ГОДИШНИК НА СОФИЙСКИЯ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ"

ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА
Том 93, 1999

ANNUAIRE DE L'UNIVERSITE DE SOFIA „ST. KLIMENT OHRIDSKI"

FACULTE DE MATHEMATIQUES ET INFORMATIQUE
Tome 93, 1999

PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATION OF CONSTANT DRIFT AND DIFFUSION PARAMETERS IN k-DIMENSIONAL DIFFUSION PROCESS OBSERVED AT DISCRETE RANDOM SAMPLING

DIMITRINKA IVANOVA VLADEVA

This article is concerned with the problem of a parameter estimation of the constant drift and diffusion coefficients: unknown vector A and unknown positive definite matrix B, respectively, of a k-dimensional diffusion type process, when the observations at the moment of random point process are given. We compute the means and variances of the maximum likelihood estimators and establish their asymptotic properties. The unbiasedness, the strong consistency and the asymptotical efficiency of the estimation for A are proved. The estimator of B is unbiased and consistent and the variance of this estimator does not depend on the distribution of the random moments of observations.

Keywords: diffusion process, k-dimensional Wiener process, discrete random sampling, maximum likelihood estimation, unbiasedness, consistency, efficiency, Fisher information

MSC 2000: 62H12, 62M05

1. INTRODUCTION

In this article we consider the diffusion process $X_t = (X_t^1, X_t^2, \dots, X_t^k)^T$, $t \geq 0$, defined by the stochastic differential equation

$$dX_t = A\,dt + B^{1/2}\,dW_t, \quad t \geq 0, \quad X_0 = 0, \tag{1}$$

where $A = (a^1, a^2, \dots, a^k)^T$ and

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ \cdots & & & \\ b_{k1} & b_{k2} & \dots & b_{kk} \end{pmatrix}$$

are unknown constant vector and positive definite symmetric matrix, respectively, and $W_t = (W_t^1, W_t^2, ..., W_t^k)^T$ is a standard Wiener process with a mean 0 and a variance I_k.

The solution of the differential equation (1) exists in a strong sense, it is unique and is represented by the process

$$X_t = At + B_1 W_t, \quad t \geq 0. \tag{2}$$

More information about this problem can be found in [1].

The maximum likelihood estimation problem for the model (1), when we observe the process X_t, $t \geq 0$, continuously in the interval $[0, T]$, is solved in [2]. In the case when we have at disposal the discrete observations at equidistant points, many close to this one problems can be found in the monography [3]. At first, the random sampling scheme has been used by J. Beutler in [4]. Recently, many authors (see [5] and [6]) consider continuous diffusion processes when the observations are provided in discrete moments belonging to the interval $[0, T]$. In [7] A. Le Breton has solved the estimation problem for the model (1) when the points of observations are determinant.

Usually, the maximum of the distance between the points of observations tends to zero, while their number tends to infinity.

Our conditions are more natural. Let us denote the observations $X_{t_1}, \ldots, X_{t_N}$, $X_{t_i} = (X_{t_i}^1, X_{t_i}^2, \ldots, X_{t_i}^k)^T$. The moments $t_1, \ldots, t_N$ are the first N points of an arbitrary point process with independent identically distributed increments. The process $\{t_i\}$, $i = 1, \ldots, N$, is independent of the process X_t, $t \geq 0$, and we compute $E(F(X_t)) = E_t(E_X(F(X_t)))$. The problem is to find the maximum likelihood estimators of the unknown constant vector A and the matrix B and to establish their properties. In the one-dimensional case this problem was solved in [8].

Let us denote $X_i = X_{t_i}$, $\Delta X_i = X_i - X_{i-1}$, $\Delta_i = t_i - t_{i-1}$, $i = 1, \ldots, N$. For simplicity we denote $\overline{B} = B^{1/2}$.

2. MAXIMUM LIKELIHOOD ESTIMATION

Using the maximum likelihood method, we can prove the following natural results.

Theorem 1. *If $N \geq 2$, the statistic*

$$\hat{A}_N = \frac{X_N}{t_N} \tag{3}$$

is a maximum likelihood estimator for an unknown vector A.

We prove this theorem using the standart maximum likelihood procedure.

Theorem 2. *If $N \geq k$, the statistic*

$$\widetilde{B_N} = \frac{1}{N} \sum_{i=1}^{N} \left\{ \frac{\Delta X_i \Delta X_i^T}{\Delta_i} - \frac{X_N X_N^T}{t_N} \right\} \tag{4}$$

is a maximum likelihood estimator for an unknown matrix B, when $A \neq 0$.

For $A = 0$ the maximum likelihood estimator is

$$\widetilde{B_N} = \frac{1}{N} \sum_{i=1}^{N} \frac{\Delta X_i \Delta X_i^T}{\Delta_i} .$$

Our approach is different from the ones used in the proofs of similar propositions. (For example, see [9, p. 75].) Our proof is based on the following lemma.

Lemma 1. *Let $y_i = (y_i^1, \dots, y_i^k)^T$, $i = 1, \dots, N$, be k-dimensional vectors such that $B = \sum_{i=1}^{N} y_i y_i^T$ is a non-singular matrix. Then:*

a) *the matrix B is symmetric and positive definite;*

b) $C = \sum_{i=1}^{N} y_i^T \left(\sum_{i=1}^{N} y_i y_i^T \right)^{-1} y_i = k.$

Let us note that for the estimation of one of the parameters (A or B) it is not necessary to know the other one.

The estimator $\widetilde{B_N}$ is not unbiased. Therefore we prefer to use the estimator

$$\hat{B}_N = \frac{1}{N-1} \sum_{i=1}^{N} \left\{ \frac{\Delta X_i \Delta X_i^T}{\Delta_i} - \frac{X_N X_N^T}{t_N} \right\} . \tag{5}$$

The proofs of these theorems can be found in [12].

Our purpose here is to establish the properties of the maximum likelihood estimators, which are given in Theorems 1 and 2.

3. PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATORS

By arbitrary distribution of the process $t_1, \dots, t_N$ and natural conditions for its moments we can compute the means and variances of the considered estimators and to obtain their asymptotic properties.

Let $\max_{1 \le i \le N} \Delta_i$ do not tend to 0 when $N \to \infty$, and $t_N \to \infty$ when $N \to \infty$.

Theorem 3. *The estimator given by (5) is unbiased, strongly consistent and asymptotically effective for the unknown vector A if the condition*

$$\sum_{N=2}^{\infty} E(t_N^{-3}) < \infty \tag{6}$$

is satisfied.

Proof. We compute

$$E\hat{A}_N = E\left(\frac{X_N}{t_N}\right) = E_t E_X \left(\frac{X_N}{t_N}\right) = E_t \frac{A t_N}{t_N} = A.$$

Hence, $\hat{A}_N$ is an unbiased estimator.

The variance of this estimation is

$$E\hat{A}_N\hat{A}_N^T = E\left(\frac{X_N X_N^T}{t_N^2}\right) = \frac{E(At_N + \overline{B}W_N)(At_N + \overline{B}W_N)^T}{t_N^2}$$

$$= AA^T + B^2 E\frac{W_N W_N^T}{t_N^2} = AA^T + BE\frac{1}{t_N} \cdot$$

As $E\left(\frac{1}{t_N}\right) \to 0$ when $N \to \infty$, the considered statistic $\hat{A}_N$ is consistent.

A sufficient condition for the strong consistency of $\hat{A}_N$ is

$$\sum_{N=2}^{\infty} P\left(|\hat{A}_N - A| > \varepsilon\right) < \infty.$$

Let $\|A\|_4 = \left(E\sum_{i=1}^{k} a_i^4\right)^{1/4}$ be the norm of the random vector A and $\lambda = \|B\|$ be the maximum eigenvalue and the norm of the positive definite matrix B. Then using Markov's inequality, we obtain

$$\sum_{N=2}^{\infty} P\left(\left|\hat{A}_N - A\right| > \varepsilon\right) \le \frac{1}{\varepsilon^4}\sum_{N=2}^{\infty} E\left|\hat{A}_N - A\right|^4 \le \frac{1}{\varepsilon^4}\sum_{N=2}^{\infty} E\left(\|\hat{A}_N - A\|_4\right)^4$$

$$= \frac{1}{\varepsilon^4}\sum_{N=2}^{\infty}\left\|\frac{\overline{B}W_N}{t_N}\right\|^4 \le \frac{\lambda^2}{\varepsilon^4}\sum_{N=2}^{\infty} E\sum_{i=1}^{k}\left(\frac{W_N^i}{t_N}\right)^4 = \frac{3k\lambda^2}{\varepsilon^4}\sum_{N=2}^{\infty} E\frac{1}{{t_N}^3} \cdot$$

From condition (6) it follows that the estimator (3) is strongly consistent.

The Fisher information matrix for the estimator $\hat{A}_N$ is

$$I_{\hat{A}_N}(l,m) = E_A\left(\left\{\frac{\partial}{\partial a^l}\ln\frac{dP_{A,B}}{d\lambda}(x)\right\}\left\{\frac{\partial}{\partial a^m}\ln\frac{dP_{A,B}}{d\lambda}(x)\right\}\right),$$

where λ is the Lebesque measure in R^k.

Let c_{jl} be the (j,l)-th element of the matrix B^{-1} and $b_{i,j}$ be the (i,j)-th element of the matrix $\overline{B}$. We compute the first factor of $I_{\hat{A}_N}(l,m)$:

$$\frac{\partial}{\partial a^l}\ln\frac{dP_{A,B}}{d\lambda}(x) = \frac{\partial}{\partial a^l}l(A,B) = \sum_{i=1}^{N}\sum_{j=1}^{k} c_{lj}\left(a^j\Delta_i - \Delta X_i^j\right).$$

We use that

$$\Delta X_i^j = a^j\Delta_i + \sum_{n=1}^{k} b_{jn}\Delta W_i^n$$

and

$$\frac{\partial}{\partial a^l}l(A,B) = -\sum_{i=1}^{N}\sum_{j=1}^{k}\sum_{n=1}^{k} b_{jn}c_{lj}\Delta W_i^n.$$

Then

$$I_{\hat{A}_N}(l,m) = E\sum_{i=1}^{N}\sum_{j=1}^{k}\sum_{n=1}^{k}\sum_{p=1}^{k}\sum_{q=1}^{k} b_{jn}c_{lj}\Delta W_i^n b_{pq}c_{mp}\Delta W_i^q.$$

For $n \neq q$ the mathematical expectation is equal to zero and

$$E\sum_{i=1}^{N}(\Delta W_i^n)^2 = \sum_{i=1}^{N} E\Delta_i = Et_N,$$

so the information matrix is

$$I_{\hat{A}_N}(l,m) = E\sum_{i=1}^{N}\sum_{j=1}^{k}\sum_{p=1}^{k}\sum_{n=1}^{k} b_{jn}c_{lj}c_{mp}b_{pn}(\Delta W_i^n)^2$$

$$= Et_N\sum_{j=1}^{k}\sum_{p=1}^{k}\sum_{n=1}^{k} b_{jn}c_{lj}c_{mp}b_{pn} = Et_N\sum_{p=1}^{k}\sum_{j=1}^{k} B_{jp}c_{lj}c_{mp}$$

$$= Et_N\sum_{p=1}^{k} D_{lp}c_{mp} = Et_N c_{lm}.$$

Here B_{jp} denotes the (jp)-th element of the matrix B and D_{lp} denotes the (lp)-th element of the diagonal matrix $B^{-1}B = I_k$.

In this way we obtained that the information matrix is $B^{-1}Et_N$. Hence

$$\text{eff } \hat{A}_N \sim \left(BE\,\frac{1}{t_N}B^{-1}Et_N\right)^{-1} \to 1$$

when $N \to \infty$ and the estimator $\hat{A}_N$ is asymptotically effective.

The distribution of $\hat{A}_N$ (for a fixed t_N) is normal with parameters A and $\dfrac{B}{t_N}$, i.e. the vector

$$\left(\hat{A}_N - A\right)\overline{B}^{-1}\sqrt{t_N}$$

has a k-dimensional standard normal distribution.

Note. The sufficient conditions for strong consistency of the estimator can be written in terms of statistical moments of Δ_i.

For the estimator $\hat{B}_N$ we can establish the following properties:

Theorem 4. *The statistic* (5) *from Theorem* 2 *is unbiased and consistent estimation for the unknown matrix* B *and the variance of this estimation is*

$$E(\hat{B}_N - B)^2 = \frac{k+1}{N-1}B^2. \tag{7}$$

To prove these properties, we calculate the moments of the estimator. We need the next lemmas.

Lemma 2. *For every integer $k \geq 2$ the identity*

$$\sum_{i_1=1}^{N} \cdots \sum_{i_k=1}^{N} \min(i_1, \ldots, i_k) = \sum_{i=1}^{N} i^k$$

holds, where i_1, i_2, ..., i_k are natural numbers between 1 and N.

Proof. We consider the sum

$$\sum_{i_1=1}^{N} \cdots \sum_{i_k=1}^{N} \min(i_1, \ldots, i_k),$$

where $k \geq 2$. The minimum can accept all integer values from 1 to N. The number of all possible values of the k-dimensional variable $(i_1, i_2, \ldots, i_k)$ is N^k. The number 1 does not appear in $(N-1)^k$ cases. For the rest $N^k - (N-1)^k$ cases, $\min(i_1, \ldots, i_k) = 1$. Thus $\min(i_1, \ldots, i_k) = 2$ exactly $(N-1)^k - (N-2)^k$ times, ..., $\min(i_1, \ldots, i_k) = N$ only one time.

Therefore

$$\sum_{i_1, \ldots, i_k=1}^{N} \min(i_1, \ldots, i_k) = \sum_{i=1}^{N} (N-i+1)[i^k - (i-1)^k]$$

$$= N.[1^k - 0^k] + \cdots + 1.[N^k - (N-1)^k] = \sum_{i=1}^{N} i^k.$$

If $k = 2$, we obtain

$$\sum_{i,j=1}^{N} \min(i,j) = \sum_{i=1}^{N} i^2 = \frac{N(N+1)}{2} \cdot$$

Some characteristics of the considered processes will be usefull for our proofs. Let us denote

$$\Delta W_i = W_{t_i} - W_{t_{i-1}} = \left(W_{t_i}^1 - W_{t_{i-1}}^1, W_{t_i}^2 - W_{t_{i-1}}^2, \ldots, W_{t_i}^k - W_{t_{i-1}}^k\right)^T,$$

$\Delta_i = t_i - t_{i-1}$ and I_k is the identity matrix of dimension k. The arbitrary reneval point process t_i, $i = 1, \ldots, N$, is independent of the process X_t.

Lemma 3. *For the moments of the Wiener process the following equalities are satisfied*:

$$E\Delta W_i \Delta W_i^T = E\Delta_i \, I_k, \quad E\left(\Delta W_i \Delta W_i^T\right)^2 = (k+2)\, E\Delta_i^2 \, I_k,$$

$$E\Delta W_i \Delta W_j^T = 0, \ i \neq j, \quad E\left(\Delta W_i \Delta W_j^T\right)^2 = E\Delta_i^2 \, I_k, \ i < j,$$

$$E\Delta W_i \Delta W_i^T \Delta W_i = 0, \quad E\Delta W_i W_N^T = E\Delta_i \, I_k,$$

$$EW_N W_N^T = Et_N \, I_k, \quad E\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\Delta_j}{t_N} = N - 1.$$

Proof. We will prove only the first two equalities.

We use the facts that B and $\hat{B}_N$ are symmetric matrices, the increments of the Wiener process are independent and their odd moments are equal to zero.

$$E\Delta W_i^l \Delta W_i^m = 0,\ l \neq m, \quad E(\Delta W_i^l)^2 = E\Delta_i,\ l, m = 1, \dots, k.$$

Then it holds:

$$E\Delta W_i \Delta W_i^T = \begin{pmatrix} E(\Delta W_i^1)^2 & E(\Delta W_i^1 \Delta W_i^2) & \dots & E(\Delta W_i^1 \Delta W_i^k) \\ \dots & & & \\ E(\Delta W_i^k \Delta W_i^1) & E(\Delta W_i^k \Delta W_i^2) & \cdots & E(\Delta W_i^k)^2 \end{pmatrix}$$

$$= \begin{pmatrix} \Delta_i & 0 & \dots & 0 \\ \cdots & & & \\ 0 & 0 & \dots & \Delta_i \end{pmatrix} = E\Delta_i \, I_k.$$

Let d_{lm}, where $l, m = 1, 2, \dots, k$, be the (l, m)-th element of the matrix $E(\Delta W_i \Delta W_i^T)^2$. Then

$$d_{lm} = E \sum_{\substack{n=1 \\ n \neq l,m}}^{k} (\Delta W_i^n)^2 \Delta W_i^l \Delta W_i^m + E \, \Delta W_i^l (\Delta W_i^m)^3$$
$$+ E \, (\Delta W_i^l)^3 \Delta W_i^m + E \, (\Delta W_i^l)^4.$$

So $d_{lm} = 0$ if $l \neq m,\ \ l, m = 1, \dots, k$, and

$$d_{ll} = E \sum_{\substack{n=1 \\ n \neq l}}^{k} (\Delta W_i^n)^2 (\Delta W_i^l)^2 + E \, (\Delta W_i^l)^4 = E \sum_{\substack{n=1 \\ n \neq l}}^{k} (\Delta W_i^n)^2 (\Delta W_i^l)^2 + E(\Delta W_i^l)^4$$

$$= (k-1) E\Delta_i^2 + 3E\Delta_i^2 = (k+2) E\Delta_i^2.$$

Hence

$$E(\Delta W_i \Delta W_i^T)^2 = (k+2) E\Delta_i^2 \, I_k, \quad \forall i = 1, \dots, N,$$
$$E\Delta W_i W_N^T = E\Delta W_i^2 I_k = E\Delta_i I_k,$$
$$EW_N W_N^T = \sum_{i=1}^{N} E\Delta W_i \Delta W_i^T I_k = \sum_{i=1}^{N} E\Delta_i I_k = Et_N I_k.$$

Proof of Theorem 4. It is easy to establish that the estimator $\hat{B}_N$ is unbiased:

$$E\hat{B}_N = \frac{1}{N-1} \left(\sum_{i=1}^{N} \left(AA^T E\Delta_i + BE\frac{\Delta_i}{\Delta_i} \right) - AA^T t_N - BE\frac{t_N}{t_N} \right) = B.$$

After substituting $\Delta X_i = A\Delta_i + \overline{B}\Delta W_i$, $X_N = At_N + \overline{B}W_N$, we find:

$$E\left(\hat{B}_N^2 - B^2\right) = E\,(N-1)^{-2}\left(\left(\sum_{i=1}^{N}\Delta X_i \Delta X_i^T {\Delta_i}^{-1}\right)^2 - 2\sum_{i=1}^{N}\Delta X_i \Delta X_i^T X_N X_N^T \Delta_i {t_N}^{-1} + \left(X_N X_N^T\right)^2 t_N^{-2}\right) - B^2$$

$$= (N-1)^{-2}\sum_{i,j=1}^{N} E\Big(\left(AA^T\right)^2 \Delta_i\Delta_j + \overline{B}\Delta W_i \Delta W_i^T \overline{B}^T AA^T \Delta_j {\Delta_i}^{-1}$$
$$+ AA^T\overline{B}\Delta W_j \Delta W_j^T \overline{B}^T \Delta_i {\Delta_j}^{-1} + \overline{B}\Delta W_i A \overline{B}\Delta W_j A$$
$$+ A\Delta W_i^T \overline{B}^T \overline{B}\Delta W_j A^T + \overline{B}\Delta W_i A^T A \Delta W_j^T \overline{B}^T + A\Delta W_i^T \overline{B}^T A \Delta W_j^T \overline{B}^T$$
$$+ \overline{B}\Delta W_i \Delta W_i^T \overline{B}^T \overline{B}\Delta W_j \Delta W_j^T \overline{B}^T \Delta_i^{-1}\Delta_j^{-1} - 2\Big(AA^T\overline{B}W_N W_N^T \overline{B}^T \Delta_i {t_N}^{-1}$$
$$+ \left(AA^T\right)^2 \Delta_i t_N + \overline{B}\Delta W_i \Delta W_i^T \overline{B}^T AA^T t_N {\Delta_i}^{-1} + A\Delta W_i A^T A W_N^T \overline{B}^T$$
$$+ A\Delta W_i^T B^T A W_N^T B^T + A\Delta W_i^T B^T B W_N A^T + \overline{B}\Delta W_i A^T \overline{B} W_N A^T$$
$$+ \overline{B}\Delta W_i \Delta W_i^T \overline{B}^T \overline{B} W_N W_N^T \overline{B}^T \Delta_i^{-1} t_N^{-1}\Big) + \overline{B}W_N W_N^T \overline{B}^T \overline{B} W_N W_N^T \overline{B}^T {t_N}^{-2}$$
$$+ \left(AA^T\right)^2 t_N^2 + \overline{B}W_N W_N^T \overline{B}^T AA^T + AA^T \overline{B}W_N W_N^T \overline{B}^T + AW_N^T \overline{B}^T A W_N^T \overline{B}^T$$
$$+ AW_N^T \overline{B}^T B W_N^T A t_N + \overline{B}W_N AA^T W_N^T \overline{B}^T + \overline{B}W_N A^T \overline{B} W_N A^T\Big).$$

Using the formulas from Lemma 3, we calculate that a part of similar terms are equal to zero. For example:

$$E\left(\sum_{i,j=1}^{N}\overline{B}\overline{B}^T AA^T\Delta_j - 2\sum_{i=1}^{N}\overline{B}\overline{B}^T AA^T t_N + \overline{B}\overline{B}^T AA^T t_N + \sum_{i,j=1}^{N} AA^T B \Delta_i - 2\sum_{i=1}^{N} AA^T \overline{B} W_N W_N^T \overline{B}^T \Delta_i t_N^{-1} + AA^T \overline{B}\overline{B}^T t_N\right)$$
$$= BAA^T E\left(Nt_N - 2Nt_N + t_N + Nt_N - 2t_N + t_N\right) = 0.$$

Deleting these terms, we get:

$$E\left(\hat{B}_N^2 - B^2\right)$$
$$= (N-1)^{-2} E\left(\sum_{i=1}^{N}\overline{B}^T \Delta W_i \Delta W_i^T \overline{B}^T {\Delta_i}^{-1}\sum_{j=1}^{N}\overline{B}^T \Delta W_j \Delta W_j^T \overline{B}^T {\Delta_j}^{-1}\right.$$

$$+\sum_{i=1}^{N} B^2 \Delta W_i \Delta W_i^T \Delta W_i \Delta W_i^T \Delta_i^{-2}$$

$$-\sum_{i=1}^{N} \overline{B} \Delta W_i \Delta W_i^T B \sum_{j=1}^{N} \Delta W_j \sum_{l=1}^{N} \Delta W_l \overline{B}^T \Delta_i^{-1} t_N^{-1}$$

$$\left. + \frac{B^2 (k+1) t_N^2}{t_N^2} - (N-1)^2 B^2 \right)$$

$$= (N-1)^{-2} \left(\sum_{i=1}^{N} \overline{B}^T E \Delta W_i \Delta W_i^T \overline{B}^T \Delta_i{}^{-1} \sum_{j=1}^{N} \overline{B}^T E \Delta W_j \Delta W_j^T \overline{B}^T \Delta_j{}^{-1} \right.$$

$$+ \sum_{i=1}^{N} B^2 E \Delta W_i \Delta W_i^T \Delta W_i \Delta W_i^T \Delta_i^{-2}$$

$$- E \left(\sum_{i=1}^{N} \overline{B} \Delta W_i \Delta W_i^T B \sum_{j=1}^{N} \Delta W_j \sum_{l=1}^{N} \Delta W_l \overline{B}^T \Delta_i^{-1} t_N^{-1} \right)$$

$$\left. + B^2 (k+1) - (N-1)^2 B^2 \right)$$

$$= (N-1)^{-2} \left(B^2 \left(N^2 - N\right) + N(k+2) B^2 - 2(k+2) B^2 \sum_{i=1}^{N} \Delta_i t_N{}^{-1} \right.$$

$$\left. - 2B^2 \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \Delta_j t_N{}^{-1} + B^2 - (N-1)^2 B^2 \right).$$

The means of some terms we calculate by the following reasoning. Let us denote

$$F(i,j,l) = E \left(\sum_{i=1}^{N} \overline{B} \Delta W_i \Delta W_i^T B \sum_{j=1}^{N} \Delta W_j \sum_{l=1}^{N} \Delta W_l \overline{B}^T \Delta_i t_N{}^{-1} \right).$$

Then $F(i,j,l) = 0$ in all cases when $i \neq j \neq l \neq i$.

If $j = l$, we obtain

$$F(i,j,j) = B^2 \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} E\left(\frac{\Delta W_i \Delta W_i^T}{\Delta_i} \right) E \left(\frac{\Delta W_j \Delta W_j^T}{t_N} \right) + F(i,i,i)$$

$$= B^2 \sum_{i=1}^{N} I_k \left(\frac{\Delta_i}{\Delta_i} \right) I_k \sum_{j=1, j\neq i}^{N} E \left(\frac{\Delta_i}{t_N} \right) = B^2 (N-1) + F(i,i,i),$$

where

$$F(i,i,i) = \sum_{i=1}^{N} B^2 E\left(\frac{\Delta W_i \Delta W_i^T \Delta W_i \Delta W_i^T}{\Delta_i t_N}\right) = B^2(k+2)\sum_{i=1}^{N} E\left(\frac{\Delta_i^2}{t_N \Delta_i}\right)$$
$$= B^2(k+2).$$

Hence

$$E\left(\sum_{i=1}^{N} \overline{B}\Delta W_i \Delta W_i^T B \sum_{j=1}^{N} \Delta W_j \sum_{l=1}^{N} \Delta W_l \overline{B}^T \Delta_i {t_N}^{-1}\right) = B^2(N+k+1).$$

In the same way we calculate the means of all terms.

Finally, we obtain

$$E\left(\hat{B}_N^2 - B^2\right) = \frac{k+1}{N-1} B^2.$$

Hence $\hat{B}_N$ is the consistent estimator for the unknown matrix B.

The estimator $\hat{B}_N$ can be represented as follows:

$$\hat{B}_N = (N-1)^{-1}\left(\sum_{i=1}^{N} (\Delta X_i - A\Delta_i)(\Delta X_i - A\Delta_i)^T \Delta_i^{-1}\right.$$
$$\left. - (X_N - At_N)(X_N - At_N)^T t_N^{-1}\right) = (N-1)^{-1} \sum_{i,j=1}^{N} \alpha_{ij} Y_i Y_j^T,$$

where

$$Y_i = (X_i - A\Delta_i)\Delta_i^{-1/2} \sim N(0,B),$$
$$\alpha_{ii} = 1 - \Delta_i \left(\sum_{j=1}^{N} \Delta_j\right)^{-1}, \qquad \alpha_{ij} = \sqrt{\Delta_i \Delta_j}\left(\sum_{j=1}^{N} \Delta_j\right)^{-1}.$$

The random variables Y_i and Y_j are independent and

$$\sum_{i=1}^{N} \alpha_{ii} = N-1, \qquad \sum_{i,j=1}^{N} \alpha_{ij}^2 = N-1.$$

There exists an orthogonal transformation $Y = CZ$ such that

$$\hat{B}_N = (N-1)^{-1} \sum_{i=1}^{N} Z_i Z_i^T,$$

which is k-dimensional Wishard distribution with $N-1$ degrees of freedom.

4. COMMENTS

It is interesting to underline the next facts.

At first, the estimator $\hat{A}_N$ depends only on the last observation, the same as the continuous time sampling, and in the case when determined moments of observation are used. It is interesting to compare the estimators given by different sampling schemes. For example, the point process t_i, $i = 1, \ldots, N$, can be Poisson's, geometric, uniform. Results of this kind can be found in [10] and [11].

The second fact is that the variance of $\hat{B}_N$ is independent of the distribution of the random point process $t_1, t_2, \ldots, t_N, \ldots$ and tends to zero as $O(N^{-1})$ by $N \to \infty$. The proved formula (7) for $k = 1$ is given in [8], i.e. the obtained results generalize the one-dimensional case.

The third fact is that the used sampling scheme is natural. We add the $(N+1)$-th observation to the first N observations and do not need new $N+1$ observations. We established good properties of the estimations without the condition $\max\limits_{1 \le i \le N} \Delta_i \to 0$ when $N \to \infty$, as in the other sampling schemes.

REFERENCES

1. Gichman, I., A. Skorochod. Stochastic Differential Equations. Naukova Dumka, Kiev, 1982.
2. Liptzer, R., A. Shiryayev. Statistics of Random Processes. Nauka, Moskow, 1974.
3. Basawa, B., B. L. M. Prakassa Rao. Statistical Inference for Stochastic Processes. Academic Press, London, 1980.
4. Beutler, P. Alias Free Randomly Time Sampling of Stochastic Processes. IEEE Trans. Inform. Theory, IT – 16, 1970, 147–154.
5. Genon-Catalot, J. Jacod. Estimation of the diffusion Coefficient for Diffusion Processes: Random Sampling. *Scand. J. of Stat.*, **21**, 2, 1994, 193–221.
6. Kessler, M. Estimation of an Ergodic Diffusion from Discrete Observations. *Scand. J. of Stat.*, **24**, 2, 1997, 211–230.
7. Le Breton, A. Estimation d'un Systemme Bilineare Autonome. *Rapport Rech.*, **166**, Univ. Grenoble, 1979.
8. Stoyanov, J., D. Vladeva. Estimation of Unknown Parameters of Continuous Time Stochastic Process by Observations at Random Moments. *Compt. rend. Bulg. Acad. Sci.*, **35**, 2, 1982, 153–156.
9. Giri, N. Multivariable Statistical Inference. Acad. Press, **4**, 1977.
10. Vladeva, D. Estimation of Unknown Parameters in a Diffusion — Geometric Model. In: *Math. and Educ. in Math.*, Proc. of 13-th Spring Conf., Sunny Beach, 1984, 265–271.
11. Vladeva, D. Parameter Estimation of k-Dimensional Wiener Process by Random Observations. In: *Conf. of Appl. Math.*, Tech. Univ. of Sofia, Sozopol, 1999, 154–156.
12. Vladeva, D. Maximum Likelihood Estimation of k-Dimensional Diffusion-Discrete Model. *Mathematica Balkanica*, **14**, 2000 (to appear).

Received February 28, 2000
Revised May 5, 2000

E-mail: `vladeva@mail.vvtu.bg`