularly noisy.

where $\alpha_j = \alpha(X_j, t_j)$ and $\sigma_j = \sigma(X_j, t_j)$. Table 3 shows the results of these calculations.

Model Parameters	500 Trials		1000 Trials	
	$\Delta t = 0.50$	$\Delta t = 1.00$	$\Delta t = 0.50$	$\Delta t = 1.00$
α	-0.038 ± 0.018	-0.031 ± 0.018	-0.022 ± 0.018	-0.029 ± 0.018
σ	0.027 ± 0.007	0.032 ± 0.008	0.025 ± 0.007	0.025 ± 0.007
γ	0.669 ± 0.019	0.747 ± 0.018	0.644 ± 0.016	0.633 ± 0.017
$\bar{f}_{\min}$	42.93	42.58	42.92	42.54

Table 3: Comparison of parameter estimates of the term-structure equation using simulated maximum likelihood estimation.

In estimating the model, $\bar{r}$, the long-term rate of interest in (22), was set at 8% reflecting a judgement about the "market" view of this quantity over the period in question. Parameter estimation of the model with $\bar{r}$ unconstrained broadly confirmed this view although it proved difficult to get good estimates of α and $\bar{r}$ simultaneously. This could be due to a number of factors[3] and the matter was not pursued. The development of better models of fundamental asset price dynamics is already recognised as a crucial area of current financial research [3].

The results of the estimation are reported in table 3. It is clear that the drift term is least well resolved indicating a fundamental problem with the specification of equation (22). One possible course of action (see [1]) suggests that the drift and diffusion coefficients might be better estimated nonparametrically. In general, however, the parameter estimates and the average value of the likelihood returned in the case of (22) are stable for different numbers of trials and different choices of discretisation in the Milstein scheme. For this particular sample, the square-root formulation of the diffusion term in [4] is probably the preferred choice among the common specifications.

5 Conclusion

The robust conclusion of this research is that both methods of estimating the parameters of SDE's from discrete time-series data are viable in practice. Furthermore, both methods have applicability beyond the scope of parameter estimation. In particular, the spectral method used here to solve the Kolmogorov equation for the probability density, may be used to solve any problem involving the solution of partial differential equations, even when they have delta-like initial conditions. An obvious application in financial econometrics would be to solve the fundamental bond-pricing equation of the term-structure of interest rates. Clearly much work remains to be done as the problem of measurement

[3]For example, the UK exit from the exchange rate mechanism of the European Community.

error needs to be addressed in situations where the entire density of stochastic process needs to be propagated as opposed to just the low order moments.

References

[1] Ait-Sohalia, Y. (1996) "Nonparametric pricing of interest rate derivative securities", *Econometrica*, **64** 527-560.

[2] Brennan, M.J. and Schwartz, E.S. (1979) "A continuous-time approach to the pricing of bonds", *Journal of Banking and Finance*, **3** 133-155.

[3] Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997) *The Econometrics of Financial Markets*, Princeton University Press: Princeton, New Jersey.

[4] Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985) "A theory of the term structure of interest rates", *Econometrica*, **53** 385-407.

[5] Canuto, C., Hussaini, M.Y., Quarteroni, A. and Zang, T.A., (1988) *Spectral Methods in Fluid Dynamics*, Springer-Verlag: Berlin Heidelberg New York.

[6] Habbema, J.D.F., Hermans, J. and Van Der Broek, K. (1974) "A stepwise discriminant analysis programme using density estimation" in G. Bruckman (ed.) *Proceedings in Computational Statistics*, 101-110, Physica-Verlag: Vienna.

[7] Hall, P., Sheather, S.J., Jones, M.C. and Marron J.S. (1991) "On optimal data-based bandwidth selection in kernel density estimation", *Biometrika* **78** 263-270.

[8] Kloeden, P.E., Platen, E. and Schurz, H. (1991) *Numerical Solution of SDE Through Computer Experiments*, Springer-Verlag: Berlin Heidelberg New York.

[9] Merton, R. (1980) "On estimating the expected return on the market: an exploratory investigation", *Journal of Financial Economics*, **8** 323-361.

[10] Scott, D.W. (1992) *Multivariate Density Estimation: Theory, Practice and Visualization* John Wiley: New York.

[11] Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall: London.

[12] Vasicek, O. (1977) "An equilibrium characterisation of the term structure", *Journal of Financial Economics*, **5** 177-188.