R-ESTIMATION FOR ARMA MODELS

By J. Allal, A. Kaaouachi and D. Paindaveine

Université Mohamed Ier, Oujda (Morocco)
Université Libre de Bruxelles (Belgium)

ABSTRACT

This paper is devoted to the R-estimation problem for the parameter of a stationary ARMA model. The asymptotic uniform linearity of a suitable vector of rank statistics leads to the asymptotic normality of $\sqrt{n}$-consistent R-estimates resulting from the minimization of the norm of this vector. By using a discretized $\sqrt{n}$-consistent preliminary estimate, we construct a new class of one-step R-estimators. We compute the asymptotic relative efficiency of the proposed estimators with respect to the LS estimator. Efficiency properties are investigated via a Monte-Carlo study in the particular case of an AR(1) model.

KEYWORDS AND PHRASES: R-estimation, ARMA models, local asymptotic normality, asymptotic linearity.

1 Introduction

R-estimation — estimation methods based on ranks — was initiated by Hodges and Lehmann (1963), who proposed the first R-estimator for location. The theory subsequently has been developed and systematized in the general framework of linear models with independent errors; we refer to the papers by Adichie (1967), Jurečková (1971), Koul (1971) and Jaeckel (1972), and the monographs by Puri and Sen (1985) and Jurečková and Sen (1996) for a detailed account and extensive bibliography.

Introducing R-estimation ideas to the context of time series analysis has been much slower. A first successfull attempt has been made by Allal (1991), who proposes a R-estimator for the parameter of a first-order autoregressive (AR(1)) model, based on the serial linear rank statistics introduced by Hallin et al. (1985) and Hallin and Puri (1988). A theory of R-estimation has been developed for higher order AR(p) models by Koul and Saleh (1993); the statistics on which their method relies however are not genuinely rank-based, as they involve both the ranks (of residuals) and the observations themselves.

The purpose of this paper is to investigate the problem of R-estimation in the more general context of ARMA(p, q) models, of the form

$$X_t - A_1 X_{t-1} - ... - A_p X_{t-p} = \varepsilon_t + B_1 \varepsilon_{t-1} + ... + B_q \varepsilon_{t-q}, \quad t \in \mathbb{Z}, \tag{1}$$

where $\{\varepsilon_t; t \in \mathbb{Z}\}$ is white noise with unspecified density g. Contrary to Koul and Saleh (1993), the statistics we are using are genuinely rank-based, in the sense that they only depend on the vector of (residual) ranks.

An important tool in our approach is the LAN property of ARMA models for given g (satisfying mild regularity assumptions). This LAN result was established in Kreiss (1987, a and b). More particularly, we will use a rank-based version of this result, as given in Hallin and Puri (1994).

The paper is organized as follows. Section 2 introduces notation and the main technical assumptions. Section 3 presents the asymptotic representation and asymptotic normality results for serial rank statistics to be used in the sequel. Section 4 studies the R-estimation problem in ARMA model. Section 5 gives a method of constructing, with the help of a preliminary $\sqrt{n}$-consistent estimate of $\boldsymbol{\theta}_0$, a class of asymptotic R-estimators. Section 6 provides the asymptotic relative efficiencies of R-estimators with respect to the least-squares estimate. In Section 7, we investigate the finite-sample performance of the proposed estimates via a Monte-Carlo study. The appendix contains details of the proof of Proposition 3.3.

2 Notation and basic assumptions

In this section we introduce a class of serial rank statistics and state the basic assumptions to be made. The ARMA(p, q) model (1) can be written under the form

$$A(L)X_t = B(L)\varepsilon_t, \quad t \in \mathbb{Z},$$

where L is the lag operator, $A(L) := 1 - \sum_{i=1}^{p} A_i L^i$ and $B(L) := 1 + \sum_{i=1}^{q} B_i L^i$. The parameter $\boldsymbol{\theta}_0 = (A_1, ..., A_p, B_1, ..., B_q) \in \mathbb{R}^{p+q}$ is chosen in such a way that both the stationarity and the invertibility conditions are fulfilled.

Let $\mathbf{X}^{(n)} := (X_1^{(n)}, \ldots, X_n^{(n)})$ be an observed series of length n, and denote by $H_g^{(n)}(\boldsymbol{\theta}_0)$ the hypothesis under which $X^{(n)}$ is generated by model (1). Denote by $R_t^{(n)}(\boldsymbol{\theta}_0)$ the rank of the residual $Z_t^{(n)}(\boldsymbol{\theta}_0)$ among $\left\{Z_1^{(n)}(\boldsymbol{\theta}_0), \ldots, Z_n^{(n)}(\boldsymbol{\theta}_0)\right\}$, where

$$Z_t^{(n)}(\boldsymbol{\theta}_0) := \frac{A(L)}{B(L)} X_t^{(n)}, \quad t = 1, \ldots, n.$$

We suppose that the vector $(\varepsilon_{-q+1}, \ldots, \varepsilon_0, X_{-p+1}, \ldots, X_0)$ is observed, or that $X_t^{(n)} = 0$, $t \leq 0$. Such assumptions have no influence on asymptotic results. Then, under $H_g^{(n)}(\boldsymbol{\theta}_0)$, $\left\{Z_1^{(n)}(\boldsymbol{\theta}_0), \ldots, Z_n^{(n)}(\boldsymbol{\theta}_0)\right\}$ is white noise, with probability density function g.

Consider the serial rank statistics of order k ($k = 1, \ldots, n-1$), known as a rank autocorrelation of order k,

$$r_k^{(n)}(\boldsymbol{\theta}_0) := \left\{(n-k)^{-1} \sum_{t=k+1}^{n} J_1\left(\frac{R_t^{(n)}(\boldsymbol{\theta}_0)}{n+1}\right) J_2\left(\frac{R_{t-k}^{(n)}(\boldsymbol{\theta}_0)}{n+1}\right) - m^{(n)}\right\} \Big/ \sigma_k^{(n)}, \tag{2}$$

where J_1 and J_2 are score functions, and $m^{(n)}$ and $\sigma_k^{(n)}$ are normalizing constants such that $(n-k)^{1/2} r_k^{(n)}(\boldsymbol{\theta}_0)$ is standardized under $H^{(n)}(\boldsymbol{\theta}_0)$ (the hypothesis under which the density g remains unspecified). See Hallin and Puri (1994), pp. 186-187 for explicit values of $m^{(n)}$ and $\sigma_k^{(n)}$.

Also denote by $r_{k,g}^{(n)}$ the g-rank autocorrelation of order k, i.e., the rank statistic (2) with $J_1 := \varphi_g \circ G^{-1}$ and $J_2 := G^{-1}$, where $\varphi_g(.) := \frac{-g'}{g}(.)$ and G^{-1} denotes the generalized inverse of the cdf G associated with g.

Throughout the paper we assume that the following assumptions hold for the error density g and the score functions J_1 and J_2.

Assumptions *A.*1**.**

(*i*) $\int xg(x)\,dx = 0$ and $0 < \sigma^2 := \int x^2 g(x)\,dx < \infty$.

(*ii*) g is absolutely continuous, with a.e. derivative g', and strongly unimodal.

1. The Fisher information $I(g) := \int \left[\frac{g'(x)}{g(x)}\right]^2 g(x)\,dx$ is finite.

(*iii*) $g(x) > 0 \quad \forall x \in \mathbb{R}$ and $(\varepsilon_{-q+1};\ldots;\varepsilon_0;X_{-p+1};\ldots;X_0)$ possesses a nowhere vanishing joint density $g^0(.,\boldsymbol{\theta})$ that satisfies $g^0(.,\boldsymbol{\theta}^{(n)}) - g^0(.,\boldsymbol{\theta}_0) = o_p(1)$, under $H_g^{(n)}(\boldsymbol{\theta}_0)$, as $\boldsymbol{\theta}^{(n)} \to \boldsymbol{\theta}_0$.

Assumptions *A.*2**.**

(*i*) J_1 and J_2 are nondecreasing and square-integrable functions such that $\int_0^1 J_i(u)\,du = 0$, $i = 1, 2$.

(*ii*) $J_1 \circ G$ and $J_2 \circ G$ are Lipschitz.

Remark 2.1

Assumptions $A.1$ are used in proving the LAN property. Assumption $A.2(ii)$ is verified, for example, if $J_i = \Phi^{-1}(.)$ and G is normal ($\Phi(.)$ stands for the standard normal distribution function), $J_i(u) = 2u - 1$ and G is normal or logistic or $J_i(u) = \ln(\frac{u}{1-u})$ and G is logistic. It easily can be weakened into a piecewise Lipschitz assumption, which also accomodates such distributions as the double exponential.

3 Asymptotic representation and asymptotic normality

Consider the sequence of local alternatives $H_g^{(n)}\left(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}\right)$, where $\boldsymbol{\tau}^{(n)} := (\boldsymbol{\gamma}^{(n)}, \boldsymbol{\delta}^{(n)}) \in \mathbb{R}^{p+q}$ is such that $\sup_n \left(\boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)}\right) < \infty$, and denote by $\left\{\psi_t^{(1)}, \ldots, \psi_t^{(p+q)}; t \in \mathbb{Z}\right\}$ a fundamental system of solutions of the homogeneous equation

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

(convenient choices of a fundamental system are given in Section 4 of Hallin and Puri (1994)).

Associated with this fundamental system, denote by $\boldsymbol{C}_\psi(\boldsymbol{\theta}_0)$ and $\boldsymbol{W}^2_\psi(\boldsymbol{\theta}_0)$ the matrices whose elements are $\psi_i^{(j)}$ and $\sum_{t=1}^{\infty} \psi_t^{(i)} \psi_t^{(j)}$ $(i,j = 1,\ldots,p+q)$, respectively. Finally, let

$$\boldsymbol{M}(\boldsymbol{\theta}_0) := \begin{pmatrix} 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ g_1 & 1 & \ldots & & h_1 & 1 & \ldots & \\ \vdots & & & \vdots & \vdots & & & \vdots \\ g_{p-1} & \ldots & \ldots & 1 & h_{q-1} & \ldots & \ldots & 1 \\ g_p & \ldots & \ldots & g_1 & h_q & \ldots & \ldots & h_1 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ g_{p+q-1} & \ldots & \ldots & g_q & h_{p+q-1} & \ldots & \ldots & h_p \end{pmatrix},$$

where g_i and h_i are the Green's functions associated with the operators $A(L)$ and $B(L)$ respectively. Note that all the above quantities are continuous functions of $\boldsymbol{\theta}_0$. In the sequel, the notation $\psi_t^{(i)}(\boldsymbol{\theta}_0)$, $\boldsymbol{C}_\psi(\boldsymbol{\theta}_0)$, $\boldsymbol{M}(\boldsymbol{\theta}_0)$ and $\boldsymbol{W}^2_\psi(\boldsymbol{\theta}_0)$ will be avoided.

Now define the vector of rank statistics

$$\sqrt{n}\,\boldsymbol{T}^{(n)}(\boldsymbol{\theta}_0) := \left(\sum_{k=1}^{n-1}(n-k)^{1/2}\psi_k^{(1)}(\boldsymbol{\theta}_0) r_{k,g}^{(n)}(\boldsymbol{\theta}_0), \ldots, \sum_{k=1}^{n-1}(n-k)^{1/2}\psi_k^{(p+q)}(\boldsymbol{\theta}_0) r_{k,g}^{(n)}(\boldsymbol{\theta}_0)\right)'. \tag{3}$$

Then we can show the following proposition :

Proposition 3.1 (Asymptotic representation) *Assume that A.1 and A.2(i) hold. Then,*

(i) $n^{1/2} r_k^{(n)}(\boldsymbol{\theta}_0) = I^{-1}_{(J_1,J_2)}\, n^{-1/2} \sum_{t=k+1}^{n} J_1 \circ G(Z_t^{(n)}(\boldsymbol{\theta}_0)) J_2 \circ G(Z_{t-k}^{(n)}(\boldsymbol{\theta}_0)) + o_p(1)$, *under* $H_g^{(n)}(\boldsymbol{\theta}_0)$, *as* $n \to \infty$, *where* $I^2_{(J_1,J_2)} := \int_0^1 [J_1(u)]^2\, du \int_0^1 [J_2(u)]^2\, du$;

(ii) the rank statistics $n^{1/2} r_k^{(n)}(\boldsymbol{\theta}_0)$ *is asymptotically standard normal under* $H^{(n)}(\boldsymbol{\theta}_0)$

Proof. (i) See Section 4 of Hallin et al. (1985).
(ii) Use (i) and the central limit theorem of k-dependent random variables (See Yoshihara (1976)). $\square$

Proposition 3.2 (Hallin and Puri 1994) Local asymptotic normality.
Assume that A.1 and A.2(i) hold. Let $\Lambda = \Lambda^{(n)}_{\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}/\boldsymbol{\theta}_0;g}$ *be the log-likelihood ratio for* $H_g^{(n)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$ *with respect to* $H_g^{(n)}(\boldsymbol{\theta}_0)$. *Then, under* $H_g^{(n)}(\boldsymbol{\theta}_0)$,

$$\Lambda = \boldsymbol{\tau}^{(n)'}\boldsymbol{\Delta}_g^{(n)}(\boldsymbol{\theta}_0) - \frac{1}{2}\sigma^2 I(g) \boldsymbol{\tau}^{(n)'}\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\boldsymbol{\tau}^{(n)} + o_p(1),$$

as $n \to \infty$, *where*

$$\boldsymbol{\Delta}_g^{(n)}(\boldsymbol{\theta}_0) := \sigma[nI(g)]^{1/2} \boldsymbol{M}'(\boldsymbol{\theta}_0) \boldsymbol{C}'^{-1}_\psi(\boldsymbol{\theta}_0) \boldsymbol{T}^{(n)}(\boldsymbol{\theta}_0)$$

and

$$\mathbf{\Gamma}(\boldsymbol{\theta}_0) := \boldsymbol{M}'(\boldsymbol{\theta}_0)\boldsymbol{C}'^{-1}_{\psi}(\boldsymbol{\theta}_0)\boldsymbol{W}^2_{\psi}(\boldsymbol{\theta}_0)\boldsymbol{C}^{-1}_{\psi}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0).$$

Moreover, the limiting distribution of $\mathbf{\Delta}^{(n)}_g(\boldsymbol{\theta}_0)$ *under* $H^{(n)}_g(\boldsymbol{\theta}_0)$ *is* $N\big(0, \sigma^2 I(g)\mathbf{\Gamma}(\boldsymbol{\theta}_0)\big)$.

Proof. See Proposition 4.1 of Hallin and Puri (1994). □

As a consequence we obtain the following corollary.

Corollary 3.1 *Assume that A.1 and A.2(i) hold. Then, the following properties hold :*
(i) The limiting distribution of $\mathbf{\Delta}^{(n)}_g(\boldsymbol{\theta}_0) - \sigma^2 I(g)\mathbf{\Gamma}(\boldsymbol{\theta}_0)\boldsymbol{\tau}^{(n)}$ *under* $H^{(n)}_g(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$ *is* $N\big(0, \sigma^2 I(g)\mathbf{\Gamma}(\boldsymbol{\theta}_0)\big)$.
(ii) The limiting distribution of $n^{1/2} r^{(n)}_k(\boldsymbol{\theta}_0)$ *under* $H^{(n)}_g(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$ *is* $N\Big(c(J_1, J_2, g)(\boldsymbol{a}^{(n)}_k + \boldsymbol{b}^{(n)}_k), 1\Big)$, *where*

$$c(J_1, J_2, g) := I^{-1}_{(J_1,J_2)} \int_0^1 J_1(u)\varphi_g \circ G^{-1}(u)\, du \int_0^1 J_2(u) G^{-1}(u)\, du.$$

and

$$a^{(n)}_k := \sum_{j=1}^{p} \gamma^{(n)}_j g_{k-j} \quad \text{and} \quad b^{(n)}_k := \sum_{j=1}^{q} \delta^{(n)}_j h_{k-j}.$$

Proof. (i) This follows by applying Le Cam's third lemma (see Hájek and Šidák (1967)) in Proposition 3.2.
(ii) The asymptotic normality under $H^{(n)}_g(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$ follows by piecing together Proposition 3.1, the asymptotic joint normality of $\Big(n^{1/2} r^{(n)}_k(\boldsymbol{\theta}_0), \Lambda^{(n)}_{\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}/\boldsymbol{\theta}_0;g}\Big)$ under $H^{(n)}_g(\boldsymbol{\theta}_0)$, and Le Cam's third lemma. □

Let us now turn to asymptotic uniform linearity of the rank autocorrelation coefficients :

Proposition 3.3 *Assume that A.1 and A.2 hold. Then, for all k and all* $c > 0$, *under* $H^{(n)}_g(\boldsymbol{\theta}_0)$, *as* $n \to \infty$,

$$\sup_{\|\boldsymbol{\tau}^{(n)}\| \le c} \left| n^{1/2} \left[r^{(n)}_k(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}) - r^{(n)}_k(\boldsymbol{\theta}_0) \right] + c(J_1, J_2, g)(a^{(n)}_k + b^{(n)}_k) \right| = o_p(1). \tag{4}$$

The proof is rather technical and we defer it to the Appendix.

4 Estimators based on ranks

In this section we introduce a class of R-estimates based on ranks for the parameter $\boldsymbol{\theta}_0$ of the ARMA model.

4.1 Definition of the R-estimators

Define the vector of rank statistics

$$\sqrt{n}\,\boldsymbol{T}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0) := \left(\sum_{k=1}^{n-1}(n-k)^{1/2}\psi^{(1)}_k(\boldsymbol{\theta}_0)r^{(n)}_k(\boldsymbol{\theta}_0),\ldots,\sum_{k=1}^{n-1}(n-k)^{1/2}\psi^{(p+q)}_k(\boldsymbol{\theta}_0)r^{(n)}_k(\boldsymbol{\theta}_0)\right)'$$

and put

$$\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0) := n^{1/2}\boldsymbol{M}'(\boldsymbol{\theta}_0)\boldsymbol{C}'^{-1}_{\psi}(\boldsymbol{\theta}_0)\boldsymbol{T}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0).$$

Obviously, when the scores $J_1(.)$ and $J_2(.)$ are $\varphi_g\circ G^{-1}(.)$ and $G^{-1}(.)$ respectively, we obtain the central sequence up to a positive factor. More precisely,

$$\boldsymbol{\Delta}^{(n)}_{g}(\boldsymbol{\theta}_0) = \sigma I^{1/2}(g)\boldsymbol{\Delta}^{(n)}_{\varphi_g\circ G^{-1},G^{-1}}(\boldsymbol{\theta}_0).$$

It easily follows from Proposition 3.1 that the vector $\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0)$ is asymptotically normal, under $H^{(n)}_g(\boldsymbol{\theta}_0)$, with mean zero and covariance matrix $\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)$. This suggests estimating the unknown parameter $\boldsymbol{\theta}_0$ by the value of $\boldsymbol{\theta}$ for which $\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta})$ is as near to zero as possible, i.e., to estimate $\boldsymbol{\theta}_0$ by

$$\bar{\boldsymbol{\theta}}^{(n)} := \arg\min_{\boldsymbol{\theta}}\|\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta})\|, \tag{5}$$

where $\|.\|$ is any standard norm in $\mathbb{R}^{p+q}$.

4.2 Asymptotic uniform linearity

In this subsection we give the asymptotic uniform linearity of $\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)})$ in $\|\boldsymbol{\tau}^{(n)}\|\le c$.

Proposition 4.1 *Assume that A.1 and A.2 hold. Then for all $c>0$, under $H^{(n)}_g(\boldsymbol{\theta}_0)$, as $n\to\infty$,*

$$\sup_{\|\boldsymbol{\tau}^{(n)}\|\le c}\left\|\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)})-\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0)+c(J_1,J_2,g)\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\boldsymbol{\tau}^{(n)}\right\| = o_p(1). \tag{6}$$

Proof. For $s=1,\ldots,p+q$, denote by $T^{(n)s}_{J_1,J_2}(.)$ the sth component of $\sqrt{n}\,\boldsymbol{T}^{(n)}_{J_1,J_2}(.)$.
We proceed by proving that

$$\sup_{\|\boldsymbol{\tau}^{(n)}\|\le c}\left|T^{(n)s}_{J_1,J_2}(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)})-T^{(n)s}_{J_1,J_2}(\boldsymbol{\theta}_0)+c(J_1,J_2,g)\sum_{k=1}^{\infty}\psi^{(s)}_k(\boldsymbol{\theta}_0)(a^{(n)}_k+b^{(n)}_k)\right| = o_p(1),$$

under $H^{(n)}_g(\boldsymbol{\theta}_0)$, as $n\to\infty$. We have

$$\begin{aligned}
&T^{(n)s}_{J_1,J_2}(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)})-T^{(n)s}_{J_1,J_2}(\boldsymbol{\theta}_0)+c(J_1,J_2,g)\sum_{k=1}^{\infty}\psi^{(s)}_k(\boldsymbol{\theta}_0)(a^{(n)}_k+b^{(n)}_k) = \qquad (7)\\
&\sum_{k=1}^{n-1}\psi^{(s)}_k(\boldsymbol{\theta}_0)\left\{(n-k)^{1/2}\left[r^{(n)}_k(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)})-r^{(n)}_k(\boldsymbol{\theta}_0)\right]+c(J_1,J_2,g)(a^{(n)}_k+b^{(n)}_k)\right\}\\
&\quad +c(J_1,J_2,g)\sum_{k=n}^{\infty}\psi^{(s)}_k(\boldsymbol{\theta}_0)(a^{(n)}_k+b^{(n)}_k)\\
&\quad +\sum_{k=1}^{n-1}(n-k)^{1/2}\left[\psi^{(s)}_k(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)})-\psi^{(s)}_k(\boldsymbol{\theta}_0)\right]r^{(n)}_k(\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)}).
\end{aligned}$$

Since the AUL result of Proposition 3.3 holds uniformly over the set $\{k = 1, \ldots, n-1\}$, the first term of the right side of (7) is $\bar{o}_p(1)$, as $n \to \infty$, where $\bar{o}_p(1)$ stands for some sequence converging to zero uniformly over the set $\{\|\boldsymbol{\tau}^{(n)}\| \le c\}$. The particular form of $a_k^{(n)}$ and $b_k^{(n)}$, and the exponential decrease in k of $\psi_k^{(s)}(.)$ imply that the second term of the right side of (7) converges to 0, uniformly in $\|\boldsymbol{\tau}^{(n)}\| \le c$ as $n \to \infty$. The exact variance of the third term, under $H_g^{(n)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$, is bounded by

$$\sum_{k=1}^{\infty}\left[\psi_k^{(s)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}) - \psi_k^{(s)}(\boldsymbol{\theta}_0)\right]^2 + \left(\sum_{k=1}^{\infty}\left|\psi_k^{(s)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}) - \psi_k^{(s)}(\boldsymbol{\theta}_0)\right|\right)^2.$$

This latter quantity is $o(1)$ as $n \to \infty$; indeed, $\sup\left\{\left|\psi_k^{(s)}(\boldsymbol{\theta})\right| : s = 1, \ldots, p+q; \boldsymbol{\theta} \in K\right\} \le A\, b^i$, where $A > 0$, $0 < b < 1$ and K is a compact set in $\mathbb{R}^{p+q}$, and $\boldsymbol{\theta} \mapsto \psi_k^{(s)}(\boldsymbol{\theta})$ is continuous. Consequently, under $H_g^{(n)}(\boldsymbol{\theta}_0)$, hence also, due to contiguity, under $H_g^{(n)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$, the third term on the right side of (7) is $\bar{o}_p(1)$, as $n \to \infty$. This completes the proof of (6).

The end of the proof is obvious, because $\boldsymbol{M}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$ and $\boldsymbol{C}_\psi(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)})$ converge respectively to $\boldsymbol{M}(\boldsymbol{\theta}_0)$ and $\boldsymbol{C}_\psi(\boldsymbol{\theta}_0)$, uniformly in $\|\boldsymbol{\tau}^{(n)}\| \le c$. □

4.3 Asymptotic normality of R-estimates

This subsection gives the limit law of R-estimates.

Proposition 4.2 *Assume that A.1 and A.2 hold. Let* $(\bar{\boldsymbol{\theta}}^{(n)})$ *be a* $\sqrt{n}$*-consistent solution of (5), then the asymptotic distribution of* $\sqrt{n}\left(\bar{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0\right)$ *under* $H_g^{(n)}(\boldsymbol{\theta}_0)$ *is* $N\big(0, c^{-2}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)\big)$.

Proof. From Proposition 3.1, the vector $\boldsymbol{\Delta}_{J_1,J_2}^{(n)}(\boldsymbol{\theta}_0)$ is asymptotically normal with mean 0 and covariance matrix $\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)$, under $H^{(n)}(\boldsymbol{\theta}_0)$. The result follows from the Cramér-Wald device and Proposition 4.1. □

Proposition 4.3 *Assume that A.1 and A.2 hold, and suppose that* $\boldsymbol{\lambda}'\boldsymbol{\Delta}_{J_1,J_2}^{(n)}(\boldsymbol{\theta}_0 + b\, n^{-1/2}\boldsymbol{\lambda})$ *is monotone in* b *for every* $\|\boldsymbol{\lambda}\| = 1$*. Then, for any solution* $(\bar{\boldsymbol{\theta}}^{(n)})$ *of (5), the asymptotic distribution of* $\sqrt{n}\left(\bar{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0\right)$ *under* $H_g^{(n)}(\boldsymbol{\theta}_0)$ *is* $N\big(0, c^{-2}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)\big)$.

Proof. The proof proceeds along the same lines as in Jurečková (1971). □

However, it is not easy to prove the $\sqrt{n}$-consistency of any solution of (5), or to find reasonable assumptions ensuring the monotonicity in b of $\boldsymbol{\lambda}'\boldsymbol{\Delta}_{J_1,J_2}^{(n)}(\boldsymbol{\theta}_0 + b\, n^{-1/2}\boldsymbol{\lambda})$. That is why we propose the following construction of R-estimators, that allows for trivial asymptotic properties and achievable local and asymptotic optimality.

5 Construction of asymptotic R-estimates

In this section we adopt the Le Cam-Hájek approach to construct a class of asymptotic R-estimators $\hat{\boldsymbol{\theta}}^{(n)}$ satisfying

$$\boldsymbol{\Delta}_{J_1,J_2}^{(n)}(\hat{\boldsymbol{\theta}}^{(n)}) = o_{P_{\boldsymbol{\theta}_0}}(1).$$

For this purpose, let $(\tilde{\boldsymbol{\theta}}^{(n)})$ denote a discretized $\sqrt{n}$-consistent preliminary estimate of $\boldsymbol{\theta}_0$, i.e., a sequence of estimates such that

(i) $\sqrt{n}\left(\tilde{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0\right) = O_{P_{\boldsymbol{\theta}_0}}(1)$, as $n \to \infty$,

(ii) the number of possible values of $\tilde{\boldsymbol{\theta}}^{(n)}$ in balls of the form

$$\left\{\boldsymbol{\theta} \in \mathbb{R}^{p+q} : \sqrt{n}\, \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq c\right\}, \quad c > 0 \ \text{ fixed}$$

remains bounded, as $n \to \infty$.

The $\sqrt{n}$-consistent condition is satisfied by all estimates that are usually considered in the context of ARMA models (see for example Fuller (1976)), and the discreteness condition goes back to Le Cam and has been useful in a variety of problems connected with the analysis of time series data.

Proposition 5.1 *Assume that A.1 and A.2 hold. Define*

$$\hat{\boldsymbol{\theta}}^{(n)} = \tilde{\boldsymbol{\theta}}^{(n)} + \frac{1}{\sqrt{n}} c^{-1}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\tilde{\boldsymbol{\theta}}^{(n)})\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\tilde{\boldsymbol{\theta}}^{(n)}). \tag{8}$$

Then,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0) = c^{-1}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0) + o_{P_{\boldsymbol{\theta}_0}}(1), \tag{9}$$

and the asymptotic distribution of $\sqrt{n}\left(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0\right)$ *under* $H^{(n)}_g(\boldsymbol{\theta}_0)$ *is* $N\big(0, c^{-2}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)\big)$. *In addition*

$$\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\hat{\boldsymbol{\theta}}^{(n)}) = o_{P_{\boldsymbol{\theta}_0}}(1).$$

Finally, the sequence of estimators $(\hat{\boldsymbol{\theta}}^{(n)})$ *associated with the scores* $J_1 := \varphi_g \circ G^{-1}$ *and* $J_2 := G^{-1}$ *is locally and asymptotically minimax under the innovation density* g.

Proof. Since $\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}^{(n)})$ is consistent for $\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)$, we have

$$\begin{aligned}
&\sqrt{n}\left(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0\right) - c^{-1}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0) \\
&\quad = \sqrt{n}\left(\tilde{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0\right) + c^{-1}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)\left(\boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\tilde{\boldsymbol{\theta}}^{(n)}) - \boldsymbol{\Delta}^{(n)}_{J_1,J_2}(\boldsymbol{\theta}_0)\right) + o_{P_{\boldsymbol{\theta}_0}}(1),
\end{aligned}$$

which, in view of Proposition 4.1, is a $o_{P_{\boldsymbol{\theta}_0}}(1)$. The end of the proof is straightforward. $\square$

The following corollary shows that $\hat{\boldsymbol{\theta}}^{(n)}$ and $\bar{\boldsymbol{\theta}}^{(n)}$ are asymptotically equivalent.

Proposition 5.2 *Assume that A.1 and A.2 hold. Let* $(\bar{\boldsymbol{\theta}}^{(n)})$ *be a* $\sqrt{n}$*-consistent solution of (5). Then*

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}^{(n)} - \bar{\boldsymbol{\theta}}^{(n)}\right) = o_P(1),$$

under $H^{(n)}_g(\boldsymbol{\theta}_0)$.

Proof. The desired result readily follows from the fact that both $\hat{\boldsymbol{\theta}}^{(n)}$ and $\bar{\boldsymbol{\theta}}^{(n)}$ satisfy (9). $\square$

The advantage of $\hat{\boldsymbol{\theta}}^{(n)}$ is that, contrary to $\bar{\boldsymbol{\theta}}^{(n)}$, it can be computed easily and its asymptotic properties are evident.

Score functions $J_1(u)J_2(v)$	Normal	Density types Logistic	Double-exp.
van der Waerden $\Phi^{-1}(u)\Phi^{-1}(v)$	1.000	1.048	1.226
Wilcoxon $(2u-1)\ln(\frac{v}{1-v})$	0.948	1.098	1.482
Laplace $sign(2u-1)F_e^{-1}(v)$	0.612	0.812	2.000

Figure 1: Asymptotic relative efficiencies of the proposed R-estimator (for van der Waerden, Wilcoxon and Laplace scores) w.r.t. the LS-estimator under normal, logistic and double-exponential innovation densities.

6 Asymptotic relative efficiencies

We now compute the ARE of the proposed estimators with respect to the usual Gaussian estimator, namely the LS estimator. These $AREs$ are obtained as the ratios of the asymptotic variances of the considered estimators. The asymptotic variance-covariance matrix of the proposed R-estimate $\hat{\boldsymbol{\theta}}^{(n)}$ is given in Proposition 5.1, and it is well-known that the asymptotic variance-covariance matrix of the LS estimate $\hat{\boldsymbol{\theta}}_{LS}^{(n)}$ is $\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)$. Therefore, the ARE of $\hat{\boldsymbol{\theta}}^{(n)}$ w.r.t. $\hat{\boldsymbol{\theta}}_{LS}^{(n)}$ is given by

$$ARE_g(\hat{\boldsymbol{\theta}}^{(n)}/\hat{\boldsymbol{\theta}}_{LS}^{(n)}) = c^2(J_1, J_2, g).$$

Figure 1 summarizes the numerical values of the AREs of the proposed R-estimator w.r.t. the least-squares estimator for various score functions and density types.

Figure 1 illustrates the gains of efficiency achievable by using R-estimator. Note that, when using van der Waerden scores, the AREs are uniformly larger than or equal to one. This result was established by Chernoff and Savage (1958) for linear models with independent errors and extended to the ARMA case by Hallin (1994).

Proceeding as above, we can check that the ARE of the proposed R-estimator w.r.t. the LAM (locally and asymptotically minimax) estimator (i.e., the one based on the scores $J_1 := \varphi_g \circ G^{-1}$ and $J_2 := G^{-1}$ under $H_g^{(n)}(\boldsymbol{\theta}_0)$) is given by

$$ARE_g(\hat{\boldsymbol{\theta}}^{(n)}/\hat{\boldsymbol{\theta}}_{LAM,g}^{(n)}) = c^2(J_1, J_2, g)\,\sigma^{-2}\,I^{-1}(g).$$

7 Simulation results

We now present the results of a Monte-Carlo study that confirms the efficiency results that were presented in the previous section. The protocol of the study is the following. We generated

Estimators	Density types		
	Normal	Logistic	Double-exp.
LS estimate	-0.0020 694e-6	-0.0009 745e-6	-0.0018 709e-6
van der Waerden R-estimate	-0.0022 735e-6	-0.0013 701e-6	-0.0018 601e-6
Wilcoxon R-estimate	-0.0022 807e-6	-0.0011 658e-6	-0.0014 481e-6
Laplace R-estimate	-0.0030 1317e-6	-0.0012 908e-6	-0.0012 410e-6
Van der Waerden KS-estimate	-0.0058 752e-6	-0.0047 719e-6	-0.0046 571e-6
Wilcoxon KS-estimate	-0.0060 813e-6	-0.0045 680e-6	-0.0040 473e-6
Laplace KS-estimate	-0.0097 1368e-6	-0.0069 994e-6	-0.0063 541e-6

Figure 2: Simulation results.

independently $M = 300$ series of $n = 500$ observations from an AR(1) model with parameter $\theta_0 = \rho_0 = 0.8$ and standard normal, logistic and double-exponential innovation densities, respectively.

For each series, we computed the following seven estimators of ρ_0 : Koul and Saleh's estimator (for van der Waerden, Wilcoxon and Laplace scores), our one-step estimator (for van der Waerden, Wilcoxon and Laplace scores), and the LS estimator. For each estimate, we report the mean deviation $\frac{1}{M}\sum_{m=1}^{M}\hat{\rho}_m^{(n)} - \rho_0$ and the mean square error $\frac{1}{M}\sum_{m=1}^{M}(\hat{\rho}_m^{(n)} - \rho_0)^2$. Results are presented in Figure 2.

Figure 2 confirms the hierarchy of the *ARE*s in Figure 1. Note, in particular, that correctly specified scores lead to the most efficient estimators. The proposed estimators compete very well w.r.t. Koul and Saleh's estimators (see Koul and Saleh (1993)).

8 Appendix

We now prove Proposition 3.3. Along the same lines as in the proof of Proposition 5.1 of Hallin and Puri (1994), we show that under $H_g^{(n)}(\boldsymbol{\theta}_0)$,

$$n^{1/2}\left[r_k^{(n)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}) - r_k^{(n)}(\boldsymbol{\theta}_0)\right] + c(J_1, J_2, g)(a_k^{(n)} + b_k^{(n)}) = o_p(1), \tag{10}$$

as $n \to \infty$.
(this result is proved in Hallin and Puri (1994) only for the scores $J_1 := \varphi_g \circ G^{-1}$ and $J_2 := G^{-1}$).

Using the asymptotic representation in Proposition 3.1, one could obtain

$$\tilde{r}_k^{(n)}(\boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^{(n)}) - \tilde{r}_k^{(n)}(\boldsymbol{\theta}_0) + Ic(J_1, J_2, g)(a_k^{(n)} + b_k^{(n)}) = o_{P_{\boldsymbol{\theta}_0}}(1), \quad \text{as} \quad n \to \infty, \tag{11}$$

where

$$\tilde{r}_k^{(n)}(.) = n^{-1/2} \sum_{t=k+1}^{n} J_1 \circ G(Z_t^{(n)}(.)) J_2 \circ G(Z_{t-k}^{(n)}(.)).$$

Let $\varepsilon > 0$. As in Kreiss (1987,b), we decompose the set $B = \{\boldsymbol{\tau} \in \mathbb{R}^{p+q} : \|\boldsymbol{\tau}\| \leq c\}$ into the hypercubes whose vertices are at the points $\varepsilon(j_1, ..., j_{p+q}), j_i = 0, \pm 1, ..., \pm N(\varepsilon), \quad i = 1, ..., p+q$. For each $\boldsymbol{\tau} \in B$, let $\boldsymbol{\tau}^*$ denote the vertex nearest to zero of the sub-cube containing $\boldsymbol{\tau}$ and we adopt the notations $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}$ and $\boldsymbol{\theta}^* = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{\tau}^*$.

From (11) we have

$$\sup_{\|\boldsymbol{\tau}\| \leq c} \left| \tilde{r}_k^{(n)}(\boldsymbol{\theta}^*) - \tilde{r}_k^{(n)}(\boldsymbol{\theta}_0) + Ic(J_1, J_2, g)(a_k^{(n)*} + b_k^{(n)*}) \right| = o_{P_{\boldsymbol{\theta}_0}}(1), \quad \text{as} \quad n \to \infty, \tag{12}$$

where $a_k^{(n)*}$ and $b_k^{(n)*}$ are associated with $\boldsymbol{\tau}^*$. Considering an arbitrary hypercube B^* in the decomposition, we have

$$\sup_{\boldsymbol{\tau} \in B^*} \left| \tilde{r}_k^{(n)}(\boldsymbol{\theta}^*) - \tilde{r}_k^{(n)}(\boldsymbol{\theta}) + Ic(J_1, J_2, g)(a_k^{(n)*} + b_k^{(n)*} - a_k^{(n)} - b_k^{(n)}) \right| \leq U^{(n)} + V^{(n)}, \tag{13}$$

where

$$V^{(n)} = \sup_{\boldsymbol{\tau} \in B^*} \left| Ic(J_1, J_2, g)(a_k^{(n)*} + b_k^{(n)*} - a_k^{(n)} - b_k^{(n)}) \right| \leq I|c(J_1, J_2, g)|(p+q)\varepsilon, \tag{14}$$

$$\begin{aligned} U^{(n)} &= \sup_{\boldsymbol{\tau} \in B^*} \left| \tilde{r}_k^{(n)}(\boldsymbol{\theta}^*) - \tilde{r}_k^{(n)}(\boldsymbol{\theta}) \right| \\ &= \sup_{\boldsymbol{\tau} \in B^*} \left| n^{-1/2} \sum_{t=k+1}^{n} J_1 \circ G(Z_t(\boldsymbol{\theta}^*)) J_2 \circ G(Z_{t-k}(\boldsymbol{\theta}^*)) - J_1 \circ G(Z_t(\boldsymbol{\theta})) J_2 \circ G(Z_{t-k}(\boldsymbol{\theta})) \right| \\ &\leq U_{k,1}^{(n)} + U_{k,2}^{(n)}, \end{aligned}$$

where

$$U_{k,1}^{(n)} = \sup_{\boldsymbol{\tau} \in B^*} \left| n^{-1/2} \sum_{t=k+1}^{n} \left[J_1 \circ G(Z_t(\boldsymbol{\theta}^*)) - J_1 \circ G(Z_t(\boldsymbol{\theta})) \right] J_2 \circ G(Z_{t-k}(\boldsymbol{\theta}^*)) \right|,$$

and

$$U_{k,2}^{(n)} = \sup_{\boldsymbol{\tau} \in B^*} \left| n^{-1/2} \sum_{t=k+1}^{n} \left[J_2 \circ G(Z_{t-k}(\boldsymbol{\theta}^*)) - J_2 \circ G(Z_{t-k}(\boldsymbol{\theta})) \right] J_1 \circ G(Z_t(\boldsymbol{\theta})) \right|.$$

Kreiss (1987b, equation (2.3)) shows that

$$Z_t(\boldsymbol{\theta}^*) - Z_t(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \boldsymbol{z}^{(\boldsymbol{\theta})}(t-1, \boldsymbol{\theta}^*),$$

where $z^{(\boldsymbol{\theta})}(t-1,\boldsymbol{\theta}^*) := \sum_{s=1}^{t} h^*_{s-1}\,(X_{t-s},...,,X_{t-s+1-p},Z_{t-s}(\boldsymbol{\theta}),...,Z_{t-s+1-q}(\boldsymbol{\theta}))'$ (here the (h^*_i)'s denote the Green's functions associated with the operator $B^*[L] := 1+\sum_{i=1}^{q} B^*_i L^i$). From this identity and assumption A.2(ii), $U^{(n)}_{k,1}$ is bounded by

$$n^{-1}H_1(p+q)\varepsilon \sup_{\boldsymbol{\tau}\in B^*} \sum_{t=k+1}^{n} \|\boldsymbol{z}^{(\boldsymbol{\theta})}(t-1,\boldsymbol{\theta}^*)\| \;\left| J_2\circ G(Z_{t-k}(\boldsymbol{\theta}^*))\right|, \tag{15}$$

where H_1 is the Lipshitz constant for $J_1\circ G$. Now consider the maximum over all hypercubes of B. We have

$$\begin{aligned} E_{\boldsymbol{\theta}_0} \max\nolimits_{B^*} \sup_{\boldsymbol{\tau}\in B^*} \left\|\boldsymbol{z}^{(\boldsymbol{\theta})}(t-1,\boldsymbol{\theta}^*)\right\|^2 \quad \leq & \quad \sum_{i=0}^{p-1} E_{\boldsymbol{\theta}_0} \max_{B^*} \sup_{\boldsymbol{\tau}\in B^*} \left[\sum_{s=1}^{t} h^*_{s-1}X_{t-s-i}\right]^2 \\ & + \sum_{i=0}^{q-1} E_{\boldsymbol{\theta}_0} \max_{B^*} \sup_{\boldsymbol{\tau}\in B^*} \left[\sum_{s=1}^{t} h^*_{s-1}Z_{t-s-i}(\boldsymbol{\theta})\right]^2. \end{aligned}$$

The first term can be bounded by $\sum_{s=1}^{t}[\max\sup h^{*2}_{s-1}]E_{\boldsymbol{\theta}_0}[X_1]^2$, which is $O(1)$ uniformly in t, because of Lemma 5.1 of Kreiss (1987,b). The second term can be bounded by

$$O(1)+E_{\boldsymbol{\theta}_0}\max\sup[\sum_{s=1}^{t} h^*_{s-1}(\boldsymbol{\theta}_0-\boldsymbol{\theta})\boldsymbol{z}^{(\boldsymbol{\theta}_0)}(t-s-k-1,\boldsymbol{\theta})]^2, \quad \text{which is} \quad O(1)+O(\frac{\varepsilon^2}{n}).$$

It follows that $E_{\boldsymbol{\theta}_0}\max_{B^*}\sup_{\boldsymbol{\tau}\in B^*}\|\boldsymbol{z}^{(\boldsymbol{\theta})}(t-1,\boldsymbol{\theta}^*)\|^2 = O(1)$ uniformly in t. Similarly, using the inequality $[J_2\circ G(Z_{t-k}(\boldsymbol{\theta}^*))]^2 \leq 2[J_2\circ G(Z_{t-k}(\boldsymbol{\theta}^*)) - J_2\circ G(Z_{t-k}(\boldsymbol{\theta}_0))]^2 + 2[J_2\circ G(Z_{t-k}(\boldsymbol{\theta}_0))]^2$, we obtain that $E_{\boldsymbol{\theta}_0}\max_{B^*}\sup_{\boldsymbol{\tau}\in B^*}[J_2\circ G(Z_{t-k}(\boldsymbol{\theta}^*))]^2 = O(1)$, uniformly in t.

The Cauchy-Schwarz inequality applied to (15) then yields

$$E_{\boldsymbol{\theta}_0}\max_{B^*} U^{(n)}_{k,1} = \varepsilon\, O(1). \tag{16}$$

Analogous to (16), one obtains

$$E_{\boldsymbol{\theta}_0}\max_{B^*} U^{(n)}_{k,2} = \varepsilon\, O(1). \tag{17}$$

Piecing together (13), (14), (16), and (17) we thus have, as $n\to\infty$,

$$\max_{B^*}\sup_{\boldsymbol{\tau}\in B^*}\left|\tilde{r}^{(n)}_k(\boldsymbol{\theta}^*) - \tilde{r}^{(n)}_k(\boldsymbol{\theta}) + Ic(J_1,J_2,g)(a^{(n)*}_k + b^{(n)*}_k - a^{(n)}_k - b^{(n)}_k)\right| = o_{P_{\boldsymbol{\theta}_0}}(1). \tag{18}$$

The result straightforwardly follows from (12) and (18) by letting $\varepsilon\to 0$. $\square$

References

Adichie, J.N. (1967). Estimation of regression parameters based on rank tests, *Ann. Math. Statist.* **38**, 894-904.

Allal, J., (1991). R-estimation dans le modèle autorégressif d'ordre un, *Thèse de Doctorat. Université libre de Bruxelles.*

Chernoff, H. and Savage, I.R., (1958). Asymptotic normality and efficiency of certain nonparametric tests. *Ann. Math. Statist.* **29**, 972-994.

Fuller, W.A., (1976). Introduction to statistical time series, *J. Wiley, New York.*

Hájek, J. and Šidák, Z., (1967). Theory of rank tests, *Academic press, New York.*

Hallin, M., (1994). On the Pitman-nonadmissibility of correlogram-based methods. *Journal of Time Series Analysis* **15**, 607-612.

Hallin M., Ingenbleek, J.F. and Puri M.L., (1985). Linear serial rank tests for randomness against ARMA alternatives, *Ann. Statist.* **13**, 1156-1181.

Hallin, M. and Puri, M.L., (1994). Aligned rank tests for linear models with autocorrelated error terms, *J. Multivariate Anal.* **50**, 175-237.

Hodges, J.L. and Lehmann, E.L., (1963). Estimates of location based on rank tests, *Ann. Math. Statist.* **34**, 589-611.

Jaeckel, L.A., (1972). Estimating regression cœfficients by minimizing the dispersion of the residuals, *Ann. Math. Statist.* **43**, 1449-1458.

Jurečková, J. (1971). Nonparametric estimate of regression cœfficients, *Ann. Math. Statist.* **42**, 1328-1338.

Jurečková, J. and Sen, P.K. (1996), Robust statistical procedures : asymptotics and interrelations.

Koul, H.L., (1971). Asymptotic behavior of a class of confidence regions based on ranks in regression, *Ann. Math. Statist.* **42**, 466-476.

Koul, H.L. and Saleh, A.K.Md.E., (1993). R-estimation of the parameters of autoregressive AR(p) models, *Ann. Statist.* **21**, 534-551.

Kreiss, J.P., (1987-a). On adaptive estimation in stationary ARMA processes, *Ann. Statist.* **15**, 112-133.

Kreiss, J.P., (1987-b). A note on M-estimation in stationary ARMA processes, *Statistics and Decisions* **3**, 317-336.

Le Cam, L., (1960). Locally asymptotically normal families of distributions. *University of California. Publications in Statistics,* **3**, 37-98.

Puri, M.L. and Sen, P.R. (1985), Nonparametric methods in general linear models.

Yoshihara, K. I. (1976). Limiting behavior of U-statistics for stationary absolutely regular process, *Z. Wahrsch. vew. Gebiete*, **35**, 237-252.

J. Allal and A. Kaaouachi,
Département de Mathématiques,
Université Mohamed Premier,
Oujda 60000. Maroc
Fax: 212 6 50 06 03
allal@sciences.univ-oujda.ac.ma
kaaouachi@sciences.univ-oujda.ac.ma

D. Paindaveine,
Université Libre de Bruxelles,
Département de Mathématique et ISRO
Campus de la Plaine, CP 210,
1050 Bruxelles, Belgium
dpaindav@ulb.ac.be