15th Triennial World Congress, Barcelona, Spain

FIXED-ACCURACY ESTIMATION OF PARAMETERS IN THE ARMAX SYSTEM

Victor Konev and Dmitry Shapovalov

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

Abstract: Autoregressive moving average systems with exogenous inputs (ARMAX) are widely used in a variety of applied problems connected with identification, adaptive control and time series analysis. This paper proposes the method which enables one to estimate parameters in the ARMAX system with a prescribed mean-square precision at the termination time. The procedure is constructed on the basis of sequential analysis approach and makes use of special modifications of estimates obtained by the method of instrumental variables. The method can be used for guaranteed estimation (in mean-square sense) of spectral density of ARMA processes. *Copyright © 2002 IFAC*

Keywords: Stopping time, instrumental variable, fixed-precision estimates, sequential plan.

1. INTRODUCTION[1]

It is well known (Anderson, 1994; Box and Jenkins, 1970; Goodwin, 1987; Kumar, 1985) that linear models are commonly used in different applied problems, connected with the design of control systems, predictors and filters despite the fact that the majority of control systems in practice are composed of nonlinear elements. One of the most popular linear model is an autoregressive moving average system with exogenous inputs (the ARMAX system) described by the linear difference equation

$$A(q^{-1})y_n = B(q^{-1})u_{n-1} + C(q^{-1})e_n, \qquad (1)$$

where $\{y_n\}, \{u_n\}$ and $\{e_n\}$ are the output, input and noise sequences respectively; $\{e_n\}$ is an unobserved sequence of independent identically distrubuted (i.i.d.) random variables with $Ee_n = 0, Ee_n^2 = 1$; the input u_n (control) is non-anticipating sequence in the sense that it involves only current and past observations $y_n, y_{n-1}, u_{n-1}, \ldots$;

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_p q^{-p}, \\ B(q^{-1}) &= b_1 + \cdots + b_r q^{-(r-1)}, \\ C(q^{-1}) &= c_0 + c_1 q^{-1} + \cdots + c_s q^{-s} \end{aligned} \qquad (2)$$

are polynomials; q^{-1} is the unit backward shift operator (i.e. $q^{-1}y_n = y_{n-1}$); a_i, b_j and c_l are constant parameters.

This model is quite general and comprises, in particular, autoregression process (AR), autoregressive moving average process (ARMA) and others which are often applied in time series analysis. The ARMAX model is more flexible as compared with the ARX model (autoregressive system with exogenous inputs) because the disturbance is modeled by the moving average and not by the i.i.d. sequence. The problems dealing with the construction of the ARMAX models (parameter es-

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

timation) are considered by the identification theory. In the classical setting, different techniques have been developed to estimate the unknown parameters in (1): Gaussian maximum likelihood, least squares, instrumental variables, Whittle estimation schemes and so on. As a rule, statistical properties of estimates obtained by various methods are investigated in terms of large samples, that is on the basis of information from series observed a long time. It is expected that asymptotic properties of the procedures will hold approximately for on-line estimates. However in practical problems a non-asymptotic behaviour of estimates may be substantially different from the asymptotic one and this discrepancy can cause undesirable effects.

The purpose of this paper is to demonstrate that parameters of the ARMAX model can be estimated with a preassigned precision (in mean-square sense) by the sample of finite size. Over the past several years there has been substantial progress with regard to guaranteed estimation of parameters in stochastic dynamic systems in discrete time. The advances have been achieved by applying the approach of sequential analysis which presumes that observations should be conducted until enough information is gathered about unknown parameters. The theoretical results were obtained for AR, ARX processes and others (see Konev and Lai, 1995; Konev and Pergamenshchikov, 1996 for details and further references). For all these models the disturbance is modeled as a sequence of i.i.d. random variables. This condition is not satisfied for the ARMAX system. Here a modified sequential estimation scheme for guaranteed identification of model (1) is proposed.

2. PROBLEM FORMULATION AND PRELIMINARIES

In general, the identification problem for model (1) consists in estimating all the parameters a_i, b_j and c_k. This paper proposes an estimation procedure for parameters a_i and b_j. Note that this procedure can be complemented to comprise the general case but the resulting procedure becomes more cumbersome and it is omitted.

Let $\boldsymbol{\theta} = (-a_1, \dots, -a_p, b_1, \dots, b_r)'$ be the vector of unknown parameters and

$$\boldsymbol{X}_n = (y_n, \dots, y_{n-p+1}, u_n, \dots, u_{n-r+1})'$$

be the vector of regressors (prime denotes the transposition). Then the model (1) can be rewritten as

$$y_n = \boldsymbol{\theta}' \boldsymbol{X}_{n-1} + C(q^{-1}) e_n. \tag{3}$$

Assume that the vector $\boldsymbol{c} = (c_0, \dots, c_s)'$ of the parameters of polinomial C is norm bounded, that is $\|\boldsymbol{c}\|^2 \leq L < \infty$, where L is a known constant, $\|\boldsymbol{c}\|^2 = \boldsymbol{c}'\boldsymbol{c}$.

The problem is to construct an estimator for vector $\boldsymbol{\theta}$ which enables us to estimate $\boldsymbol{\theta}$ with a fixed mean-square accuracy. Our approach to this problem is closely connected with the method of instrumental variables. By this method estimate for $\boldsymbol{\theta}$ has the form (Ljung, 1987)

$$\widehat{\boldsymbol{\theta}}_n = \left(\sum_{k=1}^{n} \boldsymbol{\phi}(k) \boldsymbol{X}_k' \right)^{-1} \sum_{k=1}^{n} \boldsymbol{\phi}(k) y_{k+1}, \tag{4}$$

where $\boldsymbol{\phi}(k) = (\phi_1(k), \dots, \phi_{p+r}(k))'$ is a vector of instrumental variables. It is worth noting that by itself $\widehat{\boldsymbol{\theta}}_n$ does not give a solution to the stated problem of guaranteed estimation because in contrust to the case of deterministic regression model now the inverse matrix $(\sum_{k=1}^{n} \boldsymbol{\phi}(k) \boldsymbol{X}_k')^{-1}$ is random. By this reason the estimate $\widehat{\boldsymbol{\theta}}_n$ is a nonlinear function of observations and it is a very complicated task to find the explicit formula for its variance even for specific distributions of the noise e_n. To solve this problem one can use a sequential estimation scheme based on the method of instrumental variables.

3. GUARANTEED ESTIMATORS

The key idea of sequential analysis is to sample until enough information is gathered about unknown parameters. One of the thinkable ways to construct a sequential estimate with a fixed mean-square accuracy on the basis of (4) is to choose properly a stopping time τ and to define the sequential estimate as $\boldsymbol{\theta}_\tau$. For our purpose one stopping time turns out to be insufficient and, actually, a sequence of stopping times are needed. The procedure is constructed in two steps. First, a sequence of stopping times $\{\tau_n, n \geq 1\}$ is introduced as

$$\tau_n = \tau(h_n) = inf\{m \geq 1 : \sum_{k=1}^{m} \|\boldsymbol{\phi}(k)\|^2 \geq h_n\},$$

where h_n is non-decreasing sequence of positive numbers such that $\sum_{n \geq 1} h_n^{-1} < \infty$. For each τ_n a modified version of estimate (4) is defined by the formula

$$\boldsymbol{\theta}_n = \boldsymbol{W}_n^{+} \left[\sum_{k=1}^{\tau_n - 1} \boldsymbol{\phi}(k) y_{k+1} + \alpha_n \boldsymbol{\phi}(\tau_n) y_{\tau_n + 1} \right],$$

where

$$\boldsymbol{W}_n = \sum_{k=1}^{\tau_n - 1} \boldsymbol{\phi}(k) \boldsymbol{X}_k' + \alpha_n \boldsymbol{\phi}(\tau_n) \boldsymbol{X}_{\tau_n}';$$

$\boldsymbol{W}_n^+$ denotes the inverse matrix for $\boldsymbol{W}_n$ if it exists and $\boldsymbol{W}_n^+ = 0$ otherwise; the weight multiplier α_n is determined from the equation

$$\sum_{k=1}^{\tau_n - 1} \|\boldsymbol{\phi}(k)\|^2 + \alpha_n \|\boldsymbol{\phi}(\tau_n)\|^2 = h_n.$$

The second step. For each $H > 0$ a guaranteed estimator of vector $\boldsymbol{\theta}$ is defined as the weighted average

$$\boldsymbol{\theta}^*(H) = \left(\sum_{n=1}^{\sigma(h)} v_n \right)^{-1} \sum_{n=1}^{\sigma(h)} v_n \boldsymbol{\theta}_n, \qquad (5)$$

where the weights are defined as

$$v_n = \begin{cases} h_n^{-2} \|\boldsymbol{W}_n^{-1}\|^{-2}, & \text{if } det \boldsymbol{W}_n \neq 0, \\ 0, & \text{otherwise;} \end{cases}$$

$\sigma(H)$ is a number of modified estimates by the method of instrumental variables used in the average,

$$\sigma(H) = inf\{m \geq 1 : \sum_{n=1}^{m} v_n \geq H\}.$$

The total duration of the procedure is

$$T(H) = \tau(h_{\sigma(H)}) + 1. \qquad (6)$$

The pair $(T(H), \boldsymbol{\theta}^*(H))$ will be called sequential plan. To investigate its properties the following notation will be needed. Let $\mathcal{F}_n$ denotes the σ - algebra generated by the random variables

$$\{\widetilde{X}_0, e_1, \ldots, e_n, u_1, \ldots, u_{n-1}\},$$

where

$$\widetilde{X}_0 = (y_0, \ldots, y_{1-p}, u_0, \ldots, u_{1-r}, e_0, \ldots, e_{1-s})'$$

is the initial state of system (1).

Theorem 1 *Let for the system (1) there exists a sequence of instrumental variables such that*
(i) $\boldsymbol{\phi}(k)$ *-* $\mathcal{F}_{k-s}$ *- measurable for* $k > s$ *and* $\boldsymbol{\phi}(k) = 0, k \leq s$*;*
(ii)

$$\sum_{k \geq 1} \|\boldsymbol{\phi}(k)\|^2 = +\infty, \quad \sum_{k \geq 1} v_k = +\infty \quad a.s.$$

Then for any $H > 0$ *the sequential plan* $(T(H), \boldsymbol{\theta}^*(H))$ *has the properties:*
1°. $T(H) < \infty$ *a.s.,*
2°. $E_\theta \|\boldsymbol{\theta}^*(H) - \boldsymbol{\theta}\|^2 \leq \rho / H$,
where

$$\rho = (s+1)\|\boldsymbol{c}\|^2 \sum_{n \geq 1} h_n^{-1}, \qquad (7)$$

E_θ *denotes the mean by the distribution* P_θ *of the process (1) for given* $\boldsymbol{\theta}$*.*

Remark 1 *The assertion of Theorem 1 holds true if to change condition* (i) *to the following one* (i') *sequence* $\{\boldsymbol{\phi}(k)\}$ *and* $\{e_k\}$ *are stochastically independent.*

4. ARMAX SYSTEMS WITH QUASI-STATIONARY INPUTS

In this section the problem of guaranteed estimation of parameters in the ARMAX system (1) is cosidered in the case when the input sequence is a quasi-stationary process (Ljung, 1987). This process satisfies the following conditions.

A1. $\{u_k\}$ is a second order process such that

$$\sup_{k \geq 1} |Eu_k| < \infty, \qquad \sup_{k \geq 1, l \geq 1} |Eu_k u_l| < \infty.$$

A2. With probability one for each integer l

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} u_k u_{k-l} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} Eu_k u_{k-l} = R(l).$$

A3. The spectrum function corresponding to the correlation functions $\{R(l)\}$ is strictly positive, that is

$$F(\omega) = \sum_{l=-\infty}^{\infty} R(l) e^{-il\omega} > 0, \quad -\pi \leq \omega \leq \pi.$$

A4. The input sequence $\{u_k\}$ and the noise sequence $\{e_k\}$ are stochastically independent.

Under these conditions the sequence of instrumental variables $\{\boldsymbol{\phi}(k)\}$ is customarily defined as (see, e.g., Ljung, 1987)

$$\boldsymbol{\phi}(k) = K(q^{-1})(x_k, \ldots, x_{k-p+1}, u_k, \ldots, u_{k-r+1})', \qquad (8)$$

where $K(q^{-1})$ is a linear stable filter; x_k is a sequence satisfying the linear stable system

$$N(q^{-1}) x_k = M(q^{-1}) u_k,$$

where

$$N(q^{-1}) = 1 + n_1 q^{-1} + \cdots + n_f q^{-f},$$
$$M(q^{-1}) = 1 + m_1 q^{-1} + \cdots + m_g q^{-g},$$

and the orders of polynomials N and M are such that

$$\min(p - f, r - g) \leq 0.$$

By applying the sequential estimation scheme developed in section 3 one obtains.

Theorem 2 *Let polynomials A, B and C in (2) be relatively prime and all the roots of A are greater than one in modulus. Let the input sequence of instrumental variables $\{\boldsymbol{\phi}(k)\}$ be defined by (8). Then for any $H > 0$ sequential plan (5),(6) has the properties:*
1°. $T(H) < \infty$ *a.s.,*
2°. $E_\theta\|\boldsymbol{\theta}^*(H) - \boldsymbol{\theta}\|^2 \le \rho/H$,
(ρ is the same as in (7)).

Note that, if the sequence $\{u_k\}$ is non-random, assumptions A1-A3 coincide with the requirements imposed on the set of functions in problems of determenistic regression (Anderson, 1994). The ARMAX system with a deterministic input sequence u_k may arise in the problems of deterministic regression with a stationary noise depending on nuisance parameters. The following example illustrates this and allows one to compare asymptotic and non-asymptotic estimation schemes.

Example
Consider the regression model

$$y_n = bu_{n-1} + \xi_n, \tag{9}$$

where y_n is observed process, u_n is non-random function satisfying assumptions A1-A3, ξ_n is a stationary ARMA process specified by the equation

$$\xi_n = a\xi_{n-1} + e_n + c_1 e_{n-1} + \cdots + c_s e_{n-s}, \tag{10}$$

with unknown parameters $a, c_i, |a| < 1, \{e_n\}$ is a sequence of i.i.d. random variables such that $Ee_n = 0, Ee_n^2 = 1$. Consider the least squares estimate for b by the sample $\{y_1, \ldots, y_N\}$

$$\widehat{b}_N = \frac{\sum_{k=1}^N u_{k-1}y_k}{\sum_{k=1}^N u_{k-1}^2}.$$

Its mean square accuracy is

$$E(\widehat{b}_N - b)^2 = \frac{\sum_{k=1}^N \sum_{i=1}^N u_{k-1}u_{i-1}}{(\sum_{k=1}^N u_{k-1}^2)^2} \times \frac{a^{|k-i|}}{1-a^2}\sum_{j=0}^{s}\sum_{l=0}^{s} a^{j-l}c_j c_l.$$

From here it follows that

$$\sup_{-1<a<1} E(\widehat{b}_N - b)^2 = +\infty,$$

that is for any fixed N one can not ensure mean square accuracy of $\widehat{b}_N$ if no additional prior information about a and c_i is available.

Further the sequential estimation scheme in above is developed to model (9),(10). From (9),(10) it follows that

$$y_n = \boldsymbol{\theta}'\boldsymbol{X}_{n-1} + C(q^{-1})e_n,$$

where $\boldsymbol{\theta} = (a, b, -ab)'$, $\boldsymbol{X}_n = (y_n, u_n, u_{n-1})'$. Theorem 2 implies

$$\sup_{-1<a<1} E(b^*(H) - b)^2 \le \rho/H,$$

that is the sequential estimate $b^*(H)$ for parameter b ensures a given precision at the termination time $T(H)$ under the appropriate choice of the threshold H.

5. THE CASE OF LINEAR CONTROL

In this section the identification problem for model (1) is considered in the case when the input sequence u_n (control) is a linear function of the output sequence y_n. Assume that

$$u_n = D(q^{-1})y_n, \tag{11}$$

where

$$D(q^{-1}) = d_0 + d_1q^{-1} + \ldots + d_m q^{-m}.$$

Substituting (11) in (1) yields

$$\widetilde{A}(q^{-1})y_n = C(q^{-1})e_n, \tag{12}$$

where

$$\widetilde{A}(q^{-1}) = A(q^{-1}) - q^{-1}B(q^{-1})D(q^{-1}) = 1 + \widetilde{a}_1 q^{-1} + \ldots + \widetilde{a}_l q^{-l}, \tag{13}$$

$l = \max(p, m + r)$. Denoting
$\boldsymbol{X}_n = (y_n, \ldots, y_{n-l+1})'$, $\boldsymbol{\theta} = (-\widetilde{a}_1, \ldots, -\widetilde{a}_l)'$ one can rewrite (12) as

$$y_n = \boldsymbol{\theta}'\boldsymbol{X}_{n-1} + C(q^{-1})e_n. \tag{14}$$

The sequential estimation scheme developed above can be used to estimate parameter vector $\widetilde{a}$ with a fixed mean-square accuracy by observations of y_n.

Theorem 3 *Let all the roots of polynomial (13) be greater than one in modulus and polynomials $\widetilde{A}(z)$ and $C(z), z^sC(z^{-1})$ be relatively prime. Then for any $H > 0$ sequential plan $(T(H), \boldsymbol{\theta}^*(H))$ defined by (5),(6) with $\boldsymbol{\phi}(k) = (y_{k-s}, \ldots, y_{k-s-l+1})'$ has the properties:*
1°. $T(H) < \infty$ *a.s.,*
2°. $E_\theta\|\boldsymbol{\theta}^*(H) - \boldsymbol{\theta}\|^2 \le \rho/H$,

Examine asymptotic properties of the sequential plan $(T(H), \boldsymbol{\theta}^*(H))$ as $H \to \infty$. To this end it is assumed that the sequence h_n is defined as

$$h_n = h_n(H) = \begin{cases} H, & \text{if } n \le n_0(H), \\ n^{1+\gamma}, & \text{if } n > n_0(H), \end{cases} \tag{15}$$

where γ is a positive number, $n_0(H) = [Hg(H)]$, $[a]$ denotes the whole part of number a, $g(\cdot)$ is a slowly increasing function, namely

$$\lim_{H\to\infty} g(H) = +\infty, \quad \lim_{H\to\infty} \frac{g(H)}{H^\alpha} = 0,$$

for each $\alpha > 0$.

Theorem 4 *Let $Ee_1^{2\beta} < \infty$ for some $\beta > 2$ and $0 < \gamma < \beta/2$. Then mean duration of the procedure (5),(6) satisfy the limiting relationship*

$$\lim_{H\to\infty} \frac{E_\theta T(H)}{H} = \frac{1}{tr\mathbf{\Phi}(0)},$$

where

$$\mathbf{\Phi}(i) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{\phi}(k)\boldsymbol{\phi}'(k+i), \quad i = 0, \pm 1, \ldots$$

Theorem 5 *Let $E|e_1|^{2+\delta} < \infty$ for some $\delta > 0$. Then*

$$\sqrt{H}(\boldsymbol{\theta}^*(H) - \boldsymbol{\theta}) \xrightarrow{d} N(0, \boldsymbol{W}^{-1}\mathbf{\Gamma}(\boldsymbol{W}^{-1})'),$$

where

$$\mathbf{\Gamma} = tr\mathbf{\Phi}(0) \sum_{i=-s}^{s} \mathbf{\Phi}(i) \sum_{m=0}^{s-|i|} c_m c_{m+|i|},$$

$$\boldsymbol{W} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{\phi}(k)\boldsymbol{X}'_k.$$

Thus the sequential estimator $\boldsymbol{\theta}^*(H)$ is asymptotically normal as the estimate obtained by the technique of instrumental variables.

6. CONCLUSION

The paper considers the identification problem for the ARMAX system in non-asymptotic statement. The sequential estimation scheme which enables one to estimate the unknown parameters with a prescribed mean-square precision at the termination time has been proposed. It can be applied to different problems which may be reduced to the identification of the ARMAX system. For example, the proposed procedure can be used to solve the problem of estimating the spectral density of the ARMA process with a given mean-square accuracy. The proposed procedure is based on the method of instrumental variables. It includes several stages which are needed to gather enough information about the unknown parameters.

REFERENCES

Anderson, T.W. (1994). *The statistical analysis of time series.* John Wiley & Sons, Inc., New York, London, Sydney, Toronto.

Box, G.E.P. and C.M. Jenkins (1970). *Time series analysis, forecasting and control.* Holden- Day, San Francisco.

Goodwin, G.C. (1987). *Experiment design for system identification.* In Encyclopedia of Systems and Control. (M. Singh, ed.). Pergamon Press, Oxford.

Konev, V.V. and T.L. Lai (1995). Estimators with prescribed precision in stochastic regression models. *Sequential Anal.*, **14(3)**, 179–192.

Konev, V.V. and S.M. Pergamenshchikov (1996). On guaranteed identification of stochastic dynamic systems. *Proc. of the 13th World Congress IFAC, San Francisco,* Vol.H, 533–538.

Kumar, P.R. (1985). A survey of some results in stochastic adaptive control. *SIAM J. Control and Optimization*, **23**, 329–380.

Ljung, L. (1987). *System identification. Theory for user.* Prentice-Hall, Englewood Cliffs, N.J.