Consistency of Quasi-Maximum Likelihood Estimators for the Regime-Switching GARCH Models

Yingfu Xie

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2005:3
ISSN 1651-8543

Consistency of Quasi-Maximum Likelihood Estimators for the Regime-Switching GARCH Model

Yingfu Xie[1]
Centre of Biostochastics
Swedish University of Agricultural Sciences
SE-901 83 Umea, Sweden

Abstract

Regime-switching GARCH (generalized autoregressive conditionally heteroscedastic) model incorporates the idea of Markov switching into the somehow restrictive GARCH model, which significantly extends GARCH models. However, the statistical inference for this model is rather difficult due to the dependence to the whole regime path. In this paper, we obtain the consistency of the quasi-maximum likelihood estimators, by transforming it to an infinite order ARCH model. Simulation studies to illustrate asymptotic behavior of the estimators and a model specification problem are presented.

Keywords: Regime-switching; GARCH model; Quasi-MLE; Consistency; Asymptotic normality

[1]E-mail address to the correspondence author: yingfu.xie@sekon.slu.se

1 Introduction

The regime-switching GARCH (generalized autoregressive conditionally heteroskedastic) model studied by Cai (1994), Hamilton and Susmel (1994), Gray (1996) and Francq et al. (2001) combined the seminal work of Engle (1982) and Bollerslev (1986) for GARCH and Hamilton (1989) for Markov switching models. The main idea is that the set of parameters of GARCH model is determined by some unobservable regimes, where the switch of regimes is governed by a Markov chain. Empirical results show that it gives more reasonable volatility fit than the ordinary (single-regime) GARCH model (Gray, 1996, among others). Also, the strong persistence of ordinary (single regime) GARCH processes frequently observed (e.g., Lamoureux and Lastrapes, 1990) can be possibly explained by changes of regimes instead of the integrated GARCH (IGARCH) models. We further note that the regime-switching GARCH model can also be considered as an extension of the hidden Markov model (HMM), which is popular in various fields such as engineering, genetic biology and statistics. For HMM, readers are referred to the monograph by MacDonald and Zucchini (1997).

The statistical inference for regime-switching GARCH model, however, is rather difficult mainly because observation at every time point depends all previous regimes due to the autoregressive structure of GARCH equation, which causes the likelihood intractable when the sample size is only moderately large. Gray (1996) proposed a reduced regime-switching model in which the previous regime is integrated out at every step and hence the regime path dependence problem is overcome. Klaassen (2002) modified it by making use of more information from observations and showed superiority in prediction.

Compared with the large empirical literature, there are only few theoretical results. Cai (1994) gave a sufficient and necessary condition for the stationarity of regime-switching ARCH model. This result is extended to GARCH case by Francq et al. (2001), where they also proved the consistency of the maximum-likelihood estimator (MLE) in ARCH case. Xie and Yu (2005) obtained the consistency of quasi-MLE (QMLE) for reduced regime-switching GARCH model generalized from Gray (1996). They also illustrated the asymptotic behavior of QMLE through simulation study.

In this paper, we will consider the consistency of the QMLE of general regime-switching GARCH model. Motivated from Berkes et al. (2003), we will write the GARCH as an ARCH(∞) representation and then use similar technique to Xie and Yu (2005). Simulation studies to illustrate asymptotic behaviors of QMLE will also be presented. Similar to Xie and Yu (2005), it is of interest to analyze the stationarity of estimators if we use an ordinary GARCH model to fit data generated from regime-switching models.

We give the model and the alternative form in Section 2. Our main result is stated in Section 3. In Section 4 we give simulation studies. Section 5 includes a few discussions.

2 The model and the alternative form

The general Regime-switching GARCH(p,q) process $\{Y_t\}_{t\in\mathbb{Z}}$ satisfies

$$Y_t = (h_t)^{1/2}\eta_t,$$

$$h_t = \omega(R_t) + \sum_{i=1}^{q} \alpha_i(R_t)Y_{t-i}^2 + \sum_{j=1}^{p} \beta_j(R_t)h_{t-j}, \tag{1}$$

where $\{\eta_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with zero mean and unit variance. $\{R_t\}$ is a Markov chain with finite state (regimes, in econometric literature) space $\mathbf{E} = \{1, 2, ..., d\}$. We assume that $\alpha_i(s) \geq 0, 1 \leq i \leq q$, $\beta_j(s) \geq 0, 1 \leq j \leq p$ and $\omega(s) > 0$ given $\{R_t = s\}$, $s \in \mathbf{E}$, in order to ensure an almost surely strictly positive conditional variance h_t. Suppose that $\{\eta_t\}$ and $\{R_t\}$ are independent and that the Markov chain is stationary, irreducible and aperiodic with stationary distribution $\pi(s) := P(R_1 = s), 1 \leq s \leq d$ and transition probabilities $p(k, l) := P(R_t = l|R_{t-1} = k)$. Also note that $\pi(s) > 0, s \in \mathbf{E}$, under assumptions.

From (1) we can see that the conditional variance h_t depends on the present regime R_t, and the whole regime path as well, through h_{t-j}. Thus, the number of possible regime paths grows exponentially with t. This leads to an enormous numbers of paths up to t, and the likelihood becomes intractable very quickly. To avoid this problem, we will formulate the GARCH equation as ARCH analogous to Berkes et al. (2003).

First, we will give a sufficient and necessary condition for the existence of a strictly stationary solution for our model, following from Brandt (1986), Bougerol and Picard (1992) and Francq et al. (2001, Theorem 1).

Write (assuming $\min(p, q) \geq 2$, adding zero coefficients if necessary)

$$\boldsymbol{\tau}_t = (\beta_1(R_t) + \alpha_1(R_t)\eta_t^2, \beta_2(R_t), ..., \beta_{p-1}(R_t)) \in \mathbb{R}^{p-1},$$

$$\boldsymbol{\xi}_t = (\eta_t^2, 0, ..., 0) \in \mathbb{R}^{p-1},$$

and

$$\boldsymbol{\lambda}_t = (\alpha_2(R_t), ..., \alpha_{q-1}(R_t)) \in \mathbb{R}^{q-2}.$$

Define the square matrix A_t of size $(p+q-1)$ in block form as

$$\boldsymbol{A}_t = \begin{pmatrix} \tau_t & \beta_p(R_t) & \lambda_t & \alpha_q(R_t) \\ I_{p-1} & 0 & 0 & 0 \\ \xi_t & 0 & 0 & 0 \\ 0 & 0 & I_{q-2} & 0 \end{pmatrix},$$

where I_{p-1} and I_{q-2} are the identity matrices of size $p-1$ and $q-2$, respectively.

Let

$$\boldsymbol{B}_t = (\omega(R_t), 0, ..., 0)^T \in \mathbb{R}^{p+q-1}$$

and

$$\boldsymbol{X}_t = (h_{t+1}, ..., h_{t-p+2}, Y_t^2, ..., Y_{t-q+2}^2)^T \in \mathbb{R}^{p+q-1}.$$

Then Y_t is a solution of (1) if and only if $\boldsymbol{X}_t$ is a solution of

$$\boldsymbol{X}_{t+1} = \boldsymbol{A}_{t+1}\boldsymbol{X}_t + \boldsymbol{B}_{t+1}, \quad t \in \mathbb{Z}. \tag{2}$$

The following lemma is adapted from Brandt (1986) and Francq et al. (2001), also see Bougerol and Picard (1992).

Lemma 1 *For some operator norm of matrices, suppose that* $E(\log^+ \parallel \boldsymbol{A}_0 \parallel)$ *is finite, where* $\log^+ x = \log x$ *if* $x > 1$ *and 0 otherwise. Define the top Lyapunov exponent* γ *for* $\{\boldsymbol{A}_t\}_{t\in\mathbb{Z}}$ *as*

$$\gamma = \inf\left\{E\frac{1}{n+1}\log \parallel \boldsymbol{A}_0\boldsymbol{A}_{-1}...\boldsymbol{A}_{-n} \parallel, n \in \mathbb{N}\right\}.$$

Then, equation (2) has a strictly stationary solution if and only if $\gamma < 0$. *Moreover, the stationary solution is unique, ergodic, and defined by*

$$\boldsymbol{X}_t = \sum_{k=0}^{\infty} \boldsymbol{A}_t\boldsymbol{A}_{t-1}\cdots\boldsymbol{A}_{t-k+1}\boldsymbol{B}_{t-k}.$$

We are now in a position to give a representation of GARCH equation from Berkes et al. (2003, Theorem 2.4). We will state it as a lemma.

Lemma 2 *For a strictly stationary and ergodic solution of (1), assume* $E\log h_0$ *finite and* η_0^2 *non-degenerate. Then*

$$h_t = c_0(R_t) + \sum_{1\le i\le\infty} c_i(R_t)Y_{t-i}^2, \quad \forall \ \ t \in \mathbb{Z} \tag{3}$$

with probability one and this representation is unique, where, by defining

$$\mathcal{A}_t(x) = \alpha_1(R_t)x + \alpha_2(R_t)x^2 + ... + \alpha_q(R_t)x^q$$

and

$$\mathcal{B}_t(x) = 1 - \beta_1(R_t)x - \beta_2(R_t)x^2 - ... - \beta_p(R_t)x^p,$$

the coefficients

$$c_0(R_t) = \frac{\omega(R_t)}{\mathcal{B}_t(1)}$$

and

$$c_n(R_t) = \frac{d^n}{dx^n}\left(\frac{\mathcal{A}_t(x)}{\mathcal{B}_t(x)}\right)_{x=0} \qquad 1 \le n \le \infty.$$

From now on, we will focus on model (3) and assume the process $\{Y_t\}_{t\in\mathbb{Z}}$ strictly stationary and ergodic.

Now, suppose that we are given a realization $\{y_1, ..., y_n\}$, while the Markov chain $\{R_t\}$ is not observed. The parameters to be estimated from $\{y_t\}$ are usually chosen to be

$$\boldsymbol{\theta} := \{p(k,l), \omega(s), \alpha_i(s), \beta_j(s), k \neq l, 1 \le k, l, s \le d, 1 \le i \le q, 1 \le j \le p\}.$$

The parameter space Θ, a subset of $\mathbb{R}^{d^2+pd+qd}$ contains parameters satisfied assumptions we made and also the true parameter $\boldsymbol{\theta}_0$. As Leroux (1992, pp.135) pointed out, the stationary distribution $\pi(s), 1 \le s \le d$ will not affect the estimation. Here we assume orders p, q and d are known. See Rydén (1995) and Francq et al. (2001) for order selection.

As usual we use QMLE for GARCH and regime-switching GARCH models. Assume that η_t is standardly normally distributed for this moment. We get the likelihood function (r_t denotes the value of R_t):

$$p_{\boldsymbol{\theta}}(y_1, ..., y_n) = \sum_{(r_1,...,r_n)\in\mathbf{E}^n} \pi(r_1)\left\{\prod_{t=2}^{n} p(r_{t-1}, r_t)\right\}\left\{\prod_{t=1}^{n} f_{r_t}(y_1, ..., y_t)\right\}, \quad (4)$$

where

$$f_{r_t}(y_1, ..., y_t) = \frac{1}{(2\pi h_{r_t}(y_1, ..., y_{t-1}))^{1/2}} \exp\{-\frac{y_t^2}{2h_{r_t}(y_1, ..., y_{t-1})}\},$$

and the conditional variance process follows

$$h_{r_t}(y_1, ..., y_{t-1}) = c_0(r_t) + \sum_{1\le i\le t-1} c_i(r_t)y_{t-i}^2,$$

start with

$$h_{r_1} = c_0(r_1), \quad h_{r_2}(y_1) = c_0(r_2) + c_1(r_2)y_1^2$$

and continue recursively. Recursive formula for $\{c_i\}$ are available in Berkes et al. (2003, pp.210-211).

We can write (4) as a product of matrices. Define vector $\mathbf{1} = (1, ..., 1)^T \in \mathbb{R}^d$, $\mathbf{p} = (\pi(1)f_1(y_1), ..., \pi(d)f_d(y_1))^T \in \mathbb{R}^d$ and matrix $\boldsymbol{M}_{\boldsymbol{\theta}}(y_1, ..., y_t) =$

$$\begin{pmatrix} p(1,1)f_1(y_1,...,y_t) & p(2,1)f_1(y_1,...,y_t) & ... & p(d,1)f_1(y_1,...,y_t) \\ p(1,2)f_2(y_1,...,y_t) & p(2,2)f_2(y_1,...,y_t) & ... & p(d,2)f_2(y_1,...,y_t) \\ ... & ... & ... & \\ p(1,d)f_d(y_1,...,y_t) & p(2,d)f_d(y_1,...,y_t) & ... & p(d,d)f_d(y_1,...,y_t) \end{pmatrix},$$

then the likelihood is

$$p_{\boldsymbol{\theta}}(y_1, ..., y_n) = \mathbf{1}^T \left\{ \prod_{t=2}^{n} \boldsymbol{M}_{\boldsymbol{\theta}}(y_1, ..., y_t) \right\} \mathbf{p}, \tag{5}$$

which we can maximize numerically.

Because the indices of the states of the Markov chain can be permuted without changing the law of the model, the parameters are not strictly identifiable. So we will introduce one identifiability condition.

Identifiability Condition: For all $\boldsymbol{\theta} \in \Theta$, if $p_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, Y_{t-2}, ...) = p_{\boldsymbol{\theta}_0}(Y_t|\ Y_{t-1}, Y_{t-2}, ...)$, $P_{\boldsymbol{\theta}_0} - a.s.$, then $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

Here $p_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, Y_{t-2}, ...)$ is the density of Y_t given $Y_{t-1}, Y_{t-2}, ...$. Similarly define $p_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, ..., Y_1)$. (Let $p_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, ..., Y_1) = p_{\boldsymbol{\theta}}(Y_1)$ when $t = 1$). Let $g_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, Y_{t-2}, ...)$ and $g_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, ..., Y_1)$ be their corresponding logarithms. The existences of $p_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, Y_{t-2}, ...)$ and the expectation of $g_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, Y_{t-2}, ...)$ with respect to $P_{\boldsymbol{\theta}_0}$ can be shown as in Leroux (1992).

3 The main result

We will prove the consistency of QMLE here. Our method is benefited from Francq et al. (2001) and Xie and Yu (2005) .

Lemma 3 *Let $\tilde{p}_{\boldsymbol{\theta}}(.)$ be the density of $(Y_1, ..., Y_n)$ given $\{Y_0, Y_{-1}, ...\}$. Then for all $\boldsymbol{\theta} \in \Theta$,*

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \log p_{\boldsymbol{\theta}}(Y_1, ..., Y_n) &= \lim_{n\to\infty} \frac{1}{n} \log \tilde{p}_{\boldsymbol{\theta}}(Y_1, ..., Y_n|Y_0, Y_{-1}, ...) \\ &= E_{\boldsymbol{\theta}_0} \log g_{\boldsymbol{\theta}}(Y_t|Y_{t-1}, ...). \end{aligned} \tag{6}$$

PROOF. First, note that

$$\log \tilde{p}_{\boldsymbol{\theta}}(Y_1, ..., Y_n | Y_0, Y_{-1}, ...) = \sum_{t=1}^{n} g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, Y_{t-2}, ...) \quad (7)$$

and

$$\log p_{\boldsymbol{\theta}}(Y_1, ..., Y_n) = \sum_{t=1}^{n} g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, ..., Y_1).$$

Write

$$\begin{aligned} &\frac{1}{n}\sum_{t=1}^{n} g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, ..., Y_1) \\ &= \frac{1}{n}\sum_{t=1}^{n} g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, Y_{t-2}, ...) \\ &\quad + \frac{1}{n}\sum_{t=1}^{n} \{g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, ..., Y_1) - g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, Y_{t-2}, ...)\}. \end{aligned} \quad (8)$$

Analogous to Karlin and Taylor (1975, pp.502), define

$$\phi_N(y_0, y_{-1}, ...) = \sup_{l \geq N} |g_{\boldsymbol{\theta}}(y_0 | y_{-1}, ..., y_{-l}) - g_{\boldsymbol{\theta}}(y_0 | y_{-1}, y_{-2}, ...)|,$$

and

$$Z_t^N = \phi_N(Y_t, Y_{t-1}, ...).$$

Then $\{Z_t^N\}$ is stationary, ergodic, and $E[|Z_t^N|] < \infty$. We have

$$\begin{aligned} &\limsup_{n\to\infty} \left| \frac{1}{n}\sum_{t=1}^{n} \{g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, ..., Y_1) - g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, Y_{t-2}, ...)\} \right| \\ &\leq \limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \left| g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, ..., Y_1) - g_{\boldsymbol{\theta}}(Y_t | Y_{t-1}, Y_{t-2}, ...) \right| \\ &\leq \limsup_{n\to\infty} \frac{1}{n}\sum_{t=N+1}^{n} Z_t^N = E[Z_1^N]. \end{aligned}$$

But as $N \to \infty$, $Z_1^N \to 0$, and the interchange of limit and expectation can be justified to give $\lim_{N\to\infty} E[Z_1^N] = 0$. So the second term on the right-hand side of (8) goes to zero as $n \to \infty$. And the second equality in (6) follows by applying ergodic theorem, which completes the proof. □

We will next compare the likelihood $p_{\boldsymbol{\theta}}(Y_1, ..., Y_n)$ with the one evaluated at the true parameter $\boldsymbol{\theta}_0$, $p_{\boldsymbol{\theta}_0}(Y_1, ..., Y_n)$. Define

$$O_n(\boldsymbol{\theta}) = \frac{1}{n} \log \frac{p_{\boldsymbol{\theta}}(Y_1, ..., Y_n)}{p_{\boldsymbol{\theta}_0}(Y_1, ..., Y_n)},$$

and the following lemma follows from Lemma 3, Jensen's inequality and identifiability assumption.

Lemma 4 *For all $\boldsymbol{\theta} \in \Theta$, with probability one,*

$$\lim_{n\to\infty} O_n(\boldsymbol{\theta}) \leq 0$$

and the limit is almost surely equal to zero if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

Lemma 5 *For any $\boldsymbol{\theta}_1 \in \Theta$, $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_0$, there exists a neighborhood $V(\boldsymbol{\theta}_1)$ of $\boldsymbol{\theta}_1$ such that*

$$\limsup_{n\to\infty} \sup_{\boldsymbol{\theta} \in V(\boldsymbol{\theta}_1)} O_n(\boldsymbol{\theta}) < 0 \quad \text{a.s.}$$

PROOF. The proof is similar to Francq et al. (2001, Lemma 4) and Xie and Yu (2005, Lemma 3) and omitted here.

Lemma 5, together with the identifiability condition, proved the strong consistency of the QMLE over any compact subset containing $\boldsymbol{\theta}_0$.

Theorem 1 *Suppose that Θ^* is a compact subset of Θ and $\boldsymbol{\theta}_0 \in \Theta^*$; $(\hat{\boldsymbol{\theta}}_n)$ is a QMLE sequence satisfying almost surely*

$$p_{\hat{\boldsymbol{\theta}}_n}(Y_1, ..., Y_n) = \sup_{\boldsymbol{\theta} \in \Theta^*} p_{\boldsymbol{\theta}}(Y_1, ..., Y_n) \quad \forall\, n.$$

Then $(\hat{\boldsymbol{\theta}}_n)$ tends almost surely to $\boldsymbol{\theta}_0$ as $n \to \infty$.

4 Simulation

In our simulation studies we will illustrate the consistency, investigate asymptotic normality of QMLE and analyze a model specification problem. The most commonly used two-regime switching model is applied.

Given the recursive form of likelihood (5), we can numerically maximize it directly, following the suggestion of MacDonald and Zucchini (1997, pp.78-79). During the estimation, besides the standard constraints $0 \leq p(k,l) \leq 1, \sum_{k \neq l} p(k,l) \leq 1, \alpha_i(s) \geq 0, \beta_j(s) \geq 0$, and $\sum_{i=1}^{q} \alpha_i(s) + \sum_{j=1}^{p} \beta_j(s) \leq 1$, we also impose $\omega(s) \geq 0.001$ to avoid the underflow of numerical algorithm, and $\omega(1) \leq \omega(2) \leq ... \leq \omega(d)$ for identifiability. Several randomly chosen starting point set are used to pursue the global maxima.

4.1 Consistency

In the first experiment to verify the consistency of QMLE, we generate 100 independent series from model (1), each with size 5000. The true model has two regimes. The corresponding parameters are same as in Xie and Yu (2005): $\omega(1) = 1$, $\alpha_1(1) = 0.4$, $\beta_1(1) = 0.2$ and $\omega(2) = 20$, $\alpha_1(2) = 0.2$, $\beta_1(2) = 0.4$. The transition probabilities are set to $p(1,2) = 0.1$ and $p(2,1) = 0.1$. Table 1 summarizes the true values and the mean and standard deviation of estimators.

Table 1. *True values and the mean and standard deviation of QMLE of regime-switching GARCH model*

	$\hat{\omega}(1)$	$\hat{\alpha}_1(1)$	$\hat{\beta}_1(1)$	$\hat{\omega}(2)$	$\hat{\alpha}_1(2)$	$\hat{\beta}_1(2)$	$\hat{p}(1,2)$	$\hat{p}(2,1)$
True	1	0.4	0.2	20	0.2	0.4	0.1	0.11
mean	1.015	0.419	0.213	17.112	0.219	0.424	0.096	0.093
sd.	0.11	0.058	0.039	3.13	0.038	0.085	0.009	0.011

4.2 Asymptotic normality

Although the consistency of QMLE of our model is fairly satisfying from Table 1, we should bear in mind that we need distributional property of the estimates to utilize the information of standard deviations. With the second experiment, we want to examine the asymptotic behavior through simulation. There is no theoretical result available in the literature by now, even for the regime-switching ARCH model. Bickel et al. (1998) proved the asymptotic normality of MLE for general HMM model. However, their results cannot be readily extended to regime-switching GARCH model. By their simulation study, Xie and Yu (2005) conjectured that the asymptotic distribution of QMLE is indeed normal for the reduced regime-switching GARCH model.

First, note that by taking away the GARCH effect, i.e., letting $p = 0$ in model (1), our model is same as the reduced regime-switching model of Xie and Yu (2005). So, we will only consider regime-switching GARCH model here. We adapt the same parameter values as in first experiment. The QQ-plots with 0-1 line are shown in Figure 1. They rather conform to normal distribution. The P-values of Kolmogorov-Smirnov goodness-of-fit test verify our observation from QQ-plots. They are all above 0.1,[2] and the hypothesis of normal cannot be rejected in a reasonable level. Still, we can expect better conformity to normal distribution if we increase the length of sequences to, for example, 10000, from our experience of the improvement already achieved by increasing the length from 2000 to 5000. But we compromise for the vast time-consumption on computation and give up trying.

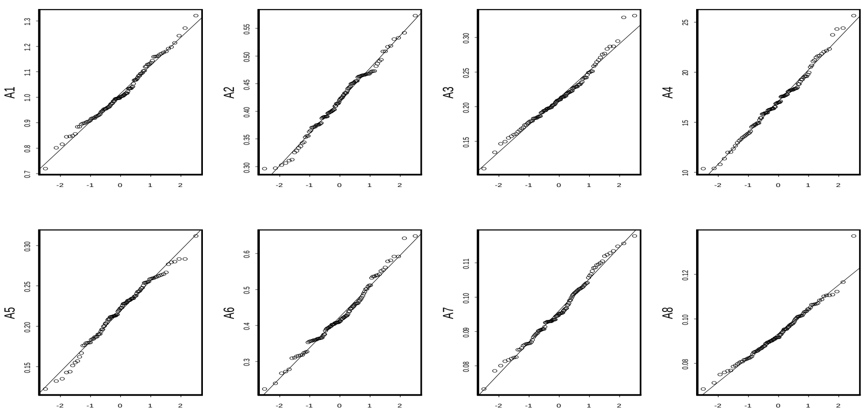

Figure 1: The QQ-plot of QMLE for regime-switching GARCH model with 0-1 line: A1–A8 represent estimators of $\omega(1)$, $\omega(2)$, $\alpha_1(1)$, $\beta_1(1)$, $\alpha_1(2)$, $\beta_1(2)$, $p(1,2)$ and $p(2,1)$, respectively.

4.3 A model specification problem

Xie and Yu (2005) observed that when we fit an ordinary GARCH to a regime-switching model, the estimated $\hat{\omega}$ and the amount of $\sum_{i=1}^{q}\hat{\alpha}_i+\sum_{j=1}^{p}\hat{\beta}_j$ change quite regularly along different scale of transition probabilities, which is inconsistent to our previous common knowledge which believes that $\sum_{i=1}^{q}\hat{\alpha}_i + \sum_{j=1}^{p}\hat{\beta}_j$ is always close to unit in that case. We want to investigate if it still holds true for our model.

[2]The P-value is calculated using the approximation of Dallal and Wilkinson (1986), which is most accurate for P-value less than 0.1. So these P-values are all set to 0.5 in statistical software S-plus®, with which we carry out our analysis.

Table 2 gives the estimate when data come from two regimes ARCH model. The true parameters have been used by Francq et al. (2001) and Xie and Yu (2005): $\omega(1) = 1$, $\alpha_1(1) = 0.6$, $\omega(2) = 100$, $\alpha_1(2) = 0$, $p(1,2) = p(2,1) = 0.01$. Then we change pair of $p(1,2)$ and $p(2,1)$ to (0.1,0.1), (0.2,0.2) and (0.5,0.5). We generate 100 independent series with size 1000 for each parameter set and apply to them the ordinary GARCH(1,1) model. The estimates for ω and sum of GARCH parameters are summarized. The conclusion is similar to Xie and Yu (2005): while $p(1,2)$ and $p(2,1)$ are small, $\hat{\alpha} + \hat{\beta}$ is close to 1, the so-called non-stationary region. However, as $p(1,2)$ and $p(2,1)$ increase, the estimates move from the non-stationary region to approximately the average of these two regimes, which is far from non-stationary. Table 3 gives us similar result when the data are from two-regime switching GARCH model.

Table 2. *Estimate result fitting ordinary GARCH(1,1) model for data from two-regime ARCH model with different transition probabilities, where $\hat{\omega}$, $\hat{\alpha}$ and $\hat{\beta}$ are estimators of parameters of the ordinary GARCH model*

$(p(1,2)\ p(2,1))$	(0.01,0.01)	(0.1,0.1)	(0.2,0.2)	(0.5,0.5)
$\hat{\omega}$	0.6695	8.5987	22.3172	41.0445
$\hat{\alpha}+\hat{\beta}$	0.9994	0.9587	0.7046	0.4205

Table 3. *Estimate result fitting ordinary GARCH(1,1) model for data from two-regime GARCH model with different transition probabilities*

$(p(1,2)\ p(2,1))$	(0.01,0.01)	(0.1,0.1)	(0.2,0.2)	(0.5,0.5)
$\hat{\omega}$	0.3192	2.621	5.8904	10.0464
$\hat{\alpha}+\hat{\beta}$	0.9918	0.8683	0.6498	0.429

5 Discussion

Regime-switching GARCH is a desirable model in that along the line of HMM it incorporates the idea of regime-switching into the somehow restrictive GARCH, which significantly extends GARCH model. Moreover, not only the parameters in GARCH equation but also the structure of GARCH model and error distribution can be specified to depend on regimes, c.f. Gray (1996) and the review by Hamilton and Raj (2002). However, because of the regime path dependence problem, only regime-switching ARCH (e.g., Cai (1994) and Hamilton and Susmel (1994)) or some kind of reduced GARCH model (Gray

(1996) and Klaassen (2002)) is feasible in practice. To author's best knowledge, there was even no algorithm available for computation of likelihood (4) in general case. In this paper, benefited from the transform proposed by Berkes et al. (2003), we rewrite the GARCH model as an ARCH(∞) form, which makes the maximization of the likelihood possible. The consistency of the QMLE is obtained.

Asymptotics is another issue we are interested in. We can see from the simulation study that the asymptotic normality for regime-switching GARCH model is quite promising. However we see no promise for the theoretical proof along the line we did for consistency. Bickel et al. (1998) obtained the asymptotic normality for general HMM, but the independence of observations given the unobserved regime is crucial in their framework and can not be relaxed to include our model. This remains an open problem.

As for the model specification problem, we observed that if the two regimes of true process switch to each other rather often, the sequence will be taken as a stationary (ordinary) GARCH model, while it is non-stationary if regimes seldom switch. This implies that even a stationary GARCH process can actually results from a regime-switching GARCH model, which we seldom realized before.

Acknowledgements

The author is grateful to Prof. Bo Ranneby and Dr. Jun Yu whose efforts are invaluable for this article.

References

Berkes, I., Horváth, L. and P.Kokoszka. (2003), GARCH processes: structure and estimation, *Bernoulli* **9**, 201-227

Bickel, P.J., Ritov, Y. and T. Rydén. (1998), Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models, *Ann. Statist.* **26**, 1614-1635.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity, *J. Econometrics* **31**, 307-327.

Bougerol, P. and N. Picard. (1992), Strict stationary of generalized autoregressive processes, *Ann. Probab.* **20**, 1714-1729.

Brandt, A. (1986), The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients, *Adv. Appl. Probab.* **18**, 221-254.

Cai, J. (1994), A Markov model of switching-regime ARCH, *J. Bus. Econom. Statist.* **12**, 309-316.

Dallal, G.E. and L. Wilkinson. (1986), An analytic approximation to the distribution of Lilliefors's test statistic for normality, *Amer. Statistician* **40**, 294-296.

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of U.K. inflation, *Econometrica* **50**, 987-1008.

Francq, C., Roussignol, M. and J. Zakoïan. (2001), Conditional heteroskedasticity driven by hidden Markov chains, *J. Time Ser. Anal.* **22**, 197-200.

Gray, S.F. (1996), Modelling the conditional distribution of interest rates as a regime-switching process, *J. Finan. Econom.* **42**, 27-62.

Hamilton, J.D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* **57**, 357-384.

Hamilton, J.D. and B. Raj. (2002), New directions in business cycle research and financial analysis, *Empirical Econom.* **27**, 149-162

Hamilton, J.D. and R. Susmel. (1994), Autoregressive conditional heteroskedasticity and changes in regime, *J. Econometrics* **64**, 307-333.

Karlin, S. and H.M. Taylor. (1975), *A First Course in Stochastic Processes* (Academic Press, New York.).

Klaassen, F. (2002), Improving GARCH volatility forecasts with regime-switching GARCH, *Empirical Econom.* **27**, 363-394.

Lamoureux, C.G. and W.D. Lastrapes. (1990), Persistence in variance, structural change and the GARCH model, *J. Bus. Econom. Statist.* **8**, 225-234

Leroux, B.G. (1992), Maximum-likelihood estimation for hidden Markov models, *Stochastic Process. Appl.* **40**, 127-143.

MacDonald, I.L., and W. Zucchini. (1997), *Hidden Markov and Other Models for Discrete-valued Time Series* (Chapman and Hall, London.).

Rydén, T. (1995), Estimating the order of hidden Markov models, *Statist.* **26**, 345-354.

Xie, Y., and J. Yu. (2005), Consistency of quasi-maximum likelihood estimators for the reduced regime-switching GARCH models, *Research Report 2005:2, Centre of Biostochastics, SLU, Umeå.*