ESTIMATION AND TESTING FOR ARCH MODELS

S. Ajay Chandra

Received February 14, 2003; revised March 7, 2003

Abstract. This paper considers the problem of estimation and testing for ARCH models under the assumption of conditional correlation. For a bivariate model with unknown volatility parameter vector, we construct an estimator for this parameter vector using the conditional least squares estimator given by Tjøstheim (1986). Such an estimation procedure is applied to more general ARCH models. Next we turn to discuss the testing problem for the multivariate model in the setting of Taniguchi and Kakizawa (2000), based on a quasi-Gaussian maximum likelihood estimator. The results show that the tests such as Gaussian likelihood ratio (GLR), Wald (W), and Langrange multiplier (LM) provide asymptotic equivalent procedures for testing a general linear hypothesis. For a composite hypothesis, the limiting distribution of such tests is derived in a parametric form. The W test is used for constructing approximate confidence intervals. As an example, the local power property is illustrated.

1. Introduction

Analysis of financial data has received a considerable amount of attention in the literature during the past two decades. Several models have been suggested to capture special features of financial data and most of these models have the property that the conditional variance depends on the past. One of the well known and most heavily used examples is the class of ARCH(p) models, introduced by Engle (1982). Since then, ARCH related models (e.g., GARCH, GARCH-M, EGARCH) have become perhaps the most popular and extensively studied financial econometric models. The implementation of these parametric models is relatively simple, and from a practical point of view, it is well known now how to identify, estimate, and test this kind of model (for a description of these models and some empirical evidence, see the survey by Bollerslev *et al.* (1992)). Moreover, Giraitis *et al.* (2000) discussed a class of ARCH(∞) models, which includes that of ARCH(p) models as a special case, and established sufficient conditions for the existence of a stationary solution and its explicit representation.

In classical time series anaylsis the asymptotic estimation and testing theory were developed for linear processes, which include the AR, MA and ARMA models. Lütkephol (1991) extended these results to the case of vector process and discussed the asymptotics of the classical testing principles such as likelihood ratio (LR), Wald (W), and Langrange multiplier (LM) under the null hypothesis. More specifically, he derived the null asymptotic distribution of a Gaussian likelihood ratio (GLR) test for a vector-valued non-Gaussian AR process. Taniguchi and Kakizawa (2000) elucidated the asymptotics of tests based on Gaussian likelihoods for non-Gaussian vector linear processes. Their approach provides a convenient framework for many testing problems in the literature.

In this paper, we focus attention on the problem of estimation for bivariate ARCH,

2000 *Mathematics Subject Classification.* 62F03, 62F05, 62F12.

Key words and phrases. ARCH model; conditional least squares estimation; quasi-Gaussian maximum likelihood estimation; testing principles; local power.

and estimation and testing for multivariate ARCH models under the assumption of conditional correlation (for a review of these models, see e.g., Bollerslev (1987); Bollerslev *et al.* (1988); Engle and Kroner (1995)). Such models allow the variances and covariances to depend on the information set in a vector ARMA manner and are particularly useful in multivariate financial models. More concretely, the paper is organized as follows. Section 2 presents the conditional least squares estimation approach in the context of nonlinear time series given by Tjøstheim (1986). Using this result, in Section 3, we first discuss the problem of estimating the bivariate ARCH model with unknown volatility parameter vector. Then we apply these results to more general ARCH models. Section 4 discusses the testing problem for the multivariate ARCH model in the setting of Taniguchi and Kakizawa (2000), based on a quasi-Gaussian maximum likelihood estimator. The testing principles GLR, W, and LM provide asymptotic equivalent procedures for a general linear hypothesis. Then the limiting distribution for a composite hypothesis of such tests is derived in a parametric form. The W test is used to construct approximate confidence intervals. Based on the results by Sakiyama and Taniguchi (2003), the local power is highlighted in Section 5.

2. Conditional Least Squares Estimation

In this section, we state Tjøstheim's result (1986) which was essentially obtained by reformulating and extending the arguments of Klimko and Nelson (1978) to nonlinear time series.

Let $\{X_t : t = 0, \pm 1, \dots\}$ be a strictly stationary and ergodic process taking values in $\mathbb{R}^p$. In addition, suppose that $E\{\|X_t\|^2\} < \infty$, where $\|\cdot\|$ denotes the Euclidean norm. Suppose that observations $(X_1, \dots, X_n)$ are available. The probability distribution of $(X_1, \dots, X_n)$ is specified by an unknown parameter vector $\beta = (\beta_1, \dots, \beta_r)^T \in \mathbb{B} \subset \mathbb{R}^r$. Its true value is denoted by β^0. Then consider a general real-valued penalty function $\mathcal{Q}_n(\beta) = \mathcal{Q}_n(X_1, \dots, X_n; \beta)$ depending on the observations and a parameter vector $\beta \in \mathbb{B}$.

Let us now specify the penalty function. Let $\mathcal{F}_t(l)$ be the σ-field generated by $\{X_s : t - l \le s \le t\}$, where l is an appropriate integer. If $\{X_t\}$ is a nonlinear autoregressive model of order k, we can take $l = k$. Write $m_\beta(t, t-1) = E_\beta\{X_t | \mathcal{F}_{t-1}(l)\}$.

Consider the penalty function

$$\mathcal{Q}_n(\beta) = \sum_{t=p+1}^{n} \{X_t - m_\beta(t, t-1)\}^T \{X_t - m_\beta(t, t-1)\}.$$

The CLS estimator $\hat{\beta}_n^{(CLS)}$ of β is defined by $\hat{\beta}_n^{(CLS)} = \arg\min_{\beta \in \mathbb{B}} \mathcal{Q}_n(\beta)$.

Then, under some regularity conditions of Cramer type, Tjøstheim (1986, pp. 254-256) showed that

$$\hat{\beta}_n^{(CLS)} \xrightarrow{a.s} \beta^0, \quad \text{and} \quad \sqrt{n}(\hat{\beta}_n^{(CLS)} - \beta^0) \xrightarrow{d} \mathscr{N}(0, \mathscr{U}^{-1}\mathscr{R}\mathscr{U}^{-1}) \quad \text{as } n \to \infty,$$

where

$$\begin{aligned} \mathscr{U} &= E\left\{\frac{\partial}{\partial \beta} m_{\beta^0}^T(t, t-1) \frac{\partial}{\partial \beta} m_{\beta^0}(t, t-1)\right\} \quad \text{and} \\ \mathscr{R} &= E\left[\frac{\partial}{\partial \beta} m_{\beta^0}^T(t, t-1)\{X_t - m_{\beta^0}(t, t-1)\}\{X_t - m_{\beta^0}(t, t-1)\}^T \frac{\partial}{\partial \beta} m_{\beta^0}(t, t-1)\right] < \infty. \end{aligned}$$

The CLS estimation approach provides a unified treatment of estimation problems for widely used classes of nonlinear time series models. Based on such an estimation approach, Tjøstheim (1986) applied it to several classes of nonlinear models.

3. Estimation of ARCH models

In the present section we will first discuss the problem of estimation for a bivariate ARCH model to illustrate the key ideas based on the preceding framework. Then we apply such an estimation procedure to more general ARCH models.

3.1 Bivariate case

To begin with, let us suppose that a bivariate time series $\{X_t = (X_{1,t}, X_{2,t})^T\}$ follows an ARCH(p) model characterized by the equations

(1)
$$X_t = \Delta_{X,t}\varepsilon_t, \quad H_{X,t} \equiv E(X_t X_t^T | \mathscr{F}_{X,t-1}) = A_{X,0} + \sum_{j=1}^{p} A_{X,j} \odot (X_{t-j}X_{t-j}^T), \quad t \geq p+1,$$

where $\odot$ stands for the Hadamard product (element-by-element multiplication), $\mathscr{F}_{X,t}$ is the σ-field generated by $\{X_t, X_{t-1}, \dots\}$, $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})^T \sim$ i.i.d. $\mathscr{N}(0, \Gamma)$,

$$\Gamma = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}, \quad |\rho_{21}| < 1,$$

with corresponding fourth-order cumulant $\kappa_\varepsilon = (\kappa_{\varepsilon,1}, \kappa_{\varepsilon,2})^T$, $H_{X,t} \equiv (h_{ij,t}) = \Delta_{X,t}\Gamma\Delta_{X,t}$ is a 2×2 symmetric matrix, $\Delta_{X,t} = \text{diag}\{\sqrt{h_{11,t}}, \sqrt{h_{22,t}}\}$, $A_{X,0}$ and $A_{X,j}$, $j = 1, \dots, p$, are unknown volatility parameter matrices, and ε_t is independent of $\mathscr{F}_{X,t-1}$.

In the sequel, we will use the following identities. Let C_1, C_2, C_3 and C_4 be matrices whose dimensions are such that the products $C_1C_2C_3$, C_1C_3 and C_2C_4 exist. Further, suppose that C_1 and C_2 are nonsingular. Then

(B1) $\text{vec}(C_1C_2C_3) = (C_3^T \otimes C_1)\text{vec}(C_2)$,

(B2) $(C_1 \otimes C_2)(C_3 \otimes C_4) = C_1C_3 \otimes C_2C_4$,

(B3) $(C_1 \otimes C_2)^{-1} = C_1^{-1} \otimes C_2^{-1}$,

(B4) $(C_1 \otimes C_2)^T = C_1^T \otimes C_2^T$,

where $\otimes$ denotes the tensor product (for the definition of vec($\cdot$), see e.g., Magnus and Neudecker (1999)).

For the ARCH specification to be sensible, $H_{X,t}$ must be a positive definite matrix for all possible realizations of X_{t-1}. Note that the time evolution of $H_{X,t}$ is governed by that of $h_{11,t}$, $h_{22,t}$ and $h_{12,t} = \rho_{12}\sqrt{h_{11,t}h_{22,t}}$. Therefore, to model the volatility of X_t, it suffices to consider the following representation

$$h_{X,t} = \begin{pmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{pmatrix} = \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} + \sum_{j=1}^{p} \begin{pmatrix} a_{11,j} & a_{12,j} & a_{13,j} \\ a_{21,j} & a_{22,j} & a_{23,j} \\ a_{31,j} & a_{32,j} & a_{33,j} \end{pmatrix} \begin{pmatrix} X_{1,t-j}^2 \\ X_{1,t-j}X_{2,t-j} \\ X_{2,t-j}^2 \end{pmatrix}$$

(2)
$$\equiv a_{X,0} + \sum_{j=1}^{p} \tilde{A}_{X,j} Z_{X,t-j}, \quad t \geq p+1,$$

which satisfies the following conditions.

Assumption 1.

(i) The volatility parameter vector $\theta_X = (a_{X,0}^T, \text{vec}^T(\tilde{A}_{X,1}), \dots, \text{vec}^T(\tilde{A}_{X,p}))^T \in \Theta_X \subset \mathbb{R}^k$ is to be estimated, where Θ_X is compact and $k = 3(3p+1)$;

(ii) $h_{jj,t} > 0$ a.e., $j = 1, 2$, and Γ is positive definite;

(iii) $\tilde{A}_{X,j}$, $j = 1, \dots, p$, satisfy $\lambda(\tilde{A}_{X,1} + \cdots + \tilde{A}_{X,p}) < 1$, where $\lambda(\phi)$ denotes the maximum eigenvalue of ϕ in modulus.

The condition (ii) guarantees positive definiteness of $H_{X,t}$, and (iii) is a sufficient condition for $\{X_t\}$ to be stationary and ergodic (see Engle and Kroner (1995)). If $\tilde{A}_{X,j}$, $j = 1, \dots, p$, are diagonal matrices, the model will be called diagonal ARCH (see e.g., Gouriéroux (1997, p. 111)).

Write $W_{X,t} = (1, Z_{X,t-1}^T, \dots, Z_{X,t-p+1}^T)^T$ and $\zeta_{X,t} = Z_{X,t} - h_{X,t}$. Then, using (B1), we have the following autoregressive representation

$$Z_{X,t} = (W_{X,t-1}^T \otimes I_3)\theta_X + \zeta_{X,t}, \tag{3}$$

where I_3 is the 3×3 identity matrix. Notice that $E[Z_{X,t}|\mathscr{F}_{X,t-1}] = h_{X,t}$, and therefore, $h_{X,t}$ is the prediction of $Z_{X,t}$ when its past is known, and $\zeta_{X,t}$ is the error term. Write

$$\Omega_t(\theta_X) \equiv E[\zeta_{X,t}\zeta_{X,t}^T|\mathscr{F}_{X,t-1}] = E[(Z_{X,t} - h_{X,t})(Z_{X,t} - h_{X,t})^T|\mathscr{F}_{X,t-1}]. \tag{4}$$

Now consider the estimation of θ_X. Suppose that an observed stretch $(X_1, \dots, X_n)$ is available. The true value of θ_X is denoted by θ_X^0. Hence, from (B2), (B3) and (3), the conditional least squares estimator for θ_X is given by

$$\hat{\theta}_{X,n} = \left[\left(\sum_{t=p+1}^{n} W_{X,t-1}W_{X,t-1}^T\right)^{-1} \otimes I_3\right]\left[\sum_{t=p+1}^{n}(W_{X,t-1} \otimes I_3)Z_{X,t}\right]. \tag{5}$$

Let us impose the following assumption.

Assumption 2.

$E\{\|Z_{X,t}\|^4\} < \infty$.

Sufficient conditions to validate this assumption is given by Giraitis *et al.* (2000).

Write

$$\mathscr{U}_X = [E(W_{X,t-1}W_{X,t-1}^T)] \otimes I_3 \quad \text{and} \quad \mathscr{R}_X = E[(W_{X,t-1} \otimes I_3)\Omega_t(\theta_X^0)(W_{X,t-1}^T \otimes I_3)].$$

Then we have the following result.

Theorem 3.1. *Suppose that the process $\{Z_{X,t}\}$ given by (3) satisfies Assumptions 1 and 2. If $\mathscr{U}_X$ and $\mathscr{R}_X$ are positive definite matrices with bounded elements, then as $n \to \infty$,*

(i) there exists a sequence of estimators $\{\hat{\theta}_{X,n}\}$ such that $\hat{\theta}_{X,n} \xrightarrow{a.s} \theta_X^0$.

(ii) $\sqrt{n}(\hat{\theta}_{X,n} - \theta_X^0) \xrightarrow{d} \mathscr{N}(0,\ \mathscr{U}_X^{-1}\mathscr{R}_X\mathscr{U}_X^{-1})$.

Proof. In view of (3), we have

$$m_{\theta_X}(t, t-1) = (W_{X,t-1}^T \otimes I_3)\theta_X \quad \text{and} \quad \frac{\partial m_{\theta_X}(t, t-1)}{\partial \theta_X} = (W_{X,t-1} \otimes I_3).$$

Now, using Assumption 2, it is not difficult to show that the regularity conditions of Cramer type given in Tjøstheim (1986) are satisfied.

Observe that

$$\begin{aligned}\hat{\theta}_{X,n} - \theta_X &= \left[\left((n-p)^{-1}\sum_{t=p+1}^{n} W_{X,t-1}W_{X,t-1}^T\right)^{-1} \otimes I_3\right] \\ &\quad \times \left[(n-p)^{-1}\sum_{t=p+1}^{n}(W_{X,t-1}\otimes I_3)\zeta_{X,t}\right].\end{aligned}$$

Since, by Assumption 2, $\{W_{X,t-1}W_{X,t-1}^T \otimes I_3\}$ and $\{(W_{X,t-1}\otimes I_3)\zeta_{X,t}\}$ are strictly stationary and ergodic. Moreover, $\mathscr{U}_X$ is finite, and $E[(W_{X,t-1}\otimes I_3)\zeta_{X,t}] = 0$. Thus, by the ergodic theorem,

$$\begin{aligned}(n-p)^{-1}\sum_{t=p+1}^{n}(W_{X,t-1}W_{X,t-1}^T)\otimes I_3 &\xrightarrow{a.s.} \mathscr{U}_X, \quad \text{and} \\ (n-p)^{-1}\sum_{t=p+1}^{n}(W_{X,t-1}\otimes I_3)\zeta_{X,t} &\xrightarrow{a.s.} 0,\end{aligned}$$

which imply $\hat{\theta}_{X,n} \xrightarrow{a.s} \theta_X$.

Next we derive the limiting distribution of $\hat{\theta}_{X,n}$. For any given vector $\alpha = (\alpha_0, \ldots, \alpha_k)^T \neq 0$, it follows from (B4), the Cramer-Wold device and Billingsley's theorem (1961) for martingales that the distribution of

$$(n-p)^{-1/2}\sum_{t=p+1}^{n}\alpha^T(W_{X,t-1}\otimes I_3)\zeta_{X,t}$$

converges to the normal distribution with mean zero and variance

$$\begin{aligned}E[(\alpha^T(W_{X,t-1}\otimes I_3)\zeta_{X,t})^2] &= E[\alpha^T(W_{X,t-1}\otimes I_3)E[\zeta_{X,t}\zeta_{X,t}^T|\mathscr{F}_{X,t-1}](W_{X,t-1}^T\otimes I_3)\alpha] \\ &= \alpha^T\mathscr{R}_X\alpha \quad \text{as } n\to\infty.\end{aligned}$$

Thus, by Slutsky's theorem, the assertion (ii) follows. Hence the proof is completed.

3.2 Multivariate case

In the preceding subsection, we described the estimation procedure for the bivariate ARCH model. Based on such an estimation procedure, it is natural to consider modeling the ARCH model in a higher dimensional situation.

Let $\{Y_t = (X_{1,t}, \ldots, X_{m,t})^T\}$ be an m-variate time series generated by

$$Y_t = \Delta_{Y,t}\tilde{\varepsilon}_t, \quad H_{Y,t} \equiv E(Y_tY_t^T|\mathscr{F}_{Y,t-1}) = A_{Y,0} + \sum_{j=1}^{p}A_{Y,j}\odot(Y_{t-j}Y_{t-j}^T), \quad t \geq p+1,$$

where $\mathscr{F}_{Y,t}$ is the σ-field generated by $\{Y_t, Y_{t-1}, \ldots\}$, $\tilde{\varepsilon}_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{m,t})^T \sim$ i.i.d. $\mathscr{N}(0, \tilde{\Gamma})$,

$$\tilde{\Gamma} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{12} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{(m-1)m} \\ \rho_{1m} & \cdots & \rho_{(m-1)m} & 1 \end{pmatrix}, \quad |\rho_{ij}| < 1,$$

with corresponding fourth-order cumulant $\kappa_{\tilde{\varepsilon}} = (\kappa_{\tilde{\varepsilon},1}, \dots, \kappa_{\tilde{\varepsilon},m})^T$, $H_{Y,t} \equiv (h_{ij,t}) = \Delta_{Y,t}\tilde{\Gamma}\Delta_{Y,t}$ is an $m \times m$ symmetric matrix, $\Delta_{Y,t} = \mathrm{diag}\{\sqrt{h_{11,t}}, \dots, \sqrt{h_{mm,t}}\}$, $A_{Y,0}$ and $A_{Y,j}$, $j = 1, \dots, p$, are unknown volatility parameter matrices, and $\tilde{\varepsilon}_t$ is independent of $\mathscr{F}_{Y,t-1}$. By analogy with (2), the corresponding representation of $H_{Y,t}$ is given by

$$\begin{aligned} h_{Y,t} = \mathrm{vech}(H_{Y,t}) &= \begin{pmatrix} a_{10} \\ \vdots \\ a_{q0} \end{pmatrix} + \sum_{j=1}^{p} \begin{pmatrix} a_{11,j} & \cdots & a_{1q,j} \\ \vdots & \ddots & \vdots \\ a_{q1,j} & \cdots & a_{qq,j} \end{pmatrix} \mathrm{vech}(Y_{t-j}Y_{t-j}^T) \\ &\equiv a_{Y,0} + \sum_{j=1}^{p} \tilde{A}_{Y,j} Z_{Y,t-j}, \quad t \geq p+1, \end{aligned}$$

where $q = m(m+1)/2$.

We now impose the following conditions.

Assumption 3.

(i) The volatility parameter vector $\theta_Y = (a_{Y,0}^T, \mathrm{vec}^T(\tilde{A}_{Y,1}), \dots, \mathrm{vec}^T(\tilde{A}_{Y,p}))^T \in \Theta_Y \subset \mathbb{R}^s$ is to be estimated, where Θ_Y is compact and $s = m(mp+1)$;

(ii) $h_{ii,t} > 0$ a.e., $i = 1, \dots, m$, and $\tilde{\Gamma}$ is positive definite;

(iii) $\tilde{A}_{Y,j}$, $j = 1, \dots, p$, satisfy $\lambda(\tilde{A}_{Y,1} + \cdots + \tilde{A}_{Y,p}) < 1$ for stationarity.

The condition (ii) ensures the positive definiteness of $H_{Y,t}$ (see e.g., Engle (1995, p. 302)).

Write $W_{Y,t} = (1, Z_{Y,t-1}^T, \dots, Z_{Y,t-p+1}^T)^T$ and $\zeta_{Y,t} = Z_{Y,t} - h_{Y,t}$. Then

$$Z_{Y,t} = (W_{Y,t-1}^T \otimes I_q)\theta_Y + \zeta_{Y,t}, \tag{6}$$

where I_q is the $q \times q$ identity matrix. Recalling (4), we write

$$\Omega_t(\theta_Y) \equiv E[\zeta_{Y,t}\zeta_{Y,t}^T | \mathscr{F}_{Y,t-1}] = E[(Z_{Y,t} - h_{Y,t})(Z_{Y,t} - h_{Y,t})^T | \mathscr{F}_{Y,t-1}].$$

Based on an observed stretch $(Y_1, \dots, Y_n)$, we shall estimate θ_Y. Its true value is denoted by θ_Y^0. Hence, by analogy with (5), the conditional least squares estimator for θ_Y is given by

$$\hat{\theta}_{Y,n} = \left[\left(\sum_{t=p+1}^{n} W_{Y,t-1}W_{Y,t-1}^T \right)^{-1} \otimes I_q \right] \left[\sum_{t=p+1}^{n} (W_{Y,t-1} \otimes I_q) Z_{Y,t} \right]. \tag{7}$$

By writing

$$\mathscr{U}_Y = [E(W_{Y,t-1}W_{Y,t-1}^T)] \otimes I_q \quad \text{and} \quad \mathscr{R}_Y = E[(W_{Y,t-1} \otimes I_q)\Omega_t(\theta_Y^0)(W_{Y,t-1}^T \otimes I_q)],$$

we can establish the asymptotics of (7) via the ergodic theorem and Billingsley's theorem (1961) for martingales similarly as in Theorem 3.1.

Theorem 3.2. *Suppose that the process $\{Z_{Y,t}\}$ given by (6) satisfies Assumption 3. If $\mathscr{U}_Y$ and $\mathscr{R}_Y$ are positive definite matrices with bounded elements, then as $n \to \infty$,*

(i) there exists a sequence of estimators $\{\hat{\theta}_{Y,n}\}$ such that $\hat{\theta}_{Y,n} \xrightarrow{a.s} \theta_Y^0$.

(ii) $\sqrt{n}(\hat{\theta}_{Y,n} - \theta_Y^0) \xrightarrow{d} \mathscr{N}(0,\ \mathscr{U}_Y^{-1}\mathscr{R}_Y\mathscr{U}_Y^{-1})$.

4. Testing problem in multivariate ARCH

An application of any statistical model to data desires to have diagnostic tests available for checking the model's ability to describe the data, once the model has been estimated. This section considers the classical testing principles such as Gaussian likelihood ratio (GLR), Wald (W) and, Langrange multiplier (LM), which provide a convenient framework for deriving such tests. The W test is used for constructing approximate confidence intervals.

Consider the multivariate ARCH model,

$$Z_{Y,t} = (W_{Y,t-1}^T \otimes I_q)\theta_Y + \zeta_{Y,t}.$$

Recalling $\Omega_t(\theta_Y) = E[\zeta_{Y,t}\zeta_{Y,t}^T | \mathscr{F}_{Y,t-1}]$, we can under the stationarity condition, set $\Omega = E\{\Omega_t(\theta_Y)\}$. The spectral density matrix of $\{Z_{Y,t}\}$ is given by

$$f(\lambda) = \frac{1}{2\pi} A(\lambda)^{-1} \Omega \{A(\lambda)^*\}^{-1},$$

where $A(\lambda) = I_q - \sum_{j=1}^p \tilde{A}_{Y,j} e^{ij\lambda}$. Henceforth we are interested in a k-dimensional unknown parameter ϕ, and suppose that ϕ is innovation free, i.e., Ω is independent of ϕ.

Let

$$\mathscr{I}_n(\lambda) = \frac{1}{2\pi n} \left\{ \sum_{t=1}^n (Z_{Y,t} - \bar{Z}_{Y,t}) e^{it\lambda} \right\} \left\{ \sum_{t=1}^n (Z_{Y,t} - \bar{Z}_{Y,t}) e^{it\lambda} \right\}^*$$

be the periodogram of the partial realization of $\{Z_{Y,t}\}$, where $\bar{Z}_{Y,t} = n^{-1} \sum_{t=1}^n Z_{Y,t}$ and $*$ is the complex conjugate transpose. We write the spectral density matrix $f(\lambda)$ as $f_\phi(\lambda)$.

Suppose that we wish to test

$$H_0 : C\phi - r = 0 \quad \text{against} \quad H_A : C\phi - r \neq 0, \tag{8}$$

where C is a specified $(k-l) \times k$ matrix of full rank, and r is a specified $k-l$ vector with $k = \dim(\phi)$.

In order to estimate ϕ, we use the results of Hosoya and Taniguchi (1982) who considered its estimation by minimizing the quantity

$$\mathscr{D}(f_\phi, \mathscr{I}_n) = \int_{-\pi}^{\pi} [\log \det f_\phi(\lambda) + \operatorname{tr}\{f_\phi(\lambda)^{-1} \mathscr{I}_n(\lambda)\}] d\lambda, \tag{9}$$

where $\det(\cdot)$ and $\operatorname{tr}(\cdot)$ stand for the determinant and trace, respectively. The motivation of this quantity stems from approximated Gaussian likelihoods although Gaussianity of $\{Z_{Y,t}\}$ is not assumed.

Let $\hat{\phi}^{(QML)}$ be the quasi-Gaussian maximum likelihood estimator of ϕ minimizing (9). Then, under general regularity conditions, they showed that

$$\sqrt{n}(\hat{\phi}^{(QML)} - \phi) \xrightarrow{d} \mathscr{N}_k(0, F^{-1}(\phi)), \tag{10}$$

where

$$F(\phi) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial \phi} \log f_\phi(\lambda) \frac{\partial}{\partial \phi^T} \log f_\phi(\lambda) d\lambda.$$

Hence, in view of (8), we have

$$\sqrt{n}(C\hat{\phi}^{(QML)} - C\phi) \xrightarrow{d} \mathscr{N}_{k-l}(0, CF^{-1}(\phi)C^T). \tag{11}$$

Let

$$\tilde{\phi}^{(QML)} = \arg \min_{\phi:\, C\phi = r} \mathscr{D}(f_\phi, \mathscr{I}_n)$$

be the restricted quasi-Gaussian maximum likelihood estimator of ϕ. Then the Gaussian likelihood ratio (GLR) test statistic is given by

$$GLR = \frac{n}{2\pi}\{\mathscr{D}(f_{\hat{\phi}^{(QML)}}, \mathscr{I}_n) - \mathscr{D}(f_{\tilde{\phi}^{(QML)}}, \mathscr{I}_n)\}.$$

Motivated by (11), the Wald (W) test statistic is given by

$$W = n(C\hat{\phi}^{(QML)} - r)^T[CF^{-1}(\hat{\phi}^{(QML)})C^T]^{-1}(C\hat{\phi}^{(QML)} - r).$$

The Langrange multiplier (LM) test statistic is of the form

$$LM = \frac{n}{16\pi^2}\left(\frac{\partial}{\partial\phi}\mathscr{D}(f_{\tilde{\phi}^{(QML)}}, \mathscr{I}_n)\right)^T F^{-1}(\tilde{\phi}^{(QML)})\left(\frac{\partial}{\partial\phi}\mathscr{D}(f_{\tilde{\phi}^{(QML)}}, \mathscr{I}_n)\right),$$

which is based on the restricted estimator $\tilde{\phi}^{(QML)}$. Hence, under appropriate regularity conditions, it can be shown similarly as in Taniguchi and Kakizawa (2000, pp. 60-63) that the limiting distribution of the tests GLR, W, and LM when H_0 is true tends to χ^2_{k-l} as $n \to \infty$.

In summary we have three test procedures with equivalent asymptotic distribution under H_0. The GLR statistic involves both the restricted and unrestricted QML estimator, the LM test statistic is based on the restricted estimator only, and the W test statistic requires just the unrestricted estimator. The choice among the three procedures is often based on computational convenience. The W test has a disavantage that it is not invariant under transformations of the restrictions.

The hypothesis (8) reduces to the following problem of testing composite hypothesis

$$H_0^* : \phi_2 = \phi_{2,0} \quad \text{against} \quad H_A^* : \phi_2 \neq \phi_{2,0}, \tag{12}$$

where $\phi = (\phi_1^T, \phi_2^T)^T$, $\phi_1 = (\phi^1, \dots, \phi^l)^T$, $\phi_2 = (\phi^{l+1}, \dots, \phi^k)^T$, and $\phi_{2,0} = (\phi_0^{l+1}, \dots, \phi_0^k)^T$, a specified vector and $(\phi_1^T, \phi_{2,0}^T)^T \in \mathrm{Int}(\Theta_X)$. Write $\hat{\phi}^{(QML)} = ((\hat{\phi}_1^{(QML)})^T, (\hat{\phi}_2^{(QML)})^T)^T$ under H_A^* and define $\tilde{\phi}_1^{(QML)}$ by

$$\tilde{\phi}_1^{(QML)} = \arg\min_{\phi_1} \mathscr{D}(f_{(\phi_1, \phi_{2,0})}, \mathscr{I}_n).$$

For (12), we use the followings:

$$\begin{aligned}
GLR^* &= \frac{n}{2\pi}\{\mathscr{D}(f_{(\hat{\phi}_1^{(QML)}, \hat{\phi}_2^{(QML)})}, \mathscr{I}_n) - \mathscr{D}(f_{(\tilde{\phi}_1^{(QML)}, \phi_{2,0})}, \mathscr{I}_n)\},\\
W^* &= n(\hat{\phi}_2^{(QML)} - \phi_{2,0})^T[GF^{-1}(\hat{\phi}^{(QML)})G^T]^{-1}(\hat{\phi}_2^{(QML)} - \phi_{2,0}),\\
LM^* &= \frac{n}{16\pi^2}\left(\frac{\partial}{\partial\phi}\mathscr{D}(f_\phi, \mathscr{I}_n)\bigg|_{\phi = \tilde{\phi}^{(QML)}}\right)^T F^{-1}(\tilde{\phi}^{(QML)})\\
&\quad \times\left(\frac{\partial}{\partial\phi}\mathscr{D}(f_\phi, \mathscr{I}_n)\bigg|_{\phi = \tilde{\phi}^{(QML)}}\right),
\end{aligned}$$

where $G = [0, I_{k-l}]$, a $(k-l) \times k$ matrix with I_{k-l}, the $(k-l) \times (k-l)$ identity matrix and $\tilde{\phi}^{(QML)} = ((\tilde{\phi}_1^{(QML)})^T, \phi_{2,0}^T)^T$ is the restricted estimator of ϕ under H_0^*.

Now we proceed to derive the limiting distribution of the tests GLR^*, W^*, and LM^* under H_0^* in the setting of Sakiyama and Taniguchi (2003). Partitioning the information matrix conformably with ϕ_1 and ϕ_2 yields

$$F(\phi) = \begin{pmatrix} F_{11} & F_{12} \\ F_{21}^T & F_{22} \end{pmatrix},$$

where the component matrices F_{11}, F_{22} and F_{12} are $l \times l$, $(k-l) \times (k-l)$ and $l \times (k-l)$, respectively. If $K = F_{22} - F_{21}F_{11}^{-1}F_{12}$, then provided F_{11}^{-1} and K^{-1} exist, the inverse of $F(\phi)$ is

$$F^{-1}(\phi) = \begin{pmatrix} F_{11}^{-1} + L & -F_{11}^{-1}F_{12}K^{-1} \\ -K^{-1}F_{21}F_{11}^{-1} & K^{-1} \end{pmatrix}, \tag{13}$$

where $L = F_{11}^{-1}F_{12}K^{-1}F_{21}F_{11}^{-1}$ (see e.g., Rencher (2000, p. 21)). Write $v = \sqrt{n}(\hat{\phi}^{(QML)} - \phi)$, $w = \sqrt{n}(\tilde{\phi}_1^{(QML)} - \phi_1)$ and $u = (w^T, 0^T)^T$. Then, under H_0^*, it is not difficult to show

$$GLR^* = (u-v)^T F(\phi)(u-v)(1 + o_p(1)). \tag{14}$$

Observe that

$$\begin{aligned} v &= -F^{-1}(\phi)\frac{\sqrt{n}}{4\pi}\frac{\partial}{\partial\phi}\mathscr{D}(f_\phi, \mathscr{I}_n)(1 + o_p(1)), \\ u &= -J(\phi)\frac{\sqrt{n}}{4\pi}\frac{\partial}{\partial\phi}\mathscr{D}(f_\phi, \mathscr{I}_n)(1 + o_p(1)), \end{aligned} \tag{15}$$

where

$$J(\phi) = \begin{pmatrix} F_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

From (15), the expression (14) becomes

$$GLR^* = \frac{\sqrt{n}}{4\pi}\frac{\partial}{\partial\phi^T}\mathscr{D}(f_\phi, \mathscr{I}_n)\{F^{-1}(\phi) - J(\phi)\}\frac{\sqrt{n}}{4\pi}\frac{\partial}{\partial\phi}\mathscr{D}(f_\phi, \mathscr{I}_n)(1 + o_p(1)). \tag{16}$$

Recall (10):

$$\frac{\sqrt{n}}{4\pi}\frac{\partial}{\partial\phi}\mathscr{D}(f_\phi, \mathscr{I}_n) \xrightarrow{d} \mathscr{N}_k(0, F(\phi)), \tag{17}$$

which, together with (13) and (16), leads to

$$GLR^* = Q_{GLR^*}^T Q_{GLR^*} + o_p(1),$$

where

$$Q_{GLR^*} = -(4\pi)^{-1}K^{-1/2}\left[\sqrt{n}\frac{\partial}{\partial\phi_2}\mathscr{D}(f_\phi, \mathscr{I}_n) - F_{21}F_{11}^{-1}\sqrt{n}\frac{\partial}{\partial\phi_1}\mathscr{D}(f_\phi, \mathscr{I}_n)\right].$$

In view of (17) we can see that $Q_{GLR^*} \xrightarrow{d} \mathscr{N}(0, I_{k-l})$, and hence,

$$GLR^* \xrightarrow{d} \chi^2_{k-l} \quad \text{under } H_0^*.$$

Next we consider the W^* test. From (13) and (15), it follows that

$$\sqrt{n}(\hat{\phi}_2^{(QML)} - \phi_{2,0}) = -(4\pi)^{-1}K^{-1}\left[\frac{\partial}{\partial \phi_2}\mathscr{D}(f_\phi, \mathscr{I}_n) - F_{21}F_{11}^{-1}\frac{\partial}{\partial \phi_1}\mathscr{D}(f_\phi, \mathscr{I}_n)\right](1 + o_p(1)).$$

Write

$$Q_{W^*} = -(4\pi)^{-1}K^{-1/2}\left[\sqrt{n}\frac{\partial}{\partial \phi_2}\mathscr{D}(f_\phi, \mathscr{I}_n) - F_{21}F_{11}^{-1}\sqrt{n}\frac{\partial}{\partial \phi_1}\mathscr{D}(f_\phi, \mathscr{I}_n)\right],$$

from which,

$$W^* = \sqrt{n}K^{1/2}(\hat{\phi}_2^{(QML)} - \phi_{2,0})^T\sqrt{n}K^{1/2}(\hat{\phi}_2^{(QML)} - \phi_{2,0}) = Q_{W^*}^T Q_{W^*} + o_p(1).$$

Then, using the fact that $Q_{W^*} \xrightarrow{d} \mathscr{N}(0, I_{k-l})$, it follows that

$$W^* \xrightarrow{d} \chi^2_{k-l}, \quad \text{under } H_0^*.$$

In the same way as in GLR^*, we get

$$LM^* \xrightarrow{d} \chi^2_{k-l}, \quad \text{under } H_0^*.$$

The above results are summarized in

Proposition 4.1. *For the testing problem (12), the limiting distribution of the tests GLR^*, W^*, and LM^* under H_0^* tends to χ^2_{k-l} as $n \to \infty$.*

The asymptotics of the W test is used to construct approximate confidence intervals. Suppose ϕ is partitioned as $\phi = (\phi_1, \phi_2^T)^T$, where ϕ_1 is the first component and also the parameter of interest, and ϕ_2 is the $(k-1)$-vector of the remaining components. By re-ordering and relabeling the components of ϕ if necessary, ϕ_1 can be taken to be any of the components of ϕ. Write $\hat{\phi}^{(QML)}(\phi_1) = (\phi_1, (\hat{\phi}_2^{(QML)}(\phi_1))^T)^T$, and denote by $\hat{\phi}_1^{(QML)}$ the QML estimator of ϕ_1. Then the approximate level $100(1-\alpha)\%$ confidence intervals for ϕ_1 is given by

$$\{\phi_1 : n(\hat{\phi}_1^{(QML)} - \phi_1)^2 F(\hat{\phi}_1^{(QML)}) \leq S_1^{-1}(1-\alpha)\},$$

where S_1 is the distribution function of the χ_1^2 distribution with one degree of freedom. The interval derived from this test reduces to

$$[\hat{\phi}_1^{(QML)} - n^{-1/2}F^{-1/2}(\hat{\phi}_1^{(QML)})\Phi^{-1}(1-\alpha/2),\ \hat{\phi}_1^{(QML)} + n^{-1/2}F^{-1/2}(\hat{\phi}_1^{(QML)})\Phi^{-1}(1-\alpha/2)],$$

where $\Phi(x) = \int_{-\infty}^x (2\pi)^{-1/2}e^{-t^2/2}dt$.

5. Local power evaluation

In this section we provide a heuristic analysis in terms of local power based on the results of Sakiyama and Taniguchi (2003), and the previous section. For this, it is natural to consider the first-order bivariate ARCH model of the form

$$\begin{aligned} h_t = \begin{pmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{pmatrix} &= \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} X_{1,t-1}^2 \\ X_{1,t-1}X_{2,t-1} \\ X_{2,t-1}^2 \end{pmatrix} \\ &\equiv a_0 + A\,Z_{t-1}, \quad 2 \leq t \leq n, \end{aligned}$$

where $a_{10} > 0$, $a_{20} > 0$, $a_{20} > 0$, and $a_{11} \geq 0$, $a_{33} \geq 0$, $a_{11}a_{33} - a_{22}^2 \geq 0$. Write $\xi_t = Z_t - h_t$. Then $Z_t = a_0 + A Z_{t-1} + \zeta_t$. Note that $E\{E[\zeta_t \zeta_t^T | \mathscr{F}_{t-1}]\} = \Omega \equiv (\omega_{ij})$, $i, j = 1, 2, 3$, and is independent of A, where $\mathscr{F}_t$ is the σ-field generated by $\{X_t, X_{t-1}, \dots\}$. Write $\phi = (\phi_1^T, \phi_2^T)^T$, $\phi_1 = (\omega_{11}, \omega_{22}, \omega_{33}, \omega_{21}, \omega_{31}, \omega_{32})^T$, $\phi_2 = (a_{11}, a_{22}, a_{33})^T$. Then the spectral density matrix is given by $f_\phi(\lambda) = (2\pi)^{-1}\mathcal{A}(\lambda)^{-1}\Omega\{\mathcal{A}(\lambda)^*\}^{-1}$, where $\mathcal{A}(\lambda) = I_3 - A e^{i\lambda}$.

Let us consider the problem of testing composite hypothesis

$$H_0 : \phi_2 = \phi_{2,0} \quad \text{against} \quad H_A : \phi_2 \neq \phi_{2,0},$$

where $\phi_{2,0} = (a_{11,0}, a_{22,0}, a_{33,0})^T$ is a specified vector. From the definition of $F(\phi)$, we have the followings:

$$\begin{aligned}
\underset{(6\times 6)}{F_{11}} &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial}{\partial \phi_1} \log f_\phi(\lambda) \frac{\partial}{\partial \phi_1^T} \log f_\phi(\lambda) d\lambda, \\
\underset{(6\times 3)}{F_{12}} &= F_{21}^T = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial}{\partial \phi_1} \log f_\phi(\lambda) \frac{\partial}{\partial \phi_2^T} \log f_\phi(\lambda) d\lambda = 0, \\
\underset{(3\times 3)}{F_{22}} &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial}{\partial \phi_2} \log f_\phi(\lambda) \frac{\partial}{\partial \phi_2^T} \log f_\phi(\lambda) d\lambda \\
&= \begin{pmatrix} (1-a_{11}^2)^{-1} & (1-a_{11}a_{22})^{-1} & (1-a_{11}a_{33})^{-1} \\ (1-a_{11}a_{22})^{-1} & (1-a_{22}^2)^{-1} & (1-a_{22}a_{33})^{-1} \\ (1-a_{11}a_{33})^{-1} & (1-a_{22}a_{33})^{-1} & (1-a_{33}^2)^{-1} \end{pmatrix}.
\end{aligned}$$

Hence, $K = F_{22} - F_{21}^T F_{11}^{-1} F_{12} = F_{22}$, and from the result by Sakiyama and Taniguchi (2003), the local power of the test with asymptotic level z_α is given by

$$\iiint_{|x_1|,|x_2|,|x_3| > z_\alpha} d\mathscr{N}(F_{22}^{1/2}h, I_3), \quad h \in \mathbb{R}^3.$$

For $\phi_{2,0} = (0.3, 0.1, 0.2), (0.9, 0.1, 0.8), (0.9, 0.7, 0.8)$, and $\alpha = 0.05$, Figure 1 provides the local power. Note that for these choice of the parameter values, the concerned conditions are satisfied. An interesting feature is that the local power increases when the components of $\phi_{2,0}$ increase, which means that $h/\sqrt{n}$ approaches zero at the rate of $1/\sqrt{n}$.

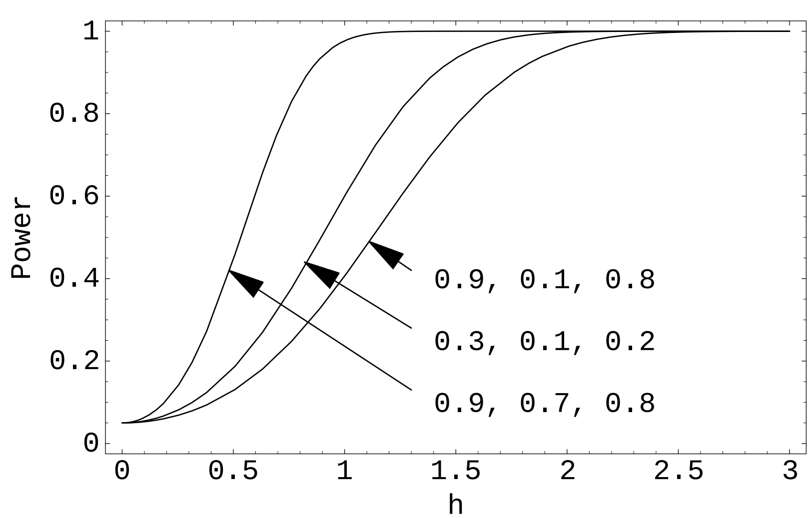

Figure 1: Local power with $\alpha = 0.05$

Acknowledgement

The author wishes to thank the referee, and Professor M. Taniguchi for his constructive suggestions.

References

[1] Billingsley, P. (1961). The Linderberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, **12**, 788-792.

[2] Bollerslev, T. P. (1987). Modeling the coherence in short run nominal exchange rates: A multivariate generalized ARCH approach. *Rev. Econ. Statist.*, **72**, 498-505.

[3] Bollerslev, T. P., Chou, R. Y. and Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *J. Econometrics*, **31**, 309-328.

[4] Bollerslev, T. P., Engle, R. F. and Wooldridge, J. M. (1988). A capitial asset pricing model with time varying covariances. *J. Political Economy*, **96**, 116-131.

[5] Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, **50**, 987-1007.

[6] Engle, R. F. (1995). *ARCH Selected Readings*. New York: Oxford University Press.

[7] Engle, R. F. and Granger, C. W. J. (1984). Combining competing forecasts of inflation using a bivariate ARCH model. *J. Econ. Dynamic Control*, **8**, 151-165.

[8] Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, **11**, 122-150.

[9] Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory*, **16**, 3-22.

[10] Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. New York: Springer.

[11] Hosoya, Y. and Taniguchi, M. (1982). A central limit thorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.*, **10**, 132-153.

[12] Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.*, **6**, 420-642.

[13] Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

[14] Magnus, J. R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Chichester: Wiley.

[15] Rencher, A. C. (2000). *Linear Models in Statistics*. New York: Wiley.

[16] Sakiyama, K. and Taniguchi, M. (2003). Testing composite hypotheses for locally stationary processes. To appear in *J. Time Ser. Anal.*.

[17] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*. New York: Springer-Verlag.

[18] Tjøstheim, D. (1986). Estimation in nonlinear time series models. *Stoch. Proc. Appl.*, **21**, 251-273.

Department of Mathematical Science
Graduate School of Engineering Science
Osaka University, Japan
E-mail: chandra@sigmath.es.osaka-u.ac.jp