Automatica 46 (2010) 637–646

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

On guaranteed parameter estimation of a multiparameter linear regression process[☆]

Uwe Küchler[a,*], Vyacheslav A. Vasiliev[b]

[a] Humboldt University, Berlin, Institute of Mathematics, Unter den Linden 6, D-10099 Berlin, Germany

[b] Tomsk State University, Department of Applied Mathematics and Cybernetics, Lenina 36, 634050 Tomsk, Russia

ARTICLE INFO

Article history:
Received 17 September 2008
Received in revised form
12 November 2009
Accepted 3 December 2009
Available online 20 March 2010

Keywords:
Identification methods
System identification
Estimation theory
Statistical analysis
Delay analysis
Sequential identification

ABSTRACT

This paper presents a sequential estimation procedure for the unknown parameters of a continuous-time stochastic linear regression process. As an example, the sequential estimation problem of two dynamic parameters in stochastic linear systems with memory and in autoregressive processes is solved. The estimation procedure is based on the least squares method with weights and yields estimators with guaranteed accuracy in the sense of the L_q-norm for fixed $q \geq 2$.

The proposed procedure works in the mentioned examples for all possible values of unknown dynamic parameters on the plane R^2 for the autoregressive processes and on the plane R^2 with the exception of some lines for the linear stochastic delay equations. The asymptotic behaviour of the duration of observations is determined.

The general estimation procedure is designed for two or more parametric models. It is shown that the proposed procedure can be applied to the sequential parameter estimation problem of affine stochastic delay differential equations and autoregressive processes of an arbitrary order.

1. Introduction

In this article we consider a linear regression model of the type

$$dx(t) = \vartheta' a(t)dt + dW(t), \quad t \geq 0 \tag{1}$$

with the initial condition $x(0) = x_0$. Here we assume that $(W(t), t \geq 0)$ is an adapted one-dimensional standard Wiener process on a filtered probability space $(\Omega, F, (F_t)_{t\geq 0}, P)$, ϑ an unknown parameter from some subset Θ of R^{p+1}, $(a(t), t \geq 0)$ an observable adapted $(p+1)$-dimensional cadlag process and $x = (x(t), t \geq 0)$ solves Eq. (1). We assume that $p > 1$.

The model described includes several more concrete cases like linear stochastic differential equations (SDE's) of first or of higher-order (CARMA-processes) linear stochastic delay differential equations (SDDE's). They can be found e.g. in Brockwell (2001), Galtchouk and Konev (2001), Konev and Pergamenshchikov (1985, 1992), Küchler and Vasiliev (2001, 2003, 2005, 2006, 2008) and Liptzer and Shiryaev (1977).

In what follows we will study the problem of sequential estimating the parameter ϑ from Θ based on the observation of $(x(t), a(t))_{t\geq 0}$.

We shall construct for every $\varepsilon > 0$ and arbitrary but fixed $q \geq 2$ a sequential procedure $\vartheta(\varepsilon)$ to estimate ϑ with ε-accuracy in the sense

$$\|\vartheta(\varepsilon) - \vartheta\|_q^2 \leq \varepsilon. \tag{2}$$

Here the L_q-norm is defined as $\|\cdot\|_q = (E_\vartheta \|\cdot\|^q)^{\frac{1}{q}}$, where $\|a\| = (\sum_{i=0}^m a_i^2)^{\frac{1}{2}}$ and E_ϑ denotes the expectation under P_ϑ for $\vartheta \in \Theta$ (the number $q \geq 2$ is fixed in what follows).

Moreover, we shall determine the rate of convergence of the duration of observations $T(\varepsilon)$ to infinity and almost surely convergence of $\vartheta(\varepsilon)$ if $\varepsilon \to 0$.

The new results presented here consist in the greater generality of the conditions on $(a(t))$ than in previous papers of Küchler and Vasiliev (2001, 2003, 2005, 2006). A similar estimation problem for a more general model was investigated in Galtchouk and Konev (2001). The authors have considered the problem of sequential estimation of parameters in multivariate stochastic regression models with martingale noise and an arbitrary finite number of unknown parameters. The estimation procedure in Galtchouk and Konev (2001) is based on the least squares method with a special choice of weight matrices. The proposed procedure enables them to estimate the parameters with any prescribed mean square accuracy under appropriate conditions on the regressors $(a(t))$. Among

☆ Research was supported by RFBR - DFG 05-01-04004 Grant. The material in this paper was presented by Vyacheslav Vasiliev, 09 July 2008, IFAC World Congress, Seoul, Korea. This paper was recommended for publication in revised form by Associate Editor Johan Schoukens under the direction of Editor Torsten Söderström.

* Corresponding author. Tel.: +49 30 2093 5851; fax: +49 30 2093 5848.
E-mail address: kuechler@mathematik.hu-berlin.de (U. Küchler).

0005-1098/$ – see front matter
doi:10.1016/j.automatica.2010.01.003

conditions on the regressors there is one limiting the growth of the maximum eigenvalue of the symmetric design matrix with respect to its minimal eigenvalue. This condition is slightly stronger than those usually imposed in asymptotic investigations and it is not possible to apply this estimation procedure to continuous-time models with essentially different behaviour of the eigenvalues (if, for example, the smallest eigenvalue growth linearly and the largest one — exponentially with the observation time).

The paper (Galtchouk & Konev, 2001) also includes extended hints to earlier works of different authors on sequential estimations for parameters of both continuous as well as discrete-time processes.

The methods applied in this paper to (1) were inspired by the following basic examples for (1):

I. SDE's of autoregressive type given by

$$\mathrm{d}x_t^{(p)} = \sum_{i=0}^{p} \vartheta_i x_t^{(p-i)}\mathrm{d}t + \mathrm{d}W(t), \quad t \geq 0. \tag{3}$$

II. SDDE's given by

$$\mathrm{d}X(t) = \sum_{i=0}^{p} \vartheta_i X(t - r_i)\mathrm{d}t + \mathrm{d}W(t), \quad t \geq 0. \tag{4}$$

The sequential parameter estimation problem of the process (3) was solved in Konev and Pergamenshchikov (1992) under some additional condition on the roots of its characteristic equation (and as follows, on the corresponding parameters). Similar to Galtchouk and Konev (2001), in Konev and Pergamenshchikov (1992) obtained the sequential estimators of the parameter ϑ with given accuracy in the mean square sense.

Our paper considers the sequential parameter estimation problem of the process (3) with $p = 1$ as an example of the general estimation procedure, elaborated for linear regression model (1).

It is shown, that the presented sequential estimation procedure works for all parameters $\tilde{\Theta} = \{\vartheta \in R^2 : \vartheta_1 \neq 0\}$. As usual, the condition $\vartheta_1 \neq 0$ means the knowledge of the order ($p = 1$) of the process (3). It should be noted that the problem of sequential estimation for the case $\tilde{\Theta} \setminus \{\vartheta \in R^2 : \vartheta_0 = 0\}$ has been solved in Konev and Pergamenshchikov (1985, 1992).

The problem of sequential parameter estimation for the process (4) was considered in Küchler and Vasiliev (2001, 2003, 2005, 2006) under some additional conditions on the underlying parameters. The general estimation procedure, presented in this paper, works under the most weakest possible assumptions on the parameters. Thus it is shown, that in the case $p = 1$ in the model (4) the constructed general estimation procedure gives the possibility to solve the parameter estimation problem with guaranteed accuracy for all parameter points $\vartheta \in R^2$ except for some curves Lebesgue of measure zero.

The estimators with such property may be used in various adaptive procedures (control, prediction, filtration).

2. The general case of a regression model

2.1. Assumptions and definitions

In this section we shall consider the linear regression model (1)

$$\mathrm{d}x(t) = \vartheta' a(t)\mathrm{d}t + \mathrm{d}W(t), \quad t \geq 0.$$

The problem is to estimate the unknown vector parameter ϑ with a given accuracy in the sense (2) from the observation of $(x(t), a(t))_{t\geq 0}$.

A natural candidate for estimating ϑ is the least squares estimator (LSE)

$$\tilde{\vartheta}(T) = \left(\int_0^T a(t)a'(t)\mathrm{d}t\right)^{-1} \int_0^T a(t)\mathrm{d}x(t), \quad T > 0.$$

It turns out in examples that the information matrix $\int_0^T a(t)a'(t)\mathrm{d}t$ has different asymptotic properties for different parameters ϑ. Thus e.g., the information matrix normalized by a scalar function may tend to a singular limit matrix.

To avoid this problem we rewrite the expression of the LSE $\tilde{\vartheta}(T)$ above in such a way, that as the inverse matrix factor there appears an appropriate chosen normalized matrix for which the asymptotic behaviour of its maximal eigenvalue for $T \to \infty$ is under control. To do this we apply a certain matrix V as a weight matrix to $a(t)$ to obtain the new process $(Va(t))$ with better asymptotic properties in the sense of Assumption (V) below (see formula (7)). The concrete form of V is determined by the kind of regressor $a(t)$ and cannot specified for the general case. Moreover V may depend on the unknown ϑ. To overcome these problem we shall construct a process $(V(t))$ based on the observations of (x, a) up to t, which estimates V and keeps the property (7) for the observed process $(b(t))_{t\geq 0}$, where $b(t) = V(t)a(t)$.

To get a first estimation of V by $V(\cdot)$ and some rates of convergence which defined below, we use the observation (x, a) from 0 to some time S. The properly estimation of the parameter ϑ starts from S.

The weighted LSE of ϑ for the given observation from S to T has the form:

$$\hat{\vartheta}(S, T) = G^{-1}(S, T)\Phi(S, T), \quad T > S > 0, \tag{5}$$

where

$$\Phi(S, T) = \int_S^T b(s)\mathrm{d}x(s), \qquad G(S, T) = \int_S^T b(s)a'(s)\mathrm{d}s,$$

$b(s) = V(s)a(s)$. Put $\Phi(T) = \Phi(0, T)$, $G(T) = G(0, T)$, $\bar{b}(s) = Va(s)$.

Let the weight process $(V(t))_{t\geq 0}$ be (F_t)-adapted and for all $T > 0$ the following integrals be finite:

$$\int_0^T E_\vartheta \|b(t)\|^q \mathrm{d}t < \infty. \tag{6}$$

We shall write in the following $f(x) \simeq C$ as $x \to \infty$ ($f_\varepsilon \simeq C$ as $\varepsilon \to 0 \ldots$) instead of the relations:

$$0 < \varliminf_{x\to\infty} f(x) \leq \varlimsup_{x\to\infty} f(x) < \infty$$

$$(0 < \varliminf_{\varepsilon\to 0} f_\varepsilon \leq \varlimsup_{\varepsilon\to 0} f_\varepsilon < \infty \ldots).$$

The rates of increase of the integrals $\int_0^T b_i^2(t)\mathrm{d}t$, $i = \overline{0, p}$ in general depend on some vector parameter $\alpha \in R^r$.

Assumption (V). Let A be a non-empty subset of R^r, such that, for every $i = \overline{0, p}$ there exists a family of unboundedly increasing positive functions $\{\varphi_i(\alpha, T),\ T > 0\}_{\alpha\in A}$ with the following properties: for every $\vartheta \in \Theta$ and $\alpha = \alpha(\vartheta) \in A$

$$\varphi_i^{-1}(\alpha, T)\int_0^T \tilde{b}_i^2(t)\mathrm{d}t \simeq C, \quad \text{as } T \to \infty \quad P_\vartheta - \text{a.s.}, \tag{7}$$

where $\tilde{b}_i(\cdot)$ equals $b_i(\cdot)$ or $\bar{b}_i(\cdot)$, $i = \overline{0, p}$.

Often we shall omit the dependence $\varphi_i(\alpha, T)$ of the parameter α in our notation. The functions $\varphi_i(\alpha, T)$ are called *rates of increase* of integrals $\int_0^T \tilde{b}_i^2(t)\mathrm{d}t$ $i = \overline{0, p}$.

Our sequential plans will be constructed by using first hitting times of the processes $\int_0^T \tilde{b}_i^2(s)\mathrm{d}s$, $i = \overline{0, p}$, $T > 0$. To investigate the asymptotic properties of these hitting times, we will use the rates $\varphi_i(T)$ of increase of these integrals.