PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 22, Fasc. 2 (2002), pp. 293–302

ON ADAPTIVE ESTIMATION BASED ON RANKS IN ARMA PROCESSES

BY

J. ALLAL, A. KAAOUACHI* AND S. MELHAOUI (OUJDA, MOROCCO)

Abstract. This paper describes the adaptive estimation problem based on ranks for the parameter of an ARMA process. The local asymptotic normality property with a ranked based central sequence allows for the construction of estimators which are locally asymptotically minimax (LAM). By using a consistent estimate of the score function, we obtain the adaptive estimators which are LAM and which do not depend on the innovation density.

Key words and phrases: Local asymptotic normality (LAN), asymptotic linearity, locally asymptotically minimax (LAM), adaptivity.

1. INTRODUCTION

The present paper is concerned with the construction of adaptive estimators based on ranks for the parameter of an ARMA process. This process is defined by

$$(1) \qquad X_t - A_1 X_{t-1} - \ldots - A_p X_{t-p} = \varepsilon_t + B_1 \varepsilon_{t-1} + \ldots + B_q \varepsilon_{t-q}, \qquad t \in \boldsymbol{Z},$$

where $(\varepsilon_t;\ t \in \boldsymbol{Z})$ is a white noise, i.e. a family of i.i.d. random variables with mean zero, variance σ^2 and density function g. As usual, it is assumed that the parameter $\theta_0 = (A_1, \ldots, A_p, B_1, \ldots, B_q) \in \boldsymbol{R}^{p+q}$ is such that the stationarity as well as the invertibility conditions are fulfilled.

The adaptive estimation problem was the subject of several searchs since 1956. At this date, Stein discussed conditions under which adaptivity may be possible. These conditions are verified in symmetrical location model (cf. Stone [16]). Fabian and Hannan [5] reformulate some Stein's results in terms of locally asymptotically normal families. A general method of constructing adaptive estimates was originally proposed by Bickel [2]. The adaptive theory is then developed for the class of i.i.d. random variables. For a bibliography, see the monograph by Bickel et al. [3] and the references therein.

* Corresponding author. *E-mail*: kaaouachi@sciences.univ-oujda.ac.ma

For dependent random variables, Kreiss [12] provides adaptive estimates for parameters in ARMA models when the innovation density is symmetric. Then this author exploits the ideas of Stone [16] to construct adaptive estimates of parameters in AR models without the symmetry assumption (cf. Kreiss [13]). Kreiss's work has recently been generalized by Drost et al. [4], Jeganathan [9] and Koul and Schick [11].

The objective of this article is to extend the approach by introducing the rank-based procedures. Our asymptotic considerations are based on a local asymptotic normality (LAN) property established by Hallin and Puri [8], and our methodology utilizes the Hájek–Le Cam theory for locally asymptotically normal families.

Our paper is organized as follows. Section 2 introduces the notation and reviews the technical assumptions. Section 3 states the LAN property of our model with a ranked based central sequence. The construction of estimators which are locally asymptotically minimax (LAM) is given in Section 4. This construction relies on the availability of discretized root-n consistent preliminary estimates of θ_0. Section 5 derives estimators which are LAM and which do not depend on the innovation density. The proof of Proposition 5.1 is given in Section 6.

2. NOTATION AND TECHNICAL ASSUMPTIONS

Let $X^{(n)} = (X_1, \ldots, X_n)$ be an observed series of length n and denote by $H_g^{(n)}(\theta_0)$ the hypothesis under which $X^{(n)}$ is generated by the model (1). Denote by $R_t^{(n)}(\theta_0)$ the rank of the residual

$$Z_t(\theta_0) = \frac{A(L)}{B(L)} X_t, \qquad t = 1, \ldots, n,$$

among $\{Z_j(\theta_0); j = 1, \ldots, n\}$, where L is the lag operator, $A(L) = 1 - \sum_{i=1}^{p} A_i L^i$ and $B(L) = 1 + \sum_{i=1}^{q} B_i L^i$.

We suppose that the vector $(\varepsilon_{-q+1}, \ldots, \varepsilon_0, X_{-p+1}, \ldots, X_0)$ is observed or that $X_t = 0$, $t \leqslant 0$. Such assumptions have no influence on asymptotic results. Then, under $H_g^{(n)}(\theta_0)$, $\{Z_1(\theta_0), \ldots, Z_n(\theta_0)\}$ is a white noise with probability density function g.

Consider the rank autocorrelation coefficient associated with the density g:

$$(2) \quad r_{i,g}^{(n)}(\theta_0) = (n-i)^{-1} \sigma^{-1} I^{-1/2}(g) \sum_{t=i+1}^{n} \varphi_g \circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right) G^{-1}\left(\frac{R_{t-i}^{(n)}(\theta_0)}{n+1}\right),$$

$i = 0, 1, \ldots, n-1$, where G^{-1} is the inverse of the distribution function G associated with g and defined by $G^{-1}(u) = \inf\{x\colon\ G(x) \geqslant u\}$; the quantities $\varphi_g(\cdot)$ and $I(g)$ will be defined bellow in Assumptions (A.1).

Throughout the paper we assume that the following assumptions hold:

ASSUMPTIONS (A.1).

(i) For all $x \in \boldsymbol{R}$, $g(x) > 0$, $\int xg(x)\,dx = 0$ and $0 < \int x^2 g(x)\,dx = \sigma^2 < \infty$.

(ii) There exists a function g' such that $g(b) - g(a) = \int_a^b g'(x)\,dx$ for all $-\infty < a < b < \infty$ (when g is differentiable, g' coincides a.e. with the derivative dg/dx).

(iii) Letting $\varphi_g = -g'/g$, the Fisher information $I(g) = \int [\varphi_g(x)]^2 g(x)\,dx < \infty$ is finite.

(iv) g is strongly unimodal.

(v) $(\varepsilon_{-q+1}; \ldots; \varepsilon_0; X_{-p+1}; \ldots; X_0)$ has a nowhere vanishing joint density $g^0(\cdot, \theta_0)$ satisfying $g^0(\cdot, \theta^{(n)}) - g^0(\cdot, \theta_0) = o_P(1)$, under $H_g^{(n)}(\theta_0)$, as $\theta^{(n)} \to \theta_0$.

(vi) φ_g is piecewise Lipschitz, i.e. there exists a finite partition of $\boldsymbol{R}$ into J non-overlapping intervals I_j such that $|\varphi_g(x) - \varphi_g(y)| \leqslant A_g |x-y|$ for all $j = 1, \ldots, J$ and all $x, y \in I_j$.

All the above assumptions are taken from Hallin and Puri [8].

3. LOCAL ASYMPTOTIC NORMALITY

In this section we shall state the LAN property of our model. For this purpose, let $H_g^{(n)}(\theta_0 + n^{-1/2}\tau^{(n)})$ be the local alternatives, where $(\tau^{(n)}) = (\gamma^{(n)}, \delta^{(n)})'$ is a sequence of real vectors in $\boldsymbol{R}^{p+q}$ such that $\sup_n (\tau^{(n)\prime}\tau^{(n)}) < \infty$.

Denote by $\{\psi_t^{(1)}, \ldots, \psi_t^{(p+q)}; t \in \boldsymbol{Z}\}$ a fundamental system of solutions of the homogeneous equation

$$A(L)B(L)\psi_t = 0, \quad t \in \boldsymbol{Z}.$$

Convenient choices of this system are given in Section 4 of Hallin and Puri [8].

Define the two matrices $C_\psi(\theta_0)$ and $W_\psi^2(\theta_0)$ whose elements are $\psi_t^{(j)}$ and $\sum_{t=1}^{\infty} \psi_t^{(i)} \psi_t^{(j)}$, $i, j = 1, \ldots, p+q$, respectively. Finally, let

$$M(\theta_0) = \begin{bmatrix} 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ g_1 & 1 & \ldots & & h_1 & 1 & \ldots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ g_{p-1} & \ldots & \ldots & 1 & h_{q-1} & \ldots & \ldots & 1 \\ g_p & \ldots & \ldots & g_1 & h_q & \ldots & \ldots & h_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ g_{p+q-1} & \ldots & \ldots & g_q & h_{p+q-1} & \ldots & \ldots & h_p \end{bmatrix},$$

where g_i and h_i are Green's functions associated with the two operators $A(L)$ and $B(L)$, respectively. Note that all the above matrices are continuous in θ_0.

We now consider the vector of rank statistics

$$(3)\quad n^{1/2}\,T_g^{(n)}(\theta_0) = \Big(\sum_{i=1}^{n-1}(n-i)^{1/2}\,\psi_i^{(1)}\,r_{i,g}^{(n)}(\theta_0), \ldots, \sum_{i=1}^{n-1}(n-i)^{1/2}\,\psi_i^{(p+q)}\,r_{i,g}^{(n)}(\theta_0)\Big).$$

Then we have the following LAN property in an ARMA model with a ranked based central sequence.

PROPOSITION 3.1. *Assume that* (A.1) (i)–(v) *hold. Let* $\Lambda = \Lambda^{(n)}_{g;\theta_0+n^{-1/2}\tau^{(n)}/\theta_0}$ *be the log-likelihood ratio for* $H_g^{(n)}(\theta_0+n^{-1/2}\tau^{(n)})$ *with respect to* $H_g^{(n)}(\theta_0)$. *Then, under* $H_g^{(n)}(\theta_0)$,

$$\Lambda = \tau^{(n)\prime}\,\Delta_g^{(n)}(\theta_0) - \tfrac{1}{2}\sigma^2 I(g)\,\tau^{(n)\prime}\,\Gamma(\theta_0)\,\tau^{(n)} + o_p(1) \quad \textit{as } n\to\infty,$$

where

$$\Delta_g^{(n)}(\theta_0) = \sigma I^{1/2}(g)\,M'(\theta_0)\,C_\psi'^{-1}(\theta_0)\,n^{1/2}\,T_g^{(n)}(\theta_0)$$

and

$$\Gamma(\theta_0) = M'(\theta_0)\,C_\psi'^{-1}(\theta_0)\,W_\psi^2(\theta_0)\,C_\psi^{-1}(\theta_0)\,M(\theta_0).$$

Moreover, the limiting distribution of $\Delta_g^{(n)}(\theta_0)$ *under* $H_g^{(n)}(\theta_0)$ *is* $N\big(0, \sigma^2 I(g)\,\Gamma(\theta_0)\big)$.

For the proof see Proposition 4.1 of Hallin and Puri [8]. ∎

As a consequence, we have the following corollary.

COROLLARY 3.1. *Under the assumptions of Proposition* 3.1 *the two sequences of hypotheses* $H_g^{(n)}(\theta_0)$ *and* $H_g^{(n)}(\theta_0+n^{-1/2}\tau^{(n)})$ *are contiguous. In addition, the limiting distribution of* $\Delta_g^{(n)}(\theta_0)-\sigma^2 I(g)\,\Gamma(\theta_0)\,\tau^{(n)}$ *under* $H_g^{(n)}(\theta_0+n^{-1/2}\tau^{(n)})$ *is* $N\big(0, \sigma^2 I(g)\,\Gamma(\theta_0)\big)$.

4. EXISTENCE AND CONSTRUCTION OF LAM ESTIMATORS

In this section, we shall construct estimators which are locally asymptotically minimax. The basic tool to derive their asymptotic normality is the asymptotic linearity of the central sequence.

PROPOSITION 4.1 (Asymptotic linearity). *Assume that* (A.1) *hold. Let* (θ_n) *be a deterministic sequence such that* $n^{1/2}(\theta_n-\theta_0)$ *is bounded by a constant* $c>0$. *Then, under* $H_g^{(n)}(\theta_0)$,

$$(4)\quad \Delta_g^{(n)}(\theta_n) - \Delta_g^{(n)}(\theta_0) + \sigma^2 I(g)\,\Gamma(\theta_0)\,n^{1/2}(\theta_n-\theta_0) = o_P(1) \quad \textit{as } n\to\infty.$$

Proof. Set $n^{1/2}(\theta_n-\theta_0) = \tau^{(n)}$. By Proposition 5.1 of Hallin and Puri [8], we have under $H_g^{(n)}(\theta_0)$,

$$(5)\quad n^{1/2}\big[r_{i,g}^{(n)}(\theta_0+n^{-1/2}\tau^{(n)}) - r_{i,g}^{(n)}(\theta_0)\big] + \sigma I^{1/2}(g)\,(a_i^{(n)}+b_i^{(n)}) = o_P(1) \quad \text{as } n\to\infty,$$

where

$$a_i^{(n)} = \sum_{j=1}^{p} \gamma_j^{(n)} g_{i-j} \quad \text{and} \quad b_i^{(n)} = \sum_{j=1}^{q} \delta_j^{(n)} h_{i-j}.$$

The proof of Proposition 4.1 follows from (5) because $T_g^{(n)}(\cdot)$ is a linear combination of $r_{i,g}^{(n)}(\cdot)$. The treatment consists in decomposing the sum $\sum_{i=1}^{n-1}$ into $\sum_{i=1}^{k}+\sum_{i=k+1}^{n-1}$ (k is a fixed integer such that $k<n-1$), and in proceeding along the same lines as in the proof of Lemma 5.11 of Hallin and Puri [8]. ∎

Let us now construct a sequence of LAM estimators which is based on ranks. For this purpose, let $(\tilde{\theta}_n)$ denote a discretized root-n consistent preliminary estimate of θ_0, i.e. a sequence of statistics such that

(i) $n^{1/2}(\tilde{\theta}_n-\theta_0) = O_P(1)$, under $H_g^{(n)}(\theta_0)$, as $n\to\infty$.

(ii) The number of possible values of $\tilde{\theta}_n$ in balls of the form

$$\{\theta\in \boldsymbol{R}^{p+q}\colon \|n^{1/2}(\theta-\theta_0)\| \leqslant c\}, \qquad c>0 \text{ fixed},$$

remains bounded as $n\to\infty$.

The root-n consistency condition is satisfied by all estimates usually considered in the context of ARMA models (see for example Fuller [6]), and the local discreteness condition goes back to Le Cam [14] and has become an important technical tool in the construction of efficient estimators in semiparametric models; see Bickel et al. [3] and references therein.

PROPOSITION 4.2 (Existence of LAM estimators). *Assume that* (A.1) *hold. Define*

$$\hat{\theta}_n = \tilde{\theta}_n+n^{-1/2}\sigma^{-2}I^{-1}(g)\Gamma^{-1}(\tilde{\theta}_n)\Delta_g^{(n)}(\tilde{\theta}_n). \tag{6}$$

Then, under $H_g^{(n)}(\theta_0)$,

$$n^{1/2}(\hat{\theta}_n-\theta_0) = \sigma^{-2}I^{-1}(g)\Gamma^{-1}(\theta_0)\Delta_g^{(n)}(\theta_0)+o_P(1) \quad \textit{as } n\to\infty,$$

and the asymptotic distribution of $n^{1/2}(\hat{\theta}_n-\theta_0)$ *is* $N(0, \sigma^{-2}I^{-1}(g)\Gamma^{-1}(\theta_0))$.

Proof. The equality follows from (4), the discreteness of $\tilde{\theta}_n$ and the fact that $\Gamma(\tilde{\theta}_n)$ is consistent for $\Gamma(\theta_0)$.

By using the Cramer–Wald device and the limiting distribution of $\Delta_g^{(n)}(\theta_0)$ under $H_g^{(n)}(\theta_0)$, one readily obtains the asymptotic normality of estimators. ∎

Remark 4.1. The formula (6) gives an explicit form of LAM estimators which are based on ranks. This result is also established with different manner in Allal et al. [1].

Remark 4.2. The construction of $\hat{\theta}_n$ in (6) supposes that σ^2 is specified. Otherwise, we replace it by a consistent estimate.

5. CONSTRUCTION OF ADAPTIVE ESTIMATORS BASED ON RANKS

In the previous section, we have constructed a sequence of LAM estimators which is based on ranks. This sequence explicitly depends on the innovation density g.

In order to obtain a sequence of estimators which is LAM and which is independent of g, we construct a central sequence $\hat{\Delta}^{(n)}(\theta_0)$ that does not depend on the unknown density g and is equivalent to $\Delta_g^{(n)}(\theta_0)$ in the sense that

$$\hat{\Delta}^{(n)}(\theta_0)-\Delta_g^{(n)}(\theta_0) = o_P(1), \text{ under } H_g^{(n)}(\theta_0), \quad \text{as } n \to \infty.$$

Our methodology consists in estimating the score function $\varphi_g \circ G^{-1}$ by a consistent estimate $\hat{\varphi}_n$. Suppose that the following additional assumption is satisfied.

Assumption (A.2). There exists an estimate $\hat{\varphi}_n$ of the score function $\varphi_g \circ G^{-1}$, measurable with respect to order statistics of the residual and consistent in the sense that

$$\lim_{n\to\infty} P_{\theta_0}\{\int [\hat{\varphi}_n(u)-\varphi_g \circ G^{-1}(u)]^2\, du > \varepsilon\} = 0. \tag{7}$$

A possible estimate is given in Section VII.1.6 of Hájek and Šidák [7] or in Kim and Van Ryzin [10]. In general, of course, many other examples are available, which in principle can be considered as long as the consistency condition is fulfilled.

We now consider the adaptive rank statistics

$$\hat{r}_i^{(n)}(\theta_0) = (n-i)^{-1}\,\hat{\sigma}_n^{-1}\,\hat{I}_n^{-1/2} \sum_{t=i+1}^{n} \hat{\varphi}_n\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right) Z_{t-i}(\theta_0),$$

where

$$\hat{I}_n = n^{-1}\sum_{t=1}^{n}\left[\hat{\varphi}_n\left(\frac{t}{n+1}\right)\right]^2 \quad \text{and} \quad \hat{\sigma}_n^2 = n^{-1}\sum_{t=1}^{n}[Z_t(\theta_0)]^2. \tag{8}$$

Proposition 5.1. *Assume that* (A.2) *holds. Then*

$$n^{1/2}[\hat{r}_i^{(n)}(\theta_0)-r_{i,g}^{(n)}(\theta_0)] = o_P(1), \tag{9}$$

under $H_g^{(n)}(\theta_0)$, *as* $n \to \infty$.

The proof is rather technical and is given in Section 6. Now letting

$$\hat{\Delta}^{(n)}(\theta_0) = \hat{\sigma}_n \hat{I}_n^{1/2} M'(\theta_0)\, C_\psi'^{-1}(\theta_0)\, n^{1/2}\, \hat{T}^{(n)}(\theta_0),$$

where $\hat{T}^{(n)}(\theta_0)$ is computed by (3) when $\hat{r}_i^{(n)}(\theta_0)$ replaces $r_{i,g}^{(n)}(\theta_0)$.

COROLLARY 5.1. *Under the same assumptions as above, let* $(\tilde{\theta}_n)$ *be a discrete and root-n consistency sequence of estimators for* θ_0. *Then*

$$\hat{\Delta}^{(n)}(\tilde{\theta}_n) - \Delta_g^{(n)}(\tilde{\theta}_n) = o_P(1),$$

under $H_g^{(n)}(\theta_0)$, *as* $n \to \infty$.

Proof. The proof follows by Proposition 5.1 and some algebras. ∎

Now we are ready to present the main result of the paper.

PROPOSITION 5.2. *Assume that* (A.1) *and* (A.2) *hold. Let* $(\tilde{\theta}_n)$ *be a discrete and root-n consistency sequence of estimators for* θ_0. *Define*

$$\hat{\hat{\theta}}_n = \tilde{\theta}_n + n^{-1/2} \hat{\sigma}_n^{-2} \hat{I}_n^{-1} \Gamma^{-1}(\tilde{\theta}_n) \hat{\Delta}^{(n)}(\tilde{\theta}_n),$$

where $\hat{I}_n$ *and* $\hat{\sigma}_n$ *are given by* (8), *when* $\tilde{\theta}_n$ *replaces* θ_0. *Then, under* $H_g^{(n)}(\theta_0)$,

$$n^{1/2}(\hat{\hat{\theta}}_n - \theta_0) = \sigma^{-2} I^{-1}(g) \Gamma^{-1}(\theta_0) \Delta_g^{(n)}(\theta_0) + o_P(1) \quad \text{as } n \to \infty,$$

and the asymptotic distribution of $n^{1/2}(\hat{\hat{\theta}}_n - \theta_0)$ *is* $N(0, \sigma^{-2} I^{-1}(g) \Gamma^{-1}(\theta_0))$.

Proof. Since $\hat{I}_n$ and $\hat{\sigma}_n$ are consistent, we have, in view of the discreteness of $\tilde{\theta}_n$ and Proposition 4.1,

$$\begin{aligned}
& n^{1/2}(\hat{\hat{\theta}}_n - \theta_0) - \sigma^{-2} I^{-1}(g) \Gamma^{-1}(\theta_0) \Delta_g^{(n)}(\theta_0) \\
&= n^{1/2}(\tilde{\theta}_n - \theta_0) + \sigma^{-2} I^{-1}(g) \Gamma^{-1}(\theta_0) (\hat{\Delta}^{(n)}(\tilde{\theta}_n) - \Delta_g^{(n)}(\theta_0)) + o_P(1) \\
&= \sigma^{-2} I^{-1}(g) \Gamma^{-1}(\theta_0) (\Delta_g^{(n)}(\tilde{\theta}_n) - \Delta_g^{(n)}(\theta_0) + \sigma^2 I(g) \Gamma(\theta_0) n^{1/2}(\tilde{\theta}_n - \theta_0)) + o_P(1) \\
&= o_P(1).
\end{aligned}$$

The rest is obvious. ∎

6. PROOF OF PROPOSITION 5.1

In view of the consistency of $\hat{\sigma}_n$ and $\hat{I}_n$, it suffices to prove that

$$n^{-1/2} \sum_{t=i+1}^{n} \left[\varphi_g \circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right) G^{-1}\left(\frac{R_{t-i}^{(n)}(\theta_0)}{n+1}\right) - \hat{\varphi}_n\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right) Z_{t-i}(\theta_0) \right] = o_P(1),$$

under $H_g^{(n)}(\theta_0)$, as $n \to \infty$. This latter summand can be decomposed into $A^{(n)} + B^{(n)}$, where

$$A^{(n)} = n^{-1/2} \sum_{t=i+1}^{n} \left[\hat{\varphi}_n\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right) - \varphi_g \circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right) \right] Z_{t-i}(\theta_0)$$

and

$$B^{(n)} = n^{-1/2} \sum_{t=i+1}^{n} \left[Z_{t-i}(\theta_0) - G^{-1}\left(\frac{R_{t-i}^{(n)}(\theta_0)}{n+1}\right) \right] \varphi_g \circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right).$$

Since, under $H_g^{(n)}(\theta_0)$, the vector of ranks and the vector of order statistics are independent, the conditional mean value and the conditional variance of $A^{(n)}$ and $B^{(n)}$, given the order statistic $Z_{(\cdot)}(\theta_0)$ of residual series, may be established. Particularly, for $A^{(n)}$ we obtain

$$E_{\theta_0}[A^{(n)2}/Z_{(\cdot)}(\theta_0)]$$

$$= n^{-1} E_{\theta_0}\left\{\left(\sum_{t=i+1}^{n}\left[\hat{\varphi}_n\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]Z_{t-i}(\theta_0)\right)^2/Z_{(\cdot)}(\theta_0)\right\}$$

$$= n^{-1}\sum_{t=i+1}^{n} E_{\theta_0}\left\{\left[\hat{\varphi}_n\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]^2 Z_{t-i}^2(\theta_0)/Z_{(\cdot)}(\theta_0)\right\}$$

$$+n^{-1}\sum_{i+1\leqslant t\neq s\leqslant n} E_{\theta_0}\left\{\left[\hat{\varphi}_n\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]\right.$$

$$\left.\times\left[\hat{\varphi}_n\left(\frac{R_s^{(n)}(\theta_0)}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{R_s^{(n)}(\theta_0)}{n+1}\right)\right]Z_{t-i}(\theta_0)Z_{s-i}(\theta_0)/Z_{(\cdot)}(\theta_0)\right\}$$

$$= A_1^{(n)}+A_2^{(n)}.$$

A simple computation yields

$$A_1^{(n)} = [n(n-1)]^{-1}\sum_{1\leqslant i\neq j\leqslant n}\left[\hat{\varphi}_n\left(\frac{i}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{i}{n+1}\right)\right]^2 Z_{(j)}^2(\theta_0)$$

$$\leqslant [n(n-1)]^{-1}\sum_{i=1}^{n}\left[\hat{\varphi}_n\left(\frac{i}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{i}{n+1}\right)\right]^2\sum_{i=1}^{n} Z_i^2(\theta_0)$$

$$\leqslant \int_0^1\left[\hat{\varphi}_n(u)-\varphi_g\circ G^{-1}\left(\frac{1+[nu]}{n+1}\right)\right]^2 du\,(n-1)^{-1}\sum_{i=1}^{n} Z_i^2(\theta_0).$$

Using the fact that

$$\int_0^1\left[\hat{\varphi}_n(u)-\varphi_g\circ G^{-1}\left(\frac{1+[nu]}{n+1}\right)\right]^2 du \to 0 \text{ in probability}$$

and the fact that

$$(n-1)^{-1}\sum_{i=1}^{n} Z_i^2(\theta_0)\to\sigma^2$$

one obtains $A_1^{(n)} = o_P(1)$, under $H_g^{(n)}(\theta_0)$.

Consider now the following quantity:

$$S_{lm} = \sum_{i=1}^{n}\left[\hat{\varphi}_n\left(\frac{i}{n+1}\right)-\varphi_g\circ G^{-1}\left(\frac{i}{n+1}\right)\right]^l Z_{(i)}^m(\theta_0).$$

The term $A_2^{(n)}$ can be decomposed into

$$2(n(n-1)(n-2))^{-1}(n-i)^{-1}(S_{11}S_{01}S_{10}-S_{20}S_{01}-2S_{11}S_{01}S_{10}$$
$$+4S_{12}S_{10}-2S_{11}S_{10}S_{01}-S_{10}^2S_{02}+S_{20}S_{02}+2S_{11}^2+6S_{22}).$$

By using the consistency of $\hat{\varphi}_n$ and some requirements, we can prove that each term of the last decomposition converges to 0 in probability under $H_g^{(n)}(\theta_0)$. Thus, $A_2^{(n)}=o_P(1)$, even under $H_g^{(n)}(\theta_0)$.

On the other hand,

$$E_{\theta_0}[B^{(n)2}/Z_{(\cdot)}(\theta_0)]$$
$$=n^{-1}E_{\theta_0}\left\{\left(\sum_{t=i+1}^{n}\varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\left[Z_{t-i}(\theta_0)-G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]\right)^2\Big/Z_{(\cdot)}(\theta_0)\right\}$$
$$=n^{-1}\sum_{t=i+1}^{n}E_{\theta_0}\left\{\left[\varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]^2\right.$$
$$\left.\times\left[Z_{t-i}(\theta_0)-G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]^2\Big/Z_{(\cdot)}(\theta_0)\right\}$$
$$+n^{-1}\sum_{i+1\leqslant t\neq s\leqslant n}E_{\theta_0}\left\{\left[Z_{t-i}(\theta_0)-\varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\right]\right.$$
$$\times\left[Z_{s-i}(\theta_0)-\varphi_g\circ G^{-1}\left(\frac{R_s^{(n)}(\theta_0)}{n+1}\right)\right]$$
$$\left.\times\ \varphi_g\circ G^{-1}\left(\frac{R_t^{(n)}(\theta_0)}{n+1}\right)\varphi_g\circ G^{-1}\left(\frac{R_s^{(n)}(\theta_0)}{n+1}\right)\Big/Z_{(\cdot)}(\theta_0)\right\}=B_1^{(n)}+B_2^{(n)}.$$

Then we have

$$B_1^{(n)}=[n(n-1)]^{-1}\sum_{1\leqslant i\neq j\leqslant n}\left[\varphi_g\circ G^{-1}\left(\frac{i}{n+1}\right)\right]^2\left[Z_{(j)}(\theta_0)-G^{-1}\left(\frac{j}{n+1}\right)\right]^2$$
$$\leqslant[n(n-1)]^{-1}\sum_{i=1}^{n}\left[\varphi_g\circ G^{-1}\left(\frac{i}{n+1}\right)\right]^2\sum_{i=1}^{n}\left[Z_{(i)}(\theta_0)-G^{-1}\left(\frac{i}{n+1}\right)\right]^2.$$

With the help of the fact that

$$(n-1)^{-1}\sum_{i=1}^{n}\left[Z_{(i)}(\theta_0)-G^{-1}\left(\frac{i}{n+1}\right)\right]^2\to 0\text{ in probability}$$

and the fact that

$$n^{-1}\sum_{i=1}^{n}\left[\varphi_g\circ G^{-1}\left(\frac{i}{n+1}\right)\right]^2\to I(g),$$

one obtains $B_1^{(n)}=o_P(1)$, under $H_g^{(n)}(\theta_0)$.

A similar decomposition as for $A_2^{(n)}$ entails that $B_2^{(n)}$ converges to 0 in probability under $H_g^{(n)}(\theta_0)$. This completes the proof of the proposition.

Acknowledgement. The present research is partially supported by the Programme d'Appui à la Recherche Scientifique (PARS).

REFERENCES

[1] J. Allal, A. Kaaouachi and D. Paindaveine, *R-estimation for ARMA models*, Journal of Nonparametric Statistics 13 (2001), pp. 815–831.
[2] P. J. Bickel, *On adaptive estimation*, Ann. Statist. 10 (1982), pp. 647–671.
[3] P. J. Bickel, C. A. J. Klaassen, Y. Ritov and J. A. Wellner, *Efficient and Adaptive Statistical Inference for Semiparametric Models*, Johns Hopkins University Press, Baltimore 1993.
[4] F. G. Drost, C. A. J. Klaassen and B. J. M. Werker, *Adaptive estimation in time-series models*, Ann. Statist. 25 (1997), pp. 786–817.
[5] V. Fabian and J. Hannan, *On estimation and adaptive estimation for locally asymptotically normal families*, Z. Wahrsch. verw. Gebiete 59 (1982), pp. 459–478.
[6] W. A. Fuller, *Introduction to Statistical Time Series*, 2nd ed., Wiley, New York 1976.
[7] J. Hájek and Z. Šidák, *Theory of Rank Tests*, Academic Press, New York 1967.
[8] M. Hallin and M. L. Puri, *Aligned rank tests for linear models with autocorrelated error terms*, J. Multivariate Anal. 50 (1994), pp. 175–237.
[9] P. Jeganathan, *Some aspects of asymptotic theory with applications to time series models*, Econometric Theory 11 (1995), pp. 818–887.
[10] B. K. Kim and J. Van Ryzin, *Estimation of the derivative of a density and related functions*, Comm. Statist. Theory Methods A 9 (7) (1985), pp. 697–709.
[11] H. L. Koul and A. Schick, *Efficient estimation in nonlinear autoregressive time series models*, Bernoulli 3 (1997), pp. 247–277.
[12] J. P. Kreiss, *On adaptive estimation in stationary ARMA processes*, Ann. Statist. 15 (1987), pp. 112–133.
[13] J. P. Kreiss, *On adaptive estimation in autoregressive models when there are nuisance functions*, Statist. Decisions 5 (1987), pp. 59–76.
[14] L. Le Cam, *Locally asymptotically normal families of distributions*, Univ. Calif. Publ. Statist. 3 (1960), pp. 37–98.
[15] C. Stein, *Efficient nonparametric testing and estimation*, Proc. Third Berkeley Sympos. Math. Statist. and Probab. (Univ. Calif. Press) I (1956), pp. 187–195.
[16] C. J. Stone, *Adaptive maximum likelihood estimators of a location parameter*, Ann. Statist. 3 (1975), pp. 267–284.

J. Allal, A. Kaaouachi and S. El Melhaoui
Département de Mathématiques
Faculté des Sciences
Université Mohamed Premier
Oujda 60000 Morocco
E-mail: allal@sciences.univ-oujda.ac.ma
kaaouachi@sciences.univ-oujda.ac.ma

Received on 5.9.2002