15th Triennial World Congress, Barcelona, Spain

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[1]

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

$$x(t+1) = \mathbf{\Lambda a}(t,x) + B\xi_{t+1},\ t = 0, 1, ...,$$

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

[1] Research supported by the Russian Foundation of Fundamental Research, grant 00-01-00880.

$\mathbf{a}(t,x)$ 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; $\mathbf{\Lambda} = (\lambda_1, ..., \lambda_m)$ is unknown parameter.

The problem is to construct a sequential estimator of unknown vector $\mathbf{\Lambda}$ 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 $\mathbf{\Lambda}^*$ of unknown vector $\mathbf{\Lambda}$ on the basis of the weighted least squares estimator, defined as

$$\mathbf{\Lambda}(N) = \Big(\sum_{l=0}^{N} x(l+1)v(l,x)\mathbf{a}^T(l,x)\Big)\mathbf{A}^{-1}(N), \quad (1)$$

$$\mathbf{A}(N) = \sum_{l=0}^{N} \mathbf{a}(l,x)v(l,x)\mathbf{a}^T(l,x).$$

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

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

where

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

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

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

Introduce a special stopping rule τ, defined as

$$\tau = \tau(H) = \inf\Big(N > 0 : \nu_{min}(N) \geq H\Big), \quad (4)$$

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

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

$$\nu_{min}(\tau) = H.$$

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

$$\mathbf{\Lambda}^*(H) = \Big(\sum_{l=0}^{\tau} x(l+1)v(l,x)\mathbf{a}^T(l,x)\Big)\mathbf{A}^{-1}(\tau), \quad (6)$$

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

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

$$\sum_{l\geq 0} v^2(l,x)f(l,x) = +\infty \quad a.s. \quad (7)$$

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

$$E_{\mathbf{\Lambda}}||\mathbf{\Lambda}^*(H) - \mathbf{\Lambda}||^2 \leq \frac{H+m-1}{H^2},$$

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

$$E_{\mathbf{\Lambda}}\Big\|\mathbf{\Lambda}^*(H) - \mathbf{\Lambda}\Big\|^2 =$$

$$= E_{\mathbf{\Lambda}}\Big\|\mathbf{A}^{-1}(\tau)\Big(\sum_{l=0}^{\tau}\mathbf{a}^T(l,\mathbf{x})v(l,x)B\xi_{l+1}\Big)\Big\|^2 \leq$$

$$\leq E_{\mathbf{\Lambda}}\Big\{\Big\|\mathbf{A}^{-1}(\tau)\Big\|^2\Big\|\sum_{l=0}^{\tau}\mathbf{a}^T(l,x)v(l,x)B\xi_{l+1}\Big\|^2\Big\} \leq$$

$$\leq \frac{B^2}{H^2}\Big(E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{\tau}v^2(l,x)\|\mathbf{a}(l,x)\|^2\xi_{l+1}^2\Big\}+$$

$$+2E_{\mathbf{\Lambda}}\Big\{\sum_{s,l=0,\ s<l}^{\tau} v(l,x)v(s,x)\xi_{l+1}\mathbf{a}^T(l,x)\cdot$$

$$\cdot\mathbf{a}(s,x)\xi_{s+1}\Big\}\Big). \quad (8)$$

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

$$E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{\tau(N)} v^2(l,x)\|\mathbf{a}(l,x)\|^2\xi_{l+1}^2\Big\} =$$

$$= E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{N} v^2(l,x)\|\mathbf{a}(l,x)\|^2\xi_{l+1}^2\chi_{(l\leq\tau)}\Big\} =$$

$$= E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{N} E\Big(v^2(l,x)\|\mathbf{a}(l,x)\|^2\xi_{l+1}^2\chi_{(l\leq\tau)}|F_l\Big)\Big\} =$$

$$= E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{N} v^2(l,\mathbf{x})\|\mathbf{a}(l,x)\|^2\chi_{(l\le\tau)}\Big\} =$$

$$= E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{\tau(N)} v^2(l,x)\|\mathbf{a}(l,x)\|^2\Big\}. \qquad (9)$$

Here $\chi_{(A)}$ is the indicator function of an event A, F_l is σ-algebra generated by $(x(0),\xi_1,...,\xi_l)$. The quantity (9) converges to

$$E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{\tau} v^2(l,x)\|\mathbf{a}(l,x)\|^2\Big\} \text{ as } N\to\infty.$$

From (2) and (4) it follows that

$$\frac{H}{B^2} \ge \sum_{l=\sigma}^{\tau} v^2(l,x)\|\mathbf{a}(l,x)\|^2.$$

From here and (3) we obtain

$$E_{\mathbf{\Lambda}}\Big\{\sum_{l=0}^{\tau} v^2(l,\mathbf{x})\|\mathbf{a}(l,x)\|^2\Big\} =$$

$$= E_{\mathbf{\Lambda}}\Big\{\Big(\sum_{l=0}^{\sigma} v^2(l,x)\|\mathbf{a}(l,x)\|^2 +$$

$$+\sum_{l=\sigma}^{\tau} v^2(l,x)\|\mathbf{a}(l,x)\|^2\Big)\Big\} \le$$

$$\le \frac{m-1+H}{H^2}.$$

Similarly one can show that the second summand in (8) equals zero.

The finiteness of the stopping time $\tau(H)$ is due to (7).

4. THE CASE OF UNKNOWN NOISE VARIANCE

Let the noise variance B^2 be unknown and the distribution function of ξ_t be known. In this case additional stage is needed for estimating the variance. Such approach was proposed by Dmitrienko and Konev, (1995). The sequential procedure (1-6) is modified as

$$\mathbf{\Lambda}^*(H) = \Big(\sum_{l=n+1}^{\tau} x(l+1)v(l,x)\mathbf{a}^T(l,x)\Big)\mathbf{A}^{-1}(\tau),$$

$$\mathbf{A}(\tau) = \sum_{l=n+1}^{\tau} \mathbf{a}(l,x)v(l,x)\mathbf{a}^T(l,x),$$

$$\frac{\nu_{min}(k)}{\Gamma_n} = \sum_{l=n+\sigma}^{k} v^2(l,x)f(l,x). \qquad (10)$$

The multiplier Γ_n serves to compensate unknown noise variance

$$\Gamma_n = C_n\sum_{l=1}^{n} x^2(l),\ C_n = E\Big(\sum_{l=1}^{n}\xi_l^2\Big)^{-1},$$

where n is the size of the initial sample for estimating the variance; C_n is a normalizing factor depending on distribution function of noises ξ_l.

The weight function $v(l,x)$ for $l\in(0,\sigma)$ is defined as

$$v(n+l,x) = \begin{cases} \dfrac{1}{\sqrt{\Gamma_n f(n+l,x)}}, & \text{if } \mathbf{a}(n+l,x) \text{ is linearly independent on } \mathbf{a}(n,x),\dots,\mathbf{a}(n+l-1,x). \\ 0, & \text{otherwise.} \end{cases} \qquad (11)$$

The properties of this procedure are given in

Theorem 2. Let the function $v(l,x)$ determined in (10), (11) be such that

$$\sum_{l\ge 0} v^2(l,x)f(l,x) = +\infty \quad a.s.$$

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

$$E_{\mathbf{\Lambda}}||\mathbf{\Lambda}^*(H)-\mathbf{\Lambda}||^2 \le \frac{H+m-1}{H^2}$$

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

This result can be proved along the lines of the proof of Theorem 1 taking into account the inequality

$$\frac{H}{\Gamma_n} \ge \sum_{l=\sigma}^{\tau} v^2(l,x)\|\mathbf{a}(l,x)\|$$

and the results, obtained by Dmitrienko and Konev, (1995).

5. NUMERICAL MODELING

Consider a stable autoregressive process of the order 2

$$x(t+1) = \lambda_1 x(t) + \lambda_2 x(t-1) + B\xi_{t+1},\ t=0,1,...,$$

where λ_1,λ_2 are unknown parameters, ξ_{t+1} is a sequence of i.i.d. standard gaussian random variables (noises). In this case it is easy to show that

$$C_n = \frac{1}{2^{n/2}\Gamma\Big(\frac{n}{2}\Big)}\int_0^{\infty} x^{\frac{n}{2}-1}e^{-x/2}\,dx = \frac{1}{n-2},$$

where $\Gamma(p)$ is gamma function.

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 ($m.s.d._i,\ i = 1, 2$) for parameters λ_1, λ_2 and different variances 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 ($\lambda_1 = 0.4,\ \lambda_2 = 0.4$)

B	H	$m.e.t.$	$m.s.d._1$	$m.s.d._2$
1	100	306	$5.08{\cdot}10^{-3}$	$4.59{\cdot}10^{-3}$
10	100	306	$5.06{\cdot}10^{-3}$	$4.56{\cdot}10^{-3}$
1	1000	2848	$3.59{\cdot}10^{-4}$	$5.21{\cdot}10^{-4}$
10	1000	2850	$3.72{\cdot}10^{-4}$	$4.78{\cdot}10^{-4}$

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

B	H	$m.e.t.$	$m.s.d._1$	$m.s.d._2$
1	100	378	$4.72{\cdot}10^{-3}$	$4.83{\cdot}10^{-3}$
10	100	378	$4.71{\cdot}10^{-3}$	$4.83{\cdot}10^{-3}$
1	1000	3586	$4.03{\cdot}10^{-4}$	$5.26{\cdot}10^{-4}$
10	1000	3586	$4.01{\cdot}10^{-4}$	$5.28{\cdot}10^{-4}$

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

B	H	$m.e.t.$	$m.s.d._1$	$m.s.d._2$
1	100	280	$4.77{\cdot}10^{-3}$	$5.47{\cdot}10^{-3}$
10	100	280	$4.67{\cdot}10^{-3}$	$5.49{\cdot}10^{-3}$
1	1000	2768	$3.77{\cdot}10^{-4}$	$3.88{\cdot}10^{-4}$
10	1000	2769	$3.81{\cdot}10^{-4}$	$4.86{\cdot}10^{-4}$

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

B	H	$m.e.t.$	$m.s.d._1$	$m.s.d._2$
1	100	719	$3.20{\cdot}10^{-3}$	$2.34{\cdot}10^{-3}$
10	100	623	$3.01{\cdot}10^{-3}$	$3.58{\cdot}10^{-3}$
1	1000	6900	$4.17{\cdot}10^{-4}$	$2.95{\cdot}10^{-4}$
10	1000	8093	$2.60{\cdot}10^{-4}$	$3.13{\cdot}10^{-4}$

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

B	H	$m.e.t.$	$m.s.d._1$	$m.s.d._2$
1	100	626	$3.21{\cdot}10^{-3}$	$2.10{\cdot}10^{-3}$
10	100	709	$6.93{\cdot}10^{-3}$	$4.45{\cdot}10^{-2}$
1	1000	6879	$3.24{\cdot}10^{-4}$	$2.32{\cdot}10^{-4}$
10	1000	6713	$3.21{\cdot}10^{-4}$	$3.83{\cdot}10^{-4}$

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

B	H	$m.e.t.$	$m.s.d._1$	$m.s.d._2$
1	100	836	$3.10{\cdot}10^{-3}$	$3.87{\cdot}10^{-3}$
10	100	1192	$4.88{\cdot}10^{-3}$	$3.20{\cdot}10^{-3}$
1	1000	10097	$1.88{\cdot}10^{-4}$	$2.60{\cdot}10^{-4}$
10	1000	9795	$2.51{\cdot}10^{-4}$	$3.01{\cdot}10^{-4}$

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 then 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

Anderson, T.W. (1971). *The statistical analysis of time series.* John Wiley & Sons,Inc., New York, London, Sydney, Toronto.

Borisov, V.Z. and V.V. Konev (1978). Sequential estimation of parameters of discrete processes. *Autom. Remote Control*, **38**, 1475-1480.

Dmitriyenko, A.A. and V.V. Konev (1995). On the sequential classification of autoregressive process with unknown nuisance parameters. *The problem of information transmission.* **31**, 51-62.

Konev, V.V. and S.M. Pergamenshchikov (1981). Sequential identification procedures for the parameters of dynamic systems. *Autom. Remote Control*, **42**, 917-924.

Lai, T.L. and C.Z. Wei (1982). Least square estimation in stohastic regression models with application to indentification and control of dynamic system. *The Annals of Statistics.*, **10**, 154-166.

Vorobejchikov, S.E. (1983). On the sequential identification of random process parameters of recursive type. *Mathematical statistics and its applications.*, **IX**, 42-47 (in Russian).