Truncated sequential estimation of the parameter of a first order autoregressive process with dependent noises

Dominique Fourdrinier[1], Victor Konev[2] and Sergey Pergamenshchikov[3]

[1] LITIS EA 4108, Université de Rouen,
Avenue de l'Université BP. 12,
F76801, Saint-Étienne-du-Rouvray, Cedex, France
Dominique.Fourdrinier@univ-rouen.fr

[2] Department of Applied Mathematics and Cybernetics,
Tomsk State University,
Lenin str. 36, 634050 Tomsk, Russia
vvkonev@vmm.tsu.ru

[3] Laboratoire de Mathématiques Raphaël Salem,
Université de Rouen,
Avenue de l'Université BP. 12,
F76801, Saint-Étienne-du-Rouvray, Cedex, France
Serge.Pergamenchtchikov@univ-rouen.fr

Abstract. For a first-order non-explosive autoregressive process with dependent noise, we propose a truncated sequential procedure with a fixed mean-square accuracy. The asymptotic distribution of the estimator depends on the type of the noise distribution: it is normal when the noise has a Kotz's distribution, while it is a mixture of normal distributions if the noise distribution is a variance mixture of normals as well. In both cases, the convergence to the limiting distribution is uniform in the unknown parameter.

Keywords. autoregression model, truncated sequential estimators, uniform normality.

1 The model

For a fixed integer $N \geq 1$, consider an autoregressive process of the first order

$$x_{k,N} = \theta\, x_{k-1,N} + \xi_{k,N}\,, \quad 1 \leq k \leq N\,, \tag{1}$$

where $x_{0,N}, x_{1,N}, \ldots, x_{N,N}$ are observations, θ is an unknown parameter and $\xi_{1,N}, \ldots, \xi_{N,N}$ are noises with $\mathbf{E}\,\xi_{j,N} = 0$ and $0 < \mathbf{E}\,\xi^2_{j,N} < \infty$. The initial state $x_{0,N}$ is a random variable which does not depend on θ and which is stochastically independent of the $\xi_{i,N}$'s.

Note that, for that model, the number N of observations is fixed, which is, in many situations, more realistic than to assume that N goes to infinity. In this context, our goal is to obtain a non-asymptotic upper bound and an asymptotic distribution for the general model (1) where the $\xi_{i,N}$'s may be dependent. The type of dependence considered here is the one implied by the class of spherically symmetric distributions. More specifically, we will assume that the random vector $\xi^{(N)} = (\xi_{1,N}, \ldots, \xi_{N,N})$ has a spherically symmetric distribution in R^N with a generating function g_N; in other words, the distribution density of ξ is written as

$$f_{\xi_{1,N}\cdots\xi_{N,N}}(x_1, \ldots, x_N) = g_N(\|x\|^2) \tag{2}$$

for $x = (x_1, \ldots, x_N)$ and $\|x\|^2 = \sum_{i=1}^N x_i^2$.

The classes of spherically symmetric distributions that we consider are the variance mixture of normal distributions with generating function g_N of the form

$$g_N(t) = \mathbf{E}\left[(2\pi\, V)^{-N/2}\, e^{-t/2V}\right]\,, \tag{3}$$

where V is the mixing random variable, and the Kotz distributions (see Kotz (1975)) where g_N is given by

$$g_N(t) = \frac{\Gamma(N/2)}{\pi^{N/2}(2\sigma^2)^{N/2+q}\Gamma(N/2+q)}\, t^q\, e^{-t/2\sigma^2}\,, \tag{4}$$

for some real number q such as $q > -N/2$ and scale parameter $\sigma > 0$.

For the model (1) with limited information (N is fixed), we estimate θ through a truncated sequential procedure in the following way. First, for each positive threshold h, we define the stopping time

$$\tau = \tau_{h,N} = \inf\left\{1 \le l \le N : \sum_{i=1}^{l} x_{i-1,N}^2 \ge h\right\}, \ (\inf\{\emptyset\} = N) \tag{5}$$

and a correction factor

$$\alpha_\tau = \alpha_{\tau_{h,N}} = \frac{h - \sum_{i=1}^{\tau-1} x_{i-1,N}^2}{x_{\tau,N}^2}. \tag{6}$$

Then the truncated sequential estimator associated to h is the couple $(\tau, \tilde{\theta}_{h,N})$ where

$$\tilde{\theta}_{h,N} = \frac{\sum_{i=1}^{\tau-1} x_{i-1,N}\, x_{i,N} + \alpha_\tau\, x_{\tau-1,N}\, x_{\tau,N}}{h} I_{\Gamma_{h,N}}, \tag{7}$$

with $I_{\Gamma_{h,N}}$ the indicator function of the set $\Gamma_{h,N} = \{\sum_{i=1}^{N} x_{i-1,N}^2 \ge h\}$.

2 Non-asymptotic estimation with a prescribed mean-square accuracy

In this section, for the autoregressive model (1), with N fixed observations under the spherically symmetric distribution (2), we show that the truncated sequential estimator $\tilde{\theta}_{h,N}$ in (7) allows to estimate the unknown parameter θ with a prescribed mean-square precision, provided that the generating function g_N meets some requirements on the conditional moment of second order. Those conditions are satisfied for the Kotz distributions (4) and for some mixtures of normal distributions (with bounded mixing variable V in (3)).

Proceeding from (2), the conditional densities are equal to

$$f_{\xi_{k,N}|\xi_{1,N},\dots,\xi_{k-1,N}}(y|\xi_{1,N},\dots,\xi_{k-1,N}) = \frac{g_{k,N}(S_{k-1,N} + y^2)}{g_{k-1,N}(S_{k-1,N})},$$

where $S_{j,N} = \sum_{i=1}^{j} \xi_{i,N}^2$, $1 \le j \le N$, and the conditional second order moments are expressed as

$$\mathbf{E}(\xi_{k,N}^2|\xi_{1,N},\dots,\xi_{k-1,N}) = \psi_{k,N}(S_{k-1,N}), \tag{8}$$

where

$$\psi_{k,N}(z) = \frac{\int_0^\infty t^{1/2}\, g_{k,N}(z+t)\, dt}{g_{k-1,N}(z)}, \quad 1 \le k \le N.$$

Here $g_{k,N}$ is the generating function of the orthogonal projector from $\mathbb{R}^N$ onto $\mathbb{R}^k$, $1 \le k < N$, given by

$$g_{k,N}(z) = \frac{\pi^{(N-k)/2}}{\Gamma((N-k)/2)} \int_0^\infty v^{(N-k)/2-1}\, g_N(z+v)\, dv. \tag{9}$$

Implicitly the cases $k = 1$ and $k = N$ are contained in (8) setting $g_{N,N}(z) = g_N(z)$ and $g_{0,N}(\cdot) \equiv 1$ so that

$$\psi_{1,N}(z) = \int_0^\infty t^{1/2}\, g_{1,N}(z+t)\, dt$$

and

$$\psi_{N,N}(z) = \frac{\int_0^\infty t^{1/2}\, g_N(z+t)\, dt}{g_{N-1,N}(z)}.$$

The properties of $\tilde{\theta}_{h,N}$ will be studied under the following condition on the generating function g_N :

$$d = \sup_{N \ge 1} \max_{1 \le k \le N} \sup_{z \ge 0} \psi_{k,N}(z) < \infty. \tag{10}$$

A first consequence of (10) is given in the following result.

Lemma 1. *For $N \geq 1$, let $\xi^{(N)} = (\xi_{1,N} \ldots, \xi_{N,N})$ be a random vector and $(\mathcal{F}_k^N)_{0 \leq k \leq N}$ be the system of σ-algebras defined by $\mathcal{F}_0^N = \{\emptyset, \Omega\}$ and $\mathcal{F}_k^N = \sigma\{\xi_{1,N} \ldots, \xi_{k,N}\}$. If $\xi^{(N)}$ has a spherically symmetric distribution in R^N with generating function g_N satisfying (10), then $(\xi_{k,N}, \mathcal{F}_k^N)_{1 \leq k \leq N}$ is a martingale-difference sequence such that, for each $1 \leq k \leq N$,*

$$\mathbf{E}\,(\xi_{k,N}^2 \,|\, \mathcal{F}_{k-1}^N) \,\leq\, d\,. \tag{11}$$

The bound d in (10) allows to provide a non-asymptotic result in terms of fixed accuracy estimation. This is the purpose of the following proposition.

Proposition 1. *Assume that the vector of noises $\xi^{(N)} = (\xi_{1,N}, \ldots, \xi_{N,N})$ in (1) has a spherically symmetric distribution in R^N with generating function g_N in (2) satisfying (10). Then, for any fixed $h > 0$ and $\theta_{\max} > 0$, the truncated sequential estimator $\tilde{\theta}_{h,N}$ in (7) satisfies the inequality*

$$\sup_{|\theta| \leq \theta_{\max}} \mathbf{E}_\theta (\tilde{\theta}_{h,N} - \theta)^2 \leq \frac{d}{h} \,+\, R(N, \theta_{\max}, h)\,, \tag{12}$$

where

$$R(N, \theta_{\max}, h) \,=\, \theta_{\max}^2 \, \frac{\pi^{(N-1)/2}}{\Gamma((N-1)/2)} \int_0^{(1+\theta_{\max})^2 h} t^{(N-3)/2} g_{N-1,N}(t)\, dt\,. \tag{13}$$

Remark 1. Inequality (12) gives a possibility to choose the value of h by minimizing the right hand side of (12) which is an upper bound of the mean-square accuracy. In practice, this can be done numerically.

We illustrate the result of Proposition 1 for some specific generating functions g_N (c.f. (2)). First, for the Kotz distribution with nonnegative integer parameter q in (4), it is easy to see, setting $\tilde{h} = h\,(1 + \theta_{\max})^2$ and using the inequality $(t + v)^q \leq 2^{q-1}(t^q + v^q)$, that an upper bound of $R(N, \theta_{\max}, h)$ is given by

$$R(N, \theta_{\max}, h) \leq \theta_{\max}^2 \, 2^q \, \frac{\Gamma(N/2)}{\Gamma(N/2 + q)} \left\{ \frac{\Gamma(q + 1/2)}{\Gamma(1/2)} \, \mathbf{P}(H_0 \leq \tilde{h}/2\sigma^2) + \frac{\Gamma((N-1)/2 + q)}{\Gamma((N-1)/2)} \, \mathbf{P}(H_q \leq \tilde{h}/2\sigma^2) \right\}$$

where, for $m \geq 0$, H_m is a random variable with gamma distribution $\mathcal{G}a((N-1)/2 + m, 1)$.

As another example, consider a variance mixture of normal distributions as in (3). Assume that the mixing random variable V has a density $f(\cdot)$ which decays with power rate at some vicinity of zero, that is, there exist constants $\alpha \geq 0$, $\beta > 0$ and $0 < \delta < 1$ such that $f(v) \leq \alpha\, v^\beta$ for $0 \leq v \leq \delta$. Then, it can be shown that, for any $0 < \epsilon < 1/2$, there exists $\kappa > 0$ such that, for $h \leq (N-1)(1-2\epsilon)\delta/(1+\theta_{max})^2$,

$$R(N, \theta_{\max}, h) \leq \theta_{\max}^2 \left(\frac{\alpha}{\beta + 1} \delta^{\beta+1} + e^{-\kappa(N-1)} \right)\,. \tag{14}$$

Note that, if $\alpha = 0$, then $V \geq \delta$ almost surely and the bound for $R(N, \theta_{\max}, h)$ decreases exponentially in N.

3 Limiting distributions for sequential estimators

The main result of this section yields the asymptotic properties of the truncated sequential estimator $\tilde{\theta}_{h,N}$ given in (7). Specifically, we consider the convergence of its distributions uniformly in $-1 \leq \theta \leq 1$. Although our goal is to deal with the dependent case, it is of interest to start with the i.i.d. case, since (7) gives rise to a specific form of the estimator. Then we present a spherical situation with dependence - the Kotz context with non-negative integer parameter q - which leads to the same uniform asymptotic normality (for that reason, we gather these two cases in Theorem 1). Finally, we consider the case of variance mixtures of normals which turns out to differ from the previous cases.

Theorem 1. *Suppose that, in (1), the vector of noises $\xi^N = (\xi_{k,N})_{1\le k\le N}$ is either an i.i.d. sequence with $\mathbf{E}\,\xi_{j,N} = 0$ and $\mathbf{E}\,\xi^2_{j,N} = \sigma^2$ or has a Kotz's distribution as in (4) where the non-negative parameter q is assumed to be an integer. Then the sequential procedure (7) satisfies*

$$\lim_{h\to\infty}\lim_{N\to\infty}\sup_{|\theta|\le 1}\sup_{t\in\mathbb{R}}\left|\mathbf{P}_\theta\left(\sqrt{h}\,(\tilde\theta_{h,N}-\theta)\le t\right)-\mathbf{\Phi}\left(\frac{t}{\sigma}\right)\right| = 0\,, \tag{15}$$

where $\mathbf{\Phi}(t)$ is the standard normal distribution function.

Note that, for a non-truncated sequential procedure, the first part of this theorem has been established in Lai and Sigmund (1983).

In the following distributional context of model (1), the dependence implied by the spherically symmetric distribution of the noise $\xi^{(N)}$ is such that the conditions of uniform asymptotic normality are broken.

Theorem 2. *Suppose that, in (1), the distribution of the vector of noises $\xi^{(N)} = (\xi_{1,N},\ldots,\xi_{N,N})$ is a variance mixture of normal distributions as in (3). Then the sequential procedure (7) satisfies*

$$\lim_{h\to\infty}\lim_{N\to\infty}\sup_{|\theta|\le 1}\sup_{t\in\mathbb{R}}\left|\mathbf{P}_\theta\left(\sqrt{h}(\tilde\theta_{h,N}-\theta)\le t\right)-\tilde{\mathbf{\Phi}}(t)\right| = 0\,, \tag{16}$$

where $\tilde{\mathbf{\Phi}}(t) = \mathbf{E}\,\mathbf{\Phi}\,(t/V)$.

Note that, in the case of a Kotz's type distribution, we obtain the same limiting theorem as in the case of classical independent noise, while the variance mixture of normals case provides another conclusion. It will be observed that Theorem 2 is a new limiting theorem in the sense that it states that the asymptotic distribution of sequential estimators is no longer normal while the convergence remains uniform in θ. The proofs of Theorem 1 and Theorem 2 can be derived thanks to the material given in Section 4.

4 Uniform weak convergence of martingales

In our study of the uniform asymptotic normality of (7), we will follow the approach developed by Lai and Sigmund in Lai and Sigmund (1983). To this end, we need to extend their basic probabilistic result, given in Proposition 2.1, which proves to be useful in different problems, to the case of dependent noises in the scheme of series.

Let $(\xi_{k,N})_{1\le k\le N}$, $N = 1,2,\ldots$ be a sequence of martingale-differences with respect to the increasing sequences of σ-algebras $(\mathcal{F}^N_k)_{0\le k\le N}$, and let $(x_{k,N})_{0\le k\le N}$ be a process adapted to $(\mathcal{F}^N_k)_{0\le k\le N}$. Let $(\mathbf{P}_\theta, \theta\in\Theta)$ be a family of probability measures and set, for any $a > 0$,

$$\eta_{a,N} \;=\; \max_{1\le k\le N}\sup_{\theta\in\Theta}\mathbf{E}_\theta\left(\xi^2_{k,N}\,\mathbb{1}_{(|\xi_{k,N}|\ge a)}\,|\,\mathcal{F}^N_{k-1}\right)\,. \tag{17}$$

Assume that

$\mathbf{H}_1)$ *the family of random variables $(\eta_{a,N})_{a>0,N\ge1}$ satisfies the relations*

$$\lim_{a\to\infty}\limsup_{N\to\infty}\eta_{a,N} \;=\; 0 \quad\text{and}\quad \sup_{\theta\in\Theta}\mathbf{E}_\theta\limsup_{N\to\infty}\eta_{a,N} \;<\;\infty\,;$$

$\mathbf{H}_2)$ *for each $\delta > 0$,*

$$\lim_{m\to\infty}\overline{\lim}_{N\to\infty}\sup_{\theta\in\Theta}\mathbf{P}_\theta\left(\max_{m\le k\le N}|v_k\;-\sigma^2|\;>\;\delta\right)\;=\;0\,,$$

where $v_k = \mathbf{E}_\theta(\xi^2_{k,N}\,|\mathcal{F}^N_{k-1})$ and σ^2 is some positive constant;

$\mathbf{H}_3)$ *for each $k = 1,2,\ldots$*

$$\lim_{a\to\infty}\overline{\lim}_{N\to\infty}\sup_{\theta\in\Theta}\mathbf{P}_\theta(x^2_{k,N}\;>\;a) = 0\,;$$

$\mathbf{H}_4)$ *for any* $h > 0$,

$$\lim_{N\to\infty} \sup_{\theta\in\Theta} \mathbf{P}_\theta \left(\sum_{k=1}^{N} x_{k-1,N}^2 < h \right) = 0 \,;$$

$\mathbf{H}_5)$ *for each* $\delta > 0$,

$$\lim_{m\to\infty} \limsup_{N\to\infty} \sup_{\theta\in\Theta} \mathbf{P}_\theta \left(x_{n,N}^2 \geq \delta \sum_{j=1}^{n} x_{j-1,N}^2 \ \text{ for some } m \leq n \leq N \right) = 0 \,.$$

The following proposition yields the limiting distribution of the random variable defined, for each $N \geq 1$ and each $h > 0$, as

$$\Delta_{h,N} \;=\; \frac{1}{\sqrt{h}} \sum_{k=1}^{\tau} x_{k-1,N}\, \xi_{k,N} \mathbb{1}_{\Gamma_{h,N}} \,, \tag{18}$$

where $\tau = \tau_{h,N}$ and $\Gamma_{h,N}$ are given in (5) and (7).

Proposition 2. *Under the conditions* $\mathbf{H}_1)$–$\mathbf{H}_5)$ *the random variables* (18) *satisfy the limiting relation*

$$\lim_{h\to\infty} \lim_{N\to\infty} \sup_{\theta\in\Theta} \sup_{t\in\mathbb{R}} \left| \mathbf{P}_\theta(\Delta_{h,N} \leq t) \,-\, \mathbf{\Phi}\left(\frac{t}{\sigma}\right) \right| \;=\; 0\,. \tag{19}$$

Note that the proof of Proposition 2 proceeds along the lines of the proof of Proposition 2.1 in Lai and Sigmund (1983), taking into account, in particular, Condition $\mathbf{H}_4$. Also Proposition 2 enables us to prove Theorem 1. In fact, we are able to establish the conclusion of this theorem under more general conditions (Conditions $\mathbf{C}_1)$–$\mathbf{C}_4)$ below) thanks to the distributional context of Theorem 1.

Assume that the process (1) is unstable (that is, $|\theta| \leq 1$) and that the distribution of noise $\xi^{(N)}$ does not depend on θ and satisfies the following conditions:

$\mathbf{C}_1)$ the random variable

$$\eta_a = \overline{\lim}_{N\to\infty} \max_{1\leq k\leq N} \mathbf{E}\left(\xi_{k,N}^2 \, \mathbb{1}_{(|\xi_{k,N}|\geq a)} \, |\mathcal{F}_{k-1}^N \right)$$

is such that $\lim_{a\to\infty} \eta_a = 0$ and $\mathbf{E}\, \eta_a < \infty$ for some $a \geq 0$;

$\mathbf{C}_2)$ for any $\delta > 0$,

$$\lim_{m\to\infty} \overline{\lim}_{N\to\infty} P(\max_{m\leq k\leq N} |v_{k,N} - \sigma^2| > \delta) = 0\,,$$

where $v_{k,N} = \mathbf{E}(\xi_{k,N}^2 \,|\mathcal{F}_{k-1}^N)$ and σ^2 is some positive number;

$\mathbf{C}_3)$ for any $\delta > 0$ and any $p = 1, 2, \ldots$,

$$\lim_{m\to\infty} \limsup_{N\to\infty} \mathbf{P}\left(A_{m,N,p}\right) \;=\; 0\,,$$

where

$$A_{m,N,p} = \cup_{n=m}^{N} \left\{ \sum_{j=0}^{p-1} \xi_{n-j,N}^2 \geq \delta \sum_{j=1}^{n-1} \xi_{j,N}^2 \right\} \;;$$

$\mathbf{C}_4)$ for any $h > 0$,

$$\lim_{N\to\infty} \mathbf{P}\left(\sum_{j=1}^{N} \xi_{j,N}^2 < h \right) = 0\;.$$

The relevance of $\mathbf{C}_1)$–$\mathbf{C}_4)$ is illustrated by the following lemma.

Lemma 2. *If* $\xi^{(N)}$ *in* (1) *has the properties as in Theorem 1, then Conditions* $\mathbf{C}_1)$–$\mathbf{C}_4)$ *are satisfied.*

The following proposition states Conditions $\mathbf{C}_1)$–$\mathbf{C}_4)$ that are sufficient to give rise to the the uniform asymptotic normality property (15).

Proposition 3. *Suppose that, in* (1)*, the distributions of noises* $\xi^{(N)}$ *satisfy conditions* $\mathbf{C}_1)$–$\mathbf{C}_4)$*. Then the truncated sequential procedure* (7) *has the uniform asymptotic normality property* (15).

5 Concluding remarks

In this paper, estimation of the parameter θ of a first order autoregressive process has been considered under a specific dependence assumption: the vector of noises had a spherically symmetric distribution. Non-asymptotic and asymptotic properties were derived for the truncated sequential estimator $\tilde{\theta}_{h,N}$ that we carried on. Thus a non-asymptotic bound for the risk of $\tilde{\theta}_{h,N}$ was provided. As for the asymptotic properties, an interesting fact was that we could show that the limiting distribution of $\tilde{\theta}_{h,N}$ depends on the specification of the spherically symmetric distribution of the noises. Two cases have been presented: when the distribution is of Kotz type, the limiting distribution of $\sqrt{h}(\tilde{\theta}_{h,N} - \theta)$ is the standard normal distribution while, when a variance mixture of normals is considered, it is also the case for the limiting distribution. The proof of these findings relies on a uniform weak convergence of martingales result.

For us, the choice of the spherically symmetric framework is mainly a first step in the consideration of dependence between the noises; other types of dependence have to be investigated. Also, we restricted ourselves to an $AR(1)$ process as a first start in the treatment of general $AR(p)$ processes; we plan to tackle these models. Finally, we already have the first steps of optimality results for $\tilde{\theta}_{h,N}$; we postpone them to a next paper.

References

Borisov, V. Z. and Konev V. V. (1977). Sequential estimation of parameters of discrete processes. *Automat. and Remote Control*, **10**, 58–64.

Kotz, S. (1975). Multivariate distributions at a cross-road. In: *Statistical Distributions in Scientific Work* (S. Kotz, and J. K. Ord and G. J. Patil, Eds). 245–270.

Lai, T. L. and Sigmund, D. Y. (1983). Fixed accuracy estimation of an autoregressive parameter. *Annals of Statistics*, **11**, 478–485.