WEIGHTED LEAST SQUARES METHOD OF GUARANTEED ESTIMATION OF ARCH(*P*) PROCESS PARAMETERS*

Julia Burkatovskaya[1], Sergei Vorobejchikov[2]

[1]Tomsk Polytechnic University, [2]Tomsk State University, Tomsk, Russia
[1]*burkatovskaya@sibmail.com*, [2]*sev@mail.tsu.ru*

A sequential procedure of estimating the ARCH(p) parameters based on the least squares method is proposed. The choice of weights and the stopping rule guarantees the accuracy of the estimation. Results of numerical simulation prove the efficiency of the suggested procedure.

Introduction. Some types of time series, for example, financial indexes possess the so-called "cluster effect" when groups of observations with large and small variances alternate. To describe such processes, R. Engle proposed an autoregressive model with conditional heteroscedasticity (ARCH). In this model, the variance of the process is a stochastic autoregressive process. In the present paper, the problem of estimating the parameters of this process is considered and the sequential method that guarantees the bounded standard deviation of the estimation from the true parameter value is proposed.

Problem statement. We consider the ARCH(p) process specified by the equations

$$x_l = \sigma_l \varepsilon_l; \qquad (1)$$
$$\sigma_l^2 = \lambda_0 + \lambda_1 x_{l-1}^2 + \ldots + \lambda_p x_{l-p}^2.$$

Here $\{\varepsilon_l\}_{l\ge1}$ is a sequence of independent identically distributed random variables with mean 0 and variance 1. Parameters μ and λ are supposed to be unknown. The problem is to construct a guaranteed estimation of the vector of unknown parameters $\Lambda = [\lambda_0, \ldots, \lambda_p]$ on the basis of observations $\{x_l\}_{l\ge1}$.

Sequential estimation. To estimate the parameters, we use the approach proposed in [1] to classify the autoregressive processes with unknown variance. To take advantage of these results, we rewrite process (1) in the form

$$x_l^2 = \sigma_l^2 + \sigma_l^2(\varepsilon_l^2 - 1).$$

Denoting $E(\varepsilon_l^2 - 1)^2$ by B^2 and $(\varepsilon_l^2 - 1)/B$ by η_l, we have

$$x_l = \lambda_0 + \lambda_1 x_{l-1}^2 + \ldots + \lambda_p x_{l-p}^2 + (\lambda_0 + \lambda_1 x_{l-1}^2 + \ldots + \lambda_p x_{l-p}^2) B \eta_l. \qquad (2)$$

Here $\{\eta_l\}_{l\ge1}$ is the sequence of independent identically distributed random variables with $E\eta_l = 0$ and $E\eta_l^2 = 1$. Here process (2) is the p-order autoregressive process. The noise variance $(\lambda_0 + \lambda_1 x_{l-1}^2 + \ldots + \lambda_p x_{l-p}^2)^2 B^2$ is unknown and moreover, unbounded from above. Denoting

$$y_{l-1}^2 = \max\{1, x_{l-1}^2, \ldots, x_{l-p}^2\}, \quad z_l = \frac{x_l^2}{y_{l-1}^2}, \quad a_{l-1} = \left[\frac{1}{y_{l-1}^2}, \frac{x_{l-1}^2}{y_{l-1}^2}, \ldots, \frac{x_{l-p}^2}{y_{l-1}^2}\right]^T, \qquad (3)$$

we rewrite (2) as

$$z_l = \Lambda a_{l-1} + \Lambda a_{l-1} B \eta_l. \qquad (4)$$

Since $\Lambda a_{l-1} \le \lambda_0 + \ldots + \lambda_p$, the noise variance in (4) is bounded from above.

To estimate the unknown parameters, we take advantage of a two-stage procedure based on the modified least squares method. The purpose of the first stage is to reduce the influence of the unknown noise variance. Let us denote

*The paper is supported by the RFBR grant № 09–01–00172a.

$$\tilde{y}_{l-1}^2 = \min\{1, x_{l-1}^2, \dots, x_{l-p}^2\}, \quad \tilde{x}_l = \frac{x_l}{\tilde{y}_{l-1}^2}, \quad \tilde{a}_{l-1} = \left[\frac{1}{\tilde{y}_{l-1}^2}, \frac{x_{l-1}^2}{\tilde{y}_{l-1}^2}, \dots, \frac{x_{l-p}^2}{\tilde{y}_{l-1}^2}\right]^T.$$

Hence, we can represent the process $\{\tilde{x}_l\}$ as

$$\tilde{x}_l = \Lambda \tilde{a}_{l-1} \varepsilon_l. \tag{5}$$

Let us define a compensating factor Γ_n as

$$\Gamma_n = C_n \sum_{l=1}^{n} \tilde{x}_l^2, \quad C_n = B^2 E\left(\sum_{l=1}^{n} \varepsilon_l^2\right)^{-1}. \tag{6}$$

We note that $\Lambda a_{l-1} \geq \lambda_0 + \dots + \lambda_p$. Taking into account this inequality from (6), we have

$$E\frac{1}{\Gamma_n} \leq \frac{1}{B^2\left(\lambda_0 + \dots + \lambda_p\right)}. \tag{7}$$

In the second stage, we construct the parameter estimation in the form

$$\Lambda^* = \Lambda^*(H) = \left(\sum_{l=n+1}^{\tau} v_l z_{l+1} a_l^T\right) A^{-1}(\tau), \quad A(k) = \sum_{l=n+1}^{k} v_l a_l a_l^T, \tag{8}$$

where τ is the random stopping time defined as

$$\tau = \tau\left(H\right) = \min\{k \geq n+1: \quad \nu_{\min}(k) \geq H\}, \tag{9}$$

$\nu_{\min}(k)$ is the minimum eigenvalue of the matrix $A(k)$. Now we define the weights v_l. Let m be the minimum value of k so that $A(n+k)$ will be regular. On the interval $[n+1, n+m-1]$, the weights are chosen as

$$v_l = \begin{cases} \dfrac{1}{\sqrt{\Gamma_n a_l^T a_l}}, & \text{if } a_l \text{ is linearly independent with } \{a_{n+1}, \dots, a_{l-1}\}; \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

The weights on the interval $[n+m, \tau-1]$ are found from the following condition:

$$\frac{\nu_{\min}(k)}{\Gamma_n} = \sum_{l=n+m}^{k} v_l^2 a_l^T a_l. \tag{11}$$

At the instant τ, the weight is found from the condition

$$\frac{\nu_{\min}(\tau)}{\Gamma_n} \geq \sum_{l=n+m}^{\tau} v_l^2 a_l^T a_l, \quad \nu_{\min}(\tau) = H. \tag{12}$$

Properties of the proposed estimation are given in Theorem 1.

Theorem 1. Let the expectation C_n in (6) exists for noise ε_l in (1) for given n. Then the stopping time $\tau(H)$ defined by (9) is finite with probability 1 and mean square accuracy of estimator (8) is bounded from above

$$M\left\|\Lambda^*(H) - \Lambda\right\|^2 \leq \frac{H+p}{H^2}. \tag{13}$$

Proof of Theorem 1. According to [2], stopping time (9) is finite with probability 1 if

$$\sum_{l=1}^{\infty} v_l^2 a_l^T a_l = \infty \text{ a.s.} \tag{14}$$

We note that according to (3), $a_l^T a_l \geq 1$. So the series in (14) converges only if $v_l \xrightarrow{P} 0$ when $l \to \infty$. However, this coefficient is a positive root of a quadratic equation. It is greater then the cosine of the angle between the eigenvector corresponding to the minimum eigenvalue of the matrix $A(l)$ and the vector a_l. This cosine converges to zero if and only if the vector a_l converges to a certain vector. However, according to (1)–(3), any component of the vector a_l

can became equal to 1 when the others do not exceed 1, so a_l does not converge. Hence v_l does not converge to zero and condition (14) holds. So stopping time (9) is finite with probability 1.

Consider the mean square accuracy with estimator (8). Using (4), the Cauchy-Bunyakovskii inequality, inequality $\| A(k) \| \geq \nu_{\min}(k)$, and (12), one can obtain

$$E\left\|\Lambda^*(H)-\Lambda\right\|^2 = E\left\|\sum_{l=n+1}^{\tau}\left(\Lambda v_l a_l a_l^T B\eta_{l+1}\right)A^{-1}(\tau)\right\|^2$$
$$\leq E\left\|\sum_{l=n+1}^{\tau}\left(\Lambda v_l a_l a_l^T B\eta_{l+1}\right)\right\|^2\left\|A^{-1}(\tau)\right\|^2 \leq \frac{1}{H^2}E\left\|\sum_{l=n+1}^{\tau}\Lambda v_l a_l a_l^T B\eta_{l+1}\right\|^2. \qquad (15)$$

Since $\Lambda a_{l-1} \leq \lambda_0 + \ldots + \lambda_p$, the following inequality holds

$$E\left\|\sum_{l=n+1}^{\tau}\Lambda v_l a_l a_l^T B\eta_{l+1}\right\|^2$$
$$\leq (\lambda_0+\ldots+\lambda_p)^2 B^2\left(E\sum_{l=n+1}^{\tau} v_l^2 a_l^T a_l \eta_{l+1}^2 + 2E\sum_{l=n+2}^{\tau}\sum_{k=n+1}^{l-1} v_k v_l a_k^T a_l \eta_{k+1}\eta_{l+1}\right). \qquad (16)$$

To estimate the first term, we now introduce the truncated stopping time $\tau(N)=\min\{\tau, N\}$. It is obvious that $\tau(N)\to\tau$ when $N\to\infty$. Let us consider the variable

$$\sum_{l=n+1}^{\tau(N)} v_l^2 a_l^T a_l \eta_{l+1}^2.$$

It differs from the first term only in the summation limit. Let $F_l=\sigma(\varepsilon_1,\ldots,\varepsilon_l)$ be σ-algebra generated by the variables $\{\varepsilon_1,\ldots,\varepsilon_l\}$. Then using the properties of conditional expectations, one can obtain

$$E\sum_{l=n+1}^{\tau(N)} v_l^2 a_l^T a_l \eta_{l+1}^2 = E\sum_{l=n+1}^{N} v_l^2 a_l^T a_l \chi_{l\leq\tau}\eta_{l+1}^2$$
$$= E\sum_{l=n+1}^{N} E\left\{v_l^2 a_l^T a_l \chi_{l\leq\tau}\eta_{l+1}^2 \middle| F_l\right\} = E\sum_{l=n+1}^{N} v_l^2 a_l^T a_l \chi_{l\leq\tau} E\left\{\eta_{l+1}^2\middle| F_l\right\} = E\sum_{l=n+1}^{\tau(N)} v_l^2 a_l^T a_l.$$

From (10)–(12), one can have

$$E\sum_{l=n+1}^{\tau(N)} v_l^2 a_l^T a_l \eta_{l+1}^2 \xrightarrow[N\to\infty]{} E\sum_{l=n+1}^{\tau} v_l^2 a_l^T a_l = E\sum_{l=n+1}^{n+m-1} v_l^2 a_l^T a_l + E\sum_{l=n+m}^{\tau} v_l^2 a_l^T a_l \leq E\left(\frac{p}{\Gamma_n}+\frac{H}{\Gamma_n}\right) = E\left(\frac{H+p}{\Gamma_n}\right).$$

Similarly, one can demonstrate that the second term in (16) is equal to zero. Using these results in (16) and (15), one gets

$$E\left\|\Lambda^*(H)-\Lambda\right\|^2 \leq \frac{H+p}{H^2}E\left(\frac{B^2(\lambda_0+\ldots+\lambda_p)^2}{\Gamma_n}\right)$$

Using (7), one gets (13).

Asymptotic properties of the estimator. Theorem 1 gives us the boundary of the mean square accuracy with estimator (8). The following theorem gives us an asymptotic boundary for $H\to\infty$.

Theorem 2. If the conditions of Theorem 1 hold and $E\varepsilon_l^4<\infty$, then for sufficiently large H

$$P\left\{\left\|\Lambda^*(H)-\Lambda\right\|^2 > x\right\} \leq (p+1)\left(1-2\Phi\left(\sqrt{\frac{xH}{p+1}}\right)\right),$$

where $\Phi(\cdot)$ is a standard normal distribution function.

Proof of Theorem 2 is based on the proof of the central martingale limit theorem proved in [3]. We have not got a short version of the proof and therefore omit it.

Numerical simulation. We considered the ARCH(1) process specified by (1) with

$\lambda_0 = 0.7$ and $\lambda_1 = 0.5$. For every H, 1000 replications of the experiment were performed. The table below presents the results of simulation of the proposed estimation procedure.

Here λ_0^* and λ_1^* are average estimations of the corresponding parameters, D_0 and D_1 are their standard deviations, Δ_0 and Δ_1 are their maximum deviations, and $\overline{\tau}$ and $\tau_{\max}$ are mean and maximum stopping times.

H	λ_0^*	D_0	Δ_0	λ_1^*	D_1	Δ_1	$\overline{\tau}$	$\tau_{\max}$
10	0.7073	0.0179	0.6670	0.4993	0.0327	0.6698	160	221
20	0.7062	0.0093	0.3736	0.5033	0.0172	0.4254	313	415
40	0.7032	0.0048	0.2581	0.5005	0.0110	0.3075	617	761

The results of numerical simulation prove the efficiency of the procedure. The standard deviations decrease with increasing H. In addition, the mean and maximum estimation times increase linearly with H, and this fact proves high efficiency of the procedure. The maximum stopping time exceeds the mean stopping time by no more than 40%, and this difference decreases with increasing H.

References

1. Dmitrienko A. A., Konev V. V. On the sequential classification of autoregression processes with unknown noise variance. – Problems of Information Transfer, 1995. – V. 31, No. 4, pp. 51–62.
2. Vorobejchikov S. E., Meder N. A. On guaranteed estimation of the parameter of random processes by the weighted least squares method. – Preprints of the 15th Triennial World Congress of the International Federation of Automatic Control Barcelona, 21-26 July 2002, № 1200.
3. Shiryaev A. N. Probability. – Moscow: Nauka, Chief Editorial Staff of Physical-Mathematical Literature, 1980.