Journal of the Korean Statistical Society (2002), **31**: 2, pp 213–228

Minimum Distance Estimation for Some Stochastic Partial Differential Equations

B. L. S. Prakasa Rao[1]

ABSTRACT

Asymptotic properties of minimum distance estimators for the parameter involved for a class of stochastic partial differential equations are investigated following the techniques in Kutoyants and Pilibossian (1994).

Keywords. Minimum distance estimation, stochastic partial differential equation.
AMS 2000 subject classifications. Primary 62M40; Secondary 60H15.

1. Introduction

In their recent monograph, Kallianpur and Xiong (1995) discuss the properties of solutions of stochastic partial differential equations (SPDE's). They indicate that SPDE's are being used for stochastic modelling for instance in the study of neuronal behaviour in neurophysiology and in building stochastic models of turbulence. The theory of SPDE's is investigated in Ito (1984) and more recently in Rozovskii (1990) and Da Prato and Zabczyk (1992). Huebner *et al.* (1993) started the investigation of maximum likelihood estimation of parameters of two types of SPDE's and extended their results for a class of parabolic SPDE's in Huebner and Rozovskii (1995). Bayes estimation of parameters for such classes of SPDE are discussed in Prakasa Rao (2000).

One can construct maxiumum likelihood estimators (MLE) in the models for SPDE discussed in Sections 2 to 4 and it is known that these estimators are consistent and asymptotically normal and asymptotically efficient as the amplitude of the noise ε decreases to zero or the time of observation T increases to infinity. In spite of having such good properties, the maximum likelihood estimators have some short comings at the same time. Their calculation is often

Received September 2001; accepted April 2002.
[1]Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi 110016, India (e-mail : blsp@isid.ac.in)

cumbersome as the expressions for MLE involve stochastic integrals which need good approximants for computational purposes. Furthermore MLE are not robust in the sense that a slight perturbation in the noise component say from a Wiener process to another Gaussian process with finite variation will change the properties of the MLE substantially. In order to circumvent these problems, the minimum distance approach is proposed. Properties of the minimum distance estimators (MDE) were discussed in Millar (1984) in a general frame work.

Our aim in this paper is to obtain the minimum distance estimators of parameters for some class of SPDE's and investigate the asymptotic properties of such estimators following the work of Kutoyants and Pilibossian (1994)(*cf.* Kutoyants, 1994) for the estimation of a parameter of the Ornstein-Uhlenbeck process.

2. Parabolic Stochastic PDE

Let $(\Omega, \mathcal{F}, P)$ be a probability space and consider a stochastic partial differential equation (SPDE) of the form

$$du_\varepsilon^\theta(t,x) = A^\theta u_\varepsilon^\theta(t,x)dt + \varepsilon dW(t,x),\ 0 \le t \le T,\ x \in G \tag{2.1}$$

where $A^\theta = \theta A_1 + A_0$, A_1 and A_0 being partial differential operators, $\theta \in \Theta \subset R$ and $W(t,x)$ is a cylindrical Brownian motion in $L_2(G), G$ being a bounded domain in R^d with the boundary ∂G as a C^∞-manifold of dimension $(d-1)$ and locally G is totally on one side of ∂G. For the definition of cylindrical Brownian motion, see, Kallianpur and Xiong (1995, p. 93).

The order $Ord(A)$ of a partial differential operator A is defined to be the order of the highest partial derivative in A. Let m_0 and m_1 be the orders of the operators A_0 and A_1 respectively. We assume that the operators A_0 and A_1 commute and m_1 is even.

Suppose the solution $u_\varepsilon^\theta(t,x)$ of (2.1) has to satisfy the boundary conditions

$$u_\varepsilon^\theta(0,x) = u_0(x), \tag{2.2}$$

$$D^\gamma u_\varepsilon^\theta(t,x)|_{\partial G} = 0 \tag{2.3}$$

for all multiindices γ such that $|\gamma| = m-1$ where $2m = \max(m_1, m_0)$. Here

$$D^\gamma f(\mathbf{x}) = \frac{\partial^{|\gamma|}}{\partial x_1^{\gamma_1} \cdots \partial x_d^{\gamma_d}} f(\mathbf{x}) \tag{2.4}$$

with $|\gamma| = \gamma_1 + \cdots + \gamma_d$. Suppose that

$$A_i(\mathbf{x})u = -\sum_{|\alpha|,|\beta| \le m_i} (-1)^{|\alpha|} D^\alpha (a_i^{\alpha\beta}(\mathbf{x}) D^\beta u) \tag{2.5}$$

where

$$a_i^{\alpha\beta}(\mathbf{x}) \in C^\infty(\bar{G}). \tag{2.6}$$

Let

$$a^{\alpha\beta}(\theta, x) = \theta a_1^{\alpha\beta}(x) + a_0^{\alpha\beta}(x). \tag{2.7}$$

Suppose θ_0 is the true parameter.

We follow the notation introduced in Huebner and Rozovskii (1995). Assume that the following conditions hold.

(H1) The operators A_0 and A_1 satisfy the condition

$$\int_G A_i uv dx = \int_G u A_i v dx,\ u, v \in C_0^\infty(G), i = 0, 1.$$

(H2) There is a compact neighbourghood Θ of θ_0 so that $\{A_\theta, \theta \in \Theta\}$ is a family of uniformly strongly elliptic operators of order $2m = \max(m_1, m_0)$.

For $s > 0$, denote the closure of $C_0^\infty(G)$ in the Sobolev space $W^{s,2}(G)$ by $W_0^{s,2}$. The operator A^θ with boundary conditions defined by (2.3) can be extended to a closed self-adjoint operator $\mathcal{L}_\theta$ on $L_2(G)$ (Shimakura, 1992). In view of the condition (H2), the operator $\mathcal{L}_\theta$ is lower semibounded, that is there exists a constant $k(\theta)$ such that $-\mathcal{L}_\theta + k(\theta)I > 0$ and the resolvent $(k(\theta)I - \mathcal{L}_\theta)^{-1}$ is compact. Let $\Lambda_\theta = (k(\theta)I - \mathcal{L}_\theta)^{\frac{1}{2m}}$. Let $h_i(\theta)$ be an orthonormal system of eigen functions of Λ_θ. We assume that

(H3) There exists a complete orthonormal system $\{h_i, i \geq 1\}$ independent of θ such that

$$\Lambda_\theta h_i = \lambda_i(\theta) h_i, \theta \in \Theta.$$

The elements of the basis $\{h_i, i \geq 1\}$ are also eigen functions for the operator $\mathcal{L}_\theta$, that is $\mathcal{L}_\theta h_i = \mu_i^\theta h_i$ where $\mu_i^\theta = -\lambda_i^{2m}(\theta) + k(\theta)$. For $s \geq 0$, define H_θ^s to be the set of all $u \in L_2(G)$ such that

$$||u||_{s,\theta} = \left(\sum_{j=1}^{\infty} \lambda_j^{2s}(\theta) |(u, h_j)_{L_2(G)}|^2 \right)^{1/2} < \infty.$$

For $s < 0$, H_θ^s is defined to be the closure of $L_2(G)$ in the norm $||u||_{s,\theta}$ given above. Then H_θ^s is a Hilbert space with respect to the inner product $(.,.)_{s,\theta}$ associated with the norm $||.||_{s,\theta}$ and the functions $h_{i,\theta}^s = \lambda_i^{-s}(\theta) h_i, i \geq 1$ form an orthonormal basis in H_θ^s. Condition (H2) imples that for every s, the spaces H_θ^s are equivalent for all θ. We identify the spaces H_θ^s and the norms $||.||_{s,\theta}$ for different $\theta \in \Theta$.

In addition to the conditions (H1)–(H3), we assume that

(H4) $u_0 \in H^{-\alpha}$ where $\alpha > \frac{d}{2}$. Note that $u_0 \in L_2(G)$.

(H5) The operator A_1 is uniformly strongly elliptic of even order m_1 and has the same system of eigen functions $\{h_i, i \geq 1\}$ as $\mathcal{L}_\theta$.

The conditons (H1)–(H5) described above are the same as those in Huebner and Rozovskii (1995).

Note that $u_0 \in H^{-\alpha}$. For $\theta \in \Theta$, define

$$u_{0i}^\theta = (u_0, h_{i\theta}^{-\alpha})_{-\alpha}. \tag{2.8}$$

Then the random field

$$u_\varepsilon^\theta(t,x) = \sum_{i=1}^{\infty} u_{i\varepsilon}^\theta(t) h_{i\theta}^{-\alpha}(x) \tag{2.9}$$

is the solution of (2.1) subject to the boundary conditions (2.2) and (2.3) where $u_{i\varepsilon}^\theta(t)$ is the unique solution of the stochastic differential equation

$$du_{i\varepsilon}^\theta(t) = \mu_i^\theta u_{i\varepsilon}^\theta(t)dt + \varepsilon\lambda_i^{-\alpha}(\theta)dW_i(t), 0 \leq t \leq T, \tag{2.10}$$

$$u_i^{(\theta)}(0) = u_{0i}^\theta. \tag{2.11}$$

Suppose that $v_i = u_{0i}^\theta > 0$. For typographical convenience, we write $\mu_i(\theta)$ for μ_i^θ in the following discussion.

Observe that the parameter θ can be estimated from the equation (2.10). We now apply the minimum distance approach adapted by Kutoyants and Pilibossian (1994) to estimate the parameter θ satisfying the equation (2.10). We define the minimum L_1-norm estimate $\tilde{\theta}_{i\varepsilon T}$ by the relation

$$\mu_i(\tilde{\theta}_{i\varepsilon T}) = \arg\inf_{\theta\in\Theta} \int_0^T \left| u_{i\varepsilon}^\theta(t) - u_i^*(t,\theta) \right| \, dt$$

where $u_i^*(t,\theta)$ is the solution of the ordinary differential equation

$$\frac{du_i^*(t)}{dt} = \mu_i(\theta)u_i^*(t), u_i^*(0,\theta) = v_i.$$

It is easy to see that

$$u_i^*(t,\theta) = v_i e^{\mu_i(\theta)t}.$$

Let

$$h_i(\delta) = \inf_{\{\theta:|\mu_i(\theta)-\mu_i(\theta_0)|>\delta\}} \int_0^T \left| u_i^*(t,\theta) - u_i^*(t,\theta_0) \right| dt.$$

The following theorem is a consequence of Theorem 1 of Kutoyants and Pilibossian (1994).

Theorem 2.1. *For any $\delta > 0$,*

$$P_{\theta_0}(|\mu_i(\tilde{\theta}_{i\varepsilon T}) - \mu_i(\theta_0)| \geq \delta) \leq 2\exp\Big\{ - q_i\lambda_i^{2\alpha}(\theta_0)h_i^2(\delta)\varepsilon^{-2}\Big\}$$

where

$$q_i = \exp\{-2|\mu_i(\theta_0)|T\}/(2T)^3.$$

Let

$$J_i(t) = e^{\mu_i(\theta_0)t}\int_0^t e^{-\mu_i(\theta_0)s}\, dW_i(s).$$

Note that the process $J_i(t)$ is a gaussian process. Define

$$\gamma_{iT} = \arg\inf_u \int_0^T \left|J_i(t) - utv_ie^{\mu_i(\theta_0)t}\right|\, dt.$$

The following theorem is again a consequence of Theorems 2 and 3 of Kutoyants and Pilibossian (1994).

Theorem 2.2. *For any fixed $T > 0$,*

$$(\varepsilon\lambda_i^{-\alpha})^{-1}(\mu_i(\tilde{\theta}_{i\varepsilon T}) - \mu_i(\theta_0)) \xrightarrow{p} \gamma_{iT} \ \ as \ \ \varepsilon \to 0$$

when θ_0 is the true parameter. Furthermore if $\mu_i(\theta_0) > 0$, then

$$\gamma_{iT}Tv_i\sqrt{2\mu_i(\theta_0)} \xrightarrow{\mathcal{L}} N(0,1) \ \ as \ \ T \to \infty.$$

We now state and prove a lemma.

Lemma 2.1. *Suppose that for every $T > 0$,*

$$X_{\varepsilon T} \xrightarrow{p} Y_T \ \ as \ \ \varepsilon \to 0$$

and further suppose that

$$Y_T \xrightarrow{\mathcal{L}} Y \ \ as \ \ T \to \infty.$$

Then

$$X_{\varepsilon T} \xrightarrow{\mathcal{L}} Y \ \ as \ \ \varepsilon \to 0 \ \ and\ then \ \ T \to \infty.$$

Proof. Let F be a closed set and $F_\delta = \{x : \rho(x, F) \leq \delta\}$ where $\rho(x, F)$ denotes the distance between the point x and the closed set F. Note that F_δ decreases to the set F as δ decreases to zero. Then

$$P(X_{\varepsilon T} \in F) \leq P(Y_T \in F_\delta) + P(|X_{\varepsilon T} - Y_T| \geq \delta). \tag{2.12}$$

Hence

$$\limsup_{\varepsilon\to 0} P(X_{\varepsilon T}\in F) \le P(Y_T\in F_\delta)+\limsup_{\varepsilon\to 0} P(|X_{\varepsilon T}-Y_T|\ge\delta) = P(Y_T\in F_\delta)$$

since $X_{\varepsilon T}\xrightarrow{p} Y_T$ as $\varepsilon\to 0$. Taking limit as $T\to\infty$ in the above inequalities, we get that

$$\limsup_{T\to\infty}\limsup_{\varepsilon\to 0} P(X_{\varepsilon T}\in F) \le \limsup_{T\to\infty} P(Y_T\in F_\delta) \le P(Y\in F_\delta)$$

since the set F_δ is closed and $Y_T\xrightarrow{\mathcal{L}} Y$ as $T\to\infty$. Let $\delta\to 0$. Then we have

$$\limsup_{T\to\infty}\limsup_{\varepsilon\to 0} P(X_{\varepsilon T}\in F) \le P(Y\in F)$$

for every closed set F. Hence, by the standard results from the theory of weak convergence, it follows that

$$X_{\varepsilon T}\xrightarrow{\mathcal{L}} Y \text{ as } \varepsilon\to 0 \text{ and then } T\to\infty. \tag{2.13}$$

□

Applying Lemma 2.1, we get the following result.

Theorem 2.3. *Under the probability measure P_{θ_0}, if $\mu_i(\theta_0)>0$, then*

$$(\varepsilon\lambda_i^{-\alpha})^{-1}v_iT(\mu_i(\tilde\theta_{i\varepsilon T})-\mu_i(\theta_0))\sqrt{2\mu_i(\theta_0)}\xrightarrow{\mathcal{L}} N(0,1) \tag{2.14}$$

as $\varepsilon\to 0$ and then $T\to\infty$.

In addition to the conditions (H1)–(H5), suppose that

(H6) The functions $\mu_i(\theta)$ are differentiable with respect to θ with nonzero derivatives.

Let $\mu_i'(\theta)$ denote the derivative of the function $\mu_i(\theta)$ with respect to θ. Applying the delta method, we obtain the following result.

Theorem 2.4. *Under the probability measure P_{θ_0}, if $\mu_i(\theta_0)>0$, then*

$$(\varepsilon\lambda_i^{-\alpha})^{-1}v_iT(\tilde\theta_{i\varepsilon T}-\theta_0)\sqrt{2\mu_i(\theta_0)}\xrightarrow{\mathcal{L}} N(0,[\mu_i'(\theta_0)]^{-2}) \tag{2.15}$$

as $\varepsilon\to 0$ and then $T\to\infty$.

In view of Theorem 2.4, the variance of the limiting normal distribution of estimator $\tilde{\theta}_{i\varepsilon T}$ is proportional to

$$\left\{2v_i^2\mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2\right\}^{-1}.$$

Note that the estimators $\tilde{\theta}_{i\varepsilon T}, i \geq 1$ are independent estimators of the parameter θ_0 since the processes $\{W_i(t), t \geq 0\}, i \geq 1$ are independent Wiener processes. We will now construct an optimum estimator out of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$ for any $N \geq 1$.

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

$$\tilde{\theta}_{\varepsilon T} \xrightarrow{p} \left[\sum_{i=1}^N \alpha_i\right]\theta_0 \text{ as } \varepsilon \to 0 \text{ and then } T \to \infty$$

by Theorem 2.4 and hence $\tilde{\theta}_{\varepsilon T}$ is a consistent estimator for θ_0 as $\varepsilon \to 0$ and then $T \to \infty$ provided $\sum_{i=1}^N \alpha_i = 1$. Furthermore

$$\varepsilon^{-1}T(\tilde{\theta}_{\varepsilon T} - \theta_0) \xrightarrow{\mathcal{L}} N\Big(0, \sum_{i=1}^N \alpha_i^2\left\{2v_i^2\mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2\right\}^{-1}\Big)$$

as $\varepsilon \to 0$ and then $T \to \infty$. This follows again by Theorem 2.4 and the independence of the estimators $\{\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N\}$. We now obtain the optimum combination of the coefficients $\{\alpha_i, 1 \leq i \leq N\}$ by minimizing the asymptotic variance

$$\sum_{i=1}^N \alpha_i^2\left\{2v_i^2\mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2\right\}^{-1}$$

subject to the condition $\sum_{i=1}^N \alpha_i = 1$. It is easy to see that α_i is proportional to

$$\left\{2v_i^2\mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2\right\}$$

and the optimal choice of $\{\alpha_i, 1 \leq i \leq N\}$ leads to the estimator

$$\theta_{\varepsilon T}^* = \frac{\sum_{i=1}^N v_i^2\mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2\tilde{\theta}_{i\varepsilon T}}{\sum_{i=1}^N v_i^2\mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2}. \tag{2.16}$$

It is easy to see that

$$\theta_{\varepsilon T}^* \xrightarrow{p} \theta_0 \text{ as } \varepsilon \to 0 \text{ and then } T \to \infty,$$

$$\varepsilon^{-1}T(\theta^*_{\varepsilon T} - \theta_0) \xrightarrow{\mathcal{L}} N\Big(0, \Big\{ \sum_{i=1}^{N} 2v_i^2 \mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2 \Big\}^{-1}\Big)$$

as $\varepsilon \to 0$ and then $T \to \infty$ again due to the independence of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$. However the random variable $\theta^*_{\varepsilon T}$ cannot be considered as an estimator of the parameter θ_0 since it depends on the unknown parameter θ_0. In order to avoid this problem, we consider a modified estimator

$$\hat{\theta}_{\varepsilon T} = \frac{\sum_{i=1}^{N} v_i^2 \mu_i(\tilde{\theta}_{i\varepsilon T})\lambda_i^{2\alpha}(\tilde{\theta}_{i\varepsilon T})[\mu_i'(\tilde{\theta}_{i\varepsilon T})]^2 \tilde{\theta}_{i\varepsilon T}}{\sum_{i=1}^{N} v_i^2 \mu_i(\tilde{\theta}_{i\varepsilon T})\lambda_i^{2\alpha}(\tilde{\theta}_{i\varepsilon T})[\mu_i'(\tilde{\theta}_{i\varepsilon T})]^2}. \tag{2.17}$$

which is obtained from $\theta^*_{\varepsilon T}$ by substituting the estimator $\tilde{\theta}_{i\varepsilon T}$ for the unknown parameter θ_0 in the i-th term in the numerator and the denominator in (2.16). In view of the independence, consistency and asymptotic normality of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$, it follows that the estimator $\hat{\theta}_{\varepsilon T}$ is consistent and asymptotically normal for the parameter θ_0 and we have the following result.

Theorem 2.5. *Under the probability measure* P_{θ_0},

$$\hat{\theta}_{\varepsilon T} \xrightarrow{p} \theta_0 \ \ as \ \ \varepsilon \to 0 \ \ and\ then \ \ T \to \infty$$

and, if $\mu_i(\theta_0) > 0,\ 1 \leq i \leq N$, *then*

$$\varepsilon^{-1}T(\hat{\theta}_{\varepsilon T} - \theta_0) \xrightarrow{\mathcal{L}} N\Big(0, \Big\{ \sum_{i=1}^{N} 2v_i^2 \mu_i(\theta_0)\lambda_i^{2\alpha}(\theta_0)[\mu_i'(\theta_0)]^2 \Big\}^{-1}\Big)$$

as $\varepsilon \to 0$ *and then* $T \to \infty$ *for any fixed* $N \geq 1$.

3. Stochastic PDE with Linear Drift (Absolutely Continuous Case)

We now consider an example illustrating the results discussed in Section 2.

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

$$du_\varepsilon(t,x) = (\triangle u_\varepsilon(t,x) + \theta u_\varepsilon(t,x))dt + \varepsilon dW_Q(t,x) \tag{3.1}$$

where $\triangle = \partial^2/\partial x^2$. Suppose that $\varepsilon \to 0$ and $\theta \in \Theta = (\alpha, \beta) \subset R$. Suppose the initial and the boundary conditions are given by

$$\begin{cases} u_\varepsilon(0,x) = f(x), f \in L_2[0,1] \\ u_\varepsilon(t,0) = u_\varepsilon(t,1) = 0, 0 \leq t \leq T \end{cases} \tag{3.2}$$

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)

$$W_Q(t,x) = \sum_{i=1}^{\infty} q_i^{1/2} e_i(x) W_i(t) \text{ a.s.} \tag{3.3}$$

where $\{W_i(t), 0 \leq t \leq T\}, i \geq 1$ are independent one-dimensional standard Wiener processes and $\{e_i\}$ is a complete orthonormal system in $L_2[0,1]$ consisting of eigen vectors of Q and $\{q_i\}$ eigen values of Q.

Let us consider a special covariance operator Q with $e_k = \sin k\pi x, k \geq 1$ and $\lambda_k = (\pi k)^2, k \geq 1$. Then $\{e_k\}$ is a complete orthonormal system with eigen values $q_i = (1+\lambda_i)^{-1}, i \geq 1$ for the operator Q and $Q = (I - \triangle)^{-1}$. Furthermore

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

We define a solution $u_\varepsilon(t,x)$ of (3.1) as a formal sum

$$u_\varepsilon(t,x) = \sum_{i=1}^{\infty} u_{i\varepsilon}(t) e_i(x) \tag{3.4}$$

(*cf.* Rozovskii, 1990). It is known that the Fourier coefficient $u_{i\varepsilon}(t)$ satisfies the stochastic differential equation

$$du_{i\varepsilon}(t) = (\theta - \lambda_i) u_{i\varepsilon}(t) dt + \frac{\varepsilon}{\sqrt{\lambda_i + 1}} dW_i(t), \; 0 \leq t \leq T \tag{3.5}$$

with the initial condition

$$u_{i\varepsilon}(0) = v_i, \; v_i = \int_0^1 f(x) e_i(x) dx. \tag{3.6}$$

We assume that $v_i > 0$.

It is known that $u_\varepsilon(t,x)$ as defined above belongs to $L_2([0,T] \times \Omega; L_2[0,1])$ together with its derivative in t. Furthermore $u_\varepsilon(t,x)$ is the only solution to (3.1) under the boundary condition (3.2). Let $P_\theta^{(\varepsilon)}$ be the measure generated by u_ε when θ is the true parameter. It can be shown that the family of probability measures $\{P_\theta, \theta \in \Theta\}$ form an equivalent family. Suppose θ_0 is the true parameter.

Observe that the parameter θ can be estimated from the equation (3.5). We now apply the minimum distance approach adapted by Kutoyants and Pilibossian (1994) to estimate the parameter θ satisfying the equation (3.5). We define the minimum L_1-norm estimate $\tilde{\theta}_{i\varepsilon T}$ by the relation

$$\tilde{\theta}_{i\varepsilon T} = \lambda_i + \arg\inf_{\theta\in\Theta} \int_0^T \Big| u_{i\varepsilon}(t) - u_i(t,\theta) \Big| \; dt$$

where $u_i(t,\theta)$ is the solution of the ordinary differential equation

$$\frac{du_i(t)}{dt} = (\theta - \lambda_i) u_i(t), u_i(0,\theta) = v_i.$$

It is easy to see that

$$u_i(t,\theta) = v_i e^{(\theta-\lambda_i)t}.$$

Let

$$g_i(\delta) = \inf_{|\theta-\theta_0|>\delta} \int_0^T \Big| u_i(t,\theta) - u_i(t,\theta_0) \Big| dt.$$

The following theorem is a consequence of Theorem 1 of Kutoyants and Pilibossian (1994).

Theorem 3.1. *For any* $\delta > 0$,

$$P_{\theta_0}^{(\varepsilon)}(|\tilde{\theta}_{i\varepsilon T} - \theta_0| \geq \delta) \leq 2\exp\Big\{ - k_i(\lambda_i + 1) g_i^2(\delta)\varepsilon^{-2} \Big\}$$

where $k_i = \exp\{-2|\theta_0 - \lambda_i|T\}/(2T)^3$.

Let

$$Y_i(t) = e^{(\theta_0-\lambda_i)t} \int_0^t e^{-(\theta_0-\lambda_i)s} \; dW_i(s).$$

Note that the process $Y_i(t)$ is a gaussian process. Define

$$\zeta_{iT} = \arg\inf_u \int_0^T \Big| Y_i(t) - utv_i e^{(\theta_0-\lambda_i)t} \Big| \; dt.$$

The following theorem is again a consequence of Theorems 2 and 3 of Kutoyants and Pilibossian (1994).

Theorem 3.2. *For any fixed* $T > 0$,

$$\Big(\frac{\varepsilon}{\sqrt{\lambda_i + 1}}\Big)^{-1} (\tilde{\theta}_{i\varepsilon T} - \theta_0) \xrightarrow{p} \zeta_{iT} \quad as\ \varepsilon \to 0$$

where θ_0 *is the true parameter. Furthermore if* $\theta_0 > \lambda_i$, *then*

$$\zeta_{iT} T v_i \sqrt{2(\theta_0 - \lambda_i)} \xrightarrow{\mathcal{L}} N(0,1) \quad as\ T \to \infty.$$

Let

$$X_{\varepsilon T} = \Big(\frac{\varepsilon}{\sqrt{\lambda_i + 1}}\Big)^{-1} (\tilde{\theta}_{i\varepsilon T} - \theta_0) T v_i \sqrt{2(\theta_0 - \lambda_i)}, \tag{3.7}$$

$$Y_T = \zeta_{iT} T v_i \sqrt{2(\theta_0 - \lambda_i)} \tag{3.8}$$

and Y be a standard normal random variable. Applying Lemma 2.1, we get the following result.

Theorem 3.3. *If $\theta_0 > \lambda_i$, then*

$$\Big(\frac{\varepsilon}{\sqrt{\lambda_i + 1}}\Big)^{-1} (\tilde{\theta}_{i\varepsilon T} - \theta_0) T v_i \sqrt{2(\theta_0 - \lambda_i)} \xrightarrow{\mathcal{L}} N(0,1) \tag{3.9}$$

as $\varepsilon \to 0$ and then $T \to \infty$.

In view of Theorem 3.3, the variance of the limiting normal distribution of estimator $\tilde{\theta}_{i\varepsilon T}$ is proportional to $[2v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)]^{-1}$. Note that the estimators $\tilde{\theta}_{i\varepsilon T}, i \geq 1$ are independent estimators of the parameter θ since the processes $\{W_i(t), t \geq 0\}, i \geq 1$ are independent Wiener processes. We will now construct an optimum estimator out of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$ for any $N \geq 1$.

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

$$\tilde{\theta}_{\varepsilon T} \xrightarrow{p} \Big[\sum_{i=1}^{N} \alpha_i\Big] \theta_0 \text{ as } \varepsilon \to 0 \text{ and then } T \to \infty$$

by Theorem 3.3 and hence $\tilde{\theta}_{\varepsilon T}$ is a consistent estimator for θ_0 as $\varepsilon \to 0$ and then $T \to \infty$ provided $\sum_{i=1}^N \alpha_i = 1$. Furthermore

$$\varepsilon^{-1} T(\tilde{\theta}_{\varepsilon T} - \theta_0) \xrightarrow{\mathcal{L}} N\Big(0, \sum_{i=1}^{N} \alpha_i^2 \Big\{2v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)\Big\}^{-1}\Big)$$

as $\varepsilon \to 0$ and then $T \to \infty$. This follows again by Theorem 3.3 and the independence of the estimators $\{\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N\}$. We now obtain the optimum combination of the coefficients $\{\alpha_i, 1 \leq i \leq N\}$ by minimizing the asymptotic variance

$$\sum_{i=1}^{N} \alpha_i^2 \Big\{2v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)\Big\}^{-1}$$

subject to the condition $\sum_{i=1}^{N} \alpha_i = 1$. It is easy to see that α_i is proportional to $[v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)]$ and the optimal choice of $\{\alpha_i, 1 \leq i \leq N\}$ leads to the "estimator"

$$\theta_{\varepsilon T}^* = \frac{\sum_{i=1}^{N} v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)\tilde{\theta}_{i\varepsilon T}}{\sum_{i=1}^{N} v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)}. \tag{3.10}$$

It is easy to see that

$$\theta_{\varepsilon T}^* \xrightarrow{p} \theta_0 \text{ as } \varepsilon \to 0 \text{ and then } T \to \infty,$$

$$\varepsilon^{-1} T(\theta_{\varepsilon T}^* - \theta_0) \xrightarrow{\mathcal{L}} N\Big(0, \Big\{\sum_{i=1}^{N} 2v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)\Big\}^{-1}\Big)$$

as $\varepsilon \to 0$ and then $T \to \infty$ again due to the independence of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$. However the random variable $\theta_{\varepsilon T}^*$ cannot be considered as estimator of the parameter θ_0 since it depends on the unknown parameter θ_0. In order to avoid this problem, we consider a modified estimator

$$\hat{\theta}_{\varepsilon T} = \frac{\sum_{i=1}^{N} v_i^2(\tilde{\theta}_{i\varepsilon T} - \lambda_i)(\lambda_i + 1)\tilde{\theta}_{i\varepsilon T}}{\sum_{i=1}^{N} v_i^2(\tilde{\theta}_{i\varepsilon T} - \lambda_i)(\lambda_i + 1)}. \tag{3.11}$$

which is obtained from $\theta_{\varepsilon T}^*$ by substituting the estimator $\tilde{\theta}_{i\varepsilon T}$ for the unknown parameter θ_0 in the i-th term in the numerator and the denominator in (3.10). In view of the independence, consistency and asymptotic normality of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$, it follows that the estimator $\hat{\theta}_{\varepsilon T}$ is consistent and asymptotically normal for the parameter θ_0 and we have the following result.

Theorem 3.4. *Under the probability measure P_{θ_0},*

$$\hat{\theta}_{\varepsilon T} \xrightarrow{p} \theta_0 \text{ as } \varepsilon \to 0 \text{ and then } T \to \infty$$

and if $\theta_0 > N^2\pi^2$, then

$$\varepsilon^{-1} T(\hat{\theta}_{\varepsilon T} - \theta_0) \xrightarrow{\mathcal{L}} N\Big(0, \Big\{\sum_{i=1}^{N} 2v_i^2(\theta_0 - \lambda_i)(\lambda_i + 1)\Big\}^{-1}\Big)$$

as $\varepsilon \to 0$ and then $T \to \infty$.

4. Stochastic PDE with Linear Drift (Singular Case)

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

$$du_\varepsilon(t,x) = \theta \, \triangle u_\varepsilon(t,x)dt + \varepsilon(I - \triangle)^{-1/2}dW(t,x) \tag{4.1}$$

where $\theta > 0$ satisfying the initial and the boundary conditions

$$\begin{aligned} u_\varepsilon(0,x) &= f(x),\ 0 < x < 1,\ f \in L_2[0,1], \\ u_\varepsilon(t,0) &= u_\varepsilon(t,1) = 0,\ 0 \leq t \leq T. \end{aligned} \tag{4.2}$$

Here I is the identity operator, $\triangle = \partial^2/\partial x^2$ as defined in Section 3 and the process $W(t,x)$ is the cylindrical Brownian motion in $L_2[0,1]$.

In analogy with (3.5), it can be checked that the Fourier coefficients $u_{i\varepsilon}(t)$ satisfy the stochastic differential equations

$$du_{i\varepsilon}(t) = -\theta\lambda_i u_{i\varepsilon}(t)dt + \frac{\varepsilon}{\sqrt{\lambda_i + 1}}dW_i(t),\ 0 \leq t \leq T, \tag{4.3}$$

with

$$u_{i\varepsilon}(0) = v_i, v_i = \int_0^1 f(x)e_i(x)dx. \tag{4.4}$$

We assume that $v_i > 0$.

Let $P_\theta^{(\varepsilon)}$ be the measure generated by u_ε when θ is the true parameter. It can be shown that the family of measures $\{P_\theta^{(\varepsilon)}, \theta \in \Theta\}$ *do not* form a family of equivalent probability measures. In fact, $P_\theta^{(\varepsilon)}$ is singular with respect to $P_{\theta'}^{(\varepsilon)}$ whenever $\theta \neq \theta'$ in Θ (*cf.* Huebner *et al.*, 1993).

Observe that the parameter θ can be estimated from the equation (4.3). We now again apply the minimum distance approach adapted by Kutoyants and Pilibossian (1994) as before to estimate the parameter θ satisfying the equation (4.3). We define the minimum L_1-norm estimate $\tilde{\theta}_{i\varepsilon T}$ by the relation

$$\tilde{\theta}_{i\varepsilon T} = -\lambda_i^{-1} \arg \inf_{\theta \in \Theta} \int_0^T \left| u_{i\varepsilon}(t) - u_i(t,\theta) \right| \, dt$$

where $u_i(t,\theta)$ is the solution of the ordinary differential equation

$$\frac{du_i(t)}{dt} = -\theta\lambda_i u_i(t), u_i(0,\theta) = v_i.$$

It is easy to see that

$$u_i(t,\theta) = v_i e^{-\theta\lambda_i t}.$$

Let

$$g_i(\delta) = \inf_{|\theta-\theta_0|>\delta\lambda_i^{-1}} \int_0^T \Big|u_i(t,\theta) - u_i(t,\theta_0)\Big| dt.$$

The following theorem is a consequence of Theorem 1 of Kutoyants and Pilibossian (1994).

Theorem 4.1. *For any* $\delta > 0$,

$$P_{\theta_0}^{(\varepsilon)}(|\tilde{\theta}_{i\varepsilon T} - \theta_0| \geq \delta\lambda_i^{-1}) \leq 2\exp\Big\{ - k_i(\lambda_i+1)g_i^2(\delta)\varepsilon^{-2}\Big\}$$

where $k_i = \exp\{-2|\theta_0|\lambda_i T\}/(2T)^3$.

Let

$$Y_i(t) = e^{-\theta_0\lambda_i t}\int_0^t e^{\theta_0\lambda_i s}\ dW_i(s).$$

Note that the process $Y_i(t)$ is a gaussian process. Define

$$\eta_{iT} = \arg\inf_u \int_0^T \Big|Y_i(t) - utv_i e^{-\theta_0\lambda_i t}\Big|\ dt.$$

The following theorem is again a consequence of Theorem 2 of Kutoyants and Pilibossian (1994).

Theorem 4.2. *For any fixed* $T > 0$,

$$\left(\frac{\varepsilon}{\sqrt{\lambda_i+1}}\right)^{-1}(\tilde{\theta}_{i\varepsilon T} - \theta_0)\lambda_i \xrightarrow{p} -\eta_{iT} \ \ as\ \ \varepsilon \to 0$$

where θ_0 *is the true parameter.*

Note that the estimators $\tilde{\theta}_{i\varepsilon T}, i \geq 1$ are independent estimators of the parameter θ since the processes $\{W_i(t), t \geq 0\}, i \geq 1$ are independent Wiener processes. Consider a linear estimator constructed out of the estimators $\tilde{\theta}_{i\varepsilon T}, 1 \leq i \leq N$ for any $N \geq 1$.

Let $\tilde{\theta}_{\varepsilon T} = \sum_{i=1}^N \alpha_i\tilde{\theta}_{i\varepsilon T}$ where $\alpha_i, 1 \leq i \leq N$ is a nonrandom sequence of coefficients. Note that

$$\tilde{\theta}_{\varepsilon T} \xrightarrow{p} \left(\sum_{i=1}^N \alpha_i\right)\theta_0 \ \ \text{as}\ \ \varepsilon \to 0$$

by Theorem 4.2 and hence $\tilde{\theta}_{\varepsilon T}$ is a consistent estimator for θ_0 as $\varepsilon \to 0$ provided $\sum_{i=1}^N \alpha_i = 1$. However no optimality properties of this estimator for suitable choice of the weights α_i's could be established.

Remark. The interesting difference in parameter estimation for SPDE as compared to SDE is that one can consider an additional parameter for asymptotics. In the case of SPDE, the dimension of finite dimensional projection N can be considered as the parameter for asymptotics in addition to the time of observation T and the amplitude of noise ε as used in the case of SDE. Here we have discussed asymptotic properties of estimators as the amplitude of noise $\varepsilon \to 0$ and then time of observation $T \to \infty$ based on the finite N-dimensional projection of the observation. It would be interesting to study the asymptotic properties of the estimators as $N \to \infty$ for fixed ε and T as was done by Huebner *et al.* (1993) in the case of maximum likelihood estimators. Another problem of interest is to obtain a lower minimax bound on the risk over all estimators for fixed N and compare it with the limiting variance of $\hat{\theta}_{\varepsilon T}$.

Acknowledgements

The author thanks the two referees for their detailed and pertinent comments on the original version of the paper.

References

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

Huebner, M., Khasminski, 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*, Springer, New York, 149–160.

Huebner, M. and Rozovskii, B. L. (1995). "On asymptotic properties of maximum likelihood estimators for parabolic stochastic SPDE's", *Probability Theory and Related 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.

Kutoyants, Y. (1994). *Identification of Dynamical Systems with Small Noise*, Kluwer Academic Publishers, Dordrecht.

Kutoyants, Y. and Pilibossian, P. (1994). "On minimum L_1-norm estimate of the parameter of the Ornstein-Uhlenbeck process", *Statistics & Probability Letters*, **20**, 117–123.

Millar, P. W. (1984). "A general approach to the optimality of the minimum distance estimators", *Transactions of the American Mathematical Society*, **286**, 249–272.

Prakasa Rao, B. L. S. (2000). "Bayes estimation for some stochastic partial differential equations", *Journal of Statistical Planning Inference*, **91**, 511–524.

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

Shimakura, N. (1992). "Partial differential equations of elliptic type", *Translations of American Mathematical Society*, **99**, Providence, Rhode Island.