isid/ms/2003/06
January 22, 2003
http://www.isid.ac.in/~statmath/eprints

Estimation for some stochastic partial differential equations

B. L. S. Prakasa Rao

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Estimation for Some Stochastic Partial Differential Equations Based on Discrete Observations II

B.L.S. PRAKASA RAO
INDIAN STATISTICAL INSTITUTE, NEW DELHI

Abstract

Stochastic partial differential equations (SPDE) are used for stochastic modelling, for instance, in the study of neuronal behaviour in neurophysiology and in building stochastic models for turbulence. Huebner, Khasminskii and Rozovskii (1993) started the investigation of the maximum likelihood estimation of the parameters involved in two types of SPDE's and extended their results for a class of parabolic SPDE's in Huebner and Rozovskii (1995). Prakasa Rao (1998,2000) obtained Bernstein - von Mises type theorems for a class of parabolic SPDE's and investigated the properties of Bayes estimators of parameters involved in such SPDE's. In all the papers cited earlier, it was assumed that a continuous observation of a random field $u_\varepsilon(x,t)$ satisfying the SPDE over the region $[0,1]\times[0,T]$ is available. It is obvious that this assumption is not tenable in practice and the problem of interest is to develop methods of estimation of parameters from the random field $u_\varepsilon(x,t)$ observed at discrete times t and at discrete positions x or from the Fourier coefficients $u_{i\varepsilon}(t)$ observed at discrete time instants. We construct consistent and asymptotically normal estimators of the parameter based on the Fourier coefficients $u_{i\varepsilon}(t)$ observed at discrete times $t_j = j\Delta, 0 \le j \le n$ where $\Delta > 0$ and n tends to infinity.

Key words: Stochastic partial differential equations ; Discrete data: Consistency: Asymptotic normality.

AMS Subject classification (2000): Primary 62 M 40; Secondary 60 H 15.

1 Introduction

In their recent monograph, Kallianpur and Xiong (1995) discussed the properties of solutions of stochastic partial differential equations (SPDE's). They indicate that SPDE's are used for stochastic modelling for instance in the study of neuronal behaviour in neurophysiology and in building stochastic models of turbulence. The probabilistic theory for SPDE's is discussed in Ito (1984) and more recently in Rozovskii (1990) and Da Prato and Zabczyk (1992) among others. Huebner et al. (1993) started the investigation of the maximum likelihood estimation of parameters for two types of SPDE's and extended their results to a class of parabolic SPDE's in Huebner and Rozovskii (1995). Prakasa Rao (1998,2000) obtained Bernstein - von Mises type theorems for a class of parabolic SPDE's and investigated the properties of Bayes estimators of the parameters involved in such SPDE's. Similar results for a class of diffusion processes have been obtained in Prakasa Rao (1981) and for diffusion fields in Prakasa Rao (1984). In all the papers cited earlier, it was assumed that a continuous observation of a random field $u_\varepsilon(x,t)$ satisfying a SPDE over the region $[0,1]\times[0,T]$ is available. It is obvious that this assumption is not tenable in practice and the problem of interest is to develop methods of estimation of the parameters from a random field $u_\varepsilon(x,t)$ observed at discrete times t and at discrete positions x or from the Fourier coefficients $u_{i\varepsilon}(t)$ observed at discrete time instants. We will discuss the latter problem in this paper for two types of SPDE's when the parameter involved occurs in the "trend" part as well as in the "forcing" part of the SPDE. In an earlier paper (Prakasa Rao (2000)), we have investigated the problem of estimation when the parameter occurs only in the "trend" part of the SPDE. Prakasa Rao (1988) discusses statistical inference from sampled data for stochastic processes in general and the methods of statistical inference for the special class of diffusion type processes is investigated extensively in Prakasa Rao (1999). Piterbarg and Rozovskii(1997) studied the properties of maximum likelihood estimators based on discrete time sampling for parameters invoved in parabolic stochastic partial differential equations when the " trend" part of the SPDE involves the parameter but not the " forcing part of the SPDE.

2 Estimation from discrete observations

2.1 Example I

Let $(\Omega,\mathcal{F},P)$ be a probability space and consider the process $u_\varepsilon(t,x), 0\le x\le 1, 0\le t\le T$ governed by the stochastic partial differential equation

$$du_\varepsilon(t,x)=(\triangle u_\varepsilon(t,x)+b(\theta)u_\varepsilon(t,x))dt+\varepsilon\sigma(\theta)dW_Q(t,x) \tag{2. 1}$$

where $\triangle=\frac{\partial^2}{\partial x^2}$. Suppose that $\theta\in\Theta\subset R$. Assume that the function $b(\theta)<0$ for all $\theta\in\Theta$. Further suppose that the functional form of the function $b(\theta)$ is known and it is differentiable with respect to θ with non zero derivative. In addition assume that the function $\sigma(\theta)>0$ is known but the parameter $\theta\in\Theta$ is unknown. Suppose the initial and the boundary conditions are given by

$$\begin{cases} u_\varepsilon(0,x)=f(x), f\in L_2[0,1] \\ u_\varepsilon(t,0)=u_\varepsilon(t,1)=0, 0\le t\le T \end{cases} \tag{2. 2}$$

and Q is the nuclear covariance operator for the Wiener process $W_Q(t,x)$ taking values in $L_2[0,1]$ so that

$$W_Q(t,x)=Q^{1/2}W(t,x)$$

and $W(t,x)$ is a cylindrical Brownian motion in $L_2[0,1]$. Then, it is known that (cf. Rozovskii (1990))

$$W_Q(t,x)=\sum_{i=1}^{\infty}q_i^{1/2}e_i(x)W_i(t) \text{ a.s.} \tag{2. 3}$$

where $\{W_i(t), 0 \le t \le T\}, i \ge 1$ are independent one - dimensional standard Wiener processes and $\{e_i\}$ is a complete orthonormal system in $L_2[0,1]$ consisting of the eigen vectors of Q and $\{q_i\}$ the eigen values of Q. Let us consider a special covariance operator Q with $e_k = \sin k\pi x, k \ge 1$ and $\lambda_k = (\pi k)^2, k \ge 1$. Then $\{e_k\}$ is a complete orthonormal system with the eigen values $q_i = (1+\lambda_i)^{-1}, i \ge 1$ for the operator Q and $Q = (I - \triangle)^{-1}$. Further more

$$dW_Q = Q^{1/2} dW.$$

We define a solution $u_\varepsilon(t,x)$ of (2.1) as a formal sum

$$u_\varepsilon(t,x) = \sum_{i=1}^{\infty} u_{i\varepsilon}(t) e_i(x) \tag{2. 4}$$

(cf. Rozovskii (1990)). It can be checked that the Fourier coefficient $u_{i\varepsilon}(t)$ satisfies the stochastic differential equation

$$du_{i\varepsilon}(t) = (b(\theta) - \lambda_i) u_{i\varepsilon}(t) dt + \frac{\varepsilon}{\sqrt{\lambda_i + 1}} \sigma(\theta) dW_i(t), \ 0 \le t \le T \tag{2. 5}$$

with the initial condition

$$u_{i\varepsilon}(0) = v_i, \ v_i = \int_0^1 f(x) e_i(x) dx. \tag{2. 6}$$

Suppose the collection of observations consists of $\{u_{i\varepsilon}(j\Delta), 0 \le j \le n, 1 \le i \le N\}$ where $\Delta > 0$. We now approach the problem following the techniques in Bibby and Sorensen (1995) using the method of estimating functions.

Note that the process $\{u_{i\varepsilon}(t), 0 \le t \le T\}$ is the Ornstein-Uhlenbeck process and it is well known that the conditional distribution of $u_{i\varepsilon}(\Delta)$ given $u_{i\varepsilon}(0)$ is normal with mean $v_i e^{(b(\theta)-\lambda_i)\Delta}$ and variance $\frac{\varepsilon^2\sigma^2(\theta)(e^{2(b(\theta)-\lambda_i)\Delta}-1)}{2(b(\theta)-\lambda_i)(\lambda_i+1)}$. It can be shown that

$$G_n(\theta) = \frac{\lambda_i + 1}{\sigma^2(\theta)\varepsilon^2} \sum_{j=1}^{n} b'(\theta) u_{i\varepsilon}((j-1)\Delta)(u_{i\varepsilon}(j\Delta) - u_{i\varepsilon}((j-1)\Delta) e^{(b(\theta)-\lambda_i)\Delta}) \tag{2. 7}$$

is proportional to the optimal estimating function for the estimation of the parameter θ (cf. Bibby and Sorensen (1995)) and an estimator for $b(\theta)$ is of the form

$$\hat{b}_{i\varepsilon} = \lambda_i + \Delta^{-1} \log \frac{\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta) u_{i\varepsilon}((j-1)\Delta)}{\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)}. \tag{2. 8}$$

Since $b(\theta) < 0$, it is well known that the solution of (2.5) is ergodic with the stationary measure with the density μ_θ given by the normal distribution with mean zero and variance $\beta_i^2(\theta) = \varepsilon^2\sigma^2(\theta)\{2(\lambda_i - b(\theta))(\lambda_i+1)\}^{-1}$. Further more we have already noted that the transition probability density π_Δ^θ of $u_{i\varepsilon}(\Delta)$ given that $u_{i\varepsilon}(0) = x$ is the normal probability density with mean $x e^{(b(\theta)-\lambda_i)\Delta}$ and variance $\frac{\varepsilon^2\sigma^2(\theta)(e^{2(b(\theta)-\lambda_i)\Delta}-1)}{2(b(\theta)-\lambda_i)(\lambda_i+1)}$. Let X be a random variable with the stationary measure μ_θ and Y be a random variable such that the conditional density of Y given $X = x$ is given by π_Δ^θ. Note that

$$E[XY] = E[XE(Y|X)] = E[XXe^{(b(\theta)-\lambda_i)\Delta}] = \beta_i^2(\theta)e^{(b(\theta)-\lambda_i)\Delta} \tag{2. 9}$$

and

$$E[X^2] = \beta_i^2(\theta). \tag{2. 10}$$

It is easy to check that the Condition 3.1 in Bibby and Sorensen (1995) holds in this case and applying the Lemma 3.1 in Bibby and Sorensen(1995) (cf. Florens-Zmirou (1989)), we obtain that

$$\frac{1}{n}\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta) \to E[XY] \text{ in probability as } n \to \infty$$

and

$$\frac{1}{n}\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta) \to E[X^2] \text{ in probability as } n \to \infty.$$

The above relations imply that

$$\frac{\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)}{\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)} \to \frac{E[XY]}{E[X^2]} \text{ in probability as } n \to \infty.$$

The following result follows as a consequence of the above observation and the relations (2.9) and (2.10).

Theorem 2.1: The estimator $\hat{b}_{i\varepsilon}$ converges in probability to $b(\theta)$ as $n \to \infty$.

Since the function $b(\theta)$ has a continuous inverse function, the following result is a consequence of Theorem 2.1.

Theorem 2.2: The estimator $\hat{\theta}_{i\varepsilon} = b^{-1}(\hat{b}_{i\varepsilon})$ converges in probability to θ as $n \to \infty$.

Let $\psi_i(\theta) = e^{\Delta(b(\theta)-\lambda_i)}$. Note that

$$\psi_i(\hat{\theta}_{i\varepsilon}) = \frac{\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)}{\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)}.$$

Hence

$$\sqrt{n}\{\psi_i(\hat{\theta}_{i\varepsilon}) - \psi_i(\theta)\} = \frac{n^{-1/2}\sum_{j=1}^{n}[u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta) - \psi_i(\theta)u_{i\varepsilon}^2((j-1)\Delta)]}{n^{-1}\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)}.$$

Since

$$E[u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)|u_{i\varepsilon}((j-1)\Delta)] = \psi_i(\theta)u_{i\varepsilon}^2((j-1)\Delta)$$

it follows by Lemma 3.1 of Bibby and Sorensen (1995) that

$$n^{-1/2}\sum_{j=1}^{n}[u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)-\psi_i(\theta)u_{i\varepsilon}^2((j-1)\Delta)]$$

converges in distribution to the normal distribution with mean zero and variance equal to $\gamma_i(\theta)=E[XY-E(XY|X)]^2$ where the random variables X and Y are as defined above. It can be checked that

$$\begin{aligned}\gamma_i(\theta) &= \frac{\varepsilon^2\sigma^2(\theta)(e^{2(b(\theta)-\lambda_i)\Delta}-1)}{2(b(\theta)-\lambda_i)(\lambda_i+1)}\beta_i^2(\theta)\\ &= (1-e^{2(b(\theta)-\lambda_i)\Delta})\beta_i^4(\theta).\end{aligned}$$

Applying the δ - method , we obtain that

$$n^{1/2}\Delta(b(\hat{\theta}_{i\varepsilon})-b(\theta))$$

converges in distribution to the normal distribution with mean zero and variance $e^{-2\Delta(b(\theta)-\lambda_i)}\gamma_i(\theta)\beta_i^{-4}(\theta)$ and we have the following theorem.

Theorem 2.3: Under the conditions stated above

$$n^{1/2}\Delta(b(\hat{\theta}_{i\varepsilon})-b(\theta))\xrightarrow{\mathcal{L}}N(0,e^{-2\Delta(b(\theta)-\lambda_i)}-1)\ \text{ as }\ n\to\infty.$$

Applying the δ-method once again, we obtain that

$$n^{1/2}\Delta(\hat{\theta}_{i\varepsilon}-\theta)$$

converges in distribution to the normal distribution with mean zero and variance $b'(\theta)^{-2}e^{-2\Delta(b(\theta)-\lambda_i)}\gamma_i(\theta)\beta_i^{-4}(\theta)$ and we have the following theorem.

Theorem 2.4: Under the conditions stated above

$$n^{1/2}\Delta(\hat{\theta}_{i\varepsilon}-\theta)\xrightarrow{\mathcal{L}}N(0,(b'(\theta))^{-2}(e^{-2\Delta(b(\theta)-\lambda_i)}-1))\ \text{ as }\ n\to\infty.$$

Remarks: Note that the estimators $\{\hat{\theta}_{i\varepsilon}, 1\le i\le N\}$ are *independent*, consistent and asymptotically normal for the parameter θ in the stochastic partial differential equation (2.1). We will now discuss a method for combining these estimators to get an improved estimator.

Let $\tilde{\theta}_\varepsilon=\sum_{i=1}^N\alpha_i\hat{\theta}_{i\varepsilon}$ where $\alpha_i, 1\le i\le N$ is a nonrandom sequence of coefficients to be chosen. Note that

$$\tilde{\theta}_\varepsilon\xrightarrow{p}[\sum_{i=1}^N\alpha_i]\theta\ \text{ as }\ n\to\infty$$

by the Theorem 2.2 and hence $\tilde{\theta}_\varepsilon$ is consistent for θ provided $\sum_{i=1}^N\alpha_i=1$. Further more

$$\sqrt{n}\Delta(\tilde{\theta}_\varepsilon-\theta)\xrightarrow{\mathcal{L}}N(0,(b'(\theta))^{-2}\sum_{i=1}^N\alpha_i^2(e^{-2\Delta(b(\theta)-\lambda_i)}-1))\ \text{ as }\ n\to\infty.$$

This follows from the Theorem 2.4 and the independence of the estimators $\{\hat{\theta}_{i\varepsilon}, 1 \leq i \leq N\}$. We now obtain the optimum combination of the coefficients $\{\alpha_i, 1 \leq i \leq N\}$ by minimising the asymptotic variance

$$\sum_{i=1}^{N} \alpha_i^2 (e^{-2\Delta(b(\theta)-\lambda_i)} - 1)$$

subject to the condition $\sum_{i=1}^{N} \alpha_i = 1$. It is easy to show that α_i is proportional to $(e^{2\Delta(\lambda_i - b(\theta))} - 1)^{-1}$ and the optimum choice of $\alpha_i, 1 \leq i \leq N$ leads to the "Estimator"

$$\theta_\varepsilon^* = \frac{\sum_{i=1}^{N} (e^{2\Delta(\lambda_i - b(\theta))} - 1)^{-1} \hat{\theta}_{i\varepsilon}}{\sum_{i=1}^{N} (e^{2\Delta(\lambda_i - b(\theta))} - 1)^{-1}}. \tag{2. 11}$$

It is easy to see that

$$\theta_\varepsilon^* \xrightarrow{p} \theta \text{ as } n \to \infty$$

and

$$\sqrt{n}\Delta(\theta_\varepsilon^* - \theta) \xrightarrow{\mathcal{L}} N(0, (b'(\theta))^{-2} (\sum_{i=1}^{N} (e^{-2\Delta(b(\theta)-\lambda_i)} - 1)^{-1})^{-1}) \text{ as } n \to \infty$$

again due to the independence of the estimators $\{\hat{\theta}_{i\varepsilon}, 1 \leq i \leq N\}$. However the random variable θ_ε^* cannot be considered as an estimator of θ in the true sense since it depends on the unknown parameter θ. In order to avoid this problem, we can consider a modified estimator

$$\hat{\theta}_\varepsilon = \frac{\sum_{i=1}^{N} (e^{2\Delta(\lambda_i - \hat{b}_{i\varepsilon})} - 1)^{-1} \hat{\theta}_{i\varepsilon}}{\sum_{i=1}^{N} (e^{2\Delta(\lambda_i - \hat{b}_{i\varepsilon})} - 1)^{-1}} \tag{2. 12}$$

which is obtained from θ_ε^* by substituting the estimator $\hat{\theta}_{i\varepsilon}$ for the unknown parameter θ in the i-th term in the numerator and denominator in (2.11). In view of the independence, consistency and asymptotic normality of the estimators $\hat{\theta}_{i\varepsilon}, 1 \leq i \leq N$, it follows that the estimator $\hat{\theta}_\varepsilon$ is consistent and asymptotically normal for the parameter θ and we have the following result.

Theorem 2.5: Under the conditions stated above,

$$\hat{\theta}_\varepsilon \xrightarrow{p} \theta \text{ as } n \to \infty$$

and

$$n^{1/2}\Delta(\hat{\theta}_\varepsilon - \theta)) \xrightarrow{\mathcal{L}} N(0, (b'(\theta))^{-2} (\sum_{i=1}^{N} (e^{-2\Delta(b(\theta)-\lambda_i)} - 1)^{-1})^{-1}) \text{ as } n \to \infty$$

for any fixed $N \geq 1$.

2.2 Example II

Let $(\Omega, \mathcal{F}, P)$ be a probability space and consider the process $u_\varepsilon(t,x), 0 \leq x \leq 1, 0 \leq t \leq T$ governed by the stochastic partial differential equation

$$du_\varepsilon(t,x) = b(\theta) \, \triangle u_\varepsilon(t,x) dt + \varepsilon\sigma(\theta)(I - \triangle)^{-1/2} dW(t,x) \tag{2. 13}$$

Suppose that $\theta \in \Theta \subset R$. Assume that the function $b(\theta) > 0$ for all $\theta \in \Theta$. Further suppose that the functional form of the function $b(\theta)$ is known and it is differentiable with respect to θ with non zero derivative. In addition assume that the function $\sigma(\theta) > 0$ is known but the parameter $\theta \in \Theta$ is unknown. Suppose further the following the initial and the boundary conditions

$$\begin{aligned} u_\varepsilon(0,x) &= f(x),\ 0 < x < 1,\ f \in L_2[0,1], \\ u_\varepsilon(t,0) &= u_\varepsilon(t,1) = 0,\ 0 \le t \le T. \end{aligned} \tag{2. 14}$$

hold. Here I is the identity operator, $\triangle = \frac{\partial^2}{\partial x^2}$ and the process $W(t,x)$ is the cylindrical Brownian motion in $L_2[0,1]$.. The Fourier coefficients $u_{i\varepsilon}(t)$ satisfy the stochastic differential equations

$$du_{i\varepsilon}(t) = -b(\theta)\lambda_i u_{i\varepsilon}(t)dt + \sigma(\theta)\frac{\varepsilon}{\sqrt{\lambda_i + 1}}dW_i(t),\ 0 \le t \le T, \tag{2. 15}$$

with

$$u_{i\varepsilon}(0) = v_i, v_i = \int_0^1 f(x)e_i(x)dx. \tag{2. 16}$$

Suppose the collection of observations consists of $\{u_{i\varepsilon}(j\Delta), 0 \le j \le n, 1 \le i \le N\}$ where $\Delta > 0$.

Note that the process $\{u_{i\varepsilon}(t), 0 \le t \le T\}$ is the Ornstein-Uhlenbeck process and it is well known that the conditional distribution of $u_{i\varepsilon}(\Delta)$ given $u_{i\varepsilon}(0)$ is normal with mean $v_i e^{-b(\theta)\lambda_i\Delta}$ and variance $\sigma^2(\theta)\frac{\varepsilon^2(e^{-2b(\theta)\lambda_i\Delta}-1)}{(-2b(\theta)\lambda_i)(\lambda_i+1)}$. It can be shown that

$$H_n(\theta) = -\frac{\lambda_i + 1}{\varepsilon^2\sigma^2(\theta)}\sum_{j=1}^{n} b'(\theta)\lambda_i u_{i\varepsilon}((j-1)\Delta)(u_{i\varepsilon}(j\Delta) - u_{i\varepsilon}((j-1)\Delta)e^{-b(\theta)\lambda_i\Delta}) \tag{2. 17}$$

is proportional to the optimal estimating function for the estimation of the parameter θ (cf. Bibby and Sorensen (1995)) and an estimator for $b(\theta)$ is of the form

$$\hat{b}_{i\varepsilon} = -\lambda_i^{-1}\Delta^{-1}\log\frac{\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)}{\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)}. \tag{2. 18}$$

Since $b(\theta) > 0$, it is well known that the solution of (2.15) is ergodic with the stationary measure with density ν_θ given by the normal distribution with mean zero and variance $\zeta_i^2(\theta) = \varepsilon^2\sigma^2(\theta)\{2\lambda_i b(\theta)(\lambda_i + 1)\}^{-1}$. Further more we have already noted that the transition probability density π_Δ^θ of $u_{i\varepsilon}(\Delta)$ given that $u_{i\varepsilon}(0) = x$ is the normal probability density with mean $xe^{-b(\theta)\lambda_i\Delta}$ and variance $\sigma^2(\theta)\frac{\varepsilon^2(e^{2(-b(\theta)\lambda_i)\Delta}-1)}{2(-b(\theta)\lambda_i)(\lambda_i+1)}$. Let X be a random variable with stationary measure ν_θ and Y be a random variable such that the conditional density of Y given $X = x$ is given by π_Δ^θ. Note that

$$\begin{aligned} E[XY] &= E[XE(Y|X)] = E[XXe^{-b(\theta)\lambda_i\Delta}] \\ &= \zeta_i^2(\theta)e^{-b(\theta)\lambda_i\Delta} \end{aligned} \tag{2. 19}$$

and

$$E[X^2] = \zeta_i^2(\theta). \tag{2. 20}$$

It is easy to check that the Condition 3.1 in Bibby and Sorensen (1995) holds in this case and applying the Lemma 3.1 in Bibby and Sorensen(1995) (cf. Florens-Zmirou (1989)), we obtain that

$$\frac{1}{n}\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta) \to E[XY] \text{ in probability as } n \to \infty$$

and

$$\frac{1}{n}\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta) \to E[X^2] \text{ in probability as } n \to \infty.$$

The above relations imply that

$$\frac{\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)}{\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)} \to \frac{E[XY]}{E[X^2]} \text{ in probability as } n \to \infty.$$

The following result follows as a consequence of the above observation and the relations (2.19) and (2.20).

Theorem 2.6: The estimator $\hat{b}_{i\varepsilon}$ converges in probability to $b(\theta)$ as $n \to \infty$.

Since the function $b(\theta)$ has a continuous inverse function, the following result is a consequence of Theorem 2.6.

Theorem 2.7: The estimator $\hat{\theta}_{i\varepsilon} = b^{-1}(\hat{b}_{i\varepsilon})$ converges in probability to θ as $n \to \infty$.

Let $\psi_i(\theta) = e^{-b(\theta)\lambda_i\Delta}$. Note that

$$\psi_i(\hat{\theta}_{i\varepsilon}) = \frac{\sum_{j=1}^{n} u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)}{\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)}.$$

Hence

$$\sqrt{n}\{\psi_i(\hat{\theta}_{i\varepsilon}) - \psi_i(\theta)\} = \frac{n^{-1/2}\sum_{j=1}^{n}[u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta) - \psi_i(\theta)u_{i\varepsilon}^2((j-1)\Delta)]}{n^{-1}\sum_{j=1}^{n} u_{i\varepsilon}^2((j-1)\Delta)}.$$

Since

$$E[u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta)|u_{i\varepsilon}((j-1)\Delta)] = \psi_i(\theta)u_{i\varepsilon}^2((j-1)\Delta),$$

it follows by the Lemma 3.1 of Bibby and Sorensen (1995) that

$$n^{-1/2}\sum_{j=1}^{n}[u_{i\varepsilon}(j\Delta)u_{i\varepsilon}((j-1)\Delta) - \psi_i(\theta)u_{i\varepsilon}^2((j-1)\Delta)]$$

converges in distribution to the normal distribution with mean zero and variance equal to $\tau_i(\theta) = E[XY - E(XY|X)]^2$ where the random variables X and Y are as defined above. It can be checked that

$$\begin{aligned}\tau_i(\theta) &= \sigma^2(\theta)\frac{\varepsilon^2(e^{-2b(\theta)\lambda_i\Delta} - 1)}{2(-b(\theta)\lambda_i)(\lambda_i+1)}\zeta_i^2(\theta) \\ &= (1 - e^{-2b(\theta)\lambda_i\Delta})\zeta_i^4(\theta).\end{aligned}$$

Applying the δ - method , we obtain that

$$n^{1/2}\Delta(b(\hat{\theta}_{i\varepsilon})-b(\theta))$$

converges in distribution to the normal distribution with mean zero and variance $\lambda_i^{-2}e^{2\Delta b(\theta)\lambda_i}\tau_i(\theta)\zeta_i^{-4}(\theta)$ and we have the following theorem.

Theorem 2.8: Under the conditions stated above

$$n^{1/2}\Delta(b(\hat{\theta}_{i\varepsilon})-b(\theta)) \xrightarrow{\mathcal{L}} N(0,\lambda_i^{-2}(e^{2\Delta b(\theta)\lambda_i}-1)) \text{ as } n\to\infty.$$

Applying the δ-method once again, we obtain that

$$n^{1/2}\Delta(\hat{\theta}_{i\varepsilon}-\theta)$$

converges in distribution to the normal distribution with mean zero and variance $(b'(\theta))^{-2}\lambda_i^{-2}(e^{2\Delta b(\theta)\lambda_i})\tau_i(\theta)\beta_i^{-4}(\theta)$ and we have the following theorem.

Theorem 2.9: Under the conditions stated above

$$n^{1/2}\Delta(\hat{\theta}_{i\varepsilon}-\theta) \xrightarrow{\mathcal{L}} N(0,(b'(\theta))^{-2}\lambda_i^{-2}(e^{2\Delta b(\theta)\lambda_i}-1)) \text{ as } n\to\infty.$$

Remarks: Note that the estimators $\{\hat{\theta}_{i\varepsilon}, 1\le i\le N\}$ are *independent*, consistent and asymptotically normal for the parameter θ in the stochastic partial differential equation (2.13). We will now discuss a method for combining these estimators to get an improved estimator.

Let $\tilde{\theta}_\varepsilon = \sum_{i=1}^N \alpha_i\hat{\theta}_{i\varepsilon}$ where $\alpha_i, 1\le i\le N$ is a nonrandom sequence of coefficients to be chosen. Note that

$$\tilde{\theta}_\varepsilon \xrightarrow{p} [\sum_{i=1}^N \alpha_i]\theta \text{ as } n\to\infty$$

by the Theorem 2.7 and hence $\tilde{\theta}_\varepsilon$ is consistent for θ provided $\sum_{i=1}^N\alpha_i=1$. Further more

$$\sqrt{n}\Delta(\tilde{\theta}_\varepsilon-\theta) \xrightarrow{\mathcal{L}} N(0,\sum_{i=1}^N \alpha_i^2(b'(\theta))^{-2}\lambda_i^{-2}(e^{2\Delta b(\theta)\lambda_i}-1)) \text{ as } n\to\infty.$$

This follows from the Theorem 2.9 and the independence of the estimators $\{\hat{\theta}_{i\varepsilon}, 1\le i\le N\}$. We now obtain the optimum combination of the coefficients $\{\alpha_i, 1\le i\le N\}$ by minimising the asymptotic variance

$$\sum_{i=1}^N \alpha_i^2(b'(\theta))^{-2}\lambda_i^{-2}(e^{2\Delta b(\theta)\lambda_i}-1)$$

subject to the condition $\sum_{i=1}^N\alpha_i=1$. It is easy to show that α_i is proportional to $\lambda_i^2(e^{2\Delta\lambda_i b(\theta)}-1)^{-1}$ and the optimum choice of $\alpha_i, 1\le i\le N$ leads to the "Estimator"

$$\theta^*_\varepsilon = \frac{\sum_{i=1}^N \lambda_i^2(e^{2\Delta\lambda_i b(\theta)}-1)^{-1}\hat{\theta}_{i\varepsilon}}{\sum_{i=1}^N \lambda_i^2(e^{2\Delta\lambda_i b(\theta)}-1)^{-1}}. \tag{2. 21}$$

It is easy to see that

$$\theta_\varepsilon^* \xrightarrow{p} \theta \text{ as } n \to \infty$$

and

$$\sqrt{n}\Delta(\theta_\varepsilon^* - \theta) \xrightarrow{\mathcal{L}} N(0, ((b'(\theta))^{-2} \sum_{i=1}^{N} \lambda_i^2 (e^{2\Delta b(\theta)\lambda_i} - 1)^{-1})^{-1}) \text{ as } n \to \infty$$

again due to the independence of the estimators $\{\hat{\theta}_{i\varepsilon}, 1 \leq i \leq N\}$. However the random variable θ_ε^* cannot be considered as an estimator of θ since it depends on the unknown parameter θ. In order to avoid this problem, we can consider a modified estimator

$$\hat{\theta}_\varepsilon = \frac{\sum_{i=1}^{N} \lambda_i^2 (e^{2\Delta\lambda_i \hat{b}_{i\varepsilon}} - 1)^{-1} \hat{\theta}_{i\varepsilon}}{\sum_{i=1}^{N} \lambda_i^2 (e^{2\Delta\lambda_i \hat{b}_{i\varepsilon}} - 1)^{-1}} \tag{2. 22}$$

which is obtained from θ_ε^* by substituting the estimator $\hat{\theta}_{i\varepsilon}$ for the unknown parameter θ in the i-th term in the numerator and denominator in (2.21). In view of the independence, consistency and asymptotic normality of the estimators $\hat{\theta}_{i\varepsilon}, 1 \leq i \leq N$, it follows that the estimator $\hat{\theta}_\varepsilon$ is consistent and asymptotically normal for the parameter θ and we have the following result.

Theorem 2.10: Under the conditions stated above,

$$\hat{\theta}_\varepsilon \xrightarrow{p} \theta \text{ as } n \to \infty$$

and

$$n^{1/2}\Delta(\hat{\theta}_\varepsilon - \theta)) \xrightarrow{\mathcal{L}} N(0, ((b'(\theta))^{-2} \sum_{i=1}^{N} \lambda_i^2 (e^{2\Delta b(\theta)\lambda_i} - 1)^{-1})^{-1}) \text{ as } n \to \infty$$

for any fixed $N \geq 1$.

3 Remarks

In the two examples discussed in the earlier sections, we have assumed that the drift coefficient and the diffusion coefficient are known except for the parameter θ which is to be estimated from the data. Since the estimating functions considered above are linear martingale estimating functions , the function $b(\theta)$ in the diffusion coefficient is only involved and not the function $\sigma(\theta)$ which makes the procedure inefficient. However we get explicit solution for the estimator if we make use of the linear martingale estimatting function. If one uses the quadratic martingale estimating functions as described by Sorensen(1997) (cf . Prakasa Rao(1999)), the estimating function involves both the functions $b(\theta)$ and $\sigma(\theta)$ but the resulting equations are too complex to solve for a user. Since the discretely observed Ornstein -Uhlenbeck processes encountered in both the examples are autoregressive processes, the likelihood function can be computed and the maximum likelihood estimator can be obtained which is asymptotically efficient. The estimators $\hat{b}_{i\varepsilon}$ described in the equations (2.8) and (2.18) *are the the maximum likelihood estimators* in the natural parametrization where the the reparametrization by means of the functions $b(\theta)$

and $\sigma(\theta)$ is done away with and the scalar parameter θ in the drift and the scalar parameter σ in the diffusion term are allowed to vary *independently* keeping the σ fixed eventually as the interest centers around the parameter θ. In such a case, the combined estimator $\hat{\theta}_\varepsilon$, described in (2.12) and (2.22), of the parameter θ in the drift term would be asymptotically efficient.

References

Basawa, I.V. and Prakasa Rao, B.L.S. 1980. *Statistical Inference for Stochastic Processes*, Academic Press, London.

Bibby, B.M. and Sorensen, M. 1995. Martingale estimating functions for discretely observed diffusion processes, *Bernoulli*, **1**, 17-39.

Da Prato, G. and Zabczyk, J. 1992. *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge.

Florens-Zmirou, D. 1989. Approximate discrete-time schemes for statistics of diffusion processes, *Statistics*, **20**, 547-557.

Huebner, M., Khasminskii, R. and Rozovskii. B.L. 1993. Two examples of parameter estimation for stochastic partial differential equations, In *Stochastic Processes : A Festschrift in Honour of Gopinath Kallianpur*, Ed. S. Cambanis, J.K. Ghosh, R.L. Karandikar and P.K. Sen, Springer, New York, pp. 149-160.

Huebner, M., and Rozovskii, B.L. 1995. On asymptotic properties of maximum likelihood estimators for parabolic stochastic SPDE's. *Prob. Theory and Relat. Fields*,**103**, 143-163.

Ito, K. 1984. *Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces*, Vol. **47** of CBMS Notes, SIAM, Baton Rouge.

Kallianpur, G., and Xiong, J. 1995. *Stochastic Differential Equations in Infinite Dimensions* , Vol. **26** of IMS Lecture Notes, Hayward, California.

Piterbarg, L., and Rozovskii, B. 1997. On asymptotic problems of parameter estimation in stochastic PDE's: Discrete time sampling. *Mathematical Methods of Statistics*, **6**, 200-223.

Prakasa Rao, B.L.S. 1981. The Bernstein - von Mises theorem for a class of diffusion processes, *Teor. Sluch. Proc.*,**9**, 95-101 (In Russian).

Prakasa Rao, B.L.S. 1984. On Bayes estimation for diffusion fields, in *Statistics : Applications and New Directions*, Ed. J.K. Ghosh and J. Roy, Statistical Publishing Society, Calcutta, pp. 504-511.

Prakasa Rao, B.L.S. 1988. Statistical inference from sampled data for stochastic processes, *Contemporary Mathematics*, American Mathematical Society, Providence, **80**, 249-284.

Prakasa Rao, B.L.S. 1998. Bayes estimation for parabolic stochastic partial differential equations, Preprint, Indian Statistical Institute, New Delhi.

Prakasa Rao, B.L.S. 1999. *Statistical Inference for Diffusion Type Processes*, Arnold, London and Oxford Unversity Press, New York.

Prakasa Rao, B.L.S. 2000. Bayes estimation for some stochastic partial differential equations, *J. Stat. Plan. Inf.* ,**91**, 511-524.

Prakasa Rao, B.L.S. 2000. Estimation for some stochastic partial differential equations based on discrete observations, *Calcutta Stat. Assoc. Bull.*, **50** , 193-206.

Rozovskii, B.L. 1990. *Stochastic Evolution Systems*, Kluwer, Dordrecht.

Sorensen, M. 1997. Estimating functions for discretely observed diffusions: A Review, in *Selected Proceedings of the Symposium on Estimating Functions* , Ed. I.V.Basawa, V.P. Godambe and R.L. Taylor, Lecture Notes-Monograph Series, Vol.32, Institute of Mathematical Statistics, pp. 305-325.