Journal of the Korean
Statistical Society
Vol. 26, No. 4, 1997

The Asymptotic Variance of the Studentized Residual Autocorrelations for a Generalized Random Coefficient Autoregressive Processes

Sangwoo Park, Sinsup Cho[1] and Sun Y. Hwang [2]

Abstract

The asymptotic distribution of residual autocorrelation functions from a generalized p-order random coefficient autoregressive process (GRCA(p)) is derived. To this end, we first describe the GRCA(p) models and then consider the normalised residuals after fitting the model. This result can be applied to the residual analysis for the diagonostic purpose.

Key Words : Asymptotic variance; GRCA(p) model, Studentized residual; Residual autocorrelation.

[1]Department of Statistics, Seoul National University, Seoul, 151-742, Korea.
[2]Department of Statistics, Sookmyung Women's University, Seoul, 140-742, Korea.

1. INTRODUCTION

From the time series modeling point of view, the autocorrelation function (ACF) of the residual is worth investigation for a diagnostic purpose. For the ARMA processes, Box and Pierce (1970) and Durbin (1970) have shown that the asymptotic variances of residuals obtained after fitting a ARMA processes are different from the ACF's based on white noise, especially for the first few lags, which need careful check for constructing the Portmanteau type lack of fit test statistic.

Within the context of nonlinear time series, Tong (1990, p.p.324) proposed the conservative use of $1/n$ for the variances of the residual ACF's. Li (1992) discussed the exact limiting variances of the residual ACF's for the fixed coefficient nonlinear autoregressive processes. Hwang et.al. (1994) obtained the similar results to Li for the random coefficient processes.

In this paper, we discuss the joint asymptotic distribution of the residual ACF's from a generalized p-order random coefficient autoregressive process (GRCA(p)) which has been introduced and studied by Hwang and Basawa (1996). This model is seen to be rich enough to cover the standard random coefficient processes of Nicholls and Quinn (1982), and the Markov bilinear models of Tong (1981) and their generalizations. Although we shall follow the similar approaches as in Li (1992) and Hwang et.al. (1994), there are some technical differences since GRCA(p) never belongs to the class of models treated in both references cited above. Furthermore, we elaborate a little on the residual ACF problems in nonlinear time series. Contrary to the usual pratice in fixed coefficient nonlinear time series context that the residuals are defined as the nonscaled (estimated) prediction errors, we here use the scaled version of residuals (which will be referred to as studentized residual) which can be justified by the fact that the conditional variance of the prediction errors given the past data is heterogenous.

This article is organized as follows. In section 2, GRCA(p) and some preliminary results which will be used throughout the paper are presented. The limiting distributions of the studentized residual ACF's, the main result of the paper, are given in section 3.

2. MODEL AND THE PRELIMINARY RESULTS

Let $\{X_t,\ t = 0, \pm 1, \cdots\}$ be a generalized random coefficient process of order p (GRCA(p)) satisfying the difference equation

$$X_t = \Phi_t' X(t-1) + \epsilon_t \tag{2.1}$$

where $\Phi_t = (\Phi_{t1}, \cdots, \Phi_{tp})'$ is a $(p \times 1)$ vector of random coefficients, $X(t-1) = (X_{t-1}, \cdots, X_{t-p})'$ is a $(p \times 1)$ vector of the recent past observations at time t and $\{\epsilon_t\}$ is a sequence of random errors. It will be assumed that $\left\{\begin{pmatrix} \Phi_t \\ \epsilon_t \end{pmatrix}\right\}$ is a sequence of *i.i.d.* $(p+1) \times 1$ random vectors with

$$E\begin{pmatrix} \Phi_t \\ \epsilon_t \end{pmatrix} = \begin{pmatrix} \phi \\ 0 \end{pmatrix} \quad \text{and} \quad V = var\begin{pmatrix} \Phi_t \\ \epsilon_t \end{pmatrix} = \begin{pmatrix} \Sigma_\phi & \sigma_{\phi\epsilon} \\ \sigma'_{\phi\epsilon} & {\sigma_\epsilon}^2 \end{pmatrix} \tag{2.2}$$

where $\phi = (\phi_1, \cdots, \phi_p)'$ is a $(p \times 1)$ parameter vector representing the mean level of the random coefficient Φ_t and $\Sigma_\phi = var(\Phi_t) : p \times p$ matrix $\sigma_{\phi\epsilon} = cov(\Phi_t, \epsilon_t) : p \times 1$ vector and ${\sigma_\epsilon}^2 = var(\epsilon_t)$. The model specified by (2.1) and (2.2) can be viewed as a generalized version of random coefficient autoregressive process of Nicholls and Quinn (1982) in the following sense :

(i) Although random coefficient is typically of additive form, namely, $\Phi_t = \phi + Z_t$ where Z_t stands for the random perturbation with zero mean and variance ${\sigma_Z}^2$, we do not assume any specific structure of Φ_t, allowing multiplicative form.

(ii) Random coefficient Φ_t is permitted to be correlated with error ϵ_t.

It will be assumed throughout that

(C.0) $\{X_t\}$ is stationary and ergodic time series and $E{X_t}^8 < \infty$.

See Hwang and Basawa (1996) for the various examples therein as special cases of GRCA(p) and for the sufficient conditions for (C.0).

Let $\mu_{t|t-1}$ and $v^2{}_{t|t-1}$ denote the conditional mean and variance of X_t given the past

$$\begin{aligned} \mu_{t|t-1} &= E(X_t|X(t-1)) \\ v^2{}_{t|t-1} &= Var(X_t|X(t-1)) \end{aligned}$$

It can then be easily seen that

$$\begin{aligned} \mu_{t|t-1} &= \phi' X(t-1) \\ v^2{}_{t|t-1} &= X'(t-1)\Sigma_\phi X(t-1) + 2X'(t-1)\sigma_{\phi\epsilon} + {\sigma_\epsilon}^2 \end{aligned} \tag{2.3}$$

For simplicity, we shall use the notation $\mu_t = \mu_{t|t-1}$ and $v_t = v_{t|t-1}$. It may be noticed that μ_t and v_t are $X(t-1)$-measurable.

Let $\beta = \text{vech}(V)$: $((p+1)\times(p+2)/2)\times 1$ vector of parameters. From (2.3), $v_t = v_{t|t-1}$ can be rewritten as in terms of a linear combination of β

$$v_t{}^2 = l_t{}'\beta \tag{2.4}$$

where l_t is an appropriate $((p+1)\times(p+2)/2)\times 1$ random vector which is a function of $X(t-1)$ only, and is free from ϕ and β.

Denote by $\alpha_t(\phi)$ the prediction errors

$$\begin{aligned}\alpha_t(\phi) &= X_t - \mu_t \\ &= X_t - \phi' X(t-1)\end{aligned} \tag{2.5}$$

Based on the given sample $\{X_{-p+1}, \cdots, X_0, X_1, \cdots, X_n\}$, the conditional least squares(CLS, c.f. Klimko and Nelson (1978)) estimators $\hat{\phi}$ and $\hat{\beta}$ of ϕ and β respectively and their limiting results are investigated by Hwang and Basawa (1996).

One can obtain $\hat{\phi}$ and $\hat{\beta}$ by minimizing

$$Q_1 = \sum_{t=1}^{n} \alpha_t{}^2(\phi)$$

and

$$Q_2 = \sum_{t=1}^{n} [\alpha_t{}^2(\phi) - v_t{}^2]^2$$

respectively. Under the condition (C.0), they are seen to be strongly consistent and asymptotically normally distributed(see, Hwang and Basawa(1996)).

3. THE ASYMPTOTIC DISTRIBUTION OF THE STUDENTIZED RESIDUAL AUTOCORRELATIONS

Consider the normalized prediction errors

$$U_t(\phi, \beta) = (X_t - \mu_t)/v_t \tag{3.1}$$

where μ_t and v_t are defined in (2.3). It is obvious that $\{U_t(\phi, \beta)\}$ is a sequence of zero mean martingale differences with unit variance.

Define the "estimated" version of $U_t(\phi, \beta)$

$$\begin{aligned} W_t &= U_t(\hat{\phi}, \hat{\beta}) \\ &= [X_t - \hat{\phi}' X(t-1)]/\sqrt{l_t'\hat{\beta}} \end{aligned} \tag{3.2}$$

which will be referred to as the studentized residual.

The sample ACF of lag i, denoted by $\hat{\rho}(i)$, is given by

$$\hat{\rho}(i) = \sum_{t=1}^{n-i}(W_t - \overline{W})(W_{t+i} - \overline{W}) / \sum_{t=1}^{n}(W_t - \overline{W})^2 \tag{3.3}$$

where $\overline{W} = n^{-1}\sum_{t=1}^{n} W_t$.

It can be shown that $\overline{W} \xrightarrow{a.s.} 0$ and $n^{-1}\sum_{t=1}^{n}(W_t - \overline{W})^2 \xrightarrow{a.s.} 1$. Consequently,

$$\begin{aligned} \sqrt{n}\hat{\rho}(i) &= n^{-1/2}\sum_{t=1}^{n-i}(W_t - \overline{W})(W_{t+i} - \overline{W})/n^{-1}\sum_{t=1}^{n}(W_t - \overline{W})^2 \\ &= n^{-1/2}\sum_{t=1}^{n-i} W_t W_{t+i} + o_p(1) \end{aligned} \tag{3.4}$$

Our main goal in this section is to derive the limiting distribution of $(\hat{\rho}(1), \cdots, \hat{\rho}(k))$ for some fixed integer k. Since the denominator $\sqrt{l_t'\hat{\beta}}$ in (3.2) appears to be difficult to handle, we go with the much more tractable term e_t (rather than W_t)

$$\begin{aligned} e_t &= [X_t - \hat{\phi}' X(t-1)]/\sqrt{l_t'\beta} \\ &= \alpha_t(\hat{\phi})/\sqrt{l_t'\beta} \\ &= \sqrt{l_t'\hat{\beta}/l_t'\beta} \cdot W_t \end{aligned} \tag{3.5}$$

We first deal with e_t and investigate the limiting distribution of ACF based on $\{e_t\}$. To this end, consider

$$\tilde{\rho}(i) = n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i}, \quad i = 1, \cdots, k$$

The joint limiting distribution of $[\tilde{\rho}(1), \cdots, \tilde{\rho}(k)]$ is presented in the following theorem.

Theorem 1. Under (C.0), we have

$$\sqrt{n}[\tilde{\rho}(1), \cdots, \tilde{\rho}(k)] \xrightarrow{d} N(0, F)$$

where F is the $(k \times k)$ matrix with (i, j)th element given by

$$\begin{aligned} F_{ij} = E[\ \alpha_t^2(\phi) \ & \{\alpha_{t-i}(\phi)/(v_{t-i}v_t) - X'(t-1)\Gamma^{-1}\triangle_i\} \\ \cdot \ & \{\alpha_{t-j}(\phi)/(v_{t-j}v_t) - X'(t-1)\Gamma^{-1}\triangle_j\}] \end{aligned}$$

where $\triangle_i, \triangle_j$ and Γ will be specified in the course of the proof and v_t and $\alpha_t(\phi)$ are defined as in (2.3) and (2.4).

Proof. For fixed $i = 1, \cdots, k$, consider $n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i}$. Split this into

$$n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i} = A - B$$

$$A = n^{-1/2}\sum_{t=1}^{n-i} \alpha_t(\phi)\{\alpha_{t+i}(\phi) - (\hat{\phi} - \phi)'X(t+i-1)\}/v_t v_{t+i}$$

and

$$B = n^{-1/2}\sum_{t=1}^{n-i}(\hat{\phi} - \phi)'X(t-1)\{\alpha_{t+i}(\phi) - (\hat{\phi} - \phi)'X(t+i-1)\}/v_t v_{t+i}$$

where $X(t+i-1)$ and $\alpha_{t+i}(\phi)$ are analogous to the previous notation, i.e.

$$\begin{aligned} X(t+i-1) &= (X_{t+i-1}, \cdots, X_{t+i-p})' \quad : p \times 1 \text{ vector} \\ \alpha_{t+i}(\phi) &= X_{t+i} - \phi' X(t+i-1) \end{aligned}$$

Using the fact that $\sqrt{n}(\hat{\phi} - \phi)$ is bounded in probability and taking into account that $\{X(t-1)\alpha_{t+i}(\phi)\}$ is a zero mean martingale differences, it can be shown that $B \xrightarrow{a.s.} 0$ via the ergodic theorem

Consequently,

$$n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i} = A + o_p(1) \tag{3.6}$$

We now turn to A which can be rewritten as

$$\begin{aligned} A \ = \ & n^{-1/2}\sum_{t=1}^{n-i} \alpha_t(\phi)\alpha_{t+i}(\phi)/v_t v_{t+i} \\ & - n^{1/2}(\hat{\phi} - \phi)' n^{-1}\sum_{t=1}^{n-i} \alpha_t(\phi) X(t+i-1)/v_t v_{t+i} \end{aligned} \tag{3.7}$$

Notice that $\hat{\phi}$ is obtained by minimizing $Q_1 = \sum_{t=1}^{n}(X_t - \phi' X(t-1))^2$ and hence

$$\hat{\phi} = [\sum_{t=1}^{n} X(t-1)X'(t-1)]^{-1} \sum_{t=1}^{n} X_t X(t-1)$$

which yields with a little more effort

$$\hat{\phi} - \phi = [\sum_{t=1}^{n} X(t-1)X'(t-1)]^{-1} \sum_{t=1}^{n} \alpha_t(\phi) X(t-1) \tag{3.8}$$

which in turn implies

$$\sqrt{n}(\hat{\phi} - \phi) = n^{-1/2}\Gamma^{-1} \sum_{t=1}^{n} \alpha_t(\phi) X(t-1) + o_p(1) \tag{3.9}$$

where $\Gamma = E(X(t-1)X'(t-1))$

Moreover, the ergodic theorem applied to the second term in the RHS of (3.7) gives the limit, say Δ_i, such that

$$\begin{aligned} \Delta_i &= \text{plim}\{n^{-1} \sum_{t=1}^{n-i} \alpha_t(\phi) X(t+i-1)/v_t v_{t+i}\} \\ &= E[\alpha_t(\phi) X(t+i-1)/v_t v_{t+i}] \end{aligned} \tag{3.10}$$

Collecting (3.7),(3.8) and (3.10), it follows that

$$A = n^{-1/2} \sum_{t=1}^{n-i} \alpha_{t+i}(\phi)\{\alpha_t(\phi)/v_t v_{t+i} - X'(t+i-1)\Gamma^{-1}\Delta_i\} + o_p(1) \tag{3.11}$$

Hence, appealing to CLT for zero mean martingale differences (Billingsley(1961)), we obtain

$$A \xrightarrow{d} N(0, F_{ii})$$

Equivalently,

$$n^{-1/2} \sum_{t=1}^{n-i} e_t e_{t+i} \xrightarrow{d} N(0, F_{ii}) \quad , i = 1, \cdots, k \tag{3.12}$$

To complete the proof, we need to derive the joint limiting distribution of $\tilde{\rho}(1), \cdots, \tilde{\rho}(k)$. Proceeding along basically the same line as above, using the Cramer-Wold device, it can be verified that for a given $(k \times 1)$ vector of constant a,

$$a'\sqrt{n}[\tilde{\rho}(1), \cdots, \tilde{\rho}(k)] \xrightarrow{d} N(0, a'Fa)$$

which in turn implies the result of the theorem.

Turning our attention to $\hat{\rho}(i)$ in (3.3), we conjecture that $\hat{\rho}(i)$, the ACF based on the studentized residual W_t is asymptotically equivalent to $\tilde{\rho}(i)$ up to terms negligible in probability, viz.,

$$n^{-1/2}\sum_{t=1}^{n-i} W_t W_{t+i} - n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i} = o_p(1) \tag{3.13}$$

In order to verify (3.13), we impose the following regularity condition on V.

(C.1) There exist $\eta > 0$ such that

$$\sigma^2_{\epsilon} - \sigma'_{\varphi\epsilon}\Sigma_{\varphi}{}^{-1}\sigma_{\varphi\epsilon} > \eta > 0 \tag{3.14}$$

Remark : (C.1) guarantees that the conditional variance of X_t given $X(t-1)$ is bounded away from zero almost surely since

$$\begin{aligned} Var(X_t|X(t-1)) &= l_t'\beta \\ &= (X(t-1)+\Sigma_{\varphi}{}^{-1}\sigma_{\varphi\epsilon})'\Sigma_{\varphi}(X(t-1)+\Sigma_{\varphi}{}^{-1}\sigma_{\varphi\epsilon}) \\ &\quad +(\sigma^2_{\epsilon} - \sigma'_{\varphi\epsilon}\Sigma_{\varphi}{}^{-1}\sigma_{\varphi\epsilon}) \\ &\geq \sigma^2_{\epsilon} - \sigma'_{\varphi\epsilon}\Sigma_{\varphi}{}^{-1}\sigma_{\varphi\epsilon} > \eta > 0 \end{aligned}$$

For the special case where $\begin{pmatrix} \Phi_t \\ \epsilon_t \end{pmatrix}$ is Gaussian, (C.1) is equivalent to $Var(\epsilon_t|\Phi_t) > \eta > 0$

Refer to the equation (3.5). We obtain

$$n^{-1/2}\sum_{t=1}^{n-i} W_t W_{t+i} = n^{-1/2}\sum_{t=1}^{n-i} c_{n,t}c_{n,t+i}e_t e_{t+i}$$

where

$$c_{n,t} = \sqrt{l_t'\beta / l_t'\hat{\beta}} \tag{3.15}$$

Following technical lemma will be needed for establishing the asymptotic distribution of $[\hat{\rho}(1), \cdots, \hat{\rho}(k)]$.

Lemma 1. Under (C.0) and (C.1), we have, for any $0 \leq \delta \leq 1/4$,

$$\sup_{1\leq t\leq n-i} |c_{n,t}c_{n,t+i} - 1| = o_p(n^{-\delta}) \tag{3.16}$$

Proof. It may be noted that it suffices to show

$$\sup_{1\le t\le n} |c_{n,t} - 1| = o_p(n^{-\delta}) \tag{3.17}$$

Consider

$$|c_{n,t} - 1| = |l_t'(\beta - \hat\beta)/l_t'\hat\beta|$$

Using the strong consistency of $\hat\beta$ for β and (C.1), it follows that for all sufficiently large n,

$$|c_{n,t} - 1| \le \eta^{-1}|l_t'(\beta - \hat\beta)|$$

Let $\epsilon > 0$ be given, and fix $0 \le \delta \le 1/4$. We then have

$$\begin{aligned} &P\{\sup_{1\le t\le n} |l_t'(\beta - \hat\beta)| > n^{-\delta}\epsilon\} \\ \le\ & nP\{|l_1'\sqrt{n}(\beta - \hat\beta)| > n^{1/2-\delta}\epsilon\} \end{aligned}$$

which goes to zero as $n \to \infty$ due to the fact that $\sqrt{n}(\beta - \hat\beta)$ is bounded in probability and the existence of $E\|l_1\|^4 < \infty$ (which is implied by the moment condition $EX_t^8 < \infty$ in (C.0)). Consequently (3.16) holds, which completes the proof.

We are now in a position to present the following theorem concerning the asymptotic distribution of ACF's from the studentized residuals.

Theorem 2. Under (C.0) and (C.1), we have, for fixed k,

$$\sqrt{n}[\hat\rho(1), \cdots, \hat\rho(k)] \xrightarrow{d} N(0, F)$$

where F is defined in Theorem 1.

Proof. First notice that

$$\begin{aligned} & n^{-1/2}\sum_{t=1}^{n-i} W_t W_{t+i} - n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i} \\ =\ & n^{-1/2}\sum_{t=1}^{n-i}(c_{n,t}c_{n,t+i} - 1)e_t e_{t+i} \end{aligned} \tag{3.18}$$

Picking up $\delta = 1/4$ in the preceding lemma, it follows that

$$n^{1/4} \sup_{1\le t\le n-i} |c_{n,t}c_{n,t+i} - 1| = o_p(1) \tag{3.19}$$

Combining (3.18) and (3.19) together with Theorem1, we obtain

$$n^{-1/2}\sum_{t=1}^{n-i} W_t W_{t+i} - n^{-1/2}\sum_{t=1}^{n-i} e_t e_{t+i} = o_p(1)$$

which essentially concludes the proof.

As an application of Theorem 2, one can construct the statistic T_n for the lack of fit test

$$T_n = \psi_n' \hat{F}^- \psi_n \tag{3.20}$$

where $\hat{F}^-$ is a consistent estimator of a generalized inverse of F and $\psi_n = \sqrt{n}[\hat{\rho}(1), \cdots, \hat{\rho}(k)]'$. It can then be easily seen that under H_0

$$T_n \xrightarrow{d} \chi_r^2$$

with r=rank(F) and hence the rejection region for the lack of fit test is given by $\{T_n \geq \chi_r^2(\alpha)\}$. It is worth indicating that the computation of T_n is not feasible in practice since F involves moments of stationary distribution up to 8-th order of which the closed forms have not yet been identified.

ACKNOWLEDGEMENT

We thank the referee for a careful reading of the paper.

REFERENCES

(1) Billingsley, P. (1961). The Lindeburg-Levy theorem for martingale. *Proceeding of American Mathematical Society*, **12**, 788-792.

(2) Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series model. *Journal of American Statistical Association*, **65**, 1509-1526.

(3) Brockwell, P.D. and Davis, R.A. (1991). *Time Series : Theory and Methods*, 2nd ed., Springer Verlag, New York.

(4) Durbin, J. (1970). Testing for serial correlation in least-squares regression when some of the regressor are lagged dependent variables. *Econometrica*, **38**, 410-421.

(5) Hwang, S.Y., Basawa, I.V., and Reeves, J. (1994). The asymptotic distributions of residual autocorrelations and related tests of fit for a class of nonlinear time series models. *Statistica Sinica*, **4**, 107-125.

(6) Hwang, S.Y. and Basawa, I.V. (1996). Parameter estimation for generalized random coefficient autoregressive processes. To appear in *Journal of Statistical Planning and Inference.*

(7) Klimko, L.A. and Nelson, P.L. (1978). On conditional least squares estimation for stochastic processes. *Annals of Statistics*, **6**, 629-642.

(8) Li, W.K. (1992). On the asymptotic standard errors of residual autocorrelations in nonlinear time series modeling. *Biometrika*, **79**, 435-437.

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

(10) Tong, H. (1981). A note on a Markov bilinear stochastic process in discrete time. *Journal of Time Series Analysis*, **2**, 279-284.

(11) Tong, H. (1990). *Non-linear Time Series.* Oxford University Press, Oxford.