ON GUARANTEED ESTIMATION OF PARAMETERS OF RANDOM PROCESSES BY THE WEIGHTED LEAST SQUARES METHOD

Nataliya Meder and Sergei Vorobejchikov

Department of Applied Mathematics and Cybernetics, Tomsk University, 36 Lenin str., 634050 Tomsk, Russia
e-mail: meder_n@mail.ru

Abstract: The problem of parameters estimation of an autoregressive process is considered. The method of guaranteed estimation is based on the least squares method with special weights and uses a special stopping rule. The properties of the procedures are studied for the case of known and unknown variance of the noise. Copyright © 2002 IFAC

Keywords: Time series analysis, autoregressive models, sequential estimation, least-square method, mean square error, numerical simulation.

1. INTRODUCTION

In the problems of control and identification the processes of autoregressive type are widely used. Such models include a small number of unknown parameters and provide the adequate description of observations. The asymptotic properties of least squares estimators are well known, such as almost sure convergence, asymptotic normality and others (Anderson, 1971; Lai and Wei, 1982). The investigation of properties of the estimators based on the sample of fixed size is very difficult. Borisov and Konev (1978) proposed the sequential estimation procedure of the autoregressive parameters, which ensures a preassigned accuracy of estimator of a scalar parameter. The guaranteed estimation procedure for multidimensional autoregressive parameter process has been constructed by Konev and Pergamenshchikov, (1981). This procedure uses a special sequence of stopping moments, for which the least squares estimators are calculated. After summation of them with some weights the guaranteed estimator of unknown parameter is obtained. This estimator has a preassigned accuracy in the mean square sense. The case of unknown noise variance was investigated by Dmitrienko and Konev, (1995). The another sequential procedure for estimating one of the unknown parameters was constructed by Vorobejchikov, (1983).

This paper proposes a procedure for estimating a multidimensional vector parameter in autoregressive process which is based on the weighted least squares method. The cases of known and unknown noise variance have been studied. The results of numerical modeling prove a good performance of the procedure.

2. PROBLEM STATEMENT

Let the observed scalar process be specified by the equation

\[ x(t + 1) = \mathbf{A}x(t) + B\xi_{t+1}, \quad t = 0, 1, \ldots, \]

where \( \xi_t \) is a sequence of independent identically distributed random variables, \( E\xi_t = 0, E\xi_t^2 = 1; \)

\[ \]
\( \mathbf{a}(t, x) \) is a column vector function of size \( m \times 1 \), with elements depending on the realization of the process \( x(t) \) up to the moment \( t \); \( B \) is a constant, which can be either known or unknown; \( \mathbf{A} = (\lambda_1, ..., \lambda_m) \) is unknown parameter.

The problem is to construct a sequential estimator of unknown vector \( \mathbf{A} \) by observations of process \( x(t) \) with a preassigned mean square accuracy.

3. THE CASE OF KNOWN NOISE VARIANCE

Assume that \( B \) is known. It is proposed to construct guaranteed estimator \( \mathbf{A}^* \) of unknown vector \( \mathbf{A} \) on the basis of the weighted least squares estimator, defined as

\[
\mathbf{A}(N) = \left( \sum_{l=0}^{N} x(l+1)v(l, x)\mathbf{a}^T(l, x) \right) \mathbf{A}^{-1}(N),
\]

(1)

Special nonnegative weight function \( v(l, x) \) is found from the equation

\[
\nu_{\min}(k) = \frac{k}{B^2} \sum_{l=\sigma}^{k} v^2(l, x)f(l, x), \tag{2}
\]

where

\[
f(l, x) = \mathbf{a}^T(l, x)\mathbf{a}(l, x),
\]

\( \nu_{\min}(N) \) is a minimal eigen value of matrix \( \mathbf{A}(N) \); \( \sigma \) is the least number \( N \) of observations for which the matrix \( \mathbf{A}(N) \) is invertible. The weight function \( v(l, x) \) on the interval \((0, \sigma)\) is given by

\[
v(l, x) = \begin{cases} \frac{1}{B \sqrt{f(l, x)}}, & \text{if } \mathbf{a}(l, x) \text{ is linearly independent on } \mathbf{a}(0, x), ..., \mathbf{a}(l-1, x), \\ 0, & \text{otherwise}. \end{cases} \tag{3}
\]

Introduce a special stopping rule \( \tau \), defined as

\[
\tau = \tau(H) = \inf \left\{ N > 0 : \nu_{\min}(N) \geq H \right\}, \tag{4}
\]

where \( H \) is a positive parameter. The last weight \( v(\tau, x) \) is chosen from the conditions

\[
\frac{\nu_{\min}(\tau)}{B^2} \geq \sum_{l=\sigma}^{\tau} v^2(l, x)f(l, x), \tag{5}
\]

\[
\nu_{\min}(\tau) = H.
\]

The guaranteed estimator \( \mathbf{A}^*(H) \) of parameter \( \mathbf{A} \) at the moment \( \tau \) is given by formula

\[
\mathbf{A}^*(H) = \left( \sum_{l=0}^{\tau} x(l+1)v(l, x)\mathbf{a}^T(l, x) \right) \mathbf{A}^{-1}(\tau), \tag{6}
\]

\[
\mathbf{A}(\tau) = \sum_{l=0}^{\tau} \mathbf{a}(l, x)v(l, x)\mathbf{a}^T(l, x).
\]

The properties of this estimator can be formulated as follows.

**Theorem 1.** Let the weights \( v(l, x) \) determined in (2), (3) be such that

\[
\sum_{l \geq 0} v^2(l, x)f(l, x) = +\infty \text{ a.s.} \tag{7}
\]

Then for any \( H > 0 \) the mean square accuracy of the estimator \( \mathbf{A}^*(H) \) satisfies the inequality

\[
E^H \| \mathbf{A}^*(H) - \mathbf{A} \|^2 \leq \frac{H + m - 1}{H^2},
\]

and the stopping time \( \tau(H) \) is finite with probability one.

**Proof of Theorem 1.** In view of the property of the minimal eigen value \( \nu_{\min}(l) \)

\[
\mathbf{A}(l) \geq \nu_{\min}(l) \cdot \mathbf{I},
\]

where \( \mathbf{I} \) is an unit matrix, and applying the Caushy - Bunyakovskii inequality and conditions (5) we obtain

\[
E^H \| \mathbf{A}^*(H) - \mathbf{A} \|^2 \leq E^H \left\| \mathbf{A}^{-1}(\tau) \left( \sum_{l=0}^{\tau} \mathbf{a}^T(l, x)v(l, x)\mathbf{B} \xi_{l+1} \right) \right\|^2.
\]

\[
\leq E^H \left\| \mathbf{A}^{-1}(\tau) \right\|^2 \left\| \left( \sum_{l=0}^{\tau} \mathbf{a}^T(l, x)v(l, x)\mathbf{B} \xi_{l+1} \right)^2 \right\|^2 \leq \frac{B^2 \tau}{H^2} \left( E^H \left\| \mathbf{a}(l, x) \right\|^2 2 \xi_{l+1}^2 \right) + \frac{2E^H \left\{ \sum_{s,l=0, s<l} v(l, x)v(s, x)\xi_{l+1}a^T(l, x)s \right\} \cdot a(s, x)\xi_{s+1}}{\tau}.
\]

Introduce the truncated moment \( \tau(N) = \min(\tau, N) \) and examine the first summand in the right-hand side of (8):

\[
E^H \left\{ \sum_{l=0}^{\tau(N)} v^2(l, x)\mathbf{a}(l, x)\mathbf{a}^T(l, x) \xi_{l+1}^2 \right\} = \sum_{l=0}^{\tau(N)} v^2(l, x)\mathbf{a}(l, x)\mathbf{a}^T(l, x) \xi_{l+1}^2 \chi_{(l \leq \tau)}
\]

\[
= E^H \left\{ \sum_{l=0}^{N} v^2(l, x)\mathbf{a}(l, x)\mathbf{a}^T(l, x) \xi_{l+1}^2 \chi_{(l \leq \tau)} \right\} = \sum_{l=0}^{N} E^H \left\{ \sum_{s,l=0, s<l} v^2(l, x)v(s, x)\xi_{l+1}a^T(l, x)s \right\} \cdot a(s, x)\xi_{s+1}.
\]
From here and \((3)\) we obtain
\[
\chi_{(A)} \text{ is modified as}
\]
\[
\text{such approach was proposed by Dmitrienko}
\]
\[
\text{additional stage is needed for estimating the variance.}
\]
\[
\text{The finiteness of the stopping time}
\]
\[
\text{Similarly one can show that the second summand}
\]
\[
\text{The finiteness of the stopping time } \tau(H) \text{ is due to}
\]
\[
5. \text{Numerical Modeling}
\]
\[
\text{Consider a stable autoregressive process of the order 2}
\]
\[
x(t+1) = \lambda_1 x(t) + \lambda_2 x(t-1) + B \xi_{t+1}, \ t = 0, 1, ...
\]
\[
\text{where } \lambda_1, \lambda_2 \text{ are unknown parameters, } \xi_{t+1} \text{ is a sequence of i.i.d. standard gaussian random variables (noises).}
\]
\[
\text{In this case it is easy to show that}
\]
\[
C_n = \frac{1}{2n/2 \Gamma \left( \frac{n}{2} \right)} \int_0^\infty x^{\frac{n}{2}-1} e^{-x/2} \, dx = \frac{1}{n - 2},
\]
The tables 1-3 present the results of the simulations for the case when the noise variance is known, and the tables 4-6 those for the case of unknown variance. By theorem 1, 2 one can find the upper bounds for mean square accuracy for $H = 100$ and $H = 1000$ respectively: $0.0101$ and $0.001$. The tables give the mean estimation times (m.e.t.) and the mean square deviations (m.s.d., $i = 1, 2$) for parameters $\lambda_1, \lambda_2$ and different sizes of noises $B^2$ and different sizes of the initial sample $n$. The results are obtained on the basis of $100$ replications of the experiment.

Table 1 The characteristics of the procedure $(\lambda_1 = 0.4, \lambda_2 = 0.4)$

<table>
<thead>
<tr>
<th>B</th>
<th>H</th>
<th>m.e.t.</th>
<th>m.s.d.1</th>
<th>m.s.d.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>306</td>
<td>5.08 $\times 10^{-3}$</td>
<td>4.59 $\times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>306</td>
<td>5.06 $\times 10^{-3}$</td>
<td>4.56 $\times 10^{-3}$</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>2848</td>
<td>3.59 $\times 10^{-4}$</td>
<td>5.21 $\times 10^{-4}$</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>2850</td>
<td>3.72 $\times 10^{-4}$</td>
<td>4.78 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 2 The characteristics of the procedure $(\lambda_1 = 0.6, \lambda_2 = 0.2)$

<table>
<thead>
<tr>
<th>B</th>
<th>H</th>
<th>m.e.t.</th>
<th>m.s.d.1</th>
<th>m.s.d.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>378</td>
<td>4.72 $\times 10^{-3}$</td>
<td>4.83 $\times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>378</td>
<td>4.71 $\times 10^{-3}$</td>
<td>4.83 $\times 10^{-3}$</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>3586</td>
<td>4.03 $\times 10^{-4}$</td>
<td>5.26 $\times 10^{-4}$</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>3586</td>
<td>4.01 $\times 10^{-4}$</td>
<td>5.28 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 3 The characteristics of the procedure $(\lambda_1 = 0.0, \lambda_2 = 0.0)$

<table>
<thead>
<tr>
<th>B</th>
<th>H</th>
<th>m.e.t.</th>
<th>m.s.d.1</th>
<th>m.s.d.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>280</td>
<td>4.77 $\times 10^{-3}$</td>
<td>5.47 $\times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>280</td>
<td>4.67 $\times 10^{-3}$</td>
<td>5.49 $\times 10^{-3}$</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>2768</td>
<td>3.77 $\times 10^{-4}$</td>
<td>3.88 $\times 10^{-4}$</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>2769</td>
<td>3.81 $\times 10^{-4}$</td>
<td>4.86 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 4 The characteristics of the procedure $(\lambda_1 = 0.4, \lambda_2 = 0.4, n = 10)$

<table>
<thead>
<tr>
<th>B</th>
<th>H</th>
<th>m.e.t.</th>
<th>m.s.d.1</th>
<th>m.s.d.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>719</td>
<td>3.20 $\times 10^{-3}$</td>
<td>2.34 $\times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>623</td>
<td>3.01 $\times 10^{-3}$</td>
<td>3.58 $\times 10^{-3}$</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>6900</td>
<td>4.17 $\times 10^{-4}$</td>
<td>2.95 $\times 10^{-4}$</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>8093</td>
<td>2.60 $\times 10^{-4}$</td>
<td>3.13 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 5 The characteristics of the procedure $(\lambda_1 = 0.4, \lambda_2 = 0.4, n = 20)$

<table>
<thead>
<tr>
<th>B</th>
<th>H</th>
<th>m.e.t.</th>
<th>m.s.d.1</th>
<th>m.s.d.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>626</td>
<td>3.21 $\times 10^{-3}$</td>
<td>2.10 $\times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>709</td>
<td>6.93 $\times 10^{-3}$</td>
<td>4.45 $\times 10^{-2}$</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>6879</td>
<td>3.24 $\times 10^{-4}$</td>
<td>2.32 $\times 10^{-4}$</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>6713</td>
<td>3.21 $\times 10^{-4}$</td>
<td>3.83 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 6 The characteristics of the procedure $(\lambda_1 = 0.6, \lambda_2 = 0.2, n = 10)$

<table>
<thead>
<tr>
<th>B</th>
<th>H</th>
<th>m.e.t.</th>
<th>m.s.d.1</th>
<th>m.s.d.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>836</td>
<td>3.10 $\times 10^{-4}$</td>
<td>3.87 $\times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>1192</td>
<td>4.88 $\times 10^{-3}$</td>
<td>3.20 $\times 10^{-3}$</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>10097</td>
<td>1.88 $\times 10^{-4}$</td>
<td>2.60 $\times 10^{-4}$</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>9795</td>
<td>2.51 $\times 10^{-4}$</td>
<td>3.01 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

The results of simulations show that the proposed estimator enables us to attain a preassigned mean-square accuracy at the termination time. The sample mean-square accuracy is less than theoretical bound and turns out to be close to it. The mean estimation time increases linearly with the growth of $H$. In the case of unknown noise variance $B$ the mean estimation time increases. The mean estimation time also depends on the size $n$ of the initial size, used estimate the variance.

6. CONCLUSION

The paper proposes one stage procedure for estimating multivariate parameter in autoregressive process, which ensures estimating unknown parameters with a prescribed mean square precision. The results can be applied to identification problems, problems of control and time series analysis.

REFERENCES


