Pergamon

Mathl. Comput. Modelling Vol. 19, No. 2, pp. 29–41, 1994
Copyright©1994 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0895-7177/94 $6.00 + 0.00

0895-7177(94)E0013-D

Parameter Estimation for Nearly Nonstationary AR(1) Processes

K. DZHAPARIDZE*
Centrum voor Wiskunde en Informatica
Kruislaa 413, 1098 SJ Amsterdam, The Netherlands

J. KORMOS*
Department of Mathematics, Lajos Kossuth University
1 Egyetem tér, Debrecen, Hungary

T. VAN DER MEER AND M. C. A. VAN ZUIJLEN
Department of Mathematics, Catholic University Nijmegen
Toernooiveld, 6525 ED Nijmegen, The Netherlands

Dedicated to Professor Lajos Tamássy on his 70$^{\text{th}}$ birthday

(Received and accepted November 1993)

Abstract—In this paper, we consider parameter estimation problems in the first order nearly nonstationary autoregression AR(1) model, which is described by formula (2.1). By allowing the most general class of innovations, we extend the result of Chan and Wei [1]. Moreover, we discuss a sequential procedure for estimating the parameter, extending the result of Lai and Siegmund [2] and Greenwood and Shiryaev [3] to the nearly nonstationary model. The results are essentially based on the preliminary Theorems 1 and 2, stating the weak convergence, as the sample size grows, of an observed nearly nonstationary AR(1) process to a corresponding AR(1) process in continuous time.

Keywords—Nearly nonstationary, Autoregressive process, Least-squares, Stochastic differential equation, Sequential estimation, Semimartingale.

1. INTRODUCTION

In the first order autoregression AR(1) model the observations X_k at time k are generated according to the scheme:

$$\begin{aligned} X_k &= \beta X_{k-1} + \varepsilon_k, \qquad k \in \{1, 2, \dots\} \\ X_0 &= 0, \end{aligned} \tag{1.1}$$

where the ϵ_k's are random disturbances ('noise' or 'innovations') and β is an unknown parameter. The traditional task is to construct an estimator for the parameter β based on the random variables $X_1, X_2, \dots, X_n$ and to characterize its exact or limiting distribution. The least-squares estimator (LSE) of β is

$$\widehat{\beta}_n := \frac{\sum_{k=1}^n X_k X_{k-1}}{\sum_{k=1}^n X_{k-1}^2}. \tag{1.2}$$

It is well-known that for $|\beta| < 1$ the standardized LSE of β is asymptotically normal in case the innovations are independent and identically distributed; (see e.g., [4–9]) indicated that in

*This work has been carried out while these authors were visiting the Department of Mathematics, University of Nijmegen.

The authors wish to thank M. Arato for his careful reading of the manuscript.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX situations where β is close to one, the limiting normal law may not be a satisfactory approximation of the exact distribution of the standardized LSE of β, even in the case where the ϵ_k's are independent standard normal random variables. In fact they showed that if $|\beta| = 1$, the limiting law of the standardized LSE of β is a functional of the Wiener process and non-normal. This observation led to the study of the so-called nearly nonstationary AR(1) model, where β in model (1.1) is replaced by $\beta_n = 1 - \gamma/n$ or by $\beta_n = e^{-\gamma/n}$, for some fixed real number γ. For example, see references [1,10–13], where the asymptotic distribution of the LSE of β is characterized as a functional of the Ornstein-Uhlenbeck process. Of course $\gamma = 0$ means $\beta_n = \beta = 1$.

The present paper is also devoted to the nearly nonstationary AR(1) model and generalizes the results of Chan and Wei [1] in several directions; see the Examples 1 and 2 in Section 2, where our general Theorem 2 is specified in cases of weakly dependent innovations and martingale differences, respectively. The conditions on the innovation process will be weakened considerably and we will allow the innovations to depend on n, so that we consider a triangular array of innovation random variables instead of a sequence. In fact, in Theorem 2a, a functional limit theorem is derived for a process generated by the standardized estimator of the parameter γ. This functional limit theorem is applied fruitfully in order to obtain asymptotic normality of our estimator in a sequential procedure, (see Theorem 3), which is a generalization of results of [2,3]. Finally, in Theorem 4, the so-called 'small noise' model is studied. It turns out that in this case our estimator is also asymptotically normal.

The results and the discussion of the results together with the applications are formulated in the Sections 2 and 3; the proofs are given in Section 4.

2. RESULTS

For every $n = 1, 2, \ldots$, let $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$ be a probability space equipped with a filtration $\mathbf{F}^n = (\mathcal{F}^n_k)_{k=0,\ldots,n}$, where $\mathcal{F}^n_0 = \{\emptyset, \Omega\}$. We will consider the model where the $X_{k,n}$'s are generated according to the scheme:

$$\begin{aligned} X_{k,n} &= \left(1 - \frac{\gamma}{n}\right) X_{k-1,n} + \varepsilon_{k,n}, \qquad k \in \{1, 2, \ldots, n\} \\ X_{0,n} &= 0, \end{aligned} \tag{2.1}$$

where γ is a fixed real number and $\varepsilon_{1,n}, \varepsilon_{2,n}, \ldots, \varepsilon_{n,n}$ are random variables adapted to the filtration $\mathbf{F}^n$. Notice that for $k = 1, \ldots, n$

$$X_{k,n} = \sum_{i=1}^{k} \left(1 - \frac{\gamma}{n}\right)^{k-i} \varepsilon_{i,n}. \tag{2.2}$$

The functions Y_n and M_n with

$$Y_n(t) := n^{-1/2} X_{[nt],n}, \qquad t \in [0,1], \tag{2.3}$$

$$M_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} \varepsilon_{i,n} \qquad t \in [0,1], \tag{2.4}$$

are right-continuous random step functions. Notice that for each fixed n, $Y_n = (Y_n(t), \mathcal{H}^n_t)$ and $M_n = (M_n(t), \mathcal{H}^n_t)$, with $\mathcal{H}^n_t = \mathcal{F}^n_{[nt]}$, are semimartingales on $[0,1]$. In this paper, we will use the following definition of the quadratic variation of a semimartingale X:

$$[X](t) := X^2(t) - X^2(0) - 2 \int_0^t X(s_-)\, dX(s).$$

Accordingly,

$$[M_n](t) = n^{-1} \sum_{i=1}^{[nt]} \varepsilon^2_{i,n}, \qquad t \in [0,1]. \tag{2.5}$$

Our underlying estimation process for γ is defined for $t \in [0,1]$ as

$$\widehat{\gamma}_{[nt],n} := \begin{cases} \dfrac{-\int_0^{[nt]/n} Y_n(s_-)\, dY_n(s)}{\int_0^{[nt]/n} Y_n^2(s)\, ds} & \text{if } \int_0^{[nt]/n} Y_n^2(s)\, ds > 0 \\ 0 & \text{elsewhere.} \end{cases} \tag{2.6}$$

Remark that the LSE of γ based on only $[nt]$ observations is given by $\widehat{\gamma}_{[nt],n}$ and of course the LSE of γ based on all the random variables $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ equals

$$\frac{n\left(\sum_{k=1}^n X_{k-1,n}^2 - \sum_{k=1}^n X_{k,n} X_{k-1,n}\right)}{\sum_{k=1}^n X_{k-1,n}^2} = \widehat{\gamma}_{n,n}.$$

In order to formulate our theorems we have to introduce some notation. $(D[0,1],\rho)$ indicates as usual the Skorokhod space of right-continuous functions on $[0,1]$ with existing left limits, endowed with the Skorokhod metric ρ. The supremum norm on $D[0,1]$ is denoted by $\|\ \|$ and $C[0,1]$ is the space of continuous functions on $[0,1]$. For a sequence of processes $(Z_n)_{n=1}^{\infty}$ and Z in $(D[0,1],\rho)$, we denote weak convergence of Z_n to Z by $Z_n(t) \xrightarrow{\mathcal{D}} Z(t)$ and for sequences of 1- or 2-dimensional random vectors we indicate weak convergence by $\xrightarrow{d}$ and $\xrightarrow{d^2}$ respectively. For M being a continuous semimartingale on $[0,1]$ let the process Y on $[0,1]$ be defined as

$$Y(t) = \int_0^t e^{\gamma(s-t)}\, dM(s). \tag{2.7}$$

Note that Y is the solution of the stochastic differential equation

$$dY(t) = -\gamma Y(t)\, dt + dM(t), \qquad Y(0) = 0.$$

Our first theorem gives sufficient conditions for the weak convergence of the process Y_n to the process Y defined by (2.3) and (2.7), respectively.

THEOREM 1. *Suppose that for some continuous semimartingale M on $D[0,1]$, with $M \not\equiv 0$, we have*

$$\begin{aligned} &M_n(t) \xrightarrow{\mathcal{D}} M(t) \\ &\big|[M_n](1) - [M](1)\big| \xrightarrow{d} 0. \end{aligned} \tag{2.8}$$

Then

$$Y_n(t) \xrightarrow{\mathcal{D}} Y(t).$$

The second theorem gives a functional law for the estimation process together with a weak convergence result of this process for fixed t.

THEOREM 2. *Suppose that for some continuous semimartingale M on $D[0,1]$, with $M \not\equiv 0$, we have*

$$\begin{aligned} &M_n(t) \xrightarrow{\mathcal{D}} M(t) \\ &\big|[M_n](t) - [M](t)\big| \xrightarrow{d} 0, \qquad \textit{for every } t \in [0,1]. \end{aligned} \tag{2.9}$$

Then the following two statements hold:

(a)

$$Z_n(t) \xrightarrow{\mathcal{D}} Z(t),$$

where

$$Z_n(t) = \left(\int_0^{[nt]/n} Y_n^2(s)\, ds\right)(\gamma - \widehat{\gamma}_{[nt],n}),$$

$$Z(t) = \left(\int_0^t Y^2(s)\, ds\right)(\gamma - \widehat{\gamma}_t)$$

with

$$\widehat{\gamma}_t = \begin{cases} \dfrac{-\int_0^t Y(s)\,dY(s)}{\int_0^t Y^2(s)\,ds} & \text{if } \int_0^t Y^2(s)\,ds > 0 \\ 0 & \text{elsewhere;} \end{cases}$$

(b) *for every* $\|\ \|$*-continuous function* $\psi_t : D[0,1] \to \mathbb{R}$,

$$\psi_t(Y_n)(\gamma - \widehat{\gamma}_{[nt],n}) \xrightarrow{d} \psi_t(Y)(\gamma - \widehat{\gamma}_t), \qquad \text{for every } t \in (0,1].$$

NOTE. We remark that with minor changes the present approach is also applicable in the stationary case, where it is not presupposed that $X_{0,n} = 0$.

REMARK 1. If the $\epsilon_{k,n}$'s are independent of n and M is a Wiener process, then for every fixed t, $\frac{n}{[nt]}[M_n](t)$ is a subsequence of $[M_n](1)$, (see (2.5)), so that (2.9) is equivalent to (2.8).

Important examples where we can apply the Theorems 1 and 2 and where M is actually a Wiener process are the following:

EXAMPLE 1. Let us assume that the $\epsilon_{k,n}$'s are independent of n and $(\epsilon_k)_{k=1}^{\infty}$ is a so-called ϕ-mixing sequence, i.e., a strictly stationary sequence which satisfies the following conditions:

$$\sum_{k=1}^{\infty} \phi_k^{1/2} < \infty, \qquad \mathbb{E}\,\epsilon_1^2 < \infty,$$

with $\phi_k = \sup|P(B|A) - P(B)|$ and where the supremum is taken over all sets $A \in \sigma(\epsilon_1)$ with $P(A) > 0$ and sets $B \in \sigma(\epsilon_{k+1}, \epsilon_{k+2}, \ldots)$. Assuming without loss of generality that $\mathbb{E}\,\epsilon_1 = 0$, Theorem 20.1 in Chapter 4 of [14] yields

$$M_n(t) \xrightarrow{\mathcal{D}} \sigma W(t) \quad \text{and} \quad [M_n](1) \xrightarrow{d} \sigma^2,$$

where $W(t)$ is a Wiener process and $\sigma^2 = \mathbb{E}\,\epsilon_1^2 + 2\sum_{k=2}^{\infty} \mathbb{E}(\epsilon_1\epsilon_k)$. Combining this result with Remark 1, we find that the Theorems 1 and 2 are applicable. We note that the conditions for weak convergence of so-called ρ-mixing or α-mixing sequences can be found in Theorem 2 in Chapter 5 of [15] and our theorems are applicable in these situations as well. (cf. [10])

EXAMPLE 2. Next let us assume that $(\epsilon_{k,n})$ is a triangular array of martingale differences with $\mathbb{E}\,\epsilon_{1,n} = 0$ and $\mathbb{E}\,\epsilon_{k,n}^2 = \sigma^2 \in (0,\infty)$ satisfying the following conditions:

$$\begin{aligned} &\forall t \in [0,1], && \frac{1}{n}\sum_{k=1}^{[nt]} \mathbb{E}\left(\varepsilon_{k,n}^2 \mid \mathcal{F}_{k-1}^n\right) \xrightarrow{d} \sigma^2 t, && \text{as } n \to \infty \\ &\forall \alpha > 0, && \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left(\varepsilon_{k,n}^2 I_{|\varepsilon_{k,n}| > n^{1/2}\alpha} \mid \mathcal{F}_{k-1}^n\right) \xrightarrow{d} 0, && \text{as } n \to \infty. \end{aligned} \tag{2.10}$$

Theorem 7.11 in [15] leads to

$$M_n(t) \xrightarrow{\mathcal{D}} M = \sigma\, W(t),$$

where $W(t)$ is a Wiener process, and

$$[M_n](t) \xrightarrow{d} \sigma^2 t, \qquad \text{for every } t \in [0,1].$$

Hence (2.9) is satisfied and we can apply the Theorems 1 and 2. If we add the condition that the $\epsilon_{k,n}$'s are independent of n, then we can conclude as in Remark 1, that the first condition of (2.10) is equivalent to

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left(\varepsilon_k^2 \mid \mathcal{F}_{k-1}\right) \xrightarrow{d} \sigma^2, \qquad \text{as } n \to \infty.$$

In this case, assuming that $\sigma = 1$, Chan and Wei [1] proved that

$$\left(\sum_{k=1}^{n} X_{k-1,n}^2\right)^{1/2} \left(\widehat{\beta}_n - \beta_n\right) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \frac{W(t)}{1+bt}\, dW(t)}{\left(\int_0^1 \frac{W^2(t)}{(1+bt)^2}\, dt\right)^{1/2}}, \tag{2.11}$$

where $\widehat{\beta}_n - \beta_n = n^{-1}(\gamma - \widehat{\gamma}_{n,n})$ and $b = e^{2\gamma} - 1$. One year later Chan [16] added that the distribution of the random variable at the right hand side of (2.11) is equal to the distribution of

$$\frac{\int_0^1 Y(t) dW(t)}{\left(\int_0^1 Y^2(t) dt\right)^{1/2}}. \tag{2.12}$$

This last result also follows from Theorem 2b with $\psi(Y_n) = \left(\int_0^1 Y_n^2(t)\, dt\right)^{1/2}$, as can be seen easily. Note that we wrote the factor $n^{-1/2}$ in (2.3) and (2.4) and the factor $1/n$ in (2.10) separately, although we allowed the $\epsilon_{k,n}$'s to depend on n. This is done to facilitate comparison of our theorems with the existing literature.

REMARK 2. Note that $\widehat{\gamma}_t$ is the limiting analogue of $\widehat{\gamma}_{[nt],n}$ and is in fact the maximum likelihood estimator of γ in the case where the above-mentioned M is a Wiener process (see [17]).

REMARK 3. Note that in the case where $M(t) = \sigma W(t)$, for some positive real number σ, we have

$$Y(t) = \sigma \int_0^t e^{\gamma(s-t)}\, dW(s)$$

and $Y(t)$ satisfies

$$dY(t) = -\gamma Y(t)\, dt + \sigma\, dW(t), \qquad Y(0) = 0.$$

The Fisher information process about γ (see [18]) equals

$$\frac{1}{\sigma^2} \int_0^1 Y^2(t)\, dt.$$

Hence, for the (total) Fisher information $I(\gamma, \sigma)$ about γ, we have

$$I(\gamma,\sigma) = \mathbb{E}\left(\frac{1}{\sigma^2}\int_0^1 Y^2(t)\, dt\right) = \frac{e^{-2\gamma} - 1 + 2\gamma}{4\gamma^2}.$$

Remark that, due to the fact that $Y(0) = 0, I(\gamma, \sigma)$ is independent of σ; (cf. Remark 5 below). Also, we note that the estimator of γ has the drawback to be biased, as follows from Theorem 17.3 in [17]. More precisely, we have

$$\mathbb{E}(\widehat{\gamma}_t - \gamma) = \int_0^\infty \frac{d}{d\gamma}\, \psi_t(\gamma, a)\, da, \tag{2.13}$$

$$\mathbb{E}\left(\widehat{\gamma}_t - \gamma\right)^2 = \int_0^t \psi_t(\gamma, a)\, da + \int_0^t a\, \frac{d^2}{d\gamma^2}\, \psi_t(\gamma, a)\, da,$$

where

$$\psi_t(\gamma, a) = \mathbb{E} \exp\left(-a \int_0^t Y^2(s)\, ds\right) = e^{(\lambda+\gamma/2)t} \left(\frac{2\lambda}{(\lambda+\gamma)(e^{2\gamma t} - 1) + 2\lambda}\right)^{1/2}$$

and

$$\lambda = \left(2a + \gamma^2\right)^{1/2}.$$

REMARK 4. The following example shows why it is useful to consider M in Theorem 1 and 2 as semimartingales instead of restricting ourselves to martingales, as in Examples 1 and 2. Let the $\epsilon_{k,n}$'s be i.i.d. random variables and let the X's be generated by

$$X_{k,n} = (1 - \gamma/n)\, X_{k-1,n} + \widehat{\epsilon}_{k,n}, \qquad X_{0,n} = 0,$$

where $\widehat{\epsilon}_{k,n} = \epsilon_{k,n} - n^{-1/2}\, \epsilon_{k-1,n}^2$. Then $M(t) = W(t) - t$, where $W(t)$ is a Wiener process and M is a semimartingale, but not a martingale.

3. PROCEDURES LEADING TO ASYMPTOTICALLY NORMAL ESTIMATORS

In Remark 3 we have seen that the considered estimator of γ has the drawback to be biased. The following theorem presents a sequential procedure in order to obtain asymptotic normality of our estimator, and hence, to achieve unbiasedness. Comparing (2.9) with (3.1) below, we see that Theorem 3 is applicable in the case where $M = W$.

THEOREM 3. *Let $W(t)$ be a Wiener process. Suppose that we have*

$$\begin{aligned} &M_n(t) \xrightarrow{\mathcal{D}} W(t) \\ &[M_n](t) \xrightarrow{d} t, \qquad \text{for every } t \in [0,1]. \end{aligned} \tag{3.1}$$

Fix $\delta > 0$ and define the stopping times τ_n and τ by

$$\tau_n := \begin{cases} \inf\{t : \int_0^t Y_n^2(s)\,ds \geq \delta^2\} & \text{if } \int_0^1 Y_n^2(s)\,ds \geq \delta^2 \\ 1 & \text{elsewhere,} \end{cases}$$
$$\tau := \begin{cases} \inf\{t : \int_0^t Y^2(s)\,ds \geq \delta^2\} & \text{if } \int_0^1 Y^2(s)\,ds \geq \delta^2 \\ 1 & \text{elsewhere.} \end{cases}$$

Then the two following statements hold:

(a) $$H_n(\tau_n) \xrightarrow{d} H(\tau), \qquad \text{as } n \to \infty;$$

(b) $$H(\tau) \xrightarrow{d} \mathcal{N}(0,1), \qquad \text{as } \delta \to 0,$$

where

$$H_n(t) := \frac{Z_n(t)}{\left(\int_0^{[nt]/n} Y_n^2(s)\,ds\right)^{1/2}}, \qquad H(t) := \frac{Z(t)}{\left(\int_0^t Y^2(s)\,ds\right)^{1/2}},$$

with the processes Z_n and Z defined in Theorem 2.

REMARK 5. Suppose that $\int_0^1 Y^2(s)\,ds \geq \delta^2$, then we can replace the above-mentioned statement (b) by $H(\tau) \stackrel{d}{=} \mathcal{N}(0,1)$.

Another possibility to obtain an asymptotically normal estimator is to consider the so-called 'small noise' model where we have a different condition on the $\epsilon_{k,n}$'s. Cf. (3.1) with (3.3) below.

THEOREM 4. *Let $W(t)$ be a Wiener process, $\sigma \in (0,\infty)$ and $X_{0,n} = c$, for some real constant c. Let*

$$U(t) = U(0)\,e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)}\,dW(s). \tag{3.2}$$

Suppose that we have

$$\begin{aligned} &n^{1/2} M_n(t) \xrightarrow{\mathcal{D}} \sigma W(t), \\ &n[M_n](t) \xrightarrow{d} \sigma^2 t, \qquad \text{for every } t \in [0,1]. \end{aligned} \tag{3.3}$$

Then, with the same notation as in Theorems 2 and 3, the two following statements hold:

(a)

$$\frac{n^{1/2}}{\sigma} H_n(t) \xrightarrow{d} \frac{\int_0^t U(s)\,dW(s)}{\left(\int_0^t U^2(s)\,ds\right)^{1/2}}, \qquad \text{as } n \to \infty, \text{ for every } t \in (0,1],$$

(b) *If $c \neq 0$, then*

$$\frac{\int_0^t U(s)\,dW(s)}{\left(\int_0^t U^2(s)\,ds\right)^{1/2}} \xrightarrow{d} \mathcal{N}(0,1), \qquad \text{as } \sigma \to 0, \text{ for every } t \in (0,1].$$

REMARK 6. In the present 'small noise' model, the (total) Fisher information $I(\gamma,\sigma)$ about γ is

$$I(\gamma,\sigma) = \mathbb{E}\left(\frac{1}{\sigma^2}\int_0^1 U^2(t)\,dt\right) = \frac{c^2(1-e^{-2\gamma})}{\sigma^2 2\gamma} + \frac{e^{-2\gamma}-1+2\gamma}{4\gamma^2},$$

which clearly reduces to the information formula in Remark 3 as $c=0$. If $c\neq 0$, as in assertion (b) above, then $I(\gamma,\sigma)$ tends to infinity as σ goes to 0.

REMARK 7. We get the desired asymptotic normality of the (normalized) estimator $\gamma_{[nt],n}$ in Theorem 4 by letting n to infinity first and $\sigma\to 0$ later. We can find this result in a similar way if we replace (3.3) by

$$nM_n(t) \xrightarrow{\mathcal{D}} W(t),$$
$$n^2[M_n](t) \xrightarrow{d} t, \qquad \text{for every } t\in[0,1].$$

In that case,

$$nH_n(t) \xrightarrow{d} \mathcal{N}(0,1), \qquad \text{as } n\to\infty, \text{ for every } t\in(0,1].$$

4. PROOFS

To simplify notations, we introduce on $[0,1]$ the continuous function h_γ, defined by

$$h_\gamma(t) := e^{\gamma t}.$$

Moreover, for a stochastic process Z on $[0,1]$ and $k=1,2,\dots,n$, let

$$\Delta_n Z\left(\frac{k}{n}\right) := Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right).$$

Observe that the difference operator Δ_n has the following properties for $k=1,2,\dots,n$

$$\Delta_n Y_n\left(\frac{k}{n}\right) = \Delta_n M_n\left(\frac{k}{n}\right) - \frac{\gamma}{n} Y_n\left(\frac{k-1}{n}\right),$$
$$\Delta_n h_\gamma\left(\frac{k}{n}\right) = (e^{\gamma/n}-1)\, h_\gamma\left(\frac{k-1}{n}\right).$$

LEMMA 4.1. *For $n\in\mathbb{N}$ and $\gamma\in\mathbb{R}$ we have (with $[M_n](1)\equiv \frac{1}{n}\sum_{i=1}^n \varepsilon_{i,n}^2$)*

$$\|Y_n\| \le e^{|\gamma|}\, n^{1/2}\left([M_n](1)\right)^{1/2}, \tag{4.1}$$

$$n^{3/2}\sup_{k\in\{1,2,\dots,n\}}\left|\Delta_n\left(h_\gamma M_n - h_\gamma Y_n\right)\left(\frac{k}{n}\right) - M_n\left(\frac{k-1}{n}\right)\Delta_n h_\gamma\left(\frac{k}{n}\right)\right|$$
$$\le \gamma^2 e^{4|\gamma|}\left([M_n](1)\right)^{1/2}, \tag{4.2}$$

$$\sup_{s\in[0,1]}\left|h_\gamma\left(\frac{[ns]}{n}\right) Y_n(s) + \int_0^s M_n(t)\,dh_\gamma(t) - M_n(s)\,h_\gamma(s)\right|$$
$$\le \gamma^2 e^{4|\gamma|}\left([M_n](1)\right)^{1/2} n^{-1/2}. \tag{4.3}$$

PROOF. For $k\in\{1,\dots,n\}$, we have

$$|X_{k,n}| = \left|\sum_{i=1}^k \beta_n^{k-i}\varepsilon_{i,n}\right|$$
$$\le \left(\sum_{i=1}^k (1-\gamma/n)^{2(k-i)}\right)^{1/2}\left(\sum_{i=1}^k \varepsilon_{i,n}^2\right)^{1/2}$$
$$\le e^{|\gamma|}\sqrt{n}\left(\sum_{i=1}^n \varepsilon_{i,n}^2\right)^{1/2} = e^{|\gamma|}\, n\left([M_n](1)\right)^{1/2}$$

so that

$$\begin{aligned}\|Y_n\| &= \max_{k\in\{1,\dots,n\}}\left|Y_n\left(\frac{k}{n}\right)\right| \\ &= \max_{k\in\{1,\dots,n\}}\frac{1}{\sqrt{n}}|X_{k,n}| \le e^{|\gamma|}\,n^{1/2}\big([M_n](1)\big)^{1/2},\end{aligned}$$

which proves (4.1). The proof of (4.2) follows from the fact that simple algebra shows that

$$\begin{aligned}n^{3/2}\sup_{k\in\{1,2,\dots,n\}}&\left|\Delta_n\left(h_\gamma M_n - h_\gamma Y_n\right)\left(\frac{k}{n}\right) - M_n\left(\frac{k-1}{n}\right)\Delta_n h_\gamma\left(\frac{k}{n}\right)\right| \\ &= \left|\frac{\gamma}{n}e^{\gamma/n} - e^{\gamma/n} + 1\right| n^{3/2}\sup_{k\in\{1,2,\dots,n\}}\left|h_\gamma\left(\frac{k-1}{n}\right)Y_n\left(\frac{k-1}{n}\right)\right| \\ &\le \left|\frac{\gamma}{n}e^{\gamma/n} - e^{\gamma/n} + 1\right| n^{3/2}\,\|h_\gamma Y_n\| \\ &\le n^{1/2}\,e^{2|\gamma|}\left|\gamma - \left(\frac{e^{-\gamma/n}-1}{-1/n}\right)\right|\,\|Y_n\| \\ &\le n^{1/2}\,\gamma^2\,e^{3|\gamma|}\,n^{-1}\,\|Y_n\| \\ &\le \gamma^2\,e^{4|\gamma|}\,\big([M_n](1)\big)^{1/2}.\end{aligned}$$

Finally we prove (4.3) by using (4.1) and (4.2). Indeed we have for $s\in[0,1]$

$$\begin{aligned}&\sup_{s\in[0,1]}\left|\left(h_\gamma M_n - h_\gamma Y_n\right)\left(\frac{[ns]}{n}\right) - \sum_{k=1}^{[ns]} M_n\left(\frac{k-1}{n}\right)\left(h_\gamma\left(\frac{k}{n}\right) - h_\gamma\left(\frac{k-1}{n}\right)\right)\right| \\ &= \sup_{s\in[0,1]}\left|\left(h_\gamma M_n - h_\gamma Y_n\right)\left(\frac{[ns]}{n}\right) - \int_0^{[ns]/n} M_n(t)\,dh_\gamma(t)\right| \\ &= \sup_{s\in[0,1]}\left|\left(h_\gamma M_n - h_\gamma Y_n\right)\left(\frac{[ns]}{n}\right) - \int_0^{s} M_n(t)\,dh_\gamma(t) + M_n(s)\left(h_\gamma(s) - h_\gamma\left(\frac{[ns]}{n}\right)\right)\right| \\ &= \sup_{s\in[0,1]}\left|h_\gamma\left(\frac{[ns]}{n}\right)Y_n(s) + \int_0^{s} M_n(t)\,dh_\gamma(t) - M_n(s)\,h_\gamma(s)\right| \\ &\le \gamma^2\,e^{4|\gamma|}\big([M_n](1)\big)^{1/2}n^{-1/2}.\end{aligned}$$

∎

We will need the following lemma, which is in fact a direct consequence of the continuous mapping theorem and the Skorokhod-construction.

LEMMA 4.2. *Suppose that $\Phi : D[0,1]\to D[0,1]$ is a $\|\ \ \|$-continuous function and $Z_n \xrightarrow{\mathcal{D}} Z$ in $(D[0,1],\rho)$, where Z is a continuous function. Then*

$$\Phi(Z_n) \xrightarrow{\mathcal{D}} \Phi(Z) \qquad \text{in } (D[0,1],\rho).$$

PROOF. Due to the Skorokhod-construction, we can find processes $\widetilde{Z}_n$ and a process $\widetilde{Z}$, such that $\widetilde{Z}_n \overset{\mathcal{D}}{=} Z_n$, $\widetilde{Z} \overset{\mathcal{D}}{=} Z$ and $\rho(\widetilde{Z}_n,\widetilde{Z})$ is a random variable with

$$\rho(\widetilde{Z}_n,\widetilde{Z}) \to 0 \qquad \text{a.s.}$$

Using the fact that $\widetilde{Z}\in C[0,1]$ a.s., we conclude that

$$\left\|\widetilde{Z}_n - \widetilde{Z}\right\| \to 0 \qquad \text{a.s.}$$

Since Φ is a $\| \quad \|$-continuous function, we have

$$\left\|\Phi(\widetilde{Z}_n) - \Phi(\widetilde{Z})\right\| \to 0 \qquad \text{a.s.}$$

and hence

$$\Phi(\widetilde{Z}_n) \xrightarrow{\mathcal{D}} \Phi(\widetilde{Z}) \qquad \text{in } (D[0,1], \rho).$$

The last relation implies the desired result. ∎

PROOF OF THEOREM 1. Let $\Phi : D[0,1] \to D[0,1]$ be a $\| \quad \|$-continuous function defined by

$$\Phi(X)(s) := X(s)\, h_\gamma(s) - \int_0^s X(t)\, dh_\gamma(t).$$

Then by assertion (4.3) of Lemma 4.1, we have

$$\sup_{s\in[0,1]} \left| h_\gamma\left(\frac{[ns]}{n}\right) Y_n(s) - \Phi(M_n)(s) \right| \xrightarrow{d} 0, \qquad \text{as } n \to \infty.$$

On the other hand Lemma 4.2, applied to $Z_n = M_n$ and $Z = M$ gives

$$\Phi(M_n) \xrightarrow{\mathcal{D}} \Phi(M) \qquad \text{in } (D[0,1], \rho).$$

Hence,

$$h_\gamma\left(\frac{[ns]}{n}\right) Y_n(s) \xrightarrow{\mathcal{D}} M(s)\, h_\gamma(s) - \int_0^s M(t)\, dh_\gamma(t) \qquad \text{in } (D[0,1], \rho).$$

By partial integration the last relation yields the desired result:

$$Y_n(s) \xrightarrow{\mathcal{D}} (h_\gamma(s))^{-1} \int_0^s h_\gamma(t)\, dM(t).$$

∎

For the proof of Theorem 2, we need the following two lemmas. The assertion of the first lemma is obvious; for completeness, however, we give a detailed proof. The second lemma is purely technical.

LEMMA 4.3. *We have* $[Y] = [M]$ *a.s.*

PROOF. Note that from the definition of Y in (2.7) it follows that for $t \in [0,1]$

$$(h_\gamma Y)(t) = \int_0^t h_\gamma(s)\, dM(s),$$

so that

$$[h_\gamma Y](t) = \int_0^t h_\gamma^2(s)\, d[M](s).$$

On the other hand by using Ito's formula we have for $t \in [0,1]$

$$\left((h_\gamma Y)(t)\right)^2 = 2\int_0^t h_\gamma(s)\, Y^2(s)\, dh_\gamma(s) + 2\int_0^t h_\gamma^2(s)\, Y(s)\, dY(s) + \int_0^t h_\gamma^2(s)\, d[Y](s),$$

which by partial integration means that

$$\left((h_\gamma Y)(t)\right)^2 = 2\int_0^t h_\gamma(s)\, Y(s)\, d\left(h_\gamma(s) Y(s)\right) + \int_0^t h_\gamma^2(s)\, d[Y](s).$$

Hence,

$$\int_0^t h_\gamma^2(s)\,d[M](s) = \int_0^t h_\gamma^2(s)\,d[Y](s),$$

from which the result follows. ∎

LEMMA 4.4. *For $n \in N$, $\gamma \in \mathbb{R}$ and $t \in [0,1]$, we have*

$$2(1-\gamma/n)\left(\int_0^{[nt]/n} Y_n^2(s)\,ds\right)(\gamma - \widehat{\gamma}_{[nt],n}) = Y_n^2(t) + \left(2\gamma - \frac{\gamma^2}{n}\right)\int_0^{[nt]/n} Y_n^2(s)\,ds - [M_n](t).$$

PROOF. Note first that by (2.1)

$$X_{k,n}^2 = \left(1-\gamma/n\right)^2 X_{k-1,n}^2 + \varepsilon_{k,n}^2 + 2(1-\gamma/n)\,X_{k-1,n}\,\varepsilon_{k,n},$$

which implies

$$\begin{aligned}
&\frac{1}{n}\sum_{k=1}^{[nt]} 2(1-\gamma/n)\,X_{k-1,n}\,\varepsilon_{k,n} \\
&= \frac{1}{n}\sum_{k=1}^{[nt]} X_{k,n}^2 - \frac{1}{n}\sum_{k=1}^{[nt]}\left(1-\gamma/n\right)^2 X_{k-1,n}^2 - \frac{1}{n}\sum_{k=1}^{[nt]}\varepsilon_{k,n}^2 \\
&= \sum_{k=1}^{[nt]} Y_n^2\left(\frac{k}{n}\right) - \left(1-\gamma/n\right)^2\sum_{k=1}^{[nt]} Y_n^2\left(\frac{k-1}{n}\right) - [M_n](t) \\
&= Y_n^2\left(\frac{[nt]}{n}\right) + \left(1-\left(1-\gamma/n\right)^2\right)\sum_{k=1}^{[nt]} Y_n^2\left(\frac{k-1}{n}\right) - [M_n](t) \\
&= Y_n^2(t) + \left(2\gamma - \frac{\gamma^2}{n}\right)\int_0^{[nt]/n} Y_n^2(s)\,ds - [M_n](t).
\end{aligned} \tag{4.4}$$

Next, we remark that for $\int_0^{[nt]/n} Y_n^2(s)\,ds > 0$, we have by definition (2.6) that

$$\begin{aligned}
\gamma - \widehat{\gamma}_{[nt],n} &= \gamma - \frac{n\left(\sum_{k=1}^{[nt]} X_{k-1,n}^2 - \sum_{k=1}^{[nt]} X_{k,n}\,X_{k-1,n}\right)}{\sum_{k=1}^{[nt]} X_{k-1,n}^2} \\
&= \frac{\frac{1}{n}\left(\gamma/n\sum_{k=1}^{[nt]} X_{k-1,n}^2 - \sum_{k=1}^{[nt]} X_{k-1,n}^2 + \sum_{k=1}^{[nt]} X_{k,n}\,X_{k-1,n}\right)}{\frac{1}{n^2}\sum_{k=1}^{[nt]} X_{k-1,n}^2} \\
&= \frac{\frac{1}{n}\sum_{k=1}^{[nt]} X_{k-1,n}\,\varepsilon_{k,n}}{\int_0^{[nt]/n} Y_n^2(s)\,ds}.
\end{aligned}$$

Combining this identity with (4.4), we obtain

$$\begin{aligned}
&2(1-\gamma/n)\left(\int_0^{[nt]/n} Y_n^2(s)\,ds\right)(\gamma - \widehat{\gamma}_{[nt],n}) \\
&= 2(1-\gamma/n)\,\frac{1}{n}\sum_{k=1}^{[nt]} X_{k-1,n}\,\varepsilon_{k,n} \\
&= Y_n^2(t) + \left(2\gamma - \frac{\gamma^2}{n}\right)\int_0^{[nt]/n} Y_n^2(s)\,ds - [M_n](t).
\end{aligned}$$

For $\int_0^{[nt]/n} Y_n^2(s)\,ds = 0$, we clearly have

$$Y_n^2(t) + \left(2\gamma - \frac{\gamma^2}{n}\right) \int_0^{[nt]/n} Y_n^2(s)\,ds - [M_n](t) = 0. \qquad \blacksquare$$

PROOF OF THEOREM 2a. Observe that from (2.9) and Problem 2 in Chapter 5 of [15] it follows that

$$\|[M_n] - [M]\| \xrightarrow{d} 0.$$

By Lemma 4.4 and Theorem 1, we have

$$\left(\int_0^{[nt]/n} Y_n^2 ds\right)(\gamma - \widehat{\gamma}_{[nt],n}) \xrightarrow{\mathcal{D}} \frac{1}{2}Y^2(t) + \gamma \int_0^t Y^2(s)\,ds - \frac{1}{2}[M](t).$$

Using the definition of $\widehat{\gamma}_t$ and Lemma 4.3, we see that the righthand side of the last expression can be rewritten as follows:

$$\begin{aligned}\frac{1}{2}Y^2(t) + \gamma \int_0^t Y^2(s)\,ds - \frac{1}{2}[Y](t)& \\ &= \int_0^t Y(s)\,dY(s) + \gamma \int_0^t Y^2(s)\,ds \\ &= \left(\int_0^t Y^2(s)\,ds\right)(\gamma - \widehat{\gamma}_t). \qquad \blacksquare\end{aligned}$$

PROOF OF THEOREM 2b. Notice that $\int_0^t Y^2(s)\,ds > 0$ a.s., so

$$(\psi(Y_n), \gamma - \widehat{\gamma}_{[nt],n}) \xrightarrow{d^2} (\psi(Y), \gamma - \widehat{\gamma}_t)$$

in $\mathbb{R}^2$, from which the result follows in virtue of the continuous mapping theorem. $\blacksquare$

Now we introduce two lemmas and some notation which will be applied in the course of proving Theorem 3. For $X(t) \in D[0,1]$, we define

$$(LX)(t) := \left(\int_0^t X^2(s)\,ds\right)^{1/2}. \tag{4.5}$$

LEMMA 4.5. *With τ, Y, and W as defined in Theorem 3, we have*

(a)
$$\frac{1}{\delta^2}\int_0^\tau Y^2(t)\,dt \xrightarrow{d} 1, \qquad \text{as } \delta \to 0;$$

(b)
$$\frac{1}{\delta}\int_0^\tau Y(t)\,dW(t) \xrightarrow{d} \mathcal{N}(0,1), \qquad \text{as } \delta \to 0.$$

PROOF OF (a). Let $\widetilde{Y}(t) := \frac{1}{\delta}\chi(t \le \tau)Y(t)$ and $A_\delta = \left\{\int_0^1 Y^2(t)\,dt < \delta^2\right\}$. We have $(L\widetilde{Y})(1) \le 1$ and $1 - (L\widetilde{Y})(1) \le I_{A_\delta}$ by definition of τ. Since $I_{A_\delta} \xrightarrow{d} 0$ as $\delta \to 0$, we have $(L\widetilde{Y})(1) \xrightarrow{d} 1$ as $\delta \to 0$.

PROOF OF (b). Define

$$V_\lambda(t) := \exp\left(i\lambda \int_0^t \widetilde{Y}(s)\,dW(s) + 1/2\lambda^2 \int_0^t \widetilde{Y}^2(s)\,ds\right), \qquad \lambda \in \mathbb{R}.$$

Observe that by the inequality

$$\exp\left(\frac{1}{2}\lambda^2\int_0^t \widetilde{Y}^2(s)\,ds\right) \le \exp\left(\frac{1}{2}\lambda^2\int_0^1 \widetilde{Y}^2(s)\,ds\right) \le e^{1/2\lambda^2},$$

Proposition 5.12 on page 198 of [19] is applicable here, so we have

$$\mathbb{E}\,V_\lambda(1) = 1. \tag{4.6}$$

Since $(L\widetilde{Y})(1) - 1 \xrightarrow{d} 0$, we have

$$V_\lambda(1) - \exp\left(i\lambda\int_0^1 \widetilde{Y}(s)\,dW(s) + \frac{1}{2}\lambda^2\right) \xrightarrow{d} 0,$$

and hence,

$$\mathbb{E}\,V_\lambda(1) - \mathbb{E}\exp\left(i\lambda\int_0^1 \widetilde{Y}(s)\,dW(s) + \frac{1}{2}\lambda^2\right) \to 0,$$

and it follows from (4.6) that, as $\delta \to 0$,

$$e^{1/2\lambda^2}\,\mathbb{E}\left(\exp\left(i\lambda\int_0^1 \widetilde{Y}(s)\,dW(s)\right)\right) = \mathbb{E}\exp\left(i\lambda\int_0^1 \widetilde{Y}(s)\,dW(s) + \frac{1}{2}\lambda^2\right) \to \mathbb{E}\,V_\lambda(1) = 1,$$

which establishes the lemma. ∎

LEMMA 4.6. *For fixed $\delta \in (0,\infty)$, let $\phi : (D[0,1],\rho) \to (\mathbb{R}, |\ \ |)$ be defined by*

$$\phi(x) := \begin{cases} \inf\{t : x(t) \ge \delta\} & \text{if } \|x\| \ge \delta \\ 1 & \textit{elsewhere.} \end{cases}$$

Then ϕ is continuous on A, where $A = \{x \in D[0,1] : x$ is continuous and strictly increasing$\}$.

PROOF. Let $x \in A$ and x_n a sequence in $D[0,1]$ with $\rho(x_n, x) \to 0$, as $n \to \infty$. Then also $\|x_n - x\| \to 0$. If $x(1) < \delta$ then $\|x_n\| < \delta$ for n sufficiently large, so $|\phi(x_n) - \phi(x)| \to 0$. If $x(1) > \delta$ then $\|x_n\| > \delta$ for n sufficiently large and

$$\phi(x_n) \le \inf\{t : x(t) - \|x_n - x\| \ge \delta\} = \inf\{t : x(t) \ge \delta + \|x_n - x\|\} = x^{-1}(\delta + \|x_n - x\|).$$

In the same way, we find that

$$\phi(x_n) \ge \inf\{t : x(t) \ge \delta - \|x_n - x\|\} = x^{-1}(\delta - \|x_n - x\|).$$

Since x^{-1} is continuous, we conclude that $|\phi(x_n) - \phi(x)| \to 0$. If $x(1) = \delta$ then

$$x^{-1}(\delta - \|x_n - x\|) \le \phi(x_n) \le 1.$$

Hence, we conclude that $|\phi(x_n) - \phi(x)| \to 0$. ∎

PROOF OF THEOREM 3. Clearly

$$Z_n(t) = \frac{1}{n}\sum_{k=1}^{[nt]} X_{k-1,n}\,\varepsilon_{k,n}.$$

Hence, from (4.4), we see that

$$Z_n(t) = G(Y_n, [M_n])(t) + R_n(t),$$

where $G : (D^2[0,1], \rho^2) \to (D[0,1], \rho)$ is a continuous mapping defined by

$$G(a,b)(t) := \frac{1}{2}a^2(t) + \gamma \int_0^t a^2(s)\,ds - \frac{1}{2}\,b(t),$$

and R_n is defined by

$$R_n(t) := \frac{\gamma}{n}\,H_n(t) - \gamma \int_{[nt]/n}^t Y_n^2(s)\,ds - \frac{\gamma^2}{2n}\int_0^{[nt]/n} Y_n^2(s)\,ds,$$

so that

$$\|R_n\| \xrightarrow{d} 0.$$

From the definition of Y, we conclude that LY is a.s. a continuous and strictly increasing function. Observe that $\tau_n = \phi(LY_n)$ and $\tau = \phi(LY)$ with ϕ defined in Lemma 4.6. Using Lemma 4.6, the continuous mapping theorem and the remarks on page 145 of [14], we find

$$H_n(\tau_n) = \frac{Z_n(\tau_n)}{LY_n\left(\frac{[n\tau_n]}{n}\right)} \xrightarrow{d} \frac{Z(\tau)}{(LY)(\tau)}, \qquad \text{as } n \to \infty.$$

Due to Lemma 4.5, we see that

$$\frac{Z(\tau)}{(LY)(\tau)} \xrightarrow{d} \mathcal{N}(0,1), \qquad \text{as } \delta \to 0.$$ ∎

PROOF OF THEOREM 4a. The proof is analogous to that of Theorem 2b. ∎

PROOF OF THEOREM 4b. Since $\sigma \to 0$, it follows from (3.2) that $U(t)$ converges to $ce^{-\gamma t}$ and the result follows. ∎

REFERENCES

1. N.H. Chan and C.Z. Wei, Asymptotic inference for nearly nonstationary AR (1) processes, *Ann. Statist.* **15**, 1050–1063 (1987).
2. T.L. Lai and D. Siegmund, Fixed accuracy estimation of an autoregressive parameter, *Ann. Statist.* **11**, 478–485 (1983).
3. P.E. Greenwood & A.N. Shiryayev, *Sequential Estimation for First Order Autoregressive Models*, (Preprint), (1989).
4. H.B. Mann and A. Wald, On the statistical treatment of linear stochastic difference equations, *Econometrica* **11**, 173–220 (1943).
5. T.W. Anderson, On asymptotic distributions of estimates of parameters of stochastic difference equations, *Ann. Math. Statist.* **30**, 676–687 (1959).
6. J.S. White, The limiting distribution of the serial correlation coefficient in the explosive case, *Ann. Math. Statist.* **29**, 1188–1197 (1958).
7. J.S. White, The limiting distribution of the serial correlation coefficient in the explosive case. II., *Ann. Math. Statist.* **30**, 831–834 (1959).
8. M. Arato, On the parameter estimation of processes satisfying a linear stochastic differential equation, *Studia Sci. Math. Hungar* (in Russian) **5**, 11–16 (1970).
9. M.M. Rao, Asymptotic distribution of an estimator of the boundary parameter of an unstable process, *Ann. Statist.* **6**, 185–190 (1978).
10. P.C.B. Phillips, Time series regression with unit roots, *Econometrica* **55**, 277–301 (1987).
11. J. Kormos, Hypothesis testing for nearly nonstationary first order autoregressive models, *J. Appl. Math.* (in Hungarian) **14**, 27–35 (1989).
12. J. Kormos, Hypothesis testing for nearly nonstationary autoregressive models, *Computers Math. Applic.* **19**, 75–82 (1990).
13. J. Kormos, Hypothesis testing for nearly nonstationary AR(1) models with Gaussian autoregressive innovation, *Computers Math. Applic.* **21** (1), 129–136 (1991).
14. P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, (1968).
15. R.S. Liptser and A.N. Shiryayev, *Theory of Martingales*, Kluwer, Dordrecht, (1989).
16. N.H. Chan, The parameter inference for nearly nonstationary time series, *J. Amer. Statist. Assoc.* **83**, 857–862 (1988).
17. R.S. Liptser and A.N. Shiryayev, *Statistics of Random Processes II*, Springer, New York, (1978).
18. Y.A. Kutoyants, *Parameter Estimation for Stochastic Processes*, Heldermann Verlag, Berlin, (1984).
19. I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, New York, (1988).