ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2014 Управление, вычислительная техника и информатика № 1 (26)

УДК 519.233.22

T.V. Dogadova, V.A. Vasiliev

GUARANTEED PARAMETER ESTIMATION OF STOCHASTIC LINEAR REGRESSION BY SAMPLE OF FIXED SIZE

The method of parameter estimation of the multivariate linear regression by sample of fixed size is proposed. This method makes possible to get the parameter estimators with guaranteed accuracy in the mean square sense. There are constructed and investigated the truncated sequential estimators of ARARCH(1,1), AR(1) and AR(2). Asymptotic efficiency of the parameter estimator AR(1) with unknown noise variance is established.
Keywords: parameter estimation; autoregressive process; ARARCH model; truncated sequential estimators; guaranteed accuracy.

Modern evolution of mathematical statistics is turned to development of data processing methods by dependent sample of finite size. One of such possibilities gives a well-known sequential estimation method, which was successfully applied to parametric and non-parametric problems. This approach for a scheme of independent observations has been primarily proposed in [1]. Then this idea has been applied to parameter estimation problem of dynamic systems in many papers and books (see [2–7, 12, 13, 15] among others). In particular, sequential estimators of the parameter AR(1) with unknown noise variance were proposed in [15].

To obtain sequential estimators with an arbitrary accuracy one needs to have a sample of unbounded size. However, in practice the observation time of a system is usually not only finite but fixed. One of the possibilities for finding estimators with the guaranteed accuracy of inference using a sample of fixed size is provided by the approach of truncated sequential estimation. The truncated sequential estimation method was developed in [8, 9] and others for parameter estimation problems in discrete-time dynamic models. In these papers, estimators of dynamic system parameters with known variance by sample of fixed size were constructed. Another but very similar approach was proposed in [10, 11].

It is known that nonlinear stochastic systems are being widely used for describing real processes in economics, technics, medicine etc. For simple models, for example, scalar first-order autoregression with discrete and continuous time, one-step sequential estimation procedure [2–5] can be constructed. In these cases one-step sequential estimators appear to be least squares estimators calculated in a special stopping time. These estimators are unbiased and simple for researching. In more complicated models such as autoregressive processes of high order and multidimensional regressive processes we can apply two-step sequential estimation procedure [4–7] etc. At that, there is a set of multidimensional models that makes possible to construct one-step procedure of estimation the unknown parameters [2, 5, 12]. In this paper we consider models of this type. There is constructed truncated sequential parameter estimation procedure of general regression. As an example we analyze models of scalar processes ARARCH(1,1), AR(1) and two-dimensional autoregression of special type.

1. General regression model

Let (Ω, F, P) be an arbitrary but fixed probability space with filtration $F^* = \{F_n\}_{n\geq 0}$. It is supposed that the observable p-dimensional process $\{x(n)\}$ satisfies the following equation:

$$\mathbf{x}(n) = A(n-1)\boldsymbol{\lambda} + B(n-1)\boldsymbol{\xi}(n), \ n \geq 1, \tag{1}$$

where $A(n), B(n)$ are F_n – adapted observable matrixes of size $p \times q$, $p \times m$ respectively. Elements of these matrixes may depend on realizations of the process $(\mathbf{x}(n))$.

Noises $\xi(n)$ form the sequence of F_n- adapted independent identically distributed (i.i.d.) random vectors with $E\xi(n)=0,\ E\xi(n)\xi'(n)=I;\quad \boldsymbol{\lambda}=(\lambda_1,...,\lambda_q)'$ – vector of unknown parameters. Here and below the prime means transposition.

The purpose is to construct the truncated sequential estimator of the parameter $\theta=\mathbf{a}'\boldsymbol{\lambda}$, where $\mathbf{a}$ is a given constant vector.

For construction the estimation procedure we introduce pseudoinverse matrixes $A^+(n)=[A'(n)A(n)]^{-1}A'(n)$ (assume all the inverse matrixes $[A'(n)A(n)]^{-1}$ are almost surely (a.s.) determined). Moreover, the F_n- adapted matrixes $\Sigma(n):=B(n)B'(n)$ are supposed to be known or uniformly bounded in sense of square forms for all $n\ge 0$:

$$\Sigma(n)\le\Sigma \quad \text{a.s.} \tag{2}$$

Truncated sequential estimators of θ will be constructed on the basis of the weighted least squares (LS) estimator:

$$\widehat{\theta}_N=\frac{\sum\limits_{n=1}^{N}c(n)\mathbf{a}'A^+(n-1)\mathbf{x}(n)}{\sum\limits_{n=1}^{N}c(n)},$$

where $c(n)=w(n)\cdot\tilde{c}(n-1)$, $\tilde{c}(n)=\{\mathbf{a}'A^+(n)\Sigma^+(n)(A^+(n))'\mathbf{a}\}^{-1}$, $\Sigma^+(n)=\Sigma(n)$ if $\Sigma(n)$ is known and I in the other case; and $w(n)$ – some non-negative weight function satisfying the inequalities $w(n)\le 1$, $n\ge 1$.

According to (1), deviation of the estimator $\widehat{\theta}_N$ has the form:

$$\widehat{\theta}_N-\theta=\frac{\sum\limits_{n=1}^{N}c(n)\mathbf{a}'A^+(n-1)\zeta(n)}{\sum\limits_{n=1}^{N}c(n)}, \tag{3}$$

where $\zeta(n)=B(n-1)\xi(n)$.

Define the truncated sequential estimator $\tilde{\theta}_{H,N}$ for the parameter θ by sample of size N as

$$\tilde{\theta}_{H,N}=\frac{1}{H}\sum_{n=1}^{\tau_{H,N}}\beta_n c(n)\mathbf{a}'A^+(n-1)\mathbf{x}(n)\cdot\chi\left[\sum_{n=1}^{N}c(n)\ge H\right], \tag{4}$$

where the stopping time

$$\tau_{H,N}=\begin{cases}\inf\{k\in[1,N]:\sum\limits_{n=1}^{k}c(n)\ge H\}, & \sum\limits_{n=1}^{N}c(n)\ge H,\\ N, & \sum\limits_{n=1}^{N}c(n)<H\end{cases} \tag{5}$$

and the weights

$$\beta_n = \begin{cases} 1, \; n < \tau_{H,N}, \\ 1, \; n = \tau_{H,N}, \; \sum_{n=1}^{N} c(n) < H, \\ \alpha_H, \; n = \tau_{H,N}, \; \sum_{n=1}^{N} c(n) \geq H, \end{cases} \qquad \alpha_H = \frac{H - \sum_{n=1}^{\tau_{H,N}-1} c(n)}{c(\tau_{H,N})}.$$

Define $\delta_{H,N} = P_\lambda\left(\sum_{n=1}^{N} c(n) < H\right)$ and $\sigma^2 = 1$ if the process $(B(n))$ is observable and $\|\Sigma\|$ in the other case. As well as we denote E_λ the expectation under the distribution P_λ with the given parameter λ.

Theorem 1. Assume the process (1), matrix functions $A(n)$ and $B(n)$ are such that the condition (2) is fulfilled and $E_\lambda c(n) < \infty, \; n \in [1, N]$. Then for every $N \geq 1$ and $H > 0$ the estimator $\tilde{\theta}_{H,N}$, defined in (4), possess the following property:

$$E_\lambda (\tilde{\theta}_{H,N} - \theta)^2 \leq \frac{\sigma^2}{H} + \theta^2 \cdot \delta_{H,N}.$$

Proof. To prove the theorem we find with (3) the deviation of truncated sequential estimator (4)

$$\tilde{\theta}_{H,N} - \theta = \frac{1}{H} \sum_{n=1}^{\tau_{H,N}} \beta_n c(n) \mathbf{a}' A^+(n-1) \zeta(n) \cdot \chi\left[\sum_{n=1}^{N} c(n) \geq H\right] - \theta \cdot \chi\left[\sum_{n=1}^{N} c(n) < H\right].$$

Estimate the mean square deviation of $\tilde{\theta}_{H,N}$. Second moment of the first summand can be estimated similar to, e.g. [3], considering the definition of $\tau_{H,N}$, $c(n)$ and properties $\beta_n \leq 1$

$$E_\lambda (\tilde{\theta}_{H,N} - \theta)^2 \leq \frac{1}{H^2} \cdot E_\lambda \sum_{n=1}^{\tau_{H,N}} \beta_n^2 c^2(n) \mathbf{a}' A^+(n-1) \zeta(n) \zeta'(n) (A^+(n-1))' \mathbf{a} + \theta^2 \cdot \delta_{H,N} \leq$$

$$\leq \frac{\sigma^2}{H^2} \cdot E_\lambda \sum_{n=1}^{\tau_{H,N}} \beta_n^2 c^2(n) \mathbf{a}' A^+(n-1) \Sigma^+(n-1) (A^+(n-1))' \mathbf{a} + \theta^2 \cdot \delta_{H,N} \leq$$

$$\leq \frac{\sigma^2}{H^2} \cdot E_\lambda \sum_{n=1}^{\tau_{H,N}} \beta_n c(n) + \theta^2 \cdot \delta_{H,N} = \frac{\sigma^2}{H} + \theta^2 \cdot \delta_{H,N}.$$

Theorem 1 is proved. Apply general estimation algorithm in the following problems.

2. ARARCH(1,1)

Let $\{x_n\}_{n \geq 0}$ be a scalar first order ARARCH process:

$$x_n = \lambda \cdot x_{n-1} + \sqrt{\sigma_0^2 + \sigma_1^2 x_{n-1}^2} \cdot \xi_n, \quad n \geq 1 \tag{6}$$

with the initial zero mean variable x_0 having the eighth moment; $\{\xi_n\}$ is a sequence of i.i.d. zero mean random variables having density which is an even function, does not increase as module of argument grows and $E\xi_n{}^2 = 1$, in addition, x_0 and $\{\xi_n\}$ are mutually independent. The parameters σ_0^2 and σ_1^2 are supposed to be known and $\sigma_1^2 > 0$.

Note that the volatility coefficients $B(n) = \sqrt{\sigma_0^2 + \sigma_1^2 x_n^2}$ in (6) are observable.

Put in the truncated sequential plan (4), (5) weights $w(n)\equiv 1$ and $H=H_N=\beta\cdot\mu_1\cdot N$, where $\mu_1=E\frac{\xi_1^2}{1+\sigma_1^2\xi_1^2}$, $\beta\in(0,1)$. Then the estimator (4) and the stopping time (5) in this case will be defined as follows:

$$\hat{\lambda}_N=\frac{1}{H_N}\sum_{n=1}^{\tau_N}\frac{\beta_n x_n\cdot x_{n-1}}{\sigma_0^2+\sigma_1^2x_{n-1}^2}\cdot\chi\left[\sum_{n=1}^{N}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}\geq H_N\right], \tag{7}$$

$$\tau_N=\begin{cases}\inf\left\{k\in[1,N]:\sum\limits_{n=1}^{k}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}\geq H_N\right\}, & \sum\limits_{n=1}^{N}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}\geq H_N,\\ N, & \sum\limits_{n=1}^{N}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}<H_N,\end{cases}$$

where the weights β_n are defined as:

$$\beta_n=\begin{cases}1,\ 1\leq n<\tau_N,\\ 1,\ n=\tau_N,\ \sum\limits_{n=1}^{N}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}<H_N,\\ \alpha_N,\ n=\tau_N,\ \sum\limits_{n=1}^{N}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}\geq H_N,\end{cases}$$

$$\alpha_N=\left(H_N-\sum_{n=1}^{\tau_N-1}\frac{x_{n-1}^2}{\sigma_0^2+\sigma_1^2x_{n-1}^2}\right)\Bigg/\frac{x_{\tau_N-1}^2}{\sigma_0^2+\sigma_1^2x_{\tau_N-1}^2}.$$

Theorem 2. Assume model (6). Then for every $0<L<\infty$ there exists the number $\rho=\rho(L)$ such that the estimator (7) satisfies the property

$$\sup_{|\lambda|\leq L}E_\lambda\left(\hat{\lambda}_N-\lambda\right)^2\leq\frac{\rho}{N}.$$

The proof of Theorem 2 is almost the same as in [9] (see Corollary 2 and Section 4) for the autoregression with drifting parameter. The exact expressions of numbers β and ρ can be found in [9, Section 4] as well.

3. Optimal parameter estimation of AR(1)

Let $\{x_n\}_{n\geq 0}$ be scalar autoregressive process:

$$x_n=\lambda\cdot x_{n-1}+\sigma\cdot\xi_n, \tag{8}$$

with the initial zero mean variable x_0 having the eighth moment; $\{\xi_n\}$ is a sequence of i.i.d. random variables with $E\xi_n=E\xi_n^3=0$, $E\xi_n{}^2=1$ and $E\xi_n^8<\infty$; in addition, x_0 and $\{\xi_n\}$ are mutually independent. The process (8) assumed to be stable, i.e. $|\lambda|<1$.

The main problem is to estimate the parameter λ with guaranteed in the mean square sense accuracy when σ is unknown.

In this section we will consider two variants of estimators – with the known principal term of the mean square error (having more simple structure) and optimal in the asymptotic minimax sense estimator. The first estimator will be used in Section 3b as the pilot estimator in the construction of the optimal one.

In both cases we construct truncated sequential estimator on the basis of the LS estimator:

$$\hat{\lambda}_N = \sum_{n=1}^{N} x_n \cdot x_{n-1} \Big/ \sum_{n=1}^{N} x_{n-1}^2 .$$

3a. Adaptive truncated sequential parameter estimation of AR (1)

Define in (4), (5) the threshold $H_N = h \cdot \sigma_m^2 \cdot N$, where $m = m(N)$ is a sequence of integer numbers satisfying the following.

Assumption 1.

a) $m(N) = o(N)$, $m(N) \to \infty$ as $N \to \infty$;

b) $\frac{\log m(N)}{m(N)} = o\left(\frac{1}{\sqrt{N}}\right)$ as $N \to \infty$;

c) for some $\delta \in (0,1)$ fulfilled $\frac{m(N)}{N - m(N)} \le \delta$

and the number $h \in \left(0, \left(\sqrt{2} - 1\right)^2 \cdot (1+\delta)^{-1}\right)$.

Define the pilot LS-type estimator of variance σ^2 as follows

$$\sigma_m^2 = \frac{1}{m} \sum_{n=1}^{m} \left[x_n - \lambda_m^* x_{n-1} \right]^2,$$

where

$$\lambda_m^* = proj_{[-1,1]} \tilde{\lambda}_m ,$$

$$\tilde{\lambda}_m = \hat{\lambda}_m \cdot \chi\left[\sum_{n=1}^{m} x_{n-1}^2 \ge m (\log m)^{-1} \right].$$

Analogously to [11] can be obtained

$$E_\mu \left(\sigma_m^2 - \sigma^2 \right)^4 \le \frac{\bar{C} \cdot (\log m)^2}{m^2}, \tag{9}$$

where $\mu = \left(\lambda, \sigma^2\right)$.

Define in the general estimation procedure (4), (5) the weight functions

$$w(n) = \begin{cases} 0, & 1 \le n \le m, \\ \chi\left[\sigma_m^2 > (\log m)^{-1} \right], & m < n \le N; \end{cases} \tag{10}$$

the truncated stopping time

$$\tau_N = \begin{cases} \inf\left\{ k \in [1, N] : \sum_{n=m+1}^{k} x_{n-1}^2 \ge H_N \right\}, & \sum_{n=m+1}^{N} x_{n-1}^2 \ge H_N, \\ N, & \sum_{n=m+1}^{N} x_{n-1}^2 < H_N; \end{cases} \tag{11}$$

the weight functions

$$\beta_n = \begin{cases} 1, \ 1 \le n < \tau_N, \\ 1, \ n = \tau_N, \ \sum_{n=m+1}^{N} x_{n-1}^2 < H_N, \\ \alpha_N, \ n = \tau_N, \ \sum_{n=m+1}^{N} x_{n-1}^2 \ge H_N, \end{cases} \tag{12}$$

where

$$\alpha_N = \left(H_N - \sum_{n=m+1}^{\tau_N - 1} x_{n-1}^2 \right) \Big/ x_{\tau_N - 1}^2 .$$

Then the truncated sequential estimator (4) has the form

$$\tilde{\lambda}_N = \frac{1}{H_N} \cdot \sum_{n=m+1}^{\tau_N} \beta_n x_n x_{n-1} \cdot \chi\left[\sum_{n=m+1}^{N} x_{n-1}^2 \ge H_N, \sigma_m^2 > (\log m)^{-1} \right]. \tag{13}$$

Denote for every N such that $\log m(N) > \sigma^{-2}$ the function

$$\varepsilon_N = \frac{2C_{22} \cdot (\log m)^2}{N^2 \cdot m^2} + \frac{2C_{21}}{N^2} + \frac{\bar{C}^{1/4} \cdot (\log m)^{3/2}}{hN\sqrt{m}} + \frac{2\bar{C} \cdot (\log m)^2}{m^2 \left(\sigma^2 - (\log m)^{-1} \right)^4},$$

where $C_{21} = \dfrac{8(1+\delta)^2 B_4 E\left(\xi_1^2 - \sigma^2\right)^4}{\left[\sigma^2 \left(1 - \dfrac{h}{C_0}(1+\delta) \right) \right]^4}$, $C_{22} = \dfrac{8h^4 \cdot \bar{C}(1+\delta)^6}{\left[\sigma^2 \left(C_0 - h(1+\delta) \right) \right]^4}$, B_4 is the coefficient from the Burkholder inequality.

According to Assumption 1 $\varepsilon_N = o\left(\dfrac{1}{N} \right)$ as $N \to \infty$.

The following theorem contains the main result of the section.

Theorem 3. Assume the model (8) with the parameter $|\lambda| < 1$. Then the truncated sequential estimator (13) has the following property:

1) $E_\mu \left(\tilde{\lambda}_N - \lambda \right)^2 \le \dfrac{1}{N \cdot h} + \varepsilon_N$;

if in addition the noises ξ_n and x_0 for some positive integer *s* have moments of the order 8*s* then there exist the numbers $C(s)$ such that as $N \to \infty$;

2) $E_\mu \left(\tilde{\lambda}_N - \lambda \right)^{2s} \le \dfrac{C(s)}{N^s} + o\left(\dfrac{1}{N^s} \right)$.

Proof. The proof of the first assertion of the theorem is based on the following representation of the estimator's deviation:

$$\tilde{\lambda}_N - \lambda = \frac{\sigma}{H_N} \cdot \sum_{n=m+1}^{\tau_N} \beta_n x_{n-1} \xi_n \cdot \chi\left[\sum_{n=m+1}^{N} x_{n-1}^2 \ge H_N, \sigma_m^2 > (\log m)^{-1} \right] -$$

$$-\lambda \cdot \left(\chi\left[\sum_{n=m+1}^{N} x_{n-1}^2 < H_N \right] + \chi\left[\sigma_m^2 > (\log m)^{-1} \right] \right) = I_1 + I_2 + I_3. \tag{14}$$

According to this formula we have

$$E_\mu\left(\tilde{\lambda}_N-\lambda\right)^2 \le E_\mu I_1^2+2E_\mu I_2^2+2E_\mu I_3^2. \quad (15)$$

Consider separately the summands in (15). By the definition of the stopping time τ_N using (9) and the technique proposed in [8, 9, 15] we can estimate

$$E_\mu I_1^2 \le \sigma^2 E_\mu \frac{1}{H_N^2} E_\mu\left(\left[\sum_{n=m+1}^{\tau_N}\beta_n x_{n-1}\xi_n\right]^2 \mid F_m\right)\cdot\chi\left[\sigma_m^2>\left(\log m\right)^{-1}\right]=$$

$$=\sigma^2 E_\mu \frac{1}{H_N^2} E_\mu\left(\sum_{n=m+1}^{\tau_N}\beta_n^2 x_{n-1}^2 \mid F_m\right)\cdot\chi\left[\sigma_m^2>\left(\log m\right)^{-1}\right]\le$$

$$\le\sigma^2 E_\mu \frac{1}{H_N^2} E_\mu\left(\sum_{n=m+1}^{\tau_N}\beta_n x_{n-1}^2 \mid F_m\right)\cdot\chi\left[\sigma_m^2>\left(\log m\right)^{-1}\right]=$$

$$=\frac{1}{hN}E_\mu\frac{\sigma^2}{\sigma_m^2}\cdot\chi\left[\sigma_m^2>\left(\log m\right)^{-1}\right]\le\frac{1}{hN}+\frac{\log m}{hN}E_\mu\left|\sigma_m^2-\sigma^2\right|\le$$

$$\le\frac{1}{hN}+\frac{\log m}{hN}\left(E_\mu\left(\sigma_m^2-\sigma^2\right)^4\right)^{1/4}\le\frac{1}{hN}+\frac{\bar{C}^{1/4}\cdot\left(\log m\right)^{3/2}}{hN\sqrt{m}}. \quad (16)$$

Estimate $E_\mu I_2^2$ from above:

$$E_\mu I_2^2 \le P_\mu\left(\sum_{n=m+1}^{N}x_{n-1}^2<H_N\right).$$

Define the number $C_0\left[\lambda\right]=\left[\sqrt{1+\lambda^2}-\left|\lambda\right|\right]^2$. According to Lemma 2 in [9] for every $k>m$ we have:

$$\sum_{n=m+1}^{k}x_n^2 \ge C_0\left(\lambda\right)\sum_{n=m+1}^{k}\xi_n^2. \quad (17)$$

Note that in the stable case $\left|\lambda\right|<1$

$$C_0\left(\lambda\right)\ge\left(\sqrt{2}-1\right)^2.$$

Using this formula and the Chebyshev inequality we get

$$P_\mu\left(\sum_{n=m+1}^{N}x_{n-1}^2<H_N\right)\le P_\mu\left(C_0\left(\lambda\right)\sum_{n=m+1}^{N}\xi_n^2<H_N\right)=$$

$$=P_\mu\left(-\frac{1}{N-m}\sum_{n=m+1}^{N}\left(\xi_n-\sigma^2\right)>\sigma^2-\frac{H_N}{C_0\cdot\left(N-m\right)}\right)\le$$

$$\le P_\mu\left(\left|\frac{1}{N-m}\sum_{n=m+1}^{N}\left(\xi_n^2-\sigma^2\right)\right|>\sigma^2-\frac{h\sigma_m^2 N}{C_0\cdot\left(N-m\right)}\right)\le$$

$$\le P_\mu\left(\left|\frac{1}{N-m}\sum_{n=m+1}^{N}\left(\xi_n^2-\sigma^2\right)\right|+\left|\frac{hN(\sigma_m^2-\sigma^2)}{C_0\cdot\left(N-m\right)}\right|>\sigma^2\left(1-\frac{hN}{C_0\cdot\left(N-m\right)}\right)\right)\le$$

$$\le \frac{8B_4E_\mu\left(\xi_1^2-\sigma^2\right)^4}{\left[\sigma^2\left(1-\frac{hN}{C_0\cdot(N-m)}\right)\right]^4\cdot(N-m)^2}+\frac{8h^4N^4E_\mu\left(\sigma_m^2-\sigma^2\right)^4}{\left[\sigma^2\left(1-\frac{hN}{C_0\cdot(N-m)}\right)\right]^4\cdot C_0^4\cdot(N-m)^6}\le$$

$$\le \frac{8(1+\delta)^2B_4E_\mu\left(\xi_1^2-\sigma^2\right)^4}{\left[\sigma^2\left(1-\frac{h}{C_0}(1+\delta)\right)\right]^4\cdot N^2}+\frac{8h^4\cdot\bar{C}(1+\delta)^6}{\left[\sigma^2\left(C_0-h(1+\delta)\right)\right]^4}\cdot\frac{(\log m)^2}{N^2\cdot m^2}=\frac{C_{21}}{N^2}+\frac{C_{22}\cdot(\log m)^2}{N^2\cdot m^2}. \qquad (18)$$

In the last inequality we used (9).

Let us estimate the third summand in (15)

$$E_\mu I_3^2\le P_\mu\left(\sigma_m^2>(\log m)^{-1}\right).$$

Using (9) and the Chebyshev inequality for N large enough we get

$$P_\mu\left(\sigma_m^2>(\log m)^{-1}\right)=P_\mu\left(\sigma^2-\sigma_m^2>\sigma^2-(\log m)^{-1}\right)\le\frac{\bar{C}\cdot(\log m)^2}{m^2\left(\sigma^2-(\log m)^{-1}\right)^4}=o\left(\frac{1}{N}\right). \qquad (19)$$

The first assertion of Theorem 3 follows from the obtained inequalities (15), (16), (18), (19).

The second assertion can be proved analogously to the first one.

Theorem 3 is proved. Similar results for the sequential estimators of λ were presented in [15].

3b. Efficiency of $\tilde{\lambda}_N$

In this section we consider a little bit more complicated modification of the estimator (13) and prove its optimality in the sense of some risk function defined below.

Put in the definition (13) the threshold $H_N=h_N\cdot\frac{\hat{\sigma}_m^2}{1-\hat{\lambda}_m^2}\cdot(N-m)$, $h_N=1-(\log N)^{-1}$ and $m=m(N)$ is a sequence of integer numbers satisfying the following.

Assumption 2.

a) $m(N)=o(N)$, $m(N)\to\infty$ as $N\to\infty$;

b) $\frac{\log m(N)}{m(N)}=o\left(\frac{1}{\sqrt{N}\cdot\log^2 N}\right)$ as $N\to\infty$;

c) for some $\delta\in(0,1)$ fulfilled $\frac{m(N)}{N-m(N)}\le\delta.$

Here the pilot estimators of variance σ^2 and λ are defined as follows

$$\hat{\sigma}_m^2=\frac{1}{m}\sum_{n=1}^m\left[x_n-\hat{\lambda}_m x_{n-1}\right]^2$$

and for some $r\in(0,1)$

$$\hat{\lambda}_m=proj_{[-r,r]}\tilde{\lambda}_m,$$

where estimator $\tilde{\lambda}_m$ is defined in (13) with H_N from Section 3a.

Similarly to (9) the following inequalities can be proved

$$E_\mu\left(\hat{\sigma}_m^2-\sigma^2\right)^4 \le \frac{\hat{C}\cdot(\log m)^2}{m^2}. \tag{20}$$

The estimators $\tilde{\lambda}_N$, times τ_N and weights $w(n)$, β_n are defined in (10)–(13) with H_N introduced above.

To prove the optimality of the estimator $\tilde{\lambda}_N$ we establish first the following inequality

$$\overline{\lim_{N\to\infty}} N\cdot E_\mu\left(\tilde{\lambda}_N-\lambda\right)^2 \le 1-\lambda^2. \tag{21}$$

Using the representation (14) for the deviation of the estimator $\tilde{\lambda}_N$ we estimate second moments of the summands in the right hand side of (14).

Similar to (16) we have

$$E_\mu I_1^2 \le \sigma^2 E_\mu \frac{1}{H_N^2} E_\mu\left(\sum_{n=m+1}^{\tau_N}\beta_n x_{n-1}^2 \mid F_m\right)\cdot\chi\left[\hat{\sigma}_m^2>(\log m)^{-1}\right]=$$

$$=h_N^{-2}\cdot\frac{1}{N-m}E_\mu\frac{\sigma^2\left(1-\hat{\lambda}_m^2\right)}{\hat{\sigma}_m^2}\cdot\chi\left[\hat{\sigma}_m^2>(\log m)^{-1}\right]\le h_N^{-2}\cdot\frac{1}{N-m}\left(1-\lambda^2\right)+$$

$$+h_N^{-2}\cdot\frac{1}{N-m}E_\mu\frac{\left|\hat{\sigma}_m^2-\sigma^2\right|}{\hat{\sigma}_m^2}\left(1-\hat{\lambda}_m^2\right)\cdot\chi\left[\sigma_m^2>(\log m)^{-1}\right]+h_N^{-2}\cdot\frac{1}{N-m}E_\mu\left|\hat{\lambda}_m^2-\lambda^2\right|\le$$

$$\le h_N^{-2}\cdot\frac{1}{N-m}\left(1-\lambda^2\right)+h_N^{-2}\cdot\frac{\log m}{N}\left(E_\mu\left(\sigma_m^2-\sigma^2\right)^4\right)^{1/4}+h_N^{-2}\cdot\frac{2}{N-m}\sqrt{E_\mu\left(\hat{\lambda}_m-\lambda\right)^2}\le$$

$$\le h_N^{-2}\cdot\frac{1}{N-m}\left(1-\lambda^2\right)+h_N^{-2}\cdot\frac{\overline{C}^{1/4}\cdot(\log m)^{3/2}}{(N-m)\sqrt{m}}+h_N^{-2}\cdot\frac{1}{N-m}\left(\frac{1}{m\cdot h}+\varepsilon_m\right)=o\left(\frac{1}{N}\right). \tag{22}$$

Estimate $E_\mu I_2^2$ from below

$$E_\mu I_2^2 \le P_\mu\left(\sum_{n=m+1}^{N}x_{n-1}^2<H_N\right).$$

Notice that when H_N is determined in Section 3a, the method of estimation of this probability used in Section 3a cannot be applied. Then we will use representation for the probability argument applying the Ito formula for the process $\left\{x_n^2\right\}$:

$$\frac{1}{N-m}\sum_{n=m+1}^{N}x_{n-1}^2-\frac{\sigma^2}{1-\lambda^2}=\frac{1}{1-\lambda^2}\left\{\frac{x_m^2-x_N^2}{N-m}+\frac{2\lambda}{N-m}\sum_{n=m+1}^{N}x_{n-1}\xi_n+\frac{1}{N-m}\sum_{n=m+1}^{N}\left(\xi_n^2-\sigma^2\right)\right\}.$$

Using this formula and the Chebyshev inequality we get

$$P_\mu\left(\sum_{n=m+1}^{N}x_{n-1}^2<H_N\right)=P_\mu\left(\frac{\sigma^2}{1-\lambda^2}-\frac{H_N}{N-m}<\frac{\sigma^2}{1-\lambda^2}-\frac{1}{N-m}\sum_{n=m+1}^{N}x_{n-1}^2\right)\le$$

$$\le P_\mu\left(\frac{1}{1-\lambda^2}\left|\frac{x_N^2-x_m^2}{N-m}-\frac{2\lambda}{N-m}\sum_{n=m+1}^{N}x_{n-1}\xi_n-\frac{1}{N-m}\sum_{n=m+1}^{N}\left(\xi_n^2-\sigma^2\right)\right|+h_N\left|\frac{\hat{\sigma}_m^2}{1-\hat{\lambda}_m^2}-\frac{\sigma^2}{1-\lambda^2}\right|\ge\left(1-h_N\right)\frac{\sigma^2}{1-\lambda^2}\right)\le$$

$$\leq \frac{8}{\left(\sigma^2\right)^4 \left(1-h_N\right)^4 \left(N-m\right)^4} \cdot E_\mu \left[x_N^2 + x_m^2 + 2|\lambda| \cdot \left| \sum_{n=m+1}^{N} x_{n-1}\xi_n \right| + \left| \sum_{n=m+1}^{N} \left(\xi_n^2 - 1\right) \right| \right]^4 +$$

$$+\frac{8}{\left(\sigma^2\right)^4 \left(1-r^2\right)^8 \left(1-h_N\right)^4} E_\mu \left[\left|\hat{\sigma}_m^2 - \sigma^2\right| + 2(1-r)\sigma^2 \left|\hat{\lambda}_m - \lambda\right| \right]^4 \leq$$

$$\leq \frac{2\cdot 4^4}{\left(\sigma^2\right)^4 \left(1-h_N\right)^4 \left(N-m\right)^4} \cdot \left[E_\mu x_N^8 + E_\mu x_m^8 + 2|\lambda| \cdot E_\mu \left| \sum_{n=m+1}^{N} x_{n-1}\xi_n \right|^4 + E \left| \sum_{n=m+1}^{N} \left(\xi_n^2 - 1\right) \right|^4 \right] +$$

$$+\frac{64}{\left(\sigma^2\right)^4 \left(1-r^2\right)^8 \left(1-h_N\right)^4} \left[E_\mu \left|\hat{\sigma}_m^2 - \sigma^2\right|^4 + 2\sigma^2 E_\mu \left|\hat{\lambda}_m - \lambda\right|^4 \right] =$$

$$= O\left(\frac{(\log N)^4}{(N-m)^2} + \frac{(\log N)^4 (\log m)^2}{m^2} + \frac{(\log N)^4}{m^2} \right) = o\left(\frac{1}{N}\right). \qquad (23)$$

The last inequality is true due to (20), the second assertion of Theorem 3, the second property in Assumption 2 and the obvious inequality $\sup_{n\geq 0} E_\mu x_n^8 < \infty$ which fulfills for the stable process (8).

Then the inequality (21) follows from (14), (19), (22) and (23).

From (21) it follows that the truncated estimator $\left\{\tilde{\lambda}_N\right\}_{N\geq 1}$ is optimal (see [11, 13]) in the asymptotic minimax sense

$$\underline{\lim}_{N\to\infty} R_{r,N}\left(\lambda_N\right) \geq \underline{\lim}_{N\to\infty} \inf_{\lambda_N} R_{r,N}\left(\lambda_N\right) = \lim_{N\to\infty} R_{r,N}\left(\tilde{\lambda}_N\right) = 1,$$

where

$$R_{r,N}\left(\lambda_N\right) = \sup_{P} \sup_{|\lambda|\leq 1-r} I(\lambda, f) N \cdot E_\mu \left(\lambda_N - \lambda\right)^2$$

and the infimum is taken over the class of all (non-randomized) estimators λ_N of the parameter λ. Here P is the class of all densities $f(.)$ of the noises $\{\xi_n\}$ having finite second moments and the Fisher information $I(\lambda, f) = \frac{\sigma^2}{1-\lambda^2} I(f) (= \left(1-\lambda^2\right)^{-1}$ for the case of Gaussian densities $f(.)$),

$$I(f) = \int \left(\frac{f'(x)}{f(x)} \right)^2 f(x)\,dx.$$

4. AR(2)

Consider two-dimensional autoregressive special type process AR(2)

$$\begin{cases} x_1(n) = \lambda_1 x_1(n-1) + \lambda_2 x_2(n-1) + \xi_1(n), \\ x_2(n) = \lambda_2 x_1(n-1) - \lambda_1 x_2(n-1) + \xi_2(n), \end{cases} \qquad (24)$$

where the parameter $\lambda = (\lambda_1, \lambda_2)'$ to be estimated belongs to the whole plane R^2 and $\xi(n) = (\xi_1(n), \xi_2(n))'$ form a sequence of i.i.d. zero mean random vectors independent on $x(0) = (x_1(0), x_2(0))'$ with $E\xi(n) = 0$, $E\xi(n)\cdot\xi'(n) = I$, $E\|\xi(1)\|^8 < \infty$, as well as $Ex(0) = 0$ and $E\|x(0)\|^8 < \infty$.

Note that parameter estimation problem for similar two-dimensional stochastic continuous time system was considered in [2].

The main aim of this section is to estimate with guaranteed accuracy the parameter $\theta = a'\lambda$, where the given vector a is such that $\|a\| = 1$.

For definition of the truncated sequential estimators we will use the following representation of (24)

$$x(n) = A(n-1)\lambda + \xi(n),$$

where

$$A(n) = \begin{pmatrix} x_1(n) & x_2(n) \\ -x_2(n) & x_1(n) \end{pmatrix}.$$

According to general notation in Section 1 in this case we have $B(n) = I$ and as follows $\Sigma = I$. Besides $A^+(n) = A^{-1}(n) = \frac{1}{\|x(n)\|^2}\cdot A(n)$ and choosing in the definition of $c(n)$ weights $w(n) \equiv 1$ we obtain $c(n) = \|x(n-1)\|^2$. Then according to (4), (5) the truncated sequential estimation plan of θ has the form

$$\tilde{\theta}_{H,N} = \frac{1}{H}\sum_{n=1}^{\tau_{H,N}} \beta_n a' A(n-1)x(n)\cdot\chi\left[\sum_{n=1}^{N}\|x(n-1)\|^2 \geq H\right],$$

$$\tau_{H,N} = \begin{cases} \inf\left\{k\in[1,N] : \sum\limits_{n=1}^{k}\|x(n-1)\|^2 \geq H\right\}, & \sum\limits_{n=1}^{N}\|x(n-1)\|^2 \geq H, \\ N, & \sum\limits_{n=1}^{k}\|x(n-1)\|^2 < H, \end{cases}$$

where

$$\beta_n = \begin{cases} 1, n < \tau_{H,N}, \\ 1, n = \tau_{H,N}, \sum\limits_{n=1}^{N}\|x(n-1)\|^2 < H, \\ \alpha_H, n = \tau_{H,N}, \sum\limits_{n=1}^{N}\|x(n-1)\|^2 \geq H. \end{cases}$$

According to Theorem 1

$$E_\lambda\left(\tilde{\theta}_{H,N} - \theta\right)^2 \leq \frac{1}{H} + \theta^2\cdot P_\lambda\left(\sum_{n=1}^{N}\|x(n-1)\|^2 < H\right). \quad (25)$$

To estimate the second summand in the right hand side of the last inequality we will use the following equation

$$\|x(n)\|^2 = \|\lambda\|^2\|x(n-1)\|^2 + 2x^T(n-1)A\xi(n) + \|\xi(n)\|^2, \quad n \geq 1 \quad (26)$$

with

$$A = \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_2 & -\lambda_1 \end{pmatrix},$$

which can be easily obtained from the following representation of (24):

$$x(n) = Ax(n-1) + \xi(n),\ n \geq 1.$$

Using (26) similar to (17) we get the inequality:

$$\sum_{n=1}^{N} \|x(n-1)\|^2 \geq \tilde{C}_0(\lambda) \sum_{n=1}^{N} \|\xi(n)\|^2,\ \tilde{C}_0(\lambda) = \left[\sqrt{1+2\|\lambda\|^2} - \sqrt{2}\|\lambda\|\right]^2$$

and $\tilde{C}_0(\lambda) \geq C_*,\ C_* = \left[\sqrt{1+2L^2} - \sqrt{2}L\right]^2$ when $\|\lambda\| \leq L.$

Suppose $H = H_N = 2hC_* N, h \in (0,1).$ Then

$$P_\lambda\left(\sum_{n=1}^{N} \|x(n-1)\|^2 < H_N\right) \leq P\left(C_* \cdot \sum_{n=1}^{N} \|\xi(n)\|^2 < H_N\right) \leq P\left(\left|\frac{1}{N}\sum_{n=1}^{N}\left(\|\xi(n)\|^2 - 2\right)\right| > 2(1-h)\right) \leq \frac{C}{N^2},$$

where C is a positive number.

From (25) and the last inequality it follows that the estimator $\theta_N^* = \tilde{\theta}_{H,N}$ with $H = H_N$ for every $0 < L < \infty$ satisfies the inequality

$$\sup_{\|\lambda\| \leq L} E_\lambda \left(\theta_N^* - \theta\right)^2 \leq \frac{1}{2hC_* N} + \frac{L^2 C}{N^2}.$$

5. Simulations

To confirm the convergence of the constructed estimators we made the simulations of the estimator $\tilde{\lambda}_N$ defined in (13) (Table 1) with $w(n) \equiv 1$ and $H_N = h \cdot N$ in case of known $\sigma^2 = 1$:

$$\tilde{\lambda}_N = \frac{1}{H_N} \cdot \sum_{n=1}^{\tau_N} \beta_n x_n x_{n-1} \cdot \chi\left[\sum_{n=m+1}^{N} x_{n-1}^2 \geq H_N\right]. \quad (27)$$

For this purpose we used the software package MATLAB. In Table 1 the average

$$\tilde{\lambda}(N) = \frac{1}{100} \cdot \sum_{k=1}^{100} \tilde{\lambda}_N(k)$$

of estimators (27) for the k-th realization $x^{(k)} = \left(x_n^{(k)}\right),\ k = 1...100$ of the process (8) and their quality characteristics

$$S_\lambda^2(N) = \frac{1}{100} \sum_{k=1}^{100} \left(\tilde{\lambda}_N(k) - \lambda\right)^2$$

for different N are given.

T a b l e 1

Estimation of the parameter λ with h = 0,2

λ	N = 100		N = 200		N = 500	
	$\tilde{\lambda}(N)$	$S_\lambda(N)$	$\tilde{\lambda}(N)$	$S_\lambda(N)$	$\tilde{\lambda}(N)$	$S_\lambda(N)$
0,2	0,2031	0,0395	0,2111	0,0240	0,2041	0,0090
–0,2	–0,1755	0,0521	0,0092	0,0257	–0,1973	0,0092
0,9	0,8836	0,0426	0,8678	0,0252	0,8967	0,0066
–0,9	–0,8874	0,0407	–0,9082	0,0222	–0,8943	0,0114
1	0,9841	0,0514	0,9722	0,0164	1,0013	0,0091
–1	–0,9730	0,0395	–0,9942	0,0162	–0,9993	0,0104
4	4,0107	0,0166	4,0183	0,0074	4,0087	0,0026
–4	–4,0060	0,0228	–3,9987	0,0071	–4,0008	0,0050

T a b l e 2

Estimation of the parameter λ with h = 0,6

λ	N = 100		N = 200		N = 500	
	$\tilde{\lambda}(N)$	$S_\lambda(N)$	$\tilde{\lambda}(N)$	$S_\lambda(N)$	$\tilde{\lambda}(N)$	$S_\lambda(N)$
0,2	0,2253	0,0149	0,1991	0,0090	0,2001	0,0029
–0,2	–0,2126	0,0141	–0,1945	0,0090	–0,2004	0,0029
0,9	0,8874	0,0145	0,8872	0,0067	0,8945	0,0027
–0,9	–0,8997	0,0127	–0,9015	0,0054	–0,9012	0,0037
1	1,0085	0,0123	0,9898	0,0077	0,9967	0,0038
–1	–0,9732	0,0171	–1,0044	0,0051	–0,9893	0,0033
4	3,9996	0,0047	3,9986	0,0027	4,0057	0,0014
–4	–3,9947	0,0068	–4,0038	0,0034	–4,0040	0,0015

Looking at the simulation results we can say that the deviation becomes less with growth of the sample size. It means that the estimator's value becomes closer to the true meaning of the parameter. This fact proves that these estimation procedures are quite effective. Moreover, the considered estimator works in the unstable case as well.

Summary

In this paper the guaranteed parameter estimation problem of a general multivariate regression model is solved. The truncated sequential estimator of the dynamic parameter is constructed by sample of fixed size. At the same time this method makes possible to obtain estimators with a given mean square accuracy. Three examples are given. In the first example the parameter estimation problem of ARARCH(1,1) is considered. It is supposed that the unknown parameter belongs to the whole line. The presented estimator has given mean square accuracy and the same rate of convergence as the least squares estimator in the stable case. Similar results were obtained for the stable AR(1) and AR(2) model of a special type. Asymptotic optimality in the minimax sense for the truncated sequential estimator of AR(1) is proved.

Results of simulation confirm the efficiency of the presented estimation procedure.

Truncated sequential estimators can be successfully used similar to sequential estimators (see, e.g., [14]) as a pilot estimators in various adaptive procedures (prediction, control, filtration etc.).

REFERENCES

1. Wald A. Sequential Analysis. N. Y.: Wiley. (1947).
2. Liptser R. Sh., Shiryaev A.N. Statistics of random processes. 1: General theory. N. Y.: Springer-Verlag. (1977). 2: Applications. N.Y.: Springer-Verlag. (1978).

3. Borisov V.Z., Konev V.V. On sequential estimation of parameters in discrete-time processes. *Automation and Remote Control.* (1977).
4. Konev V.V. Sequential parameter estimation of stochastic dynamical systems. Tomsk: Tomsk Univ. Press. (1985).
5. Dobrovidov A.V., Koshkin G.M., Vasiliev V.A. Non-parametric state space models. Heber City, UT. USA: Kendrick Press. (2012).
6. Galtchouk L., Konev V. On sequential estimation of parameters in semimartingaleregression models with continuous time parameter. *Annals of Statistics.* (2001).
7. Kuechler U., Vasiliev V. On guaranteed parameter estimation of a multiparameter linear regression process. *Automatica, Journal of IFAC,* Elsevier. No. 46 (4). P. 637-646. (2010).
8. Fourdrinier D., Konev V., Pergamenshchikov S. Truncated sequential estimation of the parameter of a first order autoregressive process with dependent noises. *Mathematical Methods of Statistics.* (2009).
9. Konev V.V., Pergamenshchikov S.M. Truncated sequential estimation of the parameters in random regression. *Sequential Analysis.* V. 9. Issue 1. P. 19-41. (1990).
10. Vasiliev V.A. One investigation method of ratio type estimators. *Preprint 5 of the Math. Inst. of Humboldt University*, Berlin. P. 1-15. (2012); http://www2.mathematik.hu-berlin.de/publ/pre/2012/p-list-12.html
11. Vasiliev V.A. A truncated estimation method with guaranteed accuracy. *Annals of the Institute of Statistical Mathematics*, V. 66. Issue 1. P. 141-163. (2014).
12. Vorobeichikov S.E., Konev V.V. About sequential indification of stochastic systems. *Technical Cybernetics*. No. 4. P. 176-182. (1980).
13. Shiryaev A.N., Spokoiny V.G. Statistical experiments and decisions. *Asymptotic theory.* Singapore: World Scientific, (2000).
14. Kuechler U., Vasiliev V. On a certainty equivalence design of continuous-time stochastic systems. *SIAM Journal of Control and Optimization.* V. 51. No. 2. P. 938-964. (2013).
15. Dmitrienko A., Konev V., Pergamenshchikov S. Sequential generalized least squares estimator for an autoregressive parameter. *Sequential Analysis.* V. 16. Issue 1. P. 25-46. (1997).

Догадова Татьяна Валерьевна. E-mail: aurora1900@mail.ru
Васильев Вячеслав Артурович. E-mail: vas@mail.tsu.ru
Томский государственный университет Поступила в редакцию 10 января 2014 г.

Догадова Т.В., Васильев В.А. (Томский государственный университет, Российская Федерация).
Гарантированное оценивание параметров стохастической линейной регрессии по выборке фиксированного размера.
Ключевые слова: оценивание параметров; процесс авторегрессии; модель ARARCH; усеченные последовательные оценки; гарантированная точность.

Предлагается метод оценки параметров многомерной линейной регрессии по выборке фиксированного объема. Этот метод позволяет получить оценки параметров с гарантированной в среднеквадратическом смысле точностью. Построены и исследованы усеченные последовательные оценки параметров процессов ARARCH(1,1), AR(1) и AR(2). Установлена асимптотическая эффективность оценки параметра AR(1) с неизвестной дисперсией шума.