ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2015 Управление, вычислительная техника и информатика № 1 (30)

UDK 519.233.22

M.I. Kusainov

ON OPTIMAL ADAPTIVE PREDICTION OF MULTIVARIATE ARMA(1,1) PROCESS

The problem of asymptotic efficiency of adaptive one-step predictors for ARMA(1,1) process with unknown dynamic parameters is considered. The predictors are based on the truncated estimators of the unknown matrix. The truncated estimation method is a modification of the truncated sequential estimation method, that yields estimators with a given accuracy by samples of fixed size. The criterion of prediction optimality is based on the loss function, defined as a linear combination of sample size and squared prediction error's sample mean. The cases of known and unknown variance of the noise model are studied. In the latter case the optimal sample size is a special stopping time.
Keywords: adaptive predictors; asymptotic risk efficiency; multivariate ARMA; optimal sample size; stopping time; truncated parameter estimators.

According to Ljung's concept of construction of complete probabilistic models of dynamic systems, the prediction is a crucial part of it (see [1, 2]). A model is said to be useful if it allows one to make predictions of high statistical quality. Models of dynamic systems often have unknown parameters, which demand estimation in order to build adaptive predictors. The quality of adaptive prediction is explicitly dependent on the chosen estimators of model parameters.

There is a wide variety of possible estimation methods. For example, the sequential estimation method makes it possible to obtain estimators with guaranteed accuracy by samples of finite but random and unbounded size (see, e.g., [3] among others). The more modern truncated sequential estimation method yields estimators with prescribed accuracy by samples of random but bounded size (see, e.g., [4]).

This work suggests predictors based upon the truncated estimators of parameters introduced in [5, 6] as a modification of the truncated sequential estimators. Truncated estimators were constructed for ratio type functionals and are designed to use samples of fixed (non-random) size and have guaranteed accuracy in the sense of the L_{2m}-norm, $m \geq 1$.

The requirement of both good quality of predictions and reasonable duration of observations needed to achieve one is formulated as a risk efficiency problem. The criterion is given by certain loss functions and optimization is performed based on it. The loss function describing sample mean of squared prediction errors and sample size as well as the corresponding risk as applied to scalar AR(1) were examined in [7]. It was shown that the least squares estimators of the dynamic parameter are asymptotically risk efficient. Later, this result was refined and extended to other stochastic models in [8], using the sequential estimators of unknown parameters.

In this paper we construct and investigate real-time predictors based on truncated estimators in the case of more general model. We consider the problem of the risk minimization associated with size of a sample and predictions of values of a stable multivariate ARMA(1,1) process with unknown dynamic matrix parameter. The proposed procedure is shown to be asymptotically risk efficient as the cost of prediction error tends to infinity.

The same problem for scalar AR(1) case was considered in [9], multivariate AR(1) in [10]. The ARMA model was studied in [1, 2] among others. A thorough review of risk efficient parameter estimation and adaptive prediction problem for autoregressive processes was recently made in [11] (see the references therein as well).

1. Problem statement

Consider the multivariate stable ARMA(1,1) process satisfying the equation

$$x(k) = \Lambda x(k-1) + \xi(k) + M\xi(k-1), \quad k \geq 1, \tag{1}$$

where Λ and M are $p \times p$ matrix parameters with eigenvalues from the unit circle to provide the process stability (henceforth we shall refer to such matrices as "stable" ones). We assume the parameter Λ to be unknown and M to be known. The random vectors $\xi(k)$ for $k \geq 1$ are independent and identically distributed (i.i.d.) with zero mean and finite variance $\sigma^2 = E \| \xi(1) \|^2$, we also assume the components $\xi_j(k), j = \overline{1,p},$ to be uncorrelated and i.d. so that the covariance matrix $\Sigma = E\xi(1)\xi'(1)$ is diagonal with elements σ^2 / p. Denote the Λ stable region $\Lambda^0 \subset \mathbb{R}^{p\times p}$.

It is known that the optimal in the mean square sense one-step predictor is the conditional expectation of the process with respect to its past, i.e.

$$x^{opt}(k) = \Lambda x(k-1) + M\xi(k-1), \quad k \geq 1.$$

Since both the parameter Λ and the value of $\xi(k-1)$ are unknown, it is natural to replace them with some estimators $\tilde{\Lambda}_k$ and $\tilde{\xi}(k-1)$, which we specify in Section 2 below.

Define adaptive predictors as the following (see, e.g., [1, 12]):

$$\tilde{x}(k) = \tilde{\Lambda}_{k-1} x(k-1) + M\tilde{\xi}(k-1), \quad k \geq 1, \tag{2}$$

for which the corresponding prediction errors have the following form

$$\tilde{e}(k) = x(k) - \tilde{x}(k) = (\Lambda - \tilde{\Lambda}_{k-1})x(k-1) + M(\xi(k-1) - \tilde{\xi}(k-1)) + \xi(k).$$

Let $e^2(n)$ denote the sample mean of squared prediction error

$$e^2(n) = \frac{1}{n}\sum_{k=1}^{n} \| \tilde{e}(k) \|^2 .$$

Define the loss function

$$L_n = \frac{A}{n} e^2(n) + n,$$

where the parameter $A(>0)$ is the cost of prediction error.

The corresponding risk function

$$R_n = E_\theta L_n = \frac{A}{n} E_\theta e^2(n) + n, \tag{3}$$

E_θ denotes expectation under the distribution P_θ with the given parameter $\theta = (\lambda_{11}, ..., \lambda_{pp}, \mu_{11}, ..., \mu_{pp}, \sigma^2)$. Define the set Θ such that for $\theta \in \Theta$ the matrices Λ and M are stable and $\sigma^2 > 0$.

The main aim is to minimize the risk R_n on the sample size n.

We consider the cases of known and unknown σ^2.

2. Main result

In this section we solve the stated optimization problem under different conditions on model parameters.

We use, similarly to [10], the truncated estimation method introduced in [5]. This method makes it possible to obtain the ratio type estimators with guaranteed accuracy using a sample of fixed size. Such quality may essentially simplify investigation of analytical properties in various adaptive procedures.

Let the truncated estimators of the autoregressive parameter Λ be based on the following Yule-Walker type estimators

$$\Lambda_k = \overline{\Phi}_k \overline{G}_k^{-1}, \quad k \geq 2, \quad \Lambda_0 = \Lambda_1 = 0, \tag{4}$$

$$\overline{\Phi}_k = \frac{1}{k-1}\sum_{i=2}^{k} x(i)x'(i-2), \quad \overline{G}_k = \frac{1}{k-1}\sum_{i=2}^{k} x(i-1)x'(i-2)$$

and have the form

$$\tilde{\Lambda}_k = \Lambda_k \chi\left(|\overline{\Delta}_k| \geq H_k\right), \quad k \geq 2. \tag{5}$$

Here $\overline{\Delta}_k = \det(\overline{G}_k)$, the notation $\chi(B)$ means the indicator function of the set B and

$$H_k = \log^{-1/2} k. \tag{6}$$

We note that according to [5], H_k can be taken as any decreasing slowly changing positive function.

We take the estimators for $\xi(k)$ in the following form

$$\tilde{\xi}(k)=\sum_{i=0}^{k-1}(-M)^i(x(k-i)-\tilde{\Lambda}_k x(k-1-i)),\quad k\geq 1. \tag{7}$$

This way the prediction error can be rewritten as

$$\tilde{e}(k)=\xi(k)+(-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i).$$

2.1. Known σ^2 case

If the noise variance σ^2 is known, instead of Λ_k in (2) we shall use the projection of estimators (5) onto a closed ball $B\in R^{p\times p}$, such that $\Lambda^0\subset B$

$$\Lambda_k^*=proj_{[-1,1]}\tilde{\Lambda}_k,$$

ensuring

$$\|\Lambda_k^*-\Lambda\|\leq d_B, \tag{8}$$

where d_B is the diameter of B. Given that σ^2 is known, the property (8) allows one to weaken the noise moment conditions compared to the more general case of unknown σ^2 (see Section 2.2 below).

Rewrite the formulae accordingly

$$\xi^*(k)=\sum_{i=0}^{k-1}(-M)^i(x(k-i)-\Lambda_k^* x(k-1-i)),\quad x^*(k)=\Lambda_{k-1}^*x(k-1)+M\xi^*(k-1),$$

$$e^*(k)=x(k)-x^*(k)=\xi(k)+(-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\Lambda_{k-1}^*-\Lambda)x(k-1-i),$$

$$e_*^2(n)=\frac{1}{n}\sum_{k=1}^{n}\|e^*(k)\|^2,\quad L_n=\frac{A}{n}e_*^2(n)+n,\quad R_n=E_\theta L_n=\frac{A}{n}E_\theta e_*^2(n)+n.$$

To minimize the risk R_n we rewrite it in the form

$$R_n=\frac{A}{n}(\sigma^2+D_n)+n, \tag{9}$$

where

$$D_n=\frac{1}{n}\sum_{k=1}^{n}E_\theta\|x^*(k)-x^{opt}(k)\|^2=\frac{1}{n}\sum_{k=1}^{n}E_\theta\left\|(-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\Lambda_{k-1}^*-\Lambda)x(k-1-i)\right\|^2.$$

We shall use the properties of the estimators $\tilde{\Lambda}_k$ given in Lemma 1 below.

Define $k_0=\max\{p,[e^{|\Delta|^{-2}}]_1\}$, where $[a]_1$ denotes the integer part of a and

$$\Delta=\underline{\lim}_{k\to\infty}\bar{\Delta}_k,\quad P_\theta-\text{a.s.}$$

Now we establish the conditions on the system parameters, under which $\Delta\neq 0$. It can be shown, similarly to, e.g., [13], that due to ergodicity of the process $(x(k))_{k\geq 0}$:

$$\bar{G}_k=\frac{1}{k-1}\sum_{i=2}^{k}x(i-1)x'(i-2)\xrightarrow[k\to\infty]{}G,\quad P_\theta-\text{a.s.},$$

where

$$G=\Lambda F+\mathrm{M}\Sigma,$$
$$F=\sum_{i\geq 0}\Lambda^i S\Lambda'^i,\quad S=\Lambda\Sigma M'+M\Sigma\Lambda'+\Sigma+M\Sigma M'. \tag{10}$$

The condition for $\Delta\neq 0$ is thus nondegeneracy of G. For example, in the scalar case $p=1$ we have

$$G=\frac{(\Lambda+M)(1+\Lambda M)}{1-\Lambda^2}\sigma^2,$$

which is the first order autocovariance; the condition is $\Lambda+M\neq 0$ as stability of the process implies $1+\Lambda M\neq 0$.

From here on C denotes those non-negative constants, the values of which are not critical.

Lemma 1. Assume the model (1) and let for some integer $m\geq 1$ the conditions

$$E\|\xi(1)\|^{4pm}<\infty,\quad E\|x(0)\|^{4pm}<\infty \tag{11}$$

be true. Assume also that the matrix G defined in (10) is nondegenerate. Then the truncated estimators $\tilde{\Lambda}_k$ satisfy

(i) for $1 \le k < k_0$

$$E_\theta \| \tilde{\Lambda}_k - \Lambda \|^{2m} \le C; \tag{12}$$

(ii) for $k \ge k_0$

$$E_\theta \| \tilde{\Lambda}_k - \Lambda \|^{2m} \le \frac{C \log^m k}{k^m}. \tag{13}$$

The proof of Lemma 1 is similar to that of the assertion (31) in [5] and Lemma 1 in [10].

Now we rewrite D_n in the form

$$D_n = \frac{\sigma^2}{n \cdot p} \sum_{k=1}^{n} \| M^k \|^2 + \frac{1}{n} \sum_{k=1}^{n} E_\theta \left\| \sum_{i=0}^{k-1} (-M)^i (\Lambda^*_{k-1} - \Lambda) x(k-1-i) \right\|^2 -$$

$$- \frac{2}{n} \sum_{k=1}^{n} E_\theta [(-M)^k \xi(0)]' \sum_{i=0}^{k-1} (-M)^i E_\theta (\Lambda^*_{k-1} - \Lambda) x(k-1-i). \tag{14}$$

Consider the first summand. It is known that $M^k = TJ^kT^{-1}$, where J is Jordan canonical form of M and the columns of T are generalized eigenvectors of M. It then can be shown that $\| M^k \| \le C \max |\mu_i|^k$, where $\mu_i,\ i = \overline{1, p}$ are the eigenvalues of M. Boundedness of the series now follows from the stability of M, so we have

$$\frac{\sigma^2}{n \cdot p} \sum_{k=1}^{n} \| M^k \|^2 \le \frac{C}{n}.$$

Consider the second summand of (14):

$$\frac{1}{n} \sum_{k=1}^{n} E_\theta \left\| \sum_{i=0}^{k-1} (-M)^i (\Lambda^*_{k-1} - \Lambda) x(k-1-i) \right\|^2 \le$$

$$\le \frac{1}{n} \sum_{k=1}^{n} \sum_{i=0}^{k-1} E_\theta \| M^i (\Lambda^*_{k-1} - \Lambda) x(k-1-i) \|^2 +$$

$$+ \frac{1}{n} \sum_{k=1}^{n} \sum_{j \ne i}^{k-1} \sum_{i=0}^{k-1} E_\theta \| \Lambda^*_{k-1} - \Lambda \|^2 \| M^j x(k-1-j) \| \cdot \| M^i x(k-1-i) \|. \tag{15}$$

If the conditions $E \| \xi(1) \|^{4p} < \infty, E \| x(0) \|^{4p} < \infty$ hold then using the Cauchy-Schwarz-Bunyakovsky inequality, (8) and (13) we get

$$\frac{1}{n} \sum_{k=1}^{n} \sum_{i=0}^{k-1} E_\theta \| M^i (\Lambda^*_{k-1} - \Lambda) x(k-1-i) \|^2 \le$$

$$\le \frac{C}{n} \sum_{k=1}^{n} \sum_{i=0}^{k-1} \sqrt{E_\theta \| \Lambda^*_{k-1} - \Lambda \|^2} \|M^i\|^2 \le \frac{C}{n} \sum_{k=1}^{n} \frac{\log^{1/2} k}{k^{1/2}} \le \frac{C \log^{1/2} n}{n^{1/2}}.$$

The second summand of (15) and the third summand of D_n itself are treated the same way and don't have impact on the result. Thus, usage of estimators Λ^*_k in adaptive predictors yields

$$D_n \le C n^{-1/2} \log^{1/2} n = o(1) \quad \text{as} \quad n \to \infty.$$

Considering (9), the stated risk minimization problem reduces to minimization of the principal term

$$R_n \approx \frac{A}{n} \sigma^2 + n \longrightarrow \min_n.$$

Since the parameter σ^2 is known, the expression can be easily minimized with the optimal sample size

$$n_A^o = A^{1/2} \sigma. \tag{16}$$

The corresponding approximate minimal risk value is

$$R_{n_A^o} = 2A^{1/2}\sigma + O(A^{1/4} \log^{1/2} A) \quad \text{as} \quad A \to \infty. \tag{17}$$

Thus, the following corollary is true.

Corollary 1. Assume that $E\|\xi(1)\|^{4p}<\infty, E\|x(0)\|^{4p}<\infty$ and the variance σ^2 is known. Then the number n_A^o defined in (16) minimizes the risk function R_n defined in (9) and the asymptotic formula (17) for $R_{n_A^o}$ holds.

2.2. Unknown σ^2 case

Since σ^2 is directly involved in the expression (9) for R_n, the optimal sample size can not be obtained as before. Similarly to [7, 8, 10], one uses the stopping time T_A as an estimator of n_A^o, replacing σ^2 in its definition with an estimator $\tilde{\sigma}_n^{\,2}$

$$T_A=\inf_{n\ge n_A}\left\{n\ge A^{1/2}\tilde{\sigma}_n\right\}, \tag{18}$$

where n_A is the initial sample size depending on A and specified below (see Theorem 1),

$$\tilde{\sigma}_n^{\,2}=\frac{p}{p+\|M\|^2}\frac{1}{n}\sum_{k=1}^{n}\|x(k)-\tilde{\Lambda}_n x(k-1)\|^2 . \tag{19}$$

The choice of estimator is motivated by the fact that using the strong law of large numbers we have

$$\frac{1}{n}\sum_{k=1}^{n}\|x(k)-\Lambda x(k-1)\|^2=\frac{1}{n}\sum_{k=1}^{n}\|\xi(k)+M\xi(k-1)\|^2\xrightarrow[n\to\infty]{}\sigma^2\left(1+\frac{\|M\|^2}{p}\right),\quad P_\theta-\text{a.s.}$$

In this section we define predictors of $x(k)$ using truncated estimators $\tilde{\Lambda}_k$ instead of Λ_k^*. Rewrite the needed formulae

$$\tilde{\xi}(k)=\sum_{i=0}^{k-1}(-M)^i(x(k-i)-\tilde{\Lambda}_k x(k-1-i)),\quad \tilde{x}(k)=\tilde{\Lambda}_{k-1}x(k-1)+M\tilde{\xi}(k), \tag{20}$$

$$\tilde{e}(k)=x(k)-\tilde{x}(k)=\xi(k)+(-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i),$$

$$\overline{e^2}(n)=\frac{1}{n}\sum_{k=1}^{n}\|\tilde{e}(k)\|^2,\quad \overline{L}_n=\frac{A}{n}\overline{e^2}(n)+n,\quad R_n=E_\theta\overline{L}_n=\frac{A}{n}E_\theta\overline{e^2}(n)+n. \tag{21}$$

Analogously to [7], we prove the asymptotic equivalence of T_A and n_A^o in the almost surely and mean senses (see (23), (24) below) and the optimality of the adaptive prediction procedure in the sense of equivalence of the obviously modified risk

$$\overline{R}_A=E_\theta\overline{L}_{T_A}=AE_\theta\frac{1}{T_A}\overline{e^2}(T_A)+E_\theta T_A \tag{22}$$

and $R_{n_A^o}$, see (8).

Theorem 1. Assume that $E\|\xi(1)\|^{16p}<\infty$, $E\|x(0)\|^{16p}<\infty$ and n_A in (18) is such that $\max\{k_0, A^r\log^2 A\}\le n_A=o(A^{1/2})$ with $r\in(2/5,1/2)$. Let the predictors $\tilde{x}(k)$ be defined by (20) and the risk functions defined by (21), (22). Then for every $\theta\in\Theta$

$$\frac{T_A}{n_A^o}\xrightarrow[A\to\infty]{}1,\quad P_\theta-\text{a.s.}, \tag{23}$$

$$\frac{E_\theta T_A}{n_A^o}\xrightarrow[A\to\infty]{}1, \tag{24}$$

$$\frac{\overline{R}_A}{R_{n_A^o}}\xrightarrow[A\to\infty]{}1. \tag{25}$$

The proof of Theorem 1 is presented in Section 3.

Remark 1. The third assertion of Theorem 1 is also true for predictors based on Λ_k^*.

3. Proof of Theorem 1

First, we prove the properties (23), (24) of the stopping time T_A.

From the conditions of Theorem 1 on noise moments for $\Lambda \in \Lambda^0$ it follows

$$\sup_{k\ge 0} E_\theta \| x(k) \|^{16p} \le C. \tag{26}$$

Denote

$$C_M = \frac{p}{p + \| M \|^2}.$$

Rewrite formula (19) for $\tilde{\sigma}_n^{\ 2}$ using (1):

$$\tilde{\sigma}_n^{\ 2} = \frac{C_M}{n} \sum_{k=1}^{n} \| \xi(k) + M\xi(k-1) + (\Lambda - \tilde{\Lambda}_n) x(k-1) \|^2 =$$

$$= \frac{C_M}{n} \sum_{k=1}^{n} \| \xi(k) + M\xi(k-1) \|^2 + W_n + \nu_n, \tag{27}$$

where

$$W_n = \frac{C_M}{n} \sum_{k=1}^{n} \| (\tilde{\Lambda}_n - \Lambda) x(k-1) \|^2, \quad \nu_n = -\frac{2C_M}{n} \sum_{k=1}^{n} [(\tilde{\Lambda}_n - \Lambda) x(k-1)]' (\xi(k) + M\xi(k-1)).$$

Now we show that

$$\tilde{\sigma}_n^2 \xrightarrow[n\to\infty]{} \sigma^2, \quad P_\theta - \text{a.s.} \tag{28}$$

Consider W_n. It follows from the definition (5) of the truncated estimators $\tilde{\Lambda}_n$ that they are asymptotically equivalent to the corresponding correlation estimators (4), see, e.g., p. 8 in [5]. Since the estimators (4) are strongly consistent we have

$$\tilde{\Lambda}_n - \Lambda \xrightarrow[n\to\infty]{} 0, \quad P_\theta - \text{a.s.}$$

Given that

$$\frac{1}{n} \sum_{k=1}^{n} x(k-1) x'(k-1) \xrightarrow[n\to\infty]{} F, \quad P_\theta - \text{a.s.},$$

where F is a constant matrix (see (10)), it follows that

$$W_n \xrightarrow[n\to\infty]{} 0, \quad P_\theta - \text{a.s.}$$

Similar arguments are used to show

$$\nu_n \xrightarrow[n\to\infty]{} 0, \quad P_\theta - \text{a.s.}$$

The relation (28) obviously follows from these facts, the representation (27) and strong law of large numbers.

From the definition (18) of T_A it follows that with P_θ-probability one $T_A \to \infty$ as $A \to \infty$. Therefore, by (28) we have $\tilde{\sigma}_{T_A}^2 \to \sigma^2$ P_θ – a.s. and hence

$$\frac{T_A}{A^{1/2}\sigma} \xrightarrow[A\to\infty]{} 1, \quad P_\theta - \text{a.s.}$$

For proof of (24) we introduce for any positive A the auxiliary sequence of numbers $\gamma_{A,n}$

$$\gamma_{A,n} = n^2 A^{-1} \frac{1}{2\log A}, \quad n \ge 1.$$

Denote

$$m_n = \frac{C_M}{n} \sum_{k=1}^{n} \left(\| (\xi(k) + M\xi(k-1) \|^2 - \left(\sigma^2 + \frac{\| M \|^2}{p} \sigma^2 \right) \right).$$

By the definition of T_A and (27) we have

$$E_\theta T_A \le n_A + \sum_{n\ge n_A} P_\theta\left(n^2 A^{-1} \le \frac{C_M}{n}\sum_{k=1}^{n} \| \xi(k) + M\xi(k-1)\|^2 + W_n + \nu_n\right) \le$$

$$\le n_A + \sum_{n\ge n_A}\left\{P_\theta\left(n^2 A^{-1} \le \frac{C_M}{n}\sum_{k=1}^{n} \| \xi(k) + M\xi(k-1)\|^2 + \gamma_{A,n}\right) + P_\theta(W_n + |\nu_n| > \gamma_{A,n})\right\} \le$$

$$\le n_A + \sum_{n\ge n_A}\{P_\theta\left(n^2 A^{-1} \le \sigma^2 + 2\gamma_{A,n}\right) +$$

$$+ P_\theta(|\nu_n| > \gamma_{A,n}/2) + P_\theta(W_n > \gamma_{A,n}/2) + P_\theta(|m_n| > \gamma_{A,n})\}. \tag{29}$$

Note that

$$n_A + \sum_{n\ge n_A} P_\theta\left(n^2 A^{-1} \le \sigma^2 + 2\gamma_{A,n}\right) = n_A + \sum_{n\ge n_A}^{n_A^*-1} 1 = n_A^*,$$

where

$$n_A^* = \inf_{n\ge n_A}\{n^2 A^{-1} > \sigma^2 + 2\gamma_{A,n}\} = \left[A^{1/2}\sigma\left(1+\frac{1}{\log A - 1}\right)^{1/2}\right]_1 + 1.$$

Therefore

$$\frac{n_A + \sum_{n\ge n_A} P_\theta\left(n^2 A^{-1} \le \sigma^2 + 2\gamma_{A,n}\right)}{A^{1/2}\sigma} \xrightarrow[A\to\infty]{} 1.$$

Now we show that the other summands in the right-hand side of (29) vanish as $A \to \infty$, when normalized appropriately.

Consider the probability $P_\theta(|\nu_n| > \gamma_{A,n})$. According to (26), the Chebyshev inequality and the Cauchy-Schwarz-Bunyakovsky inequality for $n \ge n_A$ we have

$$P_\theta(|\nu_n| > \gamma_{A,n}/2) = P_\theta\left(\frac{2C_M}{n}\left|\sum_{k=1}^{n}[(\tilde\Lambda_n - \Lambda)x(k-1)]'(\xi(k) + M\xi(k-1))\right| > \gamma_{A,n}/2\right) \le$$

$$\le \frac{C}{n\gamma_{A,n}}\sum_{k=1}^{n}\left(E_\theta \|\tilde\Lambda_n - \Lambda\|^2 \, E_\theta \|x(k-1)\|^2 \|\xi(k) + M\xi(k-1)\|^2\right)^{1/2} \le$$

$$\le \frac{C\log^{1/2} n}{\sqrt{n}\gamma_{A,n}} \le C A \log A \frac{\log^{1/2} n}{n^{5/2}}.$$

From the assumptions on n_A it follows that

$$A^{-1/2}\sum_{n\ge n_A} P_\theta(|\nu_n| > \gamma_{A,n}) \le CA^{1/2}\log A \sum_{n\ge n_A}\frac{\log^{1/2} n}{n^{5/2}} \le$$

$$\le CA^{1/2}\log^{3/2} A \cdot n_A^{-3/2} \le CA^{-\frac{3r-1}{2}}\log^{-3/2} A \xrightarrow[A\to\infty]{} 0.$$

The probability $P_\theta(W_n > \gamma_{A,n})$ is treated analogously.

As for the probability $P_\theta(|m_n| > \gamma_{A,n})$, note that m_n is sum of martingales, thus the Chebyshev inequality and the Burkholder inequality yield

$$P_\theta(|m_n| > \gamma_{A,n}) \le P_\theta\left(\frac{C_M}{n}\left|\sum_{k=1}^{n}(\|(\xi(k) + M\xi(k-1)\|^2 - (\sigma^2 + \|M\|\sigma^2))\right| > \gamma_{A,n}\right) \le$$

$$\le C\gamma_{A,n}^{-2} n^{-1} = 4CA^2\log^2 A \cdot n^{-5}.$$

Therefore, by assumptions on n_A

$$A^{-1/2}\sum_{n\ge n_A} P_\theta\left(|m_n| > \gamma_{A,n}\right) \le CA^{3/2}\log^2 A \sum_{n\ge n_A} n^{-5} \le$$

$$\le CA^{3/2}\log^2 A \cdot n_A^{-4} \le CA^{-\frac{8r-3}{2}}\log^{-6} A \xrightarrow[A\to\infty]{} 0.$$

Then from (29) it follows that

$$\overline{\lim_{A\to\infty}} \frac{E_\theta T_A}{A^{1/2}\sigma} \le 1. \tag{30}$$

Same arguments can be used to show

$$\varliminf_{A\to\infty} \frac{E_\theta T_A}{A^{1/2}\sigma} \geq 1$$

and thus, in view of (30) the assertion (24) holds.

Regarding (25), rewrite its left-hand side using (17) and (22)

$$\frac{\overline{R}_A}{R_{n_A^o}} = \frac{AE_\theta \frac{1}{T_A}\overline{e^2}(T_A) + E_\theta T_A}{2A^{1/2}\sigma + O(A^{1/4}\log^{1/2} A)}. \tag{31}$$

From (24) and (31) it follows that to prove (25) it suffices to show the convergence

$$A^{1/2} E_\theta \frac{1}{T_A \sigma}\overline{e^2}(T_A) \xrightarrow[A\to\infty]{} 1. \tag{32}$$

Define

$$N' = [(\sigma-\epsilon)A^{1/2}]_1, \quad N'' = [(\sigma+\epsilon)A^{1/2}]_1 + 1, \quad 0<\epsilon<\sigma.$$

We will need the following properties

$$P_\theta(T_A < N') = O(A^{-r}), \quad P_\theta(T_A > N'') = O(A^{-1}), \tag{33}$$

which we prove similarly to Lemma 4 of [7].

Denote $\delta_1 = \sigma^2 - (\sigma-\epsilon)^2$. Using non-negativeness of W_n, definitions of T_A and $\tilde{\sigma}_n^2$ one gets

$$\begin{aligned} P_\theta(T_A < N') &\leq P_\theta(T_A < (\sigma-\epsilon)A^{1/2}) = \\ &= P_\theta\left(\tilde{\sigma}_n^2 \leq A^{-1}n^2, \text{ for some } n_A \leq n \leq (\sigma-\epsilon)A^{1/2}\right) \leq \\ &\leq P_\theta\left(\frac{C_M}{n}\sum_{k=1}^n \|\xi(k)+M\xi(k-1)\|^2 + \nu_n \leq (\sigma-\epsilon)^2, \quad \text{for some } n \geq n_A\right) = \\ &= P_\theta\left(m_n + \nu_n \geq \delta_1, \quad \text{for some } n\geq n_A\right) \leq \\ &\leq \sum_{n\geq n_A} P_\theta\left(|m_n| \geq \frac{\delta_1}{2}\right) + \sum_{n\geq n_A} P_\theta(|\nu_n| \geq \delta_1/2). \end{aligned} \tag{34}$$

Consider the first summand. By the Chebyshev inequality, (26) and the Burkholder inequality

$$\begin{aligned} &\sum_{n\geq n_A} P_\theta\left(\frac{C_M}{n}\left|\sum_{k=1}^n (\|(\xi(k)+M\xi(k-1)\|^2 - (\sigma^2 + \|M\|\sigma^2))\right| > \frac{\delta_1}{2}\right) \leq \\ &\leq C\sum_{n\geq n_A}\frac{E_\theta |m_n|^4}{n^4} \leq C\sum_{n\geq n_A} n^{-2} \leq CA^{-r}\log^{-2} A. \end{aligned} \tag{35}$$

The following is proved analogously to how (29) is treated

$$\sup_{n\geq 1} E_\theta(n\log^{-1} n\cdot|\nu_n|)^2 < \infty.$$

Thus,

$$\sum_{n\geq n_A} P_\theta(|\nu_n| \geq \delta_1/2) \leq C\sum_{n\geq n_A} n^{-2}\log^2 n \leq Cn_A^{-1}\log^2 A \leq CA^{-r}. \tag{36}$$

The first property of (33) follows from (34)–(36).

Prove the second property of (33). Denote $\delta_2 = (\sigma+\epsilon)^2 - \sigma^2$. Then, by definition (18) of T_A and (27)

$$\begin{aligned} P_\theta(T_A > N'') &\leq P_\theta\left(\frac{C_M}{N''}\sum_{k=1}^{N''}\|\xi(k)+M\xi(k-1)\|^2 + W_{N''} + \nu_{N''} > A^{-1}(N'')^2\right) \leq \\ &\leq P_\theta\left(\frac{C_M}{N''}\sum_{k=1}^{N''}\|\xi(k)+M\xi(k-1)\|^2 + |W_{N''}+\nu_{N''}| > (\sigma+\epsilon)^2\right) \leq \\ &\leq P_\theta\left(|m_{N''}| > \delta_2/2\right) + P_\theta\left(|W_{N''}+\nu_{N''}| > \delta_2/2\right). \end{aligned}$$

By the Chebyshev and Burkholder inequalities

$$P_\theta\left(|m_{N''}| > \delta_2/2\right) \leq C(N'')^{-2} = O(A^{-1}), \quad P_\theta\left(|W_{N''}+\nu_{N''}| > \delta_2/2\right) \leq C(N'')^{-2} = O(A^{-1}).$$

Thus, the second assertion in (33) holds true.

To prove (32) we show that

$$A^{1/2}E_\theta \frac{1}{T_A}\overline{e^2}(T_A)\chi(T_A < N') \xrightarrow[A\to\infty]{} 0,\ A^{1/2}E_\theta \frac{1}{T_A}\overline{e^2}(T_A)\chi(T_A > N'') \xrightarrow[A\to\infty]{} 0, \tag{37}$$

$$A^{1/2}E_\theta \frac{1}{T_A\sigma}\overline{e^2}(T_A)\chi(N' \le T_A \le N'') \xrightarrow[A\to\infty]{} 1. \tag{38}$$

Prove the first assertion in (37). By the definition of $\overline{e^2}(k)$ we get

$$A^{1/2}E_\theta \frac{1}{T_A}\overline{e^2}(T_A)\chi(T_A < N') =$$

$$= A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\left\|(-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right\|^2 \chi(T_A < N') +$$

$$+2A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\xi'(k)\left((-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right)\chi(T_A < N') +$$

$$+A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\|\xi(k)\|^2\chi(T_A < N'). \tag{39}$$

Consider the first summand. By the Cauchy-Schwarz-Bunyakovsky inequality and the definition of T_A assumptions on n_A and $r,$ the properties (33) and Lemma 1 we have

$$A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\left\|(-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right\|^2 \chi(T_A < N') \le$$

$$\le A^{1/2}P_\theta^{1/2}(T_A < N')\frac{1}{n_A^2}\sum_{k=1}^{N'}\sqrt{E_\theta\left\|(-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right\|^4}.$$

Examine the expression under the root square. The most significant summand is treated using the Cauchy-Schwarz-Bunyakovsky inequality

$$E_\theta\left\|\sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right\|^4 \le$$

$$\le \sqrt{E_\theta\|\tilde{\Lambda}_{k-1}-\Lambda\|^8\ E_\theta\left\|\sum_{i=0}^{k-1}(-M)^i x(k-1-i)\right\|^8} \le \frac{C\log^2 k}{k^2}\sqrt{E_\theta\left\|\sum_{i=0}^{k-1}(-M)^i x(k-1-i)\right\|^8}.$$

It can be easily shown, employing the Hölder inequality, that

$$E_\theta\left\|\sum_{i=0}^{k-1}(-M)^i x(k-1-i)\right\|^8 \le E_\theta\left(\sum_{i=0}^{k-1}\left\|M^i\right\|\cdot\left\|x(k-1-i)\right\|\right)^8 \le C\left(\sum_{i=0}^{k-1}\left\|M^{i/2}\right\|^{\frac{8}{7}}\right)^7\cdot\sum_{i=0}^{k-1}\left\|M^{i/2}\right\|^8 \le C$$

and hence, by the assumptions on n_A and $r,$ the properties (33) and Lemma 1 we have

$$A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\left\|\sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right\|^2 \chi(T_A < N') \le$$

$$\le CA^{1/2}P_\theta^{1/2}(T_A < N')\frac{1}{n_A^2}\sum_{k=1}^{N'}\frac{\log k}{k} \le CA^{-\frac{5r-1}{2}}\log^{-2}A \xrightarrow[A\to\infty]{} 0.$$

Consider the second summand of (39). The Doob's maximal inequality for martingales (see, e.g., [14]) and the Cauchy-Schwarz-Bunyakovsky inequality yield

$$A^{1/2}E_\theta \frac{1}{T_A^2}\left|\sum_{k=1}^{T_A}\xi'(k)\left((-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right)\chi(T_A < N')\right| \le$$

$$\le A^{-\frac{r-1}{2}}\frac{1}{n_A^2}\sqrt{E_\theta\max_{1\le n\le N'}\left(\sum_{k=1}^{n}\xi'(k)\left((-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\right)\right)^2} \le$$

$$\le \sigma A^{-\frac{5r-1}{2}}\sqrt{\sum_{k=1}^{N'}E_\theta\left\|(-M)^k\xi(0) - \sum_{i=0}^{k-1}(-M)^i(\tilde{\Lambda}_{k-1}-\Lambda)x(k-1-i)\xi(0)\right\|^2} \le$$

$$\le CA^{-\frac{5r-1}{2}}\log A \xrightarrow[A\to\infty]{} 0.$$

Consider the last summand of (39). We have

$$A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\|\xi(k)\|^2\,\chi(T_A<N')\le A^{1/2}n_A^{-2}P_\theta^{1/2}(T_A<N')\sum_{k=1}^{N'}\sqrt{E_\theta\|\xi(k)\|^4}\le$$

$$\le CA^{-\frac{5r-1}{2}}\log^{-4}A\cdot N'\le CA^{-\frac{5r-2}{2}}\log^{-4}A\xrightarrow[A\to\infty]{}0.$$

Thus, the first part of (37) has been proved, similar arguments are applied to the second part with $\chi(T_A<N')$ replaced by $\chi(T_A>N'')$ to get

$$A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\left\|(-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\tilde\Lambda_{k-1}-\Lambda)x(k-1-i)\right\|^2\chi(T_A>N')\le CA^{-\frac{1}{2}}\xrightarrow[A\to\infty]{}0,$$

$$A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\xi'(k)\left((-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\tilde\Lambda_{k-1}-\Lambda)x(k-1-i)\right)\chi(T_A>N')\le CA^{-\frac{3}{4}}\log A\xrightarrow[A\to\infty]{}0,$$

$$A^{1/2}E_\theta \frac{1}{T_A^2}\sum_{k=1}^{T_A}\|\xi(k)\|^2\,\chi(T_A>N')\le CA^{-\frac{1}{4}}\xrightarrow[A\to\infty]{}0$$

and to (38) with $\chi(T_A<N')$ replaced by $\chi(N'\le T_A\le N'')$ to get

$$A^{1/2}E_\theta \frac{1}{T_A^2\sigma}\sum_{k=1}^{T_A}\left\|(-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\tilde\Lambda_{k-1}-\Lambda)x(k-1-i)\right\|^2\chi(N'\le T_A\le N'')\le CA^{-\frac{1}{2}}\log^2 A\xrightarrow[A\to\infty]{}0,$$

$$A^{1/2}E_\theta \frac{1}{T_A^2\sigma}\sum_{k=1}^{T_A}\xi'(k)\left((-M)^k\xi(0)-\sum_{i=0}^{k-1}(-M)^i(\tilde\Lambda_{k-1}-\Lambda)x(k-1-i)\right)\chi(N'\le T_A\le N'')\le CA^{-\frac{1}{2}}\log A\xrightarrow[A\to\infty]{}0.$$

Now we show that

$$A^{1/2}E_\theta \frac{1}{T_A^2\sigma}\sum_{k=1}^{T_A}\|\xi(k)\|^2\,\chi(N'\le T_A\le N'')\xrightarrow[A\to\infty]{}1.$$

To this end rewrite the left-hand side as follows

$$A^{1/2}E_\theta \frac{1}{T_A^2\sigma}\sum_{k=1}^{T_A}\|\xi(k)\|^2\,\chi(N'\le T_A\le N'')=$$

$$=A^{1/2}E_\theta \frac{1}{T_A^2\sigma}\sum_{k=1}^{T_A}(\|\xi(k)\|^2-\sigma^2)\chi(N'\le T_A\le N'')+A^{1/2}\sigma E_\theta\frac{1}{T_A}\chi(N'\le T_A\le N'').$$

We show that the first summand converges to 0 and the second one converges to 1. By the Doob's maximal inequality and the Cauchy-Schwarz-Bunyakovsky inequality

$$A^{1/2}E_\theta \frac{1}{T_A^2\sigma}\left|\sum_{k=1}^{T_A}(\|\xi(k)\|^2-\sigma^2)\right|\chi(N'\le T_A\le N'')\le CA^{1/2}\frac{1}{(N')^2}\left(E_\theta\max_{1\le n\le N''}\left(\sum_{k=1}^{n}(\|\xi(k)\|^2-\sigma^2)\right)^2\right)^{1/2}\le$$

$$\le CA^{-1/2}\left(\sum_{k=1}^{N''}E_\theta\|\xi(k)\|^4\right)^{1/2}\le CA^{-1/4}\xrightarrow[A\to\infty]{}0.$$

Consider the second summand. To prove its almost sure convergence to 1 it suffices (see, e.g., [15]) to show that

$$P_\theta\text{–}\lim_{A\to\infty}A^{1/2}\sigma\frac{1}{T_A}\chi(N'\le T_A\le N'')=1 \tag{40}$$

and that the family of random variables

$$Z=\left\{A^{1/2}\frac{1}{T_A}\chi(N'\le T_A\le N'')\right\}_{A\ge1} \tag{41}$$

is uniformly integrable.

The property (40) is fulfilled as, according to (23),

$$\frac{T_A}{A^{1/2}\sigma}\xrightarrow[A\to\infty]{}1,\quad P_\theta-\text{a.s.}$$

and hence

$$\lim_{A\to\infty}\chi(N'\le T_A\le N'')=\lim_{A\to\infty}\chi\left(1-\epsilon/\sigma\le\frac{T_A}{A^{1/2}\sigma}\le1+\epsilon/\sigma\right)=1,\quad P_\theta-\text{a.s.}$$

Property (41) holds true since Z is uniformly bounded.

4. Numerical simulation

To confirm theoretical results we performed numerical simulation programmed in MATLAB for 2-dimensional stable ARMA(1,1). The needed expected values are approximated by sample means of 100 realizations. E.g., the realizations of the stopping time T_A are

$$T_A^{(n)} = \inf_{k \geq n_A} \left\{ k \geq A^{1/2} \tilde{\sigma}_k^{(n)} \right\}, \quad n = \overline{1,100},$$

then its expectation $E_\theta T_A$ is computed as foolows

$$\widehat{E_\theta T_A} = \frac{1}{100} \sum_{n=1}^{100} T_A^{(n)}.$$

The initial value $x(0)$ and the noises $\xi(k)$ are generated from the multivariate Gaussian distribution

$$x(0) \sim N(0, I), \qquad \xi(k) \sim N\left(0, \frac{\sigma^2}{2} I\right),$$

Where I is the identity 2x2 matrix and thus, $E \| \xi(k) \|^2 = \sigma^2$.

We take the true value of the matrix parameters to be

$$\Lambda = \begin{pmatrix} 0.4 & 0.7 \\ 0.4 & -0.5 \end{pmatrix}, \qquad M = \begin{pmatrix} 0.1 & -0.2 \\ 0.7 & 0.4 \end{pmatrix}$$

with eigenvalues, respectively, $\lambda_1 = 0.6446$, $\lambda_2 = -0.7446$, $\mu_1 = 0.25 + 0.3428i$, $\mu_2 = 0.25 - 0.3428i$, which satisfy the stability conditions.

Tables 1(a, b) contain comparison of the estimates of $R_{n_A^o}$ and the value $2A^{1/2}\sigma$ as implied by (17) for the prediction error two values of prediction error cost A, as well as the estimates of $\frac{E_\theta T_A}{n_A^o}$ and $\frac{\overline{R}_A}{R_{n_A^o}}$.

Table 1

Estimates of the crucial values

a) The prediction error cost A = 5000

σ^2	n_A^o	$\frac{E_\theta T_A}{n_A^o}$	$R_{n_A^o}$	$\frac{R_{n_A^o}}{2A^{1/2}\sigma}$	$\frac{\overline{R}_A}{R_{n_A^o}}$
1	70.71	1.02	166.9	1.18	0.98
3	122.5	1.01	282.5	1.15	0.99
5	158.1	1.00	341.5	1.08	1

b) The prediction error cost A = 10000

σ^2	n_A^o	$\frac{E_\theta T_A}{n_A^o}$	$R_{n_A^o}$	$\frac{R_{n_A^o}}{2A^{1/2}\sigma}$	$\frac{\overline{R}_A}{R_{n_A^o}}$
1	100	1.01	222.3	1.12	0.99
3	173.2	0.99	377.5	1.09	1
5	223.6	1.00	468.1	1.04	1

As the tables imply, the ratio $\frac{E_\theta T_A}{n_A^o}$ converges to 1 with growth of the optimal sample size n_A^o, which is also reflected in the fact that the values of risks $\overline{R}_A$ and $R_{n_A^o}$, for cases of unknown and known noise variance respectively, are very close to each other. At the same time, the values of both risks are accurately approximated by $2A^{1/2}\sigma$ only if the prediction error cost and with it the optimal sample size are rather large.

5. Summary

This paper presents the problem of optimization of both one-step prediction quality and sample size for stable multivariate ARMA(1,1) process with unknown dynamic parameters. The cases of known and unknown noise variance were studied. In both cases optimization is performed based on the loss function describing the sample mean of squared prediction error. If the noise variance is unknown, the risk function depends on the mean of the duration of observations, defined as a stopping time in this case. It was shown that the risk functions are equivalent to each other asymptotically.

The adaptive predictors were constructed upon the basis of truncated estimators of the dynamic matrix parameter. The mentioned estimators have given statistical properties on a sample of fixed size. Usage of such estimators essentially simplifies analytical investigation of statistical properties of adaptive predictors and can be applied in various adaptive procedures (control, filtration, etc.).

Acknowledgements

The author wishes to thank his scientific adviser professor Vasiliev V.A., who provided valuable comments and ideas regarding the research theme.

REFERENCES

1. *Ljung L., Söderström T.* System Identification Theory for the User. Prentice Hall. Upper Saddle River, 1983.
2. *Ljung L.* Theory and Practice of Recursive Identification. Cambridge : The MIT Press, 1987.
3. *Konev V., Pergamenshchikov S.* On the Duration of Sequential Estimation of Parameters of Stochastic Processes in Discrete Time // Stochastics. 1986. V. 18. Is. 2. P. 133–154.
4. *Konev V., Pergamenshchikov S.* Truncated Sequential Estimation of the Parameters in Random Regression // Sequential Analysis. 1990. V. 9, issue 1. P. 19–41.
5. *Vasiliev V.A.* Truncated Estimation Method with Guaranteed Accuracy // Annals of Institute of Statistical Mathematics. 2014. V. 66. P. 141–163.
6. *Vasiliev V.* Guaranteed Estimation of Logarithmic Density Derivative By Dependent Observations // Topics in Nonparametric Statistics: Proceedings of the First Conference of the International Society for Non-parametric Statistics / eds. by M.G. Akritas et al. New York : Springer, 2014.
7. *Sriram T.* Sequential Estimation of the Autoregressive Parameter in a First Order Autoregressive Process // Sequential Analysis. 1988. V. 7. Is. 1. P. 53–74.
8. *Konev V., Lai T.* Estimators with Prescribed Precision in Stochastic Regression Models // Sequential Analysis. 1995. V. 14. Is. 3. P. 179–192.
9. *Vasiliev V., Kusainov M.* Asymptotic Risk-Efficiency of One-Step Predictors of a Stable AR(1) // Proceedings of XII All-Russian Conference on Control Problems. Moscow, 2014.
10. *Kusainov M., Vasiliev V.* On Optimal Adaptive Prediction of Multivariate Autoregression // Sequential Analysis. 2015. V. 34. Is. 2. (to be published).
11. *Sriram T., Iaci R.* Sequential Estimation for Time Series Models. Sequential Analysis. 2014. V. 33. Is. 2. P. 136–157.
12. *Box G., Jenkins G., Reinsel G.* Time Series Analysis: Forecasting and Control. Wiley, Hoboken, 2008.
13. *Pergamenshchikov S.* Asymptotic Properties of the Sequential Plan for Estimating the Parameter of an Autoregression of the First Order // Theory of Probability and Its Applications. 1991. V. 36. Is. 1. P. 42–53.
14. *Liptser R., Shiryaev A.* Statistics of Random Processes. New York : Springer, 1977.
15. *Gikhman I., Skorokhod A.* Introduction to the Theory of Random Processes. Saunders, Philadelphia, 1969.

Kusainov Marat Islambekovich. E-mail: rjrltsk@gmail.com
Tomsk State University, Tomsk, Russian Federation

Поступила в редакцию 10 декабря 2014 г.

Кусаинов Марат И. (Томский государственный университет. Россия).
Об оптимальном адаптивном прогнозе многомерного процесса АРМА(1,1).
Ключевые слова: адаптивные прогнозы; асимптотическая риск-эффекитвность; многомерный АРМА; момент остановки; оптимальный размер выборки; усечённое оценивание.

Рассматривается проблема асимптотической эффективности адаптивных одношаговых прогнозов многомерного устойчивого процесса АРМА(1,1) с неизвестными параметрами динамики. Прогнозирование основано на методе усечённого оценивания матрицы. Усечённые оценки являются модификацией усечённых последовательных оценок, позволяющей достичь заданной точности на выборках фиксированного размера. Критерий оптимальности прогнозов основан на функции потерь, определённой как линейная комбинация размера выборки и выборочного среднего квадрата ошибки прогноза. Изучены случаи известной и неизвестной дисперсии шума. В последнем случае оптимальный объём наблюдения записывается как момент остановки.

REFERENCES

1. Ljung L., Söderström T. *System Identification Theory for the User*. Prentice Hall, Upper Saddle River, 1983.
2. Ljung L. *Theory and Practice of Recursive Identification*. Cambridge: The MIT Press, 1987.
3. Konev V., Pergamenshchikov S. On the Duration of Sequential Estimation of Parameters of Stochastic Processes in Discrete Time. *Stochastics*, 1986, vol. 18., issue 2, pp. 133-154. DOI: 10.1080/17442508608833405
4. Konev V., Pergamenshchikov S. Truncated Sequential Estimation of the Parameters in Random Regression. *Sequential Analysis*, 1990, vol. 9, issue 1, pp. 19-41. DOI: 10.1080/07474949008836194
5. Vasiliev V. A Truncated Estimation Method with Guaranteed Accuracy. *Annals of Institute of Statistical Mathematics*, 2014, vol. 66, pp. 141-163. DOI: 10.1007/s10463-013-0409-x
6. Vasiliev V. Guaranteed Estimation of Logarithmic Density Derivative By Dependent Observations. *Topics in Nonparametric Statistics: Proceedings of the First Conference of the International Society for Non-parametric Statistics.* New York: Springer, 2014. DOI: 10.1007/978-1-4939-0569-0_31
7. Sriram T. Sequential Estimation of the Autoregressive Parameter in a First Order Autoregressive Process. *Sequential Analysis*, 1988, vol. 7, issue 1, pp. 53-74. DOI: 10.1080/07474948808836142
8. Konev V., Lai T. Estimators with Prescribed Precision in Stochastic Regression Models. *Sequential Analysis*, 1995, vol. 14, issue 3, pp. 179-192. DOI: 10.1080/07474949508836330
9. Vasiliev V., Kusainov M. Asymptotic Risk-Efficiency of One-Step Predictors of a Stable AR(1). *Proceedings of XII All-Russian Conference on Control Problems*, Moscow, 2014.
10. Kusainov M., Vasiliev V. On Optimal Adaptive Prediction of Multivariate Autoregression. *Sequential Analysis*, 2015, vol. 34, issue 2. 23 p. (to be published).
11. Sriram T., Iaci R. Sequential Estimation for Time Series Models. *Sequential Analysis*, 2014, vol. 33, issue 2, pp. 136-157.
12. Box G., Jenkins G., Reinsel G. *Time Series Analysis: Forecasting and Control*. Wiley, Hoboken, 2008.
13. Pergamenshchikov S. Asymptotic Properties of the Sequential Plan for Estimating the Parameter of an Autoregression of the First Order. *Theory of Probability and Its Applications*, 1991, vol. 36, issue 1, pp. 42-53.
14. Liptser R., Shiryaev A. *Statistics of Random Processes*. New York: Springer, 1977.
15. Gikhman I., Skorokhod A. *Introduction to the Theory of Random Processes*. Philadelphia: Saunders, 1969.