Biometrika (2008), **95**, 1, *pp.* 257–263
© 2008 Biometrika Trust
Printed in Great Britain

doi: 10.1093/biomet/asm084
Advance Access publication 24 January 2008

Asymptotic inference for a nonstationary double AR(1) model

BY SHIQING LING AND DONG LI

Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong
maling@ust.hk malidong@ust.hk

SUMMARY

We investigate the nonstationary double AR(1) model,

$$y_t = \phi y_{t-1} + \eta_t \surd \left(\omega + \alpha y_{t-1}^2\right),$$

where $\omega > 0$, $\alpha > 0$, the η_t are independent standard normal random variables and $E \log |\phi + \eta_t \surd \alpha| \geqslant 0$. We show that the maximum likelihood estimator of (ϕ, α) is consistent and asymptotically normal. Combination of this result with that in Ling (2004) for the stationary case gives the asymptotic normality of the maximum likelihood estimator of ϕ for any ϕ in the real line, with a root-n rate of convergence. This is in contrast to the results for the classical AR(1) model, corresponding to $\alpha = 0$.

Some key words: Asymptotic distribution; Double AR(1) model; Maximum likelihood estimator.

1. INTRODUCTION

Consider the so-called first-order double autoregressive, DAR(1), model,

$$y_t = \phi y_{t-1} + \eta_t \surd \left(\omega + \alpha y_{t-1}^2\right), \tag{1}$$

where $\omega > 0$, $\alpha > 0$, $t \in \{1, 2, \ldots\}$, $\{\eta_t\}$ is a sequence of independent standard normal random variables and η_t is independent of $\{y_j : j < t\}$. Let $\mathcal{F}_t$ be the σ-field generated by $\{y_t, \ldots, y_1, y_0\}$. Model (1) is a special case of the ARMA-ARCH models in Weiss (1986) and an example of the weak ARMA models in Francq & Zakoïan (1998, 2000), but it differs from the ARCH model if $\phi \neq 0$. Real examples of DAR(p) models can be found in Weiss (1984) and Ling (2004).

The condition for weak stationarity of model (1) is that $\phi^2 + \alpha < 1$, which was proved by Guégan & Diebolt (1994) for sufficiency and by Borkovec & Klüppelberg (2001) for necessity. The condition for strict stationarity is that $E \log |\phi + \eta_t \surd \alpha| < 0$, which was proved by Borkovec & Klüppelberg (2001) for sufficiency and by Ling (2007a) for necessity. Figure 1, copied from Ling (2004), shows the stationary and nonstationary regions. Ling (2007a) obtained the necessary and sufficient condition for stationarity and ergodicity of the higher-order DAR model via a connection to the random coefficient AR model.

The asymptotic normality of the maximum likelihood estimator for model (1) was proved first by Weiss (1986) when $Ey_t^4 < \infty$; see also Tsay (1987). Ling (2004, 2007a) showed that the maximum likelihood estimator of (ϕ, ω, α) is asymptotically normal only if $\{y_t\}$ is strictly stationary. Its least absolute deviation estimator was studied by Chan & Peng (2005). In this paper, we further show that the maximum likelihood estimator of (ϕ, α) is still consistent and asymptotically normal even if $\{y_t\}$ is not strictly stationary. When $\phi = 0$, model (1) reduces to the ARCH(1) model and this case was investigated by Jensen & Rahbek (2004a, 2004b); see Remark 3.

2. MAIN RESULTS

Assume that $y_1, \ldots, y_n$ are generated by model (1). For simplicity, we assume that the initial value $y_0 = 0$. The conditional loglikelihood function, with an additive constant omitted, can be written as

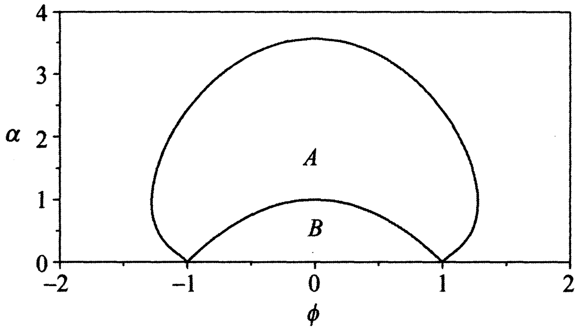

Fig. 1. Stationary and nonstationary regions for the double AR(1) model. $E\log|\phi+\eta_t\sqrt{\alpha}|<0$ for $(\phi,\alpha)\in A\cup B$ and $\phi^2+\alpha<1$ for $(\phi,\alpha)\in B$.

$$L_n(\omega,\theta)=\sum_{t=1}^{n}\ell_t(\omega,\theta),\quad \text{with } \ell_t(\omega,\theta)=-\frac{1}{2}\left\{\log\left(\omega+\alpha y_{t-1}^2\right)+\frac{(y_t-\phi y_{t-1})^2}{\omega+\alpha y_{t-1}^2}\right\}, \tag{2}$$

where $\theta=(\phi,\alpha)'$. Here $(\omega,\theta')'$ is the unknown parameter and its true value is denoted by $(\omega_0,\theta_0')'$. The maximizer, $(\hat{\omega}_n,\hat{\theta}_n')'$, of $L_n(\omega,\theta)$ is called the maximum likelihood estimator of $(\omega_0,\theta_0')'$. Ideally, we should study the asymptotic behaviour of $(\hat{\omega}_n,\hat{\theta}_n')'$. However, as $\{y_t\}$ is not stationary, Lemma A1 in the Appendix shows that $y_t^2\to\infty$ in probability as $t\to\infty$ so that ω is not identifiable in the limit of $L_n(\omega,\theta)/n$. Thus, we cannot obtain a consistent estimator of ω_0. This phenomenon was observed by Jensen & Rahbek (2004a) when they studied the nonstationary ARCH(1) model. We fix ω and study the asymptotic distribution of $\hat{\theta}_n=(\hat{\phi}_n,\hat{\alpha}_n)'$. The parameter space is given as follows.

Assumption 1. The set $\Theta=\{\theta:\alpha>\alpha_L>0\}$ for some fixed $\alpha_L>0$ is compact, θ_0 is an interior point in Θ and $\gamma\equiv E\log|\phi+\eta_t\sqrt{\alpha}|\geqslant 0$ for each $\theta\in\Theta$.

The lower bound α_L here is used because $\omega+\alpha y_{t-1}^2$ can control the loglikelihood and the score functions in such a way that they are bounded; see Lemmas A2 and A3 in the Appendix. For each fixed $\omega>0$, let $\hat{\theta}_n=\arg\max_\Theta L_n(\omega,\theta)$. The following theorem gives the asymptotic properties of the maximum likelihood estimator and its proof is given in the Appendix.

THEOREM 1. *If Assumption* 1 *holds, then, as* $n\to\infty$, *the following hold*:
(i) $\hat{\theta}_n\to\theta_0$ *in probability for any fixed* ω;
(ii) $\sqrt{n}(\hat{\theta}_n-\theta_0)\to N(0,\Omega)$, *in distribution, for any fixed* ω *when* $\gamma>0$ *and for* $\omega=\omega_0$ *when* $\gamma=0$, *where* $\Omega=\operatorname{diag}\left(\alpha_0,\,2\alpha_0^2\right)$.

Remark 1. It is surprising that the covariance matrix Ω is irrelevant to the autoregressive coefficient ϕ rather than fully determined by the coefficient in the volatility. In the usual AR(1) model, corresponding to $\alpha=0$, it is well known that the usual maximum likelihood estimator, least-squares estimator, *M*- and least absolute deviation estimators of ϕ_0 are asymptotically functions of standard Brownian motion with a rate n of convergence when $|\phi_0|=1$. When $|\phi_0|>1$, the least-squares estimator of ϕ_0 is distributed as a mixed normal with a rate ϕ_0^n of convergence; see Jeganathan (1988). With this result combined with that in Ling (2004) for the stationary case, the asymptotic normality of the maximum likelihood estimator of ϕ holds for any ϕ in the real line, with a root-n rate of convergence. Thus, the so-called unit root problem does not exist in model (1). This is an entirely new phenomenon in the field of time series. Let $u_t=\eta_t\sqrt{\left(\omega+\alpha y_{t-1}^2\right)}$. When $\phi=1$, model (1) can be written as

$$y_t=\sum_{j=1}^{t-1}u_j+\eta_t\sqrt{\left\{\omega+\alpha\left(\sum_{j=1}^{t-1}u_j\right)^2\right\}}.$$

We call this the mean-variance integration. Intuitively, Theorem 1 results from the interaction of the integration in mean and variance.

Remark 2. For the ARCH(1) model, corresponding to $\phi = 0$, Jensen & Rahbek (2004a, 2004b) showed that there exists a neighbourhood U of α_0 such that, with probability tending to 1 as $n \to \infty$, $L_n(\omega, \alpha)$ admits a unique maximizer $\hat{\alpha}_n$ in U and $\{\hat{\alpha}_n\}$ is consistent in probability and asymptotically normal. Here, we study the global maximizer $\hat{\theta}_n$ of $L_n(\omega, \theta)$ on Θ. As mentioned by a referee, Theorem 1(i) and (ii) with $\omega > 0$ should hold uniformly in $\omega \in [\omega_L, \omega_U] \subset (0, \infty)$. Similarly to Jensen & Rahbek (2004a, 2004b), we do not pursue this here since ω is not estimated.

Remark 3. A key to the proof of Theorem 1 is to use the fact that y_t from model (1) with $(\omega, \theta) = (\omega_0, \theta_0)$ has the same distribution as that of $\tilde{y}_t$, which is generated by the following random-coefficient AR model:

$$\tilde{y}_t = (\phi_0 + \tilde{\phi}_t)\tilde{y}_{t-1} + \tilde{\omega}_t, \tag{3}$$

where $(\tilde{\phi}_t, \tilde{\omega}_t)$ are independent bivariate and normal with mean 0 and covariance matrix $\mathrm{diag}(\alpha_0, \omega_0)$. This is the idea used by Ling (2007a) for the stationarity condition of the higher-order DAR model. Model (3) is a special random-coefficient AR model of Nicholls & Quinn (1982). The normality of η_t in model (1) is essential in order to use the equivalence of $\{y_t\}$ and $\{\tilde{y}_t\}$ in distribution for proving that $y_t \to \infty$ in probability as $t \to \infty$ in Lemma A1. How to relax the normality assumption remains an open problem.

Since the asymptotic property of the maximum likelihood estimator of ω_0 is not clear, we consider an alternative method to estimate ω_0. Let $\xi_t = u_t^2 - \left(\omega + \alpha y_{t-1}^2\right)$ and $x_t = u_t^2 - \alpha y_{t-1}^2$. Then $x_t = \omega + \xi_t$. Let $\hat{x}_t = (y_t - \hat{\phi}_n y_{t-1})^2 - \hat{\alpha}_n y_{t-1}^2$ and $w_t = \left[t^2 \max\left\{1, y_{t-1}^2\right\}\right]^{-1}$. Using $\hat{x}_t$ as the artificial observation of x_t, we obtain the minimizer of the sum of the self-weighted squared error, $\sum_{t=1}^n w_t(\hat{x}_t - \omega)^2$, as

$$\hat{\omega}_n = \frac{\sum_{t=1}^n w_t \hat{x}_t}{\sum_{t=1}^n w_t}.$$

This is called the self-weighted least-squares estimator of ω_0. This self-weighted estimator was proposed by Ling (2005, 2007b). The weight w_t is used to control large values of x_t such that $\hat{\omega}_n$ has a limit distribution; if we put $w_t \equiv 1$, $\hat{\omega}_n$ reduces to the ordinary least-squares estimator and we can show that it diverges to infinity in probability. By simple calculus, we obtain

$$\hat{\omega}_n - \omega_0 = (I_1 + I_2 + I_3) \Big/ \left(\sum_{t=1}^n w_t\right),$$

where

$$I_1 = 2 \sum_{t=1}^n w_t y_{t-1} \eta_t \surd\left(\omega_0 + \alpha_0 y_{t-1}^2\right)(\phi_0 - \hat{\phi}_n),$$

$$I_2 = \sum_{t=1}^n w_t y_{t-1}^2 \{(\phi_0 - \hat{\phi}_n)^2 + (\alpha_0 - \hat{\alpha}_n)\},$$

$$I_3 = \sum_{t=1}^n w_t \left(\omega_0 + \alpha_0 y_{t-1}^2\right)\left(\eta_t^2 - 1\right).$$

By Lemma A1, it is not difficult to show that $I_1 = o_p(1)$ and $I_2 = o_p(1)$. Furthermore, we can show that

$$\hat{\omega}_n - \omega_0 \to \frac{\sum_{t=1}^\infty w_t \left(\omega_0 + \alpha_0 y_{t-1}^2\right)\left(\eta_t^2 - 1\right)}{\sum_{t=1}^\infty w_t},$$

in probability. Thus, $\hat{\omega}_n$ is a self-normalized estimator of ω_0. This self-normalized phenomenon is not new here; it appears in the estimated autocorrelation of the ARCH(1) process in Davis & Mikosch (1998).

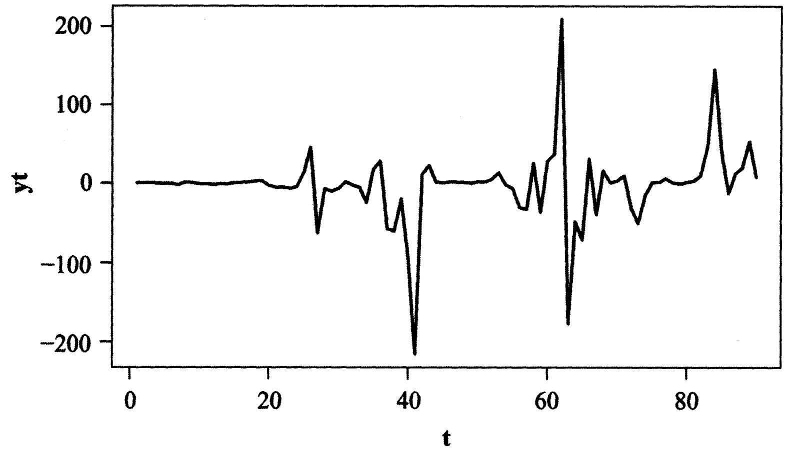

Fig. 2. A realization series $\{y_1, \ldots, y_{90}\}$ from the model $y_t = 1{\cdot}0 y_{t-1} + \eta_t \surd\left(1{\cdot}0 + 3{\cdot}0 y_{t-1}^2\right)$.

3. Simulation studies

This section examines the performance of the asymptotic results in finite samples through Monte Carlo experiments.The true observations were generated from model (1) with $\omega_0 = 1$ and $(\phi_0, \alpha_0) = (1{\cdot}0, 3{\cdot}0), (1{\cdot}0, 4{\cdot}0), (2{\cdot}0, 2{\cdot}0), (2{\cdot}0, 3{\cdot}0), (-2{\cdot}0, 3{\cdot}0)$ and $(-2{\cdot}0, 4{\cdot}0)$. Figure 2 shows one realization with $n = 90$ when $(\phi_0, \alpha_0) = (1{\cdot}0, 3{\cdot}0)$ and reveals the drastic variations of this nonstationary system. Simulations were generated for $n = 100$ and $n = 200$ and 1000 replications are used. In the likelihood function (2), we fix $\omega = 2{\cdot}0$. Table 1 summarizes the empirical means, empirical standard deviations and asymptotic

Table 1. *Simulation study. Mean and standard deviations of the maximum likelihood estimators for nonstationary* DAR *models with $\omega_0 = 1$, $y_0 = 0$ and 1000 replications*

			$n = 100$		$n = 200$	
ϕ_0	α_0		$\hat{\phi}_n$	$\hat{\alpha}_n$	$\hat{\phi}_n$	$\hat{\alpha}_n$
1·0	3·0	Mean	1·0035	2·9420	1·0040	2·9798
		SD	0·1811	0·4569	0·1265	0·3191
		AD	0·1732	0·4243	0·1225	0·3000
1·0	4·0	Mean	1·0094	3·9453	1·0079	3·9850
		SD	0·2049	0·5831	0·1438	0·4141
		AD	0·2000	0·5657	0·1414	0·4000
2·0	2·0	Mean	2·0039	1·9788	2·0045	1·9943
		SD	0·1490	0·2838	0·1026	0·2031
		AD	0·1414	0·2828	0·1000	0·2000
2·0	3·0	Mean	2·0103	2·9685	2·0089	2·9907
		SD	0·1752	0·4258	0·1242	0·3048
		AD	0·1732	0·4243	0·1225	0·3000
−2·0	3·0	Mean	−1·9939	2·9584	−1·9951	2·9863
		SD	0·1696	0·4228	0·1204	0·3047
		AD	0·1732	0·4243	0·1225	0·3000
−2·0	4·0	Mean	−1·9990	3·9448	−1·9999	3·9814
		SD	0·2002	0·5626	0·1437	0·4061
		AD	0·2000	0·5657	0·1414	0·4000

SD, standard deviation; AD, asymptotic standard deviation.

standard deviations of the maximum likelihood estimators of (ϕ_0, α_0). The asymptotic standard deviations are obtained from the asymptotic covariance matrix Ω in Theorem 1. Table 1 shows that all biases are very small. The standard deviations and asymptotic standard deviations are very similar. When α_0 is large, all standard deviations and asymptotic standard deviations become large. In general, the results indicate that the maximum likelihood estimators perform very well in the finite samples.

We also carried out experiments when the fixed ω was 4 and 10, obtaining results that were almost the same as those in Table 1. When we estimate all the parameters $(\omega_0, \phi_0, \alpha_0)$, all the estimators of (ϕ_0, α_0) are also almost the same as those in Table 1. The estimator of ω_0 always has a bias and the empirical standard deviation is quite large, particularly when α_0 is large. The self-weighted least-squares estimator of ω_0 is also biased, but it has a smaller bias and a smaller empirical standard deviation than its maximum likelihood estimator. A future project of interest is to seek a consistent or unbiased estimator of ω_0.

Acknowledgement

We thank two referees and Professor D. M. Titterington for their very helpful comments, and thank the Hong Kong Research Grants Council for partial support.

Appendix

Technical details

We first give one lemma, the proof of which is given in a longer version of this paper, available from the authors. It plays a key role in the proof of the other lemmas.

Lemma A1. (i) *If* $\gamma \geqslant 0$, *then* $|y_t| \to \infty$ *in probability when* $t \to \infty$ *and* (ii) *if* $\gamma > 0$, *then there exists an* $s_0 \in (0, 1)$ *such that* $E|y_t|^{-s_0} = O(\rho^t)$ *with* $\rho \in (0, 1)$.

We now give the first and the second partial derivatives of $\ell_t(\omega, \theta)$ as follows:

$$\frac{\partial \ell_t(\omega, \theta)}{\partial \phi} = \frac{y_{t-1}(y_t - \phi y_{t-1})}{\omega + \alpha y_{t-1}^2},$$

$$\frac{\partial \ell_t(\omega, \theta)}{\partial \alpha} = \frac{y_{t-1}^2}{2(\omega + \alpha y_{t-1}^2)} \left\{ \frac{(y_t - \phi y_{t-1})^2}{\omega + \alpha y_{t-1}^2} - 1 \right\},$$

$$\frac{\partial^2 \ell_t(\omega, \theta)}{\partial \phi^2} = -\frac{y_{t-1}^2}{\omega + \alpha y_{t-1}^2},$$

$$\frac{\partial^2 \ell_t(\omega, \theta)}{\partial \phi \partial \alpha} = -\frac{y_{t-1}^3(y_t - \phi y_{t-1})}{(\omega + \alpha y_{t-1}^2)^2},$$

$$\frac{\partial^2 \ell_t(\omega, \theta)}{\partial \alpha^2} = -\frac{y_{t-1}^4}{2(\omega + \alpha y_{t-1}^2)^2} \left\{ \frac{2(y_t - \phi y_{t-1})^2}{\omega + \alpha y_{t-1}^2} - 1 \right\}.$$

Lemma A2. *If Assumption* 1 *holds, then we have the following results*:

(i) $\sup_{\theta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial \ell_t(\omega, \theta)}{\partial \theta} \right\| = O_p(1)$;

(ii) $\sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^3 \ell_t(\omega, \theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| = O_p(1)$, *where* $i, j, k = 1, 2$, $\theta_1 = \phi$ *and* $\theta_2 = \alpha$.

Proof. Since Θ is compact, there exists a constant M such that $|\phi| \leqslant M$. Thus,

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial \ell_t(\omega,\theta)}{\partial\phi}\right| \leqslant \frac{1}{\alpha_L n}\sum_{t=1}^{n}|\eta_t|\surd\left(\frac{\alpha_0}{\alpha_L}+\frac{\omega_0}{\omega+\alpha_L y_{t-1}^2}\right)+2M\alpha_L^{-1}=O_p(1),$$

by Lemma A1(i). Similarly, $\sup_{\theta\in\Theta}|n^{-1}\partial L_n(\omega,\theta)/\partial\alpha| = O_p(1)$ and (ii) holds. □

LEMMA A3. *Let* $f(\theta)=\{\log(\alpha_0/\alpha)+1-\alpha_0/\alpha-(\phi-\phi_0)^2/\alpha\}/2$. *If Assumption* 1 *holds, it follows that* $\sup_{\theta\in\Theta}|\sum_{t=1}^{n}\{\ell_t(\omega,\theta)-\ell_t(\omega_0,\theta_0)-f(\theta)\}|/n=o_p(1)$.

Proof. First, since $\left(c+y_{t-1}^2\right)^{-1}\leqslant c^{-1}$, by Lemma A1(i) and the dominated convergence theorem, we have that $E\left(c+y_{t-1}^2\right)^{-1}\to 0$ as $t\to\infty$ for any fixed constant $c>0$. Using this, we can show that, for each $\theta\in\Theta$,

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\omega,\theta)-\ell_t(\omega_0,\theta_0)\} &= \frac{1}{2n}\sum_{t=1}^{n}\left\{\log\frac{\omega_0+\alpha_0 y_{t-1}^2}{\omega+\alpha y_{t-1}^2}+\eta_t^2-\eta_t^2\frac{\omega_0+\alpha_0 y_{t-1}^2}{\omega+\alpha y_{t-1}^2}\right.\\ &\qquad\left.-\frac{(\phi_0-\phi)^2 y_{t-1}^2}{\omega+\alpha y_{t-1}^2}-\frac{2(\phi_0-\phi)\eta_t y_{t-1}\surd\left(\omega_0+\alpha_0 y_{t-1}^2\right)}{\omega+\alpha y_{t-1}^2}\right\}\\ &= \frac{1}{2}\left\{\log\frac{\alpha_0}{\alpha}+1-\frac{\alpha_0}{\alpha}-\frac{(\phi-\phi_0)^2}{\alpha}\right\}+o_p(1)=f(\theta)+o_p(1). \quad \text{(A1)}\end{aligned}$$

The remainder of the proof, available in the longer version of the paper, amounts to showing that this convergence holds uniformly on Θ. □

Proof of Theorem 1(i). Since $f(\theta)$ has a unique maximizer at $\theta=\theta_0$ on the compact set Θ, for any $\varepsilon>0$, we have $c\equiv\sup_{\|\theta-\theta_0\|\geqslant\varepsilon}f(\theta)\in(-\infty,0)$. Thus,

$$\begin{aligned}\mathrm{pr}(\|\hat\theta_n-\theta_0\|\geqslant\varepsilon) &= \mathrm{pr}\left(\|\hat\theta_n-\theta_0\|\geqslant\varepsilon,\ \frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\omega,\hat\theta_n)-\ell_t(\omega_0,\theta_0)\}\geqslant 0\right)\\ &\leqslant \mathrm{pr}\left(\sup_{\|\theta-\theta_0\|\geqslant\varepsilon}\frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\omega,\theta)-\ell_t(\omega_0,\theta_0)-f(\theta)\}+c\geqslant 0\right)\\ &\leqslant \mathrm{pr}\left(\sup_{\Theta}\frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\omega,\theta)-\ell_t(\omega_0,\theta_0)-f(\theta)\}\geqslant -c\right)\\ &\leqslant \mathrm{pr}\left(\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\{\ell_t(\omega,\theta)-\ell_t(\omega_0,\theta_0)-f(\theta)\}\right|>-c\right)\to 0,\end{aligned}$$

by Lemma A3, as $n\to\infty$. □

LEMMA A4. *If Assumption* 1 *is satisfied then the following hold*:
(i) $n^{-1/2}\sum_{t=1}^{n}\partial\ell_t(\omega,\theta_0)/\partial\theta\to N(0,\Omega^{-1})$, *in distribution, for any fixed* ω *when* $\gamma>0$ *and for* $\omega=\omega_0$ *when* $\gamma=0$;
(ii) $n^{-1}\sum_{t=1}^{n}\partial^2\ell_t(\omega,\theta_0)/\partial\theta\partial\theta' = -\Omega^{-1}+o_p(1)$ *for any fixed* ω.

Proof. As for (A1), by Lemma A1(i), we can show that

$$\frac{1}{n}\sum_{t=1}^{n}E\left[\left\{\frac{\partial\ell_t(\omega_0,\theta_0)}{\partial\phi}\right\}^2\middle|\mathcal{F}_{t-1}\right]=\frac{1}{n}\sum_{t=1}^{n}\frac{y_{t-1}^2\left(\omega_0+\alpha_0 y_{t-1}^2\right)}{\left(\omega_0+\alpha_0 y_{t-1}^2\right)^2}=\frac{1}{\alpha_0}+o_p(1). \quad \text{(A2)}$$

Similarly, we can show that $n^{-1}\sum_{t=1}^{n} E[\{\partial \ell_t(\omega_0,\theta_0)/\partial\alpha\}^2|\mathcal{F}_{t-1}] = (2\alpha_0^2)^{-1} + o_p(1)$ and $n^{-1}\sum_{t=1}^{n} E[\{\partial \ell_t(\omega_0,\theta_0)/\partial\phi\}\{\partial \ell_t(\omega_0,\theta_0)/\partial\alpha\}|\mathcal{F}_{t-1}] = o_p(1)$. Thus,

$$n^{-1}\sum_{t=1}^{n} E[\{\partial \ell_t(\omega_0,\theta_0)/\partial\theta\}\{\partial \ell_t(\omega_0,\theta_0)/\partial\theta\}'|\mathcal{F}_{t-1}] = \Omega^{-1} + o_p(1).$$

By the martingale central limit theorem in Brown (1971), we can show that

$$n^{-1/2}\sum_{t=1}^{n} \partial \ell_t(\omega_0,\theta_0)/\partial\theta \to N(0,\Omega^{-1})$$

in distribution when $\gamma \geqslant 0$. For any fixed $\omega > 0$, by Lemma A1(ii), we can show that $n^{-1/2}\sum_{t=1}^{n} \|\partial \ell_t(\omega,\theta_0)/\partial\theta - \partial \ell_t(\omega_0,\theta_0)/\partial\theta\| \leqslant O(1)n^{-1/2}\sum_{t=1}^{n}(1+\eta_t^2)(|y_t|^{-1}+y_t^{-2}) = o_p(1)$ when $\gamma > 0$. Thus, (i) holds. Similarly to (A2), we can show that (ii) holds. □

Proof of Theorem 1(ii). By Taylor's expansion and Lemma A2(ii), we have

$$0 = \frac{\partial L_n(\omega,\hat\theta_n)}{\partial\theta} = \frac{\partial L_n(\omega,\theta_0)}{\partial\theta} + \left\{\frac{\partial^2 L_n(\omega,\theta_0)}{\partial\theta\partial\theta'} + n\|\hat\theta_n - \theta_0\|O_p(1)\right\}(\hat\theta_n - \theta_0).$$

By part (i) of this theorem and Lemma A4, the conclusion holds. □

References

Borkovec, M. & Klüppelberg, C. (2001). The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Ann. Appl. Prob.* **11**, 1220–41.

Brown, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42**, 59–66.

Chan, N. H. & Peng, L. (2005). Weighted least absolute deviation estimation for an AR(1) process with ARCH(1) errors. *Biometrika* **92**, 477–84.

Davis, R. A. & Mikosch, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26**, 2049–80.

Francq, C. & Zakoïan, J. M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plan. Infer.* **68**, 145–65.

Francq, C. & Zakoïan, J. M. (2000). Covariance matrix estimation for estimators of mixing weak ARMA models. *J. Statist. Plan. Infer.* **83**, 369–94.

Guégan, D. & Diebolt, J. (1994). Probabilistic properties of the β-ARCH-model. *Statist. Sinica* **4**, 71–87.

Jeganathan, P. (1988). On the strong approximation of the distributions of estimators in linear stochastic models: I, II. Stationary and explosive AR models. *Ann. Statist.* **16**, 1283–314.

Jensen, S. T. & Rahbek, A. (2004a). Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica* **72**, 641–6.

Jensen, S. T. & Rahbek, A. (2004b). Asymptotic inference for nonstationary GARCH. *Economet. Theory* **20**, 1203–26.

Ling, S. (2004). Estimation and testing stationarity for double autoregressive models. *J. R. Statist. Soc.* B **66**, 63–78.

Ling, S. (2005). Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *J. R. Statist. Soc.* B **67**, 381–93.

Ling, S. (2007a). A double AR(p) model: structure and estimation. *Statist. Sinica* **17**, 161–75.

Ling, S. (2007b). Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. *J. Economet.* **140**, 849–73.

Nicholls, D. F. & Quinn, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction*. Lecture Notes in Statistics, **11**. New York: Springer-Verlag.

Tsay, R. S. (1987). Conditional heteroscedastic time series models, *J. Am. Statist. Assoc.* **82**, 590–604.

Weiss, A. A. (1984). ARMA models with ARCH errors. *J. Time Ser. Anal.* **5**, 129–43.

Weiss, A. A. (1986). Asymptotic theory for ARCH models: estimation and testing. *Economet. Theory* **2**, 107–31.

[*Received June* 2006. *Revised June* 2007]