On sequential parameter estimation for some linear stochastic differential equations with time delay *

Uwe Küchler*
Humboldt University, Berlin
Institute of Mathematics

Vjatscheslav A. Vasil'iev**
Tomsk State University,
Dept. of Applied Mathematics
and Cybernetics

November 3, 1998

Abstract

We consider the parameter estimation problem for the scalar diffusion type process described by the stochastic equation with time delay

$$dX(t) = \sum_{i=0}^{m} \vartheta_i X(t - r_i)dt + dW(t).$$

The asymptotic behavior of the classical maximum likelihood estimator (MLE) very depends on the true values of parameter $\vartheta = (\vartheta_0, \vartheta_1, \ldots, \vartheta_m)'$.
Here we construct a sequential MLE with preassigned least square accuracy for the so-called stationary and the periodic cases of the solution $X(\cdot)$. The limit behaviour of the duration of the procedure with given accuracy is obtained.

Keywords: stochastic differential equations; time delay; maximum likelihood estimator; sequential analysis; least square accuracy.

* This work was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 373 "Quantifikation und Simulation ökonomischer Prozesse", Berlin, Germany

1 Introduction

Assume $(W(t), \mathcal{F}_t, t \geq 0)$ is a realvalued Wiener process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t,\ t \geq 0), P)$ and $(X(t), t \geq -r)$ satisfies the following differential equation with time delay

---

* Unter den Linden 6, D-10099 Berlin, Germany

** Lenina 36, 634050 Tomsk, Russia.

$$\left.\begin{array}{l} dX(t) = \sum\limits_{i=0}^{m} \vartheta_i X(t-r_i)dt + dW(t),\ t \geq 0, \\ X(s) = X_0(s), s \in [-r, 0]. \end{array}\right\} \tag{1}$$

The parameters $r_i,\ \vartheta_i,\ i = 0, \ldots, m$ are real numbers with $0 = r_0 < r_1 < \ldots < r_m =: r$, if $m \geq 1$ and $r_0 = r = 0$ if $m = 0$. The initial process $(X_0(s),\ s \in [-r, 0])$ is supposed to be cadlag and all $X_0(s),\ s \in [-r, 0]$ are assumed to be $\mathcal{F}_0-$measurable. Moreover assume that

$$E \int_{-1}^{0} X_0^2(s)ds < \infty.$$

The equation (1) is a special case of so-called affine stochastic differential equation studied in detail e.g. in [Mo/Sch] and [Mo]. In particular it holds, that (1) has a uniquely determined solution $(X(t), t \geq -r)$ having the representation

$$\left.\begin{array}{lll} X(t) = & \sum\limits_{j=0}^{m} \vartheta_j \int\limits_{-r_j}^{0} x_0(t-s-r_j)X_0(s)ds + & \\ & +x_0(t)X_0(0) + \int\limits_{0}^{t} x_0(t-s)dW(s), & t > 0 \\ X(t) = & X_0(t), & t \in [-r, 0] \end{array}\right\} \tag{2}$$

and satisfying $E \int_0^T X^2(s)ds < \infty$ for every T with $0 < T < \infty$. Here the function $x_0(\cdot)$ denotes the fundamental solution of the corresponding to (1) linear deterministic equation

$$x_0(t) = 1 + \sum_{j=0}^{m} \int_{0}^{t} \vartheta_j x_0(t - r_j), \qquad t \geq 0, \tag{3}$$

$$x_0(s) = 0,\ s \in [-r, 0);\ x_0(0) = 1.$$

(see [Ha/Ve] for details on (3)).
Fix a subject Θ of R^{m+1} and assume the vector $\vartheta = (\vartheta_0, \vartheta_1, \ldots, \vartheta_m)' \in \Theta$ is unknown and has to be estimated based on the observation $(X(t))$. The delay times r_i are supposed to be known.
The measures $P_\vartheta, \vartheta \in R^{m+1}$ generated by the solutions of (1) form an exponential family in the sense of [Ku/So]. Thus, one possibility to estimate ϑ is to use the maximum likelihood method. The corresponding log-likelihood-function is given by

$$\ell_t(\vartheta) = \vartheta' \Phi(t) - \frac{1}{2} \vartheta' G(t) \vartheta \quad , \quad \vartheta \in \Theta,\ t > 0, \tag{4}$$

where

$$\Phi(t) = (\int_0^t X(s-r_i)dX(s),\ i=0,\dots,m)',$$

and

$$G(t) = (\int_0^t X(s-r_i)X(s-r_j)ds,\ i,j=0,...,m)$$

denotes the Fisher information matrix (for details see [Gu/Ku] and [Ku/So]). Another method is provided by sequential estimation. Sequential estimation of one-dimensional parameters in exponential families of processes have been studied e.g. in [Li/Sh] and [Nov], see also [Ku/So] (1997), Chapter 10. The more-dimensional parameter case cannot be treated in the same way. Indeed, the construction of the stopping time for the observation in these papers very uses the one-dimensionality of the Fisher information. For processes arising from linear stochastic differential equations without time delay having more-dimensional parameters, sequential methods have been developed in [Ko/Pe] (1985), (1987), (1992).
Here we shall extend these results to equations of the type (1). We shall construct for every $\varepsilon > 0$ a sequential procedure ϑ^* to estimate ϑ with $\varepsilon-$accuracy in the square mean sense, i.e. with $E[\vartheta^* - \vartheta]^2 \le \varepsilon$.
The method used below is a two step construction of a random time, where the first step uses the trace of the Fisher information matrix and follows the line of the one-dimensional case mentioned above.
A generalization of the sequential estimators, constructed in the sequel, to differential equations of the type (1) but based on noisy observations, will be presented in a subsequent paper.

2 Results

Consider the process $(X(t), t \ge -r)$ described by equation (1) above.
Throughout this paper we suppose that the following assumption holds.

Assumption (A) : *For every $\vartheta \in \Theta$ there exist a (deterministic scalar) positive increasing function $\varphi(\cdot)$ on $[0,\infty)$ with $\lim_{T\to\infty} \varphi(T) = \infty$ and a possibly random $(m+1)\times(m+1)-$matrix function $I_\infty(T)$, $T \in [0,\infty)$, being continuous periodic with period $\Delta \ge 0$ ($\Delta = 0$ means $I_\infty(T) \equiv I_\infty(0)$ and positive definite for every T. Moreover, it holds*

$$\lim_{T\to\infty} \left| \frac{G(T)}{\varphi(T)} - I_\infty(T) \right| = 0 \quad a.s. \tag{5}$$

The assumption (A) is satisfied under further restrictions on Θ only. For example, if $m = 1$ then it holds exactly in the following two cases.
Consider the set Λ of all complex roots of the socalled characteristic equation

$$\lambda - \vartheta_0 - \vartheta_1 e^{-\lambda r} = 0$$

and put $v_0 = v_0(\vartheta) = \max\{Re\lambda | \lambda \in \Lambda\}$. It can be easily shown that $v_0 < \infty$. Then (A) holds for $\Theta = \{\vartheta \in R^2 \mid v_0(\vartheta) < 0$ or $[v_0(\vartheta) > 0$ and $v_0(\vartheta) \notin \Lambda]\}$, see [Gu/Ku] (1998) for details. If $v_0 < 0$ then the equation (1) admits a stationary solution and every solution tends to it in distribution, moreover we have $\Delta = 0$, we call this case the "stationary case". If $v_0 > 0$ and $v_0 \notin \Lambda$ the equality (5) is valid with some $\Delta > 0$. We denote this case as the "periodic" one.
A similar picture appears in the classical moredimensional linear equation

$$dX(t) = AX(t)dt + dW(t),\ t \geq 0, \quad X(0) = X_0$$

with the Fisher information matrix

$$\Gamma(T) = \int_0^T X(T)X'(T)dt.$$

Here $W(\cdot)$ is a $d-$dimensional standard Wiener process and A a given $d \times d$ matrix. Let $\lambda_{\max}$ and $\lambda_{\min}$ be eigenvalues of A having the maximal and minimal absolute value under all of eigenvalues of e^A, respectively. It is well known that the limiting matrix $\lim_{T\to\infty} T^{-1}\Gamma(T)$ exist and is a positive definite deterministic matrix in the stable case $(Re\lambda_{\max} < 0)$ and $\Gamma(T)$ increase exponentially in the unstable case $(Re\lambda_{\min} > 0)$. Note that for stable case the sequential parameter estimation problem of matrix A was considered in [Ko/Pe] (1985), for the scalar model in [Nov] and [Li/Sh], for unstable case in [Ko/Pe] (1987) and in mixed case $(Re\lambda_{\max} > 0,\ Re\lambda_{\min} < 0$ and $\lambda + \mu \neq 0$ for all eigenvalues λ, μ of $A)$ in [Ko/Pe] (1992).
The sequential estimation problem for the matrix A in the stable case by noisy observations was studed in [Va/Ko] (1987) and [Va/Ko] (1990).
Let us return to the study of (1) and let Assumption (A) be true.
To estimate ϑ with pressigned accuracy $\varepsilon > 0$ we shall start with the maximum likelihood estimator of ϑ for the given lenght T of observation defined by the equality

$$\hat{\vartheta}(T) = G^{-1}(T)\Phi(T), T > 0. \tag{6}$$

From (1) and (6) we find the deviation of the estimator $\hat{\vartheta}(T)$ from ϑ :

$$\hat{\vartheta}(T) - \vartheta = G^{-1}(T)\zeta(T) \tag{7}$$

with

$$\zeta(T) = \int_0^T Z(t)dW(t),\ Z(t) = (X(t), X(t-r_1), \ldots, X(t-r_m))'.$$

Now we make a time substitution which enables us to control the second moments of the noise ζ.
Fix an arbitrary increasing sequence $(c_n)_{n\geq 1}$ of reals tending to infinity. Let us define the sequence of $(\mathcal{F}_t)-$stopping times $(\tau_\varepsilon(n), n \geq 1)$ as follows

$$\tau_\varepsilon(n) = \inf\{T > 0 : trG(T) = \varepsilon^{-1}c_n\}, \tag{8}$$

These moments are finite a.s. due to the condition (5).
One can easily verify that for any $\varepsilon > 0$ the sequence $(\zeta(\tau_\varepsilon(n)), n \geq 1)$ satisfies the equalities

$$E_\vartheta \|\zeta(\tau_\varepsilon(n))\|^2 = \varepsilon^{-1} c_n, \ n \geq 1. \tag{9}$$

(Throughout this paper $||\cdot||$ denotes the Euclidian norm.)
The equalities (9) suggest that the estimation of the parameter ϑ should be performed at the moments $\tau_\varepsilon(n)$:

$$\vartheta_n(\varepsilon) = \hat{\vartheta}(\tau_\varepsilon(n)), \ n \geq 1. \tag{10}$$

According to (7) in order to obtain the estimates with fixed least square deviation now one should control the behaviour of the sequence of random matrices $(G^{-1}(\tau_\varepsilon(n)), \ n \geq 1)$. It can be achieved by conducting the observations up to the moment $\tau_\varepsilon(n)$ with a specially choosen number n. Let

$$\nu_\varepsilon = \inf\{N \geq 1 : S_N(\varepsilon) \geq \varrho\}, \tag{11}$$

where $S_N(\varepsilon) = \sum\limits_{n=1}^{N} \beta_n^2(\varepsilon), \ \beta_n^2(\varepsilon) = (\varepsilon \cdot c_n^{-1})^2 . \|G^{-1}(\tau_\varepsilon(n))\|^{-2}, \ \varrho = \sum\limits_{n\geq 1} 1/c_n.$
The sequential plan $(T(\varepsilon), \ \vartheta_\varepsilon^*)$ of estimation of the vector ϑ will be defined by

$$T(\varepsilon) = \tau_\varepsilon(\nu_\varepsilon) \ , \ \vartheta_\varepsilon^* = S_{\nu_\varepsilon}^{-1} \sum_{n=1}^{\nu_\varepsilon} \beta_n^2(\varepsilon) \cdot \vartheta_n(\varepsilon). \tag{12}$$

Obviously, ν_ε is a $(\mathcal{F}_{\tau_n(\varepsilon)})-$stopping time, and therefore, by construction, $T(\varepsilon)$ turns out to be an $(\mathcal{F}_t)-$stopping time.
In such a way the sequential estimate ϑ_ε^* is a random weighted mean of the maximum likelihood estimates, calculated at the stopping times $\tau_\varepsilon(n), n \geq 1$.
The following theorem summarizes the main result.

Theorem 1. *Assume that Assumption A holds. Then for any $\varepsilon > 0$ and any $\vartheta \in \Theta$ the sequential estimation plan (12) of ϑ possesses the properties:*

$1^\circ . T(\varepsilon) < \infty \quad P_\vartheta - a.s.,$

$2^\circ . E_\vartheta \|\vartheta_\varepsilon^* - \vartheta\|^2 \leq \varepsilon$

and the following inequalities hold $P_\vartheta - a.s.$

$3^\circ . 0 < \varliminf\limits_{\varepsilon\to 0} \varepsilon \cdot \varphi(T(\varepsilon)) \leq \varlimsup\limits_{\varepsilon\to 0} \varepsilon \cdot \varphi(T(\varepsilon)) < \infty.$

Proof. 1°. Let us verify the finiteness of $T(\varepsilon) = \tau_\varepsilon(\nu_\varepsilon)$. While the moments $\tau_\varepsilon(n)$ are finite for all $n \geq 1$, it suffices to establish the finiteness of the moment ν_ε. Making use of the definition (9) of $\tau_\varepsilon(n)$ and the condition (5) we have

$$\lim_{n\to\infty}\left|\frac{\varepsilon^{-1}\cdot c_n}{\varphi(\tau_\varepsilon(n))}-trI_\infty(\tau_\varepsilon(n))\right|=0 \quad a.s. \tag{13}$$

and as follows by the definition of $\beta_n^2(\varepsilon)$

$$\lim_{n\to\infty}|\beta_n^2(\varepsilon)-\beta^2(\tau_\varepsilon(n))|=0 \quad a.s., \tag{14}$$

where

$$\beta^2(u)=[trI_\infty(u)\cdot\|I_\infty^{-1}(u)\|]^{-2}. \tag{15}$$

Note that by the conditions on the matrix function $I_\infty(u)$ we have

$$\inf_{u\in R^1}\beta^2(u)>0.$$

Then $\sum\limits_{n=1}^{\infty}\beta_n^2(\varepsilon)=\infty$ a.s. and for all $\varepsilon>0$ the moments ν_ε and $T(\varepsilon)$ are finite a.s.

2°. Now we estimate the mean square deviation of ϑ_ε^*. From (7), (9), (12) and by definitions of ν_ε, β_n and ϱ it follows that

$$\begin{aligned}
&E_\vartheta\|\vartheta_\varepsilon^*-\vartheta\|^2=E_\vartheta S_{\nu_\varepsilon}^{-2}\|\sum_{n=1}^{\nu_\varepsilon}\beta_n^2(\varepsilon)(\vartheta_n(\varepsilon)-\vartheta)\|^2\le\\
&\le E_\vartheta S_{\nu_\varepsilon}^{-1}\sum_{n\ge1}\beta_n^2(\varepsilon)\|\vartheta_n(\varepsilon)-\vartheta\|^2\le\varrho^{-1}\sum_{n\ge1}E_\vartheta\beta_n^2(\varepsilon)\cdot\\
&\cdot\|G^{-1}(\tau_\varepsilon(n))\|^2\cdot\|\zeta(\tau_\varepsilon(n))\|^2=\varepsilon^2\varrho^{-1}\sum_{n\ge1}\frac{1}{c_n^2}E_\vartheta\|\zeta(\tau_\varepsilon(n))\|^2=\\
&=\varepsilon\varrho^{-1}\sum_{n\ge1}\frac{1}{c_n}=\varepsilon.
\end{aligned}$$

For the first inequality we used the Cauchy-Bunjakovsky inequality.

3°. In order to establish the limiting relationships for $T(\varepsilon)$ we note that as in (14) for all $n\ge1$ it holds

$$\lim_{\varepsilon\to0}|\beta_n^2(\varepsilon)-\beta^2(\tau_\varepsilon(n))\,|=0 \quad \text{a.s.} \tag{16}$$

According to (16) and by the definition of the moment ν_ε for small but positive ε we have the inequalities

$$\nu'\le\nu_\varepsilon\le\nu'' \quad \text{a.s.} \tag{17}$$

with

$$\begin{aligned}
\nu' &= \inf\{N\ge1:\ N>\varrho[\sup_{u\in[0,\Delta)}\beta^2(u)]^{-1}\}-1,\\
\nu'' &= \inf\{N\ge1:\ N>\varrho[\inf_{u\in[0,\Delta)}\beta^2(u)]^{-1}\}.
\end{aligned}$$

Similar (13) we can obtain

$$\lim_{\varepsilon\to 0}\left|\frac{\varepsilon^{-1}c_{\nu_\varepsilon}}{\varphi(T(\varepsilon))}-trI_\infty(T(\varepsilon))\right|=0 \quad \text{a.s.} \tag{18}$$

From (17) and (18) follows the assertion 3° of the Theorem 1

$$0<\mu'\le\lim_{\varepsilon\to 0}\varepsilon\cdot\varphi(T(\varepsilon))\le\mu''<\infty,$$

where

$$\mu'=c_{\nu'}\cdot[\sup_{u\in[0,\Delta)}I_\infty(u)]^{-1},\quad \mu''=c_{\nu''}\cdot[\inf_{u\in[0,\Delta)}I_\infty(u)]^{-1}. \tag{19}$$

Theorem 1 is proved.

3 Example

Consider system (1) with $m=1,\ r_1=1$

$$\left.\begin{array}{l} dX(t)=\vartheta_0 X(t)dt+\vartheta_1 X(t-1)dt+dW(t),\ t\ge 0,\\ X(s)=X_0(s), s\in[-1,0]. \end{array}\right\} \tag{20}$$

Assume for reasons of of citation, that X_0 is continuous.
The sequential plan $(T(\varepsilon),\vartheta_\varepsilon^*)$ of estimation $\vartheta=(\vartheta_0,\vartheta_1)'$ will be defined as (12) with the Fisher information matrix

$$G(T)=\begin{pmatrix} \int_0^T X^2(t)dt & \int_0^T X(t)X(t-1)dt \\ \int_0^T X(t)X(t-1)dt & \int_0^T X^2(t-1)dt \end{pmatrix} \tag{21}$$

[Ku/So].
We can reformulated Theorem 1 for this cases as follows.

<u>Theorem 2.</u> *Let the parameters ϑ_0 and ϑ_1 in (20) such that we have the stationary or periodic case (for the notation see chapter one). Then the sequential plan (12) of estimation $\vartheta=(\vartheta_0,\vartheta_1)'$ possesses the properties:*

$1^o. T(\varepsilon)<\infty\ \ P_\vartheta-$a.s.,

$2^o. E_\vartheta\|\vartheta_\varepsilon^*-\vartheta\|^2\le\varepsilon.$

3^o. *Besides the following limit inequalities*

$$0<\underline{\lim}_{\varepsilon\to 0}\gamma(\varepsilon)T(\varepsilon)\le\overline{\lim}_{\varepsilon\to 0}\gamma(\varepsilon)T(\varepsilon)<\infty\ \ P_\vartheta-\text{a.s.} \tag{22}$$

are fulfilled, where $\gamma(\varepsilon) = \varepsilon$ in the stationary case and $\gamma(\varepsilon) = (\ln\varepsilon^{-1})^{-1}$ in the periodic case. Moreover, in the periodic case the limiting inequality

$$\lim_{\varepsilon\to 0} |T(\varepsilon) - \frac{1}{2v_0}\ln\varepsilon^{-1}| < \infty \quad a.s. \tag{23}$$

holds.

Proof of $1^o - 2^o$: According to Theorem 1 the assertions 1^o and 2^o of Theorem 2 will be proved if the matrix $G(T)$ (21) satisfies the condition (5).
Now we establish the auxiliary equalities

$$\lim_{T\to\infty} T^{-1}G(T) = I_\infty \quad \text{a.s.} \tag{24}$$

for the stationary case and

$$\lim_{T\to\infty} |\, e^{-2v_0 T} G(T) - I_\infty(T)\, | = 0 \quad \text{a.s.} \tag{25}$$

for the periodic case, $v_0 > 0$.
Here

$$I_\infty = \begin{pmatrix} \int_0^\infty x_0^2(t)dt & \int_0^\infty x_0(t)x_0(t+1)dt \\ \int_0^\infty x_0(t)x_0(t+1)dt & \int_0^\infty x_0^2(t)dt \end{pmatrix}$$

and $I_\infty(T)$ is a periodic matrix

$$I_\infty(T) = \begin{pmatrix} g_{11}(T) & g_{12}(T) \\ g_{12}(T) & g_{22}(T) \end{pmatrix},$$

$$g_{ij}(T) \;=\; \int_0^\infty e^{-2v_0 t} U_i(T-t)U_j(T-t)dt, \; i,j = 0,2,$$

$$U_i(t) \;=\; \phi_i(t)X_0(0) + b\int_{-1}^0 \phi_i(t-s-1)e^{-v_0(s+1)}X_0(s)ds + \int_0^\infty \phi_i(t-s)e^{-v_0 s}dW(s),$$

$$\phi_i(t) \;=\; A_i\cdot\cos(\xi_0 t) + B_i\cdot\sin(\xi_o t), \quad i = 0,2,$$

$$A_0 \;=\; \frac{2(v_0-a+1)}{(v_0-a+1)^2+\xi_0^2}\;,\; B_0 = \frac{2\xi_0}{(v_0-a+1)^2+\xi_0^2},$$

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \;=\; e^{-v_0}\begin{pmatrix} \cos\xi_0 & -\sin\xi_o \\ \sin\xi_0 & \cos\xi_0 \end{pmatrix}\begin{pmatrix} A_0 \\ B_0 \end{pmatrix},$$

$$\xi_o = \arg\{Im\lambda \mid \lambda \in \Lambda,\ Re\lambda = v_0,\ Im\lambda > 0\}.$$

Taking into account the representation

$$X(t) = x_0(t)X_0(0) + b\int_{-1}^{0} x_0(t-s-1)X_0(s)ds + \int_0^t x_0(t-s)dW(s) \qquad (26)$$

for the solution $(X(t), t \geq -1)$ of (21) [Gu/Ku], [Ku/So] and the fact that in the stationary case

$$\int_0^{\infty} x_0^2(t)dt < \infty,$$

we can see that

$$\lim_{t\to\infty} |X(t) - Z(t)| = 0 \quad a.s.,$$

where $Z(t) = \int_{-\infty}^{t} x_0(t-s)dW(s)$ is a stationary process with the correlation matrix I_∞, which is ergodic [Gu/Ku], [Ku/So]. Then the equality (24) hold.
In the periodic case according to [Gu/Ku]

$$x_0(t) = \phi_0(t)\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t})$$

and

$$x_0(t-1) = \phi_2(t)\mathrm{e}^{v_o t} + o(\mathrm{e}^{\gamma t})$$

for some γ with $\gamma < v_0$. Similar to Lemma 4.8 in [Gu/Ku] we can prove the equality

$$\lim_{t\to\infty} |\,\mathrm{e}^{-\vartheta_0 t}X(t) - U_0(t)\,| = 0 \quad \text{a.s.}$$

From here we have

$$\lim_{t\to\infty} |\,\mathrm{e}^{-2v_0 T}\int_0^T X^2(t)dt - \int_0^\infty \mathrm{e}^{-2v_0 t}U^2(T-t)dt\,| =$$

$$= \lim_{T\to\infty} |\int_0^T \mathrm{e}^{-2v_0(T-t)}[\mathrm{e}^{-2v_0 t}X^2(t) - U^2(t)]dt + \int_0^T \mathrm{e}^{-2v_0(T-t)}U^2(t)dt -$$

$$- \int_0^\infty \mathrm{e}^{-2v_0 t}U^2(T-t)dt\,| = \lim_{T\to\infty}\int_T^\infty \mathrm{e}^{-2v_0 t}U^2(T-t)dt = 0 \quad \text{a.s.}$$

The other equations in (25) may be proved analogously. Note that according to [Gu/Ku] $I_\infty(u) > 0$ for $u \in [0,\Delta)$ and the matrix function $I_\infty(u)$ is continuous on R^1. It follows $I_\infty(u) > 0$ for $u \in [0,\Delta]$.Then (24), (25) and the conditions (5) for the matrix $G(T)$ defined by (21) are established.

3^o. In order to obtain the exact limiting relationships for $T(\varepsilon)$ in the stationary case it suffices to note that by the definition of stopping times $\tau_\varepsilon(n)$ and (24) we get for all $n \geq 1$

$$\lim_{\varepsilon\to 0} \varepsilon \cdot \tau_\varepsilon(n) = c_n \cdot (trI_\infty)^{-1} > 0 \text{ a.s.}, \quad (27)$$
$$\lim_{\varepsilon\to 0} \varepsilon \cdot G(\tau_\varepsilon(n)) = c_n \cdot (trI_\infty)^{-1} \cdot I_\infty > 0 \text{ a.s.}$$

and as follows

$$\lim_{\varepsilon\to 0} \beta_n^2(\varepsilon) = (trI_\infty \cdot \|I_\infty^{-1}\|)^{-2} > 0 \quad \text{a.s.} \quad (28)$$

Take into account that in this case $\varphi(T) = T$, from (8), (11), (24) and (28) we have

$$\mu_1 \leq \underline{\lim}_{\varepsilon\to 0} \varepsilon \cdot T(\varepsilon) \leq \overline{\lim}_{\varepsilon\to 0} \varepsilon \cdot T(\varepsilon) \leq \mu_2 \quad (29)$$

with

$$\mu_1 = c_{\nu-1} \cdot (trI_\infty)^{-1}, \mu_2 = c_\nu (trI_\infty)^{-1}, \quad (30)$$

$$\nu = \inf\{N \geq 1 \, : N > \varrho(trI_\infty \cdot \|I_\infty^{-1}\|)^2\}.$$

Then the inequalities (22) for the stationary case hold.
Now we establish the assertion 3^o of Theorem 2 for the periodic case.
By the definition (8) and according to (25) we have

$$\lim_{\varepsilon\to 0} |\varepsilon^{-1} c_{\nu_\varepsilon} \cdot \mathrm{e}^{-2v_0 T(\varepsilon)} - trI_\infty(T(\varepsilon))| = 0 \quad \text{a.s.} \quad (31)$$

Since $\inf_u trI_\infty(u) > 0$ we can rewrite (31) in the form

$$\lim_{\varepsilon\to 0} \left[T(\varepsilon) - \frac{1}{2v_o} \ln \varepsilon^{-1} - \frac{1}{2v_o} \ln c_{\nu_\varepsilon} + \frac{1}{2v_0} \ln trI_\infty(T(\varepsilon)) \right] = 0 \quad \text{a.s.}$$

From here and (17) we can obtain the relationships

$$\lim_{\varepsilon\to 0} (\ln \varepsilon^{-1})^{-1} \cdot T(\varepsilon) = \frac{1}{2v_0} \quad \text{a.s.}$$

and

$$\tilde{\mu}_1 \leq \underline{\lim}_{\varepsilon\to 0} \left[T(\varepsilon) - \frac{1}{2v_0} \ln \varepsilon^{-1} \right] \leq \overline{\lim}_{\varepsilon\to 0} \left[T(\varepsilon) - \frac{1}{2v_0} \ln \varepsilon^{-1} \right] \leq \tilde{\mu}_2$$

with

$$\tilde{\mu}_1 = \frac{1}{2v_0} \ln c_{\nu'} (\sup_{u\in[0,\Delta)} trI_\infty(u))^{-1},$$
$$\tilde{\mu}_2 = \frac{1}{2v_0} \ln c_{\nu''} (\inf_{u\in[0,\Delta)} trI_\infty(u))^{-1}.$$

The assertion 3^o of Therem 2 is established. Theorem 2 is proved.

From Theorem 2 it follows that the duration $T(\varepsilon)$ of the sequential estimation has a nonrandom lower and upper bounds $\gamma^{-1}(\varepsilon)\tilde{\mu}_1$ and $\gamma^{-1}(\varepsilon)\tilde{\mu}_2$ respectively asymptotically. These bounds have the same increasing rate with $\varepsilon \to 0$. From assertions 2 and 3 of Theorem 2 follows that the convergence rate of the mean square deviation of the sequential estimator ϑ_ε^* corresponds with the rate of convergence of the MLE in stationary and periodic cases [Gu/Ku].
According to the inequalities (29) the duration of observations $T(\varepsilon)$ in stationary case is approximately not great than $\varepsilon^{-1}\mu_2$ with μ_2 defined by (30) when ε is small. Note that in this case one can obtain the following limiting equalities

$$\lim_{\varepsilon\to 0} \nu_\varepsilon = \nu \text{ a.s.} \tag{32}$$

and

$$\lim_{\varepsilon\to 0} \varepsilon T(\varepsilon) = \mu_2 \text{ a.s.}$$

Here ν is defined by (30). To obtain (32) we change the definition of ν_ε a little bit. Replace the magnitudes $\beta_n^{-2}(\varepsilon)$ in the definition of ν_ε in (11) by the nearest integer from above and choose (c_n) in such a way that the constant ϱ in (11) is irrational. In this case, the limit $\lim\limits_{\varepsilon\to 0} S_N(\varepsilon)$ is stricly greate thn ϱ, and this implies (32).
From (32) it follows that by small ε the moments $\nu_\varepsilon = \nu$ a.s. and by the property (28) it is obvious that the sequential estimate ϑ_ε^* may be represented in stationary case as the mean of finite numbers ν of maximum likelihood estimates $\hat{\vartheta}$ which are calculated at the moments $\tau_\varepsilon(n)$:

$$\vartheta_\varepsilon^* \sim \frac{1}{\nu}\sum_{n=1}^{\nu} \hat{\vartheta}(\tau_\varepsilon(n)). \tag{33}$$

The number ν may be asymptotically estimated with help of the property (24) and by the definition (30) of the moment ν.
It should be pointed out also that by known bound α for $\inf\limits_{u\in[0,\Delta)} \beta^2(u) \geq \alpha > 0$ with $\beta^2(u)$ defined by (15), according to (18) we obtain

$$\nu_\varepsilon \leq \inf\{N \geq 1 : N > \varrho\alpha^{-1}\} = 1$$

by small ε if the sequence (c_n) is such that $\varrho < \alpha$. Then for the sequential estimate ϑ_ε^* defined by (12) for small ε we have

$$\vartheta_\varepsilon^* = \hat{\vartheta}(\tau_1(\varepsilon)) \text{ a.s.}$$

<u>Remark.</u> From Theorem 2 we can see that the sequential estimators ϑ_ε^* converge to the true value ϑ in mean square as $\varepsilon \to 0$ in stationary and periodic cases. Moreover, for any sequence $(\varepsilon_n, n \geq 1)$ of positive integers such that $\sum\limits_{n\geq 1} \varepsilon_n < \infty$, we can define the sequence of estimators $(\tilde{\vartheta}_n, n \geq 1)$, $\tilde{\vartheta}_n = \vartheta_{\varepsilon_n}^*$, $n \geq 1$. Then the sequence $(\tilde{\vartheta}_n)$ of estimators for ϑ is strongly consistent. It follows from the assertion 2 of Theorem 2 and the Borel - Cantelli lemma.

References

[Gu/Ku] Gushchin, A.A. and Küchler, U. Asymptotic Inference for a linear stochastic differential equation with time delay, to appear in Bernoulli.

[Ha/Ve] Hale, J.K. and Verduyn Lunel, S.M. (1993) *Introduction to functional-differential equations,* New York; Springer-Verlag.

[Ko/Pe] Konev, V.V. and Pergamenshchikov, S.M. (1985) Sequential estimation of the parameters of diffusion processes. *Problems of Inform. Trans.*, **21**, 1, 48-62.

[Ko/Pe] Konev, V.V. and Pergamenshchikov, S.M. (1987) Sequential estimation of the parameters of unstable dynamical systems in continuous time. *Math. Stat. and Appl.*, Publishing House of Tomsk University, Tomsk, **11**, 85-94.

[Ko/Pe] Konev, V.V. and Pergamenshchikov, S.M. (1992) Sequential estimation of the parameters of linear unstable stochastic systems with guaranteed accuracy. *Problems of Inform. Trans.*, **28**, 4, 35-48.

[Ku/Kut] Küchler and Kutoyants, Yu. A. (1998) Delay estimation for stationary diffustion-type process. Discussion Paper 47 of the SFB 373, Humboldt-University of Berlin, 1998.

[Ku/Me] Küchler, U. and Mensch, B. (1991) Langevins stochastic differential equations extended by a time-delayed term. *Stochastics and Stochastic Reports*, **40**, 23-42.

[Ku/So] Küchler, U. and Sorensen, M. (1997) *Exponential Families of Stochastic Processes*, New York, Heidelberg; Springer Verlag.

[Li/Sh] Liptzer, R.S. and Shiryaev A.N. (1977) *Statistics of Random Processes*, Vol 1,2. New York, Heidelberg; Springer Verlag.

[Mo] Mohammed, S.E-A. (1984) *Stochastic Functional Differential Equations,* Pitman, London.

[Mo/Sch] Mohammed, S.E-A. and Scheutzow, M.K.R. (1990) Lyapunov exponents and stationary solutions for affine stochastic delay equations. *Stochastics and Stochastic Reports*, **29**, 259-283.

[Nov] Novikov, A.A. (1971) The sequential parameter estimation in the process of diffusion type. *Probab. Theory and its Appl.*, **16**, 2, 394-396.

[Va/Ko] Vasiliev, V.A. and Konev, V.V. (1987) On sequential identification of linear dynamic systems in continuous time by noisy observations. *Probl. of Contr. and Inform. Theory*, **16**, 2, 101-112.

[Va/Ko] Vasiliev, V.A. and Konev, V.V. (1990) On sequential parameter estimation of continuous dynamic systems by discrete time observations. *Probl. of Contr. and Inform. Theory*, **19**, 3, 197-207.