JIRSS (2003)
Vol. 2, No. 2, pp 207-222

Estimation in ARMA models based on signed ranks

J. Allal[1], **A. Kaaouachi**[2]

[1]Département de Mathématiques, Faculté des Sciences, Université Mohamed Ier, Oujda 60 000 Morocco. (allal@sciences.univ-oujda.ac.ma)
[2]Département d'Informatique, EST-ENSA, Université Mohamed Ier, Oujda 60 000 Morocco. (kaaouachi@sciences.univ-oujda.ac.ma)

Abstract. In this paper we develop an asymptotic theory for estimation based on signed ranks in the ARMA model when the innovation density is symmetrical. We provide two classes of estimators and we establish their asymptotic normality with the help of the asymptotic properties for serial signed rank statistics. Finally, we compare our procedure to the one of least-squares, and we illustrate the performance of the proposed estimators via a Monte Carlo study.

1 Introduction

Consider the ARMA(p, q) model

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

Received: November 2002, Revised: September 2003
Key words and phrases: ARMA models, asymptotic relative efficiency, asymptotic uniform linearity, local asymptotic normality, rank and signed-rank procedures, R-estimation, serial signed rank statistics.

with $(\varepsilon_t; t \in \mathbb{Z})$ is a white noise with unspecified symmetric (with respect to some known median, which can be set equal to zero without any loss of generality) density g and cumulative distribution function G, and $\theta_0 = (A_1, \ldots, A_p, B_1, \ldots, B_q) \in \mathbb{R}^{p+q}$ is the vector of unknown parameters which ensures both the stationarity and the invertibility conditions.

The commonly used methods of estimating θ_0 include the method of moments, the least-squares and the maximum likelihood methods. These estimators are not very efficient in some situations, that is why many researchers are focused to other procedures which are either more robust, or more efficient. The important special classes are : M-estimation, L-estimation and R-estimation.

The R-estimation class represents the estimation methods based on ranks or signed ranks. These methods inherit all the desirable features resulting from their *invariance* properties. They are robust, and even may outperform the more classical methods based on Gaussian quasi-likelihoods. Historically, the R-estimation class was the subject of a vast literature in linear regression models. Major contributions include Hodges and Lehmann (1963), Adichie (1967), Jurečková (1971), Koul (1971), Jaeckel (1972), and Heiler and Willers (1988). The monographs by Puri and Sen (1985) and Jurečková and Sen (1996) give details and extensive bibliography. However, there are a few attempts extending the above works to time series context. Koul and Saleh (1993) developed the theory for R-estimation in AR models by minimizing dispersions based on non-decreasing right continuous score functions of ranks, using the "weak convergence technique". Such a technique involves both residual ranks and the observations themselves, and works only for bounded score functions. More recently, Allal, Kaaouachi and Paindaveine (2001) proposed a new class of R-estimators for an ARMA model, based on an appropriate function which is entirely measurable with respect to the residual ranks.

In this paper, we develop the theory of R-estimation in ARMA models when the innovation density is symmetric. In this situation, it is well known that the vector *(rank, sign)* is maximal invariant with respect to the group of continuous, even order-preserving transformations. An important tool in our approach is the LAN property, with the signed-rank based version established by Hallin and Puri (1994), and the proposed estimators are measurable with respect to the maximal invariant. The treatments put forward in Allal, Kaaouachi and Paindaveine (2001) are still valid, with some adaptations. They represent a general methodology allowing for the construction of efficient estimates in time series context under LAN property. Relevant papers in this direction are Kreiss (1987), Jeganathan (1995), Drost et al. (1997), and Koul and Schich (1997), to quote only a few.

The paper is organized as follows. Definitions and main technical assumptions are stated in Section 2. Section 3 presents the LAN result with a signed rank version and asymptotic properties of serial signed rank statistics. Section 4 gives the first class of estimators based on signed ranks. Section 5 gives the second class of estimators based on signed ranks. Section 6 provides the asymptotic relative efficiencies of estimators with respect to the least-squares estimate. In Section 7, we illustrate the performance of the proposed estimators via a Monte Carlo study.

2 Definitions and main technical assumptions

The ARMA(p,q) model (1) can be written as

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

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

Let $X^{(n)} = (X_1^{(n)}, \ldots, X_n^{(n)})$ be an observed series of length n, and denote by $H_g^{(n)}(\theta_0)$ the hypothesis under which $X^{(n)}$ is generated by model (1). Denote by $R_{+,t}^{(n)}(\theta_0)$ the rank of the residual $|Z_t^{(n)}(\theta_0)|$ among the absolute values $|Z_1^{(n)}(\theta_0)|, \ldots, |Z_n^{(n)}(\theta_0)|$, where

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

and let $s_t^{(n)}(\theta_0)$ be the sign of $Z_t^{(n)}(\theta_0)$.

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

Consider the serial signed rank statistic of order k ($k = 1, \ldots, n-1$), known as a signed rank autocorrelation of order k and is fruitfully used to provide efficient tests (see Hallin and Puri (1991, 1994)),

$$\begin{aligned} r_k^{(n)+}(\theta_0) &:= \left[(n-k)\sigma_+^{(n)}\right]^{-1} \sum_{t=k+1}^{n} s_t^{(n)}(\theta_0) s_{t-k}^{(n)}(\theta_0) \\ & J_1\left(\frac{1}{2} + \frac{R_{+,t}^{(n)}(\theta_0)}{2(n+1)}\right) J_2\left(\frac{1}{2} + \frac{R_{+,t-k}^{(n)}(\theta_0)}{2(n+1)}\right), \quad (2) \end{aligned}$$

where J_1 and J_2 are two score functions defined on $(0,1)$, and $\sigma_+^{(n)}$ is a normalizing constant such that $(n-k)^{1/2} r_k^{(n)+}(\theta_0)$ is exactly standardized (mean zero and variance one) under $H_g^{(n)}(\theta_0)$. The explicit form of $\sigma_+^{(n)}$ is given by (3.15) of Hallin and Puri (1994), with minor modifications.

Particularizing the score functions, we obtain the van der Waerden, Wilcoxon and Laplace signed rank autocorrelations (see Hallin and Puri (1991)).

Define now the vector of signed rank statistics

$$\begin{aligned} \sqrt{n}\, \boldsymbol{T}_{J_1;J_2}^{(n)+}(\theta_0) &:= \left(\sum_{k=1}^{n-1} (n-k)^{1/2} \psi_k^{(1)}(\theta_0) r_k^{(n)+}(\theta_0), \ldots, \right. \\ & \left. \sum_{k=1}^{n-1} (n-k)^{1/2} \psi_k^{(p+q)}(\theta_0) r_k^{(n)+}(\theta_0)\right)', \quad (3) \end{aligned}$$

where $\left\{\psi_t^{(1)}(\theta_0), \ldots, \psi_t^{(p+q)}(\theta_0); t \in \mathbb{Z}\right\}$ is an arbitrary fundamental system of solutions of the homogeneous equation

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

(Convenient choices of this system are given in Section 4 of Hallin and Puri (1994)).

Associated with this fundamental system, denote by $\boldsymbol{C}_\psi(\theta_0)$ and $\boldsymbol{W}_\psi^2(\theta_0)$ the matrices whose elements are $\psi_i^{(j)}(\theta_0)$ and $\sum_{t=1}^{\infty} \psi_t^{(i)}(\theta_0)\psi_t^{(j)}(\theta_0)$

$(i, j = 1, \ldots, p+q)$, respectively. Finally, let

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

where g_i and h_i are the Green's functions associated with the operators $A(L)$ and $B(L)$, respectively. They are characterized by

$$A^{-1}(L) = \sum_{i=0}^{\infty} g_i L^i \quad \text{and} \quad B^{-1}(L) = \sum_{i=0}^{\infty} h_i L^i.$$

Next putting

$$\boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta_0) := n^{1/2} \boldsymbol{M}'(\theta_0) \boldsymbol{C}'^{-1}_{\psi}(\theta_0) \boldsymbol{T}^{(n)+}_{J_1;J_2}(\theta_0).$$

The latter function will play a fundamental role in our treatment.

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

Assumptions A.1. (borrowed from Hallin and Puri (1994))

(*i*) $g(x) > 0 \ \ \forall x \in \mathbb{R}$, $\displaystyle\int x g(x)\, dx = 0$ and $0 < \sigma^2 := \displaystyle\int x^2 g(x)\, dx < \infty$;

(*ii*) g is symmetric, absolutely continuous and strongly unimodal;

(*iii*) letting $\varphi_g := -g'/g$, where g' stands for the first derivative of g, the Fisher information $I(g) := \int [\varphi_g(x)]^2 g(x)\, dx$ is finite;

(*iv*) the vector $(\varepsilon_{-q+1}; \ldots; \varepsilon_0; X_{-p+1}; \ldots; X_0)$ possesses a nowhere vanishing joint density $g^0(., \theta)$ that satisfies $g^0(., \theta^{(n)}) - g^0(., \theta_0) = o_P(1)$, under $H^{(n)}_g(\theta_0)$, as $\theta^{(n)} \to \theta_0$.

Assumptions A.2. (borrowed from Allal, Kaaouachi and Paindaveine (2001))

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

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

Remark 2.1

Assumptions A.1 are used in proving the LAN result. 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 $J_i(u) = \ln(u/1-u)$ and G is logistic. It easily can be weakened into a piecewise Lipschitz assumption, which also accomodates such distributions as the double exponential.

3 LAN result and asymptotic properties of serial signed rank statistics

This section of the paper is devoted to the LAN result with the signed-rank based version for the ARMA model, and to the asymptotic properties of serial signed rank statistics. For this purpose, consider the sequence of local alternatives $H_g^{(n)}\left(\theta_0 + n^{-1/2}\boldsymbol{\tau}\right)$, where $\tau := (\boldsymbol{\gamma}, \boldsymbol{\delta}) \in \mathbb{R}^p \times \mathbb{R}^q$. Let $\Lambda = \Lambda^{(n)}_{\theta_0+n^{-1/2}\boldsymbol{\tau}/\theta_0}$ be the log-likelihood ratio for $H_g^{(n)}(\theta_0 + n^{-1/2}\boldsymbol{\tau})$ with respect to $H_g^{(n)}(\theta_0)$, then we have

Proposition 3.1. (Local asymptotic normality)
Assume that Assumptions A.1 hold. Then, under $H_g^{(n)}(\theta_0)$,

$$\Lambda = \boldsymbol{\tau}'\sigma[I(g)]^{1/2}\boldsymbol{\Delta}^{(n)+}_{\varphi_g\circ G^{-1};G^{-1}}(\theta_0) - \frac{1}{2}\sigma^2 I(g)\boldsymbol{\tau}'\boldsymbol{\Gamma}(\theta_0)\boldsymbol{\tau} + o_P(1),$$

as $n \to \infty$, where

$$\boldsymbol{\Gamma}(\theta_0) := \boldsymbol{M}'(\theta_0)\boldsymbol{C}'^{-1}_{\psi}(\theta_0)\boldsymbol{W}^2_{\psi}(\theta_0)\boldsymbol{C}^{-1}_{\psi}(\theta_0)\boldsymbol{M}(\theta_0).$$

Moreover, the limiting distribution of $\boldsymbol{\Delta}^{(n)+}_{\varphi_g\circ G^{-1};G^{-1}}(\theta_0)$ under $H_g^{(n)}(\theta_0)$ is $N(0, \boldsymbol{\Gamma}(\theta_0))$.

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

By applying LeCam's third Lemma (we adopt here Hájek and Šidák (1967)'s terminology), we obtain the following corollary.

Corollary 3.1. *Assume that Assumptions A.1 hold. Then, the limiting distribution of $\boldsymbol{\Delta}^{(n)+}_{\varphi_g\circ G^{-1};G^{-1}}(\theta_0) - \sigma[I(g)]^{1/2}\boldsymbol{\Gamma}(\theta_0)\boldsymbol{\tau}$ under $H_g^{(n)}(\theta_0 + n^{-1/2}\boldsymbol{\tau})$ is $N\left(0, \boldsymbol{\Gamma}(\theta_0)\right)$.*

The following proposition gives some asymptotic properties of serial signed rank statistics.

Proposition 3.2. (Asymptotic properties of serial signed rank statistics)
Assume that A.1 and A.2 hold. Then,
(i) $n^{1/2} r_k^{(n)+}(\theta_0) = I^{-1}\, n^{-1/2} \sum_{t=k+1}^{n} J_1 \circ G(Z_t^{(n)}(\theta_0)) J_2 \circ G(Z_{t-k}^{(n)}(\theta_0)) + o_P(1)$, *under* $H_g^{(n)}(\theta_0)$, *as* $n \to \infty$, *where*
$I^2 := \int_0^1 [J_1(u)]^2 \, du \int_0^1 [J_2(u)]^2 \, du;$
(ii) the signed rank statistic $n^{1/2} r_k^{(n)+}(\theta_0)$ *is asymptotically normal under* $H_g^{(n)}(\theta_0)$, *with mean zero and variance one ;*
(iii) the signed rank statistic $n^{1/2} r_k^{(n)+}(\theta_0)$ *is asymptotically normal under* $H_g^{(n)}(\theta_0 + n^{-1/2}\boldsymbol{\tau})$, *with mean* $c(J_1, J_2, g)(\boldsymbol{a}_k + \boldsymbol{b}_k)$ *and variance one, where*

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

and

$$a_k := \sum_{j=1}^{p} \gamma_j g_{k-j} \quad \text{and} \quad b_k := \sum_{j=1}^{q} \delta_j h_{k-j};$$

(iv) for all $k = 1, \ldots, n-1$ *and all* $c > 0$,

$$\sup_{\|\boldsymbol{\tau}\| \le c} \Big| n^{1/2} \left[r_k^{(n)+}(\theta_0 + n^{-1/2}\boldsymbol{\tau}) - r_k^{(n)+}(\theta_0) \right] + c(J_1, J_2, g)(a_k + b_k) \Big| = o_P(1), \tag{4}$$

under $H_g^{(n)}(\theta_0)$, *as* $n \to \infty$.
Here, $\|.\|$ *stands for the Euclidean norm in* $\mathbb{R}^{p+q}$.

Proof. (i) direct application of Proposition 3.1 (i) of Hallin and Puri (1991) ; called the asymptotic representation result for serial signed rank statistics ;
(ii) use (i) and the central limit theorem for U-statistic under k-dependent random variables (See Yoshihara (1976)'s Theorem 1) ;
(iii) use LeCam's third Lemma ;
(iv) use (i) and Proposition 3.3 of Allal, Kaaouachi and Paindaveine (2001). □

4 A first class of estimators based on signed ranks

In this section we introduce a first class of estimators based on signed ranks for the parameter θ_0 of the ARMA model.

From Proposition 3.2 (ii), we can show that the vector $\boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta_0)$ is asymptotically normal under $H^{(n)}_g(\theta_0)$, with mean zero and covariance matrix $\Gamma(\theta_0)$. This suggests estimating the unknown parameter θ_0 by the value of θ for which $\boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta)$ is as near to zero as possible, i.e. to estimate θ_0 by

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

The following proposition gives the asymptotic uniform linearity of $\boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta_0 + n^{-1/2}\tau)$ in $\|\tau\| \leq c$, for $c > 0$.

Proposition 4.1. *Assume that A.1 and A.2 hold. Then, for all $c > 0$,*

$$\sup_{\|\boldsymbol{\tau}\|\leq c} \Big\| \boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta_0 + n^{-1/2}\boldsymbol{\tau}) - \boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta_0) + c(J_1, J_2, g)\boldsymbol{\Gamma}(\theta_0)\boldsymbol{\tau} \Big\| = o_P(1), \tag{6}$$

under $H^{(n)}_g(\theta_0)$, as $n \to \infty$.

Proof. The claim readily follows from Proposition 3.2 (iv), using standard arguments. □

Now we are ready to give the limit law of signed rank estimators.

Proposition 4.2. *Assume that A.1 and A.2 hold. Let $(\bar{\theta}^{(n)})$ be a $\sqrt{n}$-consistent solution of (5), i.e. satisfying*

$$\sqrt{n}\left(\bar{\theta}^{(n)} - \theta_0\right) = O_P(1), \quad \text{under} \quad H^{(n)}_g(\theta_0), \quad \text{as} \quad n \to \infty.$$

Then, the asymptotic distribution of $\sqrt{n}\left(\bar{\theta}^{(n)} - \theta_0\right)$ under $H^{(n)}_g(\theta_0)$ is $N\big(0, c^{-2}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\theta_0)\big)$.

Proof. The result follows from the asymptotic normality of $\boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\theta_0)$ under $H^{(n)}_g(\theta_0)$ together with the Cramér-Wold device and Proposition 4.1. □

Proposition 4.3. *Assume that A.1 and A.2 hold, and suppose that* $\boldsymbol{\lambda}'\boldsymbol{\Delta}^{(n)}_{J_1;J_2}(\theta_0 + b\,n^{-1/2}\boldsymbol{\lambda})$ *is monotone in b for every* $\|\boldsymbol{\lambda}\| = 1$*. Then, for any solution* $(\bar{\theta}^{(n)})$ *of (5), the asymptotic distribution of* $\sqrt{n}\left(\bar{\theta}^{(n)} - \theta_0\right)$ *under* $H^{(n)}_g(\theta_0)$ *is* $N\big(0, c^{-2}(J_1, J_2, g)\boldsymbol{\Gamma}^{-1}(\theta_0)\big)$.

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

Note that the above two propositions give the asymptotic normality of any solution of (5) with some additional assumptions, and the proposed estimators have an implicit form via the minimization problem. To resolve these difficulties, we propose in the following section a new class of signed rank estimators possessing an explicit form and asymptotic normality without any assumption.

5 A second class of estimators based on signed ranks

In this section we adopt the Hájek-LeCam approach to construct a second class of estimators $\hat{\theta}^{(n)}$based on signed ranks and satisfying

$$\boldsymbol{\Delta}^{(n)+}_{J_1;J_2}(\hat{\theta}^{(n)}) = o_P(1), \quad \text{under} \quad H^{(n)}_g(\theta_0), \quad \text{as} \quad n \to \infty.$$

For this purpose, let $(\tilde{\theta}^{(n)})$ denote a discretized $\sqrt{n}$-consistent preliminary estimate of θ_0, i.e. a sequence of estimates such that
(i) $\sqrt{n}$**-consistency:** $\sqrt{n}\left(\tilde{\theta}^{(n)} - \theta_0\right) = O_P(1)$, under $H^{(n)}_g(\theta_0)$, as $n \to \infty$;
(ii) **Local discreteness:** the number of possible values of $\tilde{\theta}^{(n)}$ in balls of the form

$$\left\{\theta \in \mathbb{R}^{p+q} : \sqrt{n}\,\|\theta - \theta_0\| \leq c\right\}, \quad c > 0 \;\text{ fixed},$$

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

The $\sqrt{n}$-consistency condition is satisfied by all estimates usually considered in the context of ARMA models (see for example Fuller (1976)), and the local discreteness condition goes back to LeCam (1960) and has become an important technical tool in the construction of efficient estimators in semiparametric models ; see Bickel, Klaassen, Ritov and Wellner (1993) and references therein.

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

$$\hat{\theta}^{(n)} = \tilde{\theta}^{(n)} + \frac{1}{\sqrt{n}} c^{-1}(J_1, J_2, g)\mathbf{\Gamma}^{-1}(\tilde{\theta}^{(n)})\mathbf{\Delta}^{(n)+}_{J_1;J_2}(\tilde{\theta}^{(n)}). \qquad (7)$$

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

$$\sqrt{n}(\hat{\theta}^{(n)} - \theta_0) = c^{-1}(J_1, J_2, g)\mathbf{\Gamma}^{-1}(\theta_0)\mathbf{\Delta}^{(n)+}_{J_1;J_2}(\theta_0) + o_P(1), \qquad (8)$$

and the asymptotic distribution of $\sqrt{n}\left(\hat{\theta}^{(n)} - \theta_0\right)$ under $H_g^{(n)}(\theta_0)$ is $N\big(0, c^{-2}(J_1, J_2, g)\mathbf{\Gamma}^{-1}(\theta_0)\big)$.
In addition

$$\mathbf{\Delta}^{(n)+}_{J_1;J_2}(\hat{\theta}^{(n)}) = o_P(1).$$

Proof. Note that by the consistency of $\mathbf{\Gamma}(\tilde{\theta}^{(n)})$ for $\mathbf{\Gamma}(\theta_0)$, we have

$$\sqrt{n}\left(\hat{\theta}^{(n)} - \theta_0\right) - c^{-1}(J_1, J_2, g)\mathbf{\Gamma}^{-1}(\theta_0)\mathbf{\Delta}^{(n)+}_{J_1;J_2}(\theta_0) =$$
$$\sqrt{n}\left(\tilde{\theta}^{(n)} - \theta_0\right) + c^{-1}(J_1, J_2, g)\mathbf{\Gamma}^{-1}(\theta_0)\left(\mathbf{\Delta}^{(n)+}_{J_1;J_2}(\tilde{\theta}^{(n)}) - \mathbf{\Delta}^{(n)+}_{J_1;J_2}(\theta_0)\right)$$
$$+o_P(1),$$

which, in view of Proposition 4.1, is an $o_P(1)$, under $H_g^{(n)}(\theta_0)$, as $n \to \infty$. The rest is obvious. □

Remark 5.1

It follows from the Hájek-LeCam theory for LAN experiments that an estimator $\hat{\theta}^{(n)}$ from (7) is asymptotically optimal in the LAM sense (local asymptotic minimaxity, concept introduced by Hájek (1972)) if the two score functions J_1 and J_2 are respectively $\varphi_g \circ G^{-1}$ and G^{-1}. In this case, a LAM estimator $\hat{\theta}^{(n)}_{LAM}$ takes the form

$$\hat{\theta}^{(n)}_{LAM} = \tilde{\theta}^{(n)} + \frac{1}{\sqrt{n}}\sigma^{-1} I^{-1/2}(g)\Gamma^{-1}(\tilde{\theta}^{(n)})\Delta^{(n)+}_{\varphi_g \circ G^{-1};G^{-1}}(\tilde{\theta}^{(n)}).$$

Consequently, the limit law of $\sqrt{n}\left(\hat{\theta}^{(n)}_{LAM} - \theta_0\right)$ under $H_g^{(n)}(\theta_0)$ is $N\big(0, \sigma^{-2} I^{-1}(g)\mathbf{\Gamma}^{-1}(\theta_0)\big)$.

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

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

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

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

Proof. The desired result readily follows from the fact that both $\hat{\theta}^{(n)}$ and $\bar{\theta}^{(n)}$ satisfy (8). □

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

6 Asymptotic relative efficiencies

Pitman asymptotic relative efficiencies of the proposed estimators with respect to least-squares estimate are obtained as ratios of the asymptotic variance matrices. The asymptotic variance matrix of the proposed signed rank estimator $\hat{\theta}^{(n)}$ is given in Proposition 5.1, and it is well-known that the asymptotic variance matrix of the least-squares estimate $\hat{\theta}^{(n)}_{LS}$ is $\mathbf{\Gamma}^{-1}(\theta_0)$ (see Whittle (1962)). Therefore, the ARE of $\hat{\theta}^{(n)}$ with respect to $\hat{\theta}^{(n)}_{LS}$ is given by

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

The following table summarizes the numerical values of ARE of the proposed signed rank estimators with respect to the least-squares estimate for various score functions (van der Waerden, Wilcoxon and Laplace) under some usual densities (normal, logistic and double-exponential).

Score functions $J_1(u)J_2(v)$	Normal	**Density types** Logistic	Double-exponential
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 $sgn(2u-1)F_e^{-1}(v)$	0.612	0.812	2.000

Table 1: ARE of $\hat{\theta}^{(n)}$ with respect to $\hat{\theta}^{(n)}_{LS}$.

An inspection of Table 1 reveals the excellent asymptotic performances of the proposed signed rank estimator: the van der Waerden estimator performs uniformly and strictly better that the corresponding least-squares except of course under normal density, when they perform equally well.

7 Simulation results

In this section we report the result of a small simulation study performed using S-plus. We considered the following AR(1) model

$$X_t - \theta_0 X_{t-1} = \varepsilon_t, \quad t \in \mathbb{Z},$$

where $|\theta_0| < 1$ and ε_t has an unspecified density g symmetric with respect to zero.

The parameter θ_0 was taken to be 0.4 and the three innovation density (normal, logistic and double-exponential) have been investigated. In each case, we generated independently $M = 1000$ replications for sample sizes $n = 40$ (small) and $n = 200$ (large), and constructed seven estimators for θ_0 : the least-squares estimate, rank estimator described in Allal, Kaaouachi and Paindaveine (2001) (for van der Waerden, Wilcoxon and Laplace scores) and our signed rank estimator (for van der Waerden, Wilcoxon and Laplace scores). For each estimate, we report the sample mean $\frac{1}{M}\sum_{i=1}^{M} \hat{\theta}_i^{(n)} - \theta_0$ and the mean square error $\frac{1}{M}\sum_{i=1}^{M}(\hat{\theta}_i^{(n)} - \theta_0)^2$. Results are presented in Tables 2 and 3. The simulations show that the signed rank estimators perform slightly better than the least-squares estimate, even when the latter is used as the initial estimate, except the case when the innovation density is normal.

On the other hand, observe in Table 3 that both the rank estimator and the signed rank estimator give approximately the same results without detecting the advantage of one of the two types. But, Table 2 indicates the performance of signed rank estimators relatively to rank estimators. This latter numerical result is also established in Hallin et al. (1990) who showed that the signed-rank procedures perform better than the unsigned ones, for short series lengths.

Estimators	**Density types**		
	Normal	Logistic	Double-exponential
least-squares estimate	0.035 171e-3	0.023 452e-3	0.021 283e-3
van der Waerden sigend rank estimator	0.041 246e-3	0.029 458e-3	0.035 275e-3
Wilcoxon sigend rank estimator	0.045 258e-3	0.028 421e-3	0.029 257e-3
Laplace sigend rank estimator	0.051 337e-3	0.032 467e-3	0.028 245e-3
Van der Waerden rank estimator	0.043 248e-3	0.031 469e-3	0.031 278e-3
Wilcoxon rank estimator	0.046 259e-3	0.027 431e-3	0.030 261e-3
Laplace rank estimator	0.053 349e-3	0.037 469e-3	0.021 246e-3

Table 2: Simulation results with series length $n = 40$.

Estimators	**Density types**		
	Normal	Logistic	Double-exponential
least-squares estimate	0.017 542e-5	0.008 745e-6	0.015 681e-5
van der Waerden sigend rank estimator	0.019 561e-5	0.011 586e-5	0.016 503e-5
Wilcoxon sigend rank estimator	0.018 613e-5	0.009 508e-5	0.012 571e-5
Laplace sigend rank estimator	0.024 715e-5	0.010 604e-5	0.011 413e-5
Van der Waerden rank estimator	0.020 560e-5	0.010 582e-5	0.014 500e-5
Wilcoxon rank estimator	0.017 617e-5	0.008 507e-5	0.013 569e-5
Laplace rank estimator	0.024 712e-5	0.012 605e-5	0.009 416e-5

Table 3: Simulation results with series length $n = 200$.

Acknowledgments

The authors express their deep gratitude to the Editor Ahmad Parsian and the referees. Their suggestions led to a considerable improvement of the presentation of the initial version of this paper. This research was partially supported by the "Programme d'Appui à la Recherche Scientifique (PARS)".

References

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

Allal, J., Kaaouachi, A. and Paindaveine, D. (2001), R-estimation for ARMA models. Journal of Nonparametric Statistics, **13**, 815-831.

Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. (1993), Efficient and Adaptive Estimation for Semiparametric Models.Baltimore: Johns Hopkins University Press.

Drost, F. G., Klaassen, C. A. J. and Werker, B. J. M. (1997), Adaptive estimation in time-series models. Ann. Statist., **25**, 786-817.

Fuller, W. A. (1976), Introduction to Statistical Time Series.New York: John Wiley.

Hájek, J. (1972), Local asymptotic minimax and admissibility in estimation. Proc. Sixth. Berkeley Sympo. Math. Statist. Prob., **1**, 175-194. Univ. of California Press.

Hájek, J. and Šidák, Z. (1967), Theory of Rank Tests. New York: Academic Press.

Hallin, M., Laforet, A. and Mélard, G. (1990), Distribution-free tests against serial dependence: signed or unsigned ranks? Jour. Statist. Plan. Inf., **24**, 151-165.

Hallin, M. and Puri, M. L. (1991), Time series analysis via rank-order theory: signed-rank tests for ARMA models. J. Multivariate Anal., **39**, 1-29.

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

Heiler, S. and Willers, R. (1988), Asymptotic normality of R-estimates in the linear model. Statistics, **19** , 173-184.

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.

Jeganathan, P. (1995), Some aspects of asymptotic theory with applications to time series models. Econometric Theory, **11**, 818-887.

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. New York: John Wiley.

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.

Koul, H. L. and Schick, A. (1997), Efficient estimation in nonlinear autoregressive time series models. Bernoulli, **3**, 247-277.

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

LeCam, 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. New York: John Wiley.

Whittle, P. (1962), Gaussian estimation in stationary time series. Bull. Int. Statist., **39**, 105-129.

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