5.1. Self-weighted QMLE

1. Introduction

2. Model and Assumptions

3. Self-weighted maximum Likelihood Estimators

1. Introduction

The AR-ARCH model in Engle (1982):

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t, \\ \varepsilon_t &= \eta_t \sqrt{h_t}, \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_r \varepsilon_{t-r}^2, \end{aligned}$$

where $\alpha_0 > 0$ and $\alpha_i \geq 0$, and $\eta_t \sim$ i.i.d (0,1)

Bollerslev (1986): GARCH model

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t, \\ \varepsilon_t &= \eta_t \sqrt{h_t}, \\ h_t &= \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}, \end{aligned}$$

Model:

$$
\begin{aligned}
y_t &= \phi y_{t-1} + \varepsilon_t, \\
\varepsilon_t &= \eta_t \sqrt{h_t}, \\
h_t &= ARCH/GARCH.
\end{aligned}
$$

Weiss (1986):

Consistency
asymptotic normality of MLE.

Tsay (1987) and Pantula (1989) for AR-ARCH models.

Ling and Li (1997), Ling and McAleer (2003a) and Francq and Zakoian (2004) for ARMA-GARCH models.

Condition: $E\varepsilon_t^4 < \infty$.

Moment condition of ε_t links directly to the **restriction on parameters** in h_t.

The ARCH(1) model with $\eta_t \sim N(0,1)$:

ε_t	Strict stat.	2nd m.	4th m.	8th m.
α_1	(0, 3.56)	(0, 1)	(0, 0.57)	(0, 0.33)

One is interested in the asymptotic theory for ARCH -type models under weak moment conditions

Model:

$$
\begin{aligned}
y_t &= 0 \text{ or constant} + \varepsilon_t, \\
\varepsilon_t &= \eta_t \sqrt{h_t}, \\
h_t &= ARCH/GARCH.
\end{aligned}
$$

Lee and Hansen (1994) and Lumsdaine (1996):

GARCH(1,1) model: $h_t = \alpha_0 + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$

(1) $\alpha + \beta < 1$, $E\varepsilon_t^2 < \infty$.

(2) $\alpha + \beta = 1$, called IGARCH(1,1)models.

(3) $E \ln(\alpha\eta_t^2 + \beta) < 0$.

Under (3), QMLE is consistent and asymptotically normal.

GARCH(r, s): $h_t = \alpha_0 + \sum_{i=1}^r \alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^s \beta_i h_{t-i}$

(1) $\sum_{i=1}^r \alpha_i + \sum_{i=1}^s \beta_s < 1$, $E\varepsilon_t^2 < \infty$.

(2) $\sum_{i=1}^r \alpha_i + \sum_{i=1}^s \beta_s = 1$, called IGARCH(r,s).

(3) $E \ln \|A_{t-1} \cdots A_{t-k}\| < 0$ for some $k > 0$.

Is QMLE consistent and asymptotically normal under (3)?

Berkes, Horvath and Kokoszka (2003):

(3) + condition

$$\lim_{s\to 0} s^{-\iota} P(\eta_t^2 \le s) = 0 \text{ and } E|\eta_t|^{4+\iota_1} < \infty,$$

for some positive ι and ι_1.

Hall and Yao (2003) under (1)

[discussed the case when $E\eta_t^4 = \infty$].

Francq and Zakoian (2004) under (3).

Ling and Li (1997) under $E\varepsilon_t^4 < \infty$.

Difficulty is normality and Key is

$$E \sup_{\Theta} \Big[\frac{\varepsilon_t^2}{h_t} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta'}\Big] < \infty.$$

Consistency of MLE: Jeantheau (1998) under (3)
Ling and McAleer (2003) under (2)

However,

(1) the zero or constant conditional mean is not been supported in applications.

(2) The condition for $E\varepsilon_t^4 < \infty$ is not satisfied.

Examples are in Engle (1982), Weiss (1984), Bollerslev (1986), Tsay (1987), Li and Li (1996), Ding, Granger and Engle (1993), Baillie, Chung and Tiles (1996) and Baillie (1996), among many others.

Important to develop the asymptotic inference theory for GARCH models with dynamic conditional means with $E\varepsilon_t^4 = \infty$.

2 Model and Assumptions

ARMA-GARCH model:

$$
\begin{aligned}
y_t &= \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t, \\
\varepsilon_t &= \eta_t \sqrt{h_t}, \\
h_t &= \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i},
\end{aligned}
$$

where $\alpha_i \geq 0$ and $\beta_j \geq 0$, $i = 0, \cdots, r$, $j = 1, \cdots, s$, and η_t is a sequence of i.i.d.(0,1).

Denote

$\gamma = (\mu, \phi_1, \cdots, \phi_p, \psi_1, \cdots, \psi_q)'$,

$\delta = (\alpha_0, \alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s)'$,

$\theta = (\gamma', \delta')'$.

$\Theta_\gamma \subset R^{p+q+1}$ and $\Theta_\delta \subset R_0^{r+s+1}$ are compact.

$\Theta = \Theta_\gamma \times \Theta_\delta$ and $m = p + q + r + s + 2$.

True parameter θ_0 of θ is an interior point in Θ

Assumption 2.1 *θ_0 is an interior point in Θ and for each $\theta \in \Theta$, $\phi(z) \neq 0$ and $\psi(z) \neq 0$ when $|z| \leq 1$, and $\phi(z)$ and $\psi(z)$ have no common root with $\phi_p \neq 0$ or $\psi_q \neq 0$.*

Assumption 2.2 *$\alpha(z)$ and $\beta(z)$ have no common root, $\alpha(1) \neq 0$, $\alpha_r + \beta_s \neq 0$, and $\sum_{i=1}^{s} \beta_i < 1$ for each $\theta \in \Theta$.*

Assumption 2.3 *η_t^2 has a non-degenerate distribution with $E\eta_t^2 = 1$.*

Assumption 2.4 *$E|\varepsilon_t|^{2\iota} < \infty$ for some $\iota > 0$.*

When $\iota = 1$, the necessary and sufficient condition for Assumption 2.4 is

$$\sum_{i=1}^{r} \alpha_{0i} + \sum_{i=1}^{s} \beta_{0i} < 1.$$

When $\iota \in (0, 1]$, we have the following theorem:

Theorem 2.1 *Let* $\iota \in (0, 1]$ *and* $\tilde{\beta} = \min\{\alpha_{0i}, \beta_{0j} : i = 0, 1, \cdots, r, j = 1, \cdots, s\}$. *(i) If*

there exists an integer i_0 *such that*

$$E\|\prod_{k=0}^{i_0-1} A_k\|^{\iota} < 1, \tag{2.7}$$

then $\{\varepsilon_t\}$ *is strictly stationary and ergodic with* $E|\varepsilon_t|^{2\iota} < \infty$*;*

(ii) If $\tilde{\beta} > 0$ *and* $\{\varepsilon_t\}$ *is strictly stationary with* $E|\varepsilon_t|^{2\iota} < \infty$*, then (2.7) holds;*

(iii) if $\tilde{\beta} > 0$*,*

$$\sum_{i=1}^{r} \alpha_{0i} + \sum_{j=1}^{s} \beta_{0j} = 1,$$

and η_t *has density* $f(x) > 0$ *on* R *s.t.* $E|\eta_t|^{\tau} < \infty$ *for all* $\tau < \tau_0$ *and* $E|\eta_t|^{\tau_0} = \infty$ *for some* $\tau_0 \in (0, \infty]$*, then* $\lim_{x\to\infty} x^2 P(|\varepsilon_t| > x)$ *exists and* > 0*.*

By (iii), the tail index of IGARCH(r, s) process is always 2. When $r = s = 1$, this tail index was obtained by Basrak, Davis and Mikosch (2002).

3 Self-weighted QMLE

Model:

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t, \\ \varepsilon_t &= \eta_t \sqrt{h_t}, \\ h_t &= ARCH/GARCH. \end{aligned}$$

Let the LSE of ϕ be $\widehat{\phi}_n$.

$$\sqrt{n}(\widehat{\phi}_n - \phi) = \Big(\frac{1}{n}\sum_{t=1}^{n} y_{t-1}^2\Big)^{-1}\Big(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} y_{t-1}\varepsilon_t\Big).$$

We need that $E(y_{t-1}\varepsilon_t)^2 = E(y_{t-1}^2 h_t) < \infty$.

Condition $E\varepsilon_t^4 < \infty$ should be minimal for the asymptotic normality of $\widehat{\phi}_n$.

It is similar to the AR model with i.i.d errors.

The minimal condition for the asymptotic normality of LSEs is $Ey_t^2 < \infty$.

When $Ey_t^2 = \infty$, all the existing estimators, such as LSE, LAD and M-estimators, do not have a closed form, see Davis, Knight and Liu (1992).

Ling (2005) introduced a self-weighted LAD estimator which is asymptotically normal even if $Ey_t^2 = \infty$.

The purpose of the weighting in Ling (2005) is to downweight the covariance matrix such that asymptotic normality can be recovered.

We try to use this idea to the LR function:

$$L_n(\theta) = -\frac{1}{2n}\sum_{t=1}^{n}\Big[\log h_t(\theta) + \frac{\varepsilon_t^2(\gamma)}{h_t(\theta)}\Big],$$

where

$$\varepsilon_t(\gamma) \;=\; y_t - \mu - \sum_{i=1}^{p}\phi_i y_{t-i} - \sum_{i=1}^{q}\psi_i\varepsilon_{t-i}(\gamma).$$

Main issue is that we cannot prove:

$$E\sup_{\Theta} A_t(\theta) < \infty,$$

where

$$A_t(\theta) = \frac{\varepsilon_t^2(\gamma)}{h_t(\theta)}\Big[\frac{1}{h_t(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\Big]\Big[\frac{1}{h_t(\theta)}\frac{\partial h_t(\theta)}{\partial\theta'}\Big].$$

We can show that

$$\sup_{\Theta} A_t(\theta) \le C\xi_{\rho t-1}^4(1+\eta_t^2),$$

where $\xi_{\rho t} = 1 + \sum_{i=0}^{\infty} \rho^i |y_{t-i}|$, and $\rho \in (0,1)$ and C are some constants not depending on t.

This is why we generally need $E\varepsilon_t^4 < \infty$. We try to downweight $\xi_{\rho t-1}^4$.

Assumption 3.1 $w_t = w(y_{t-1}, y_{t-2}, \cdots) > 0$ and w is a measurable, and bounded function on R^{Z_0} with $E(w_t\xi_{\rho t-1}^4) < \infty$ and $Z_0 = \{0, 1, 2, \cdots\}$.

Thus, we introduce the weighted log-quasi-likelihood function:

$$L_{sn}(\theta) = \frac{1}{n}\sum_{t=1}^{n} w_t l_t(\theta).$$

In practice, we do not have the initial values y_i when $i \le 0$ and hence they have to be replaced by some constants.

Denote $\varepsilon_t(\gamma)$, $h_t(\theta)$ and w_t as $\tilde{\varepsilon}_t(\gamma)$, $\tilde{h}_t(\theta)$ and $\tilde{w}_t$, respectively, when y_i is a constant not depending on parameters when $i \leq 0$.

Assumption 3.2 $E|w_t - \tilde{w}_t|^{\iota_0/4} = O(t^{-2})$, where $\iota_0 = \min\{\iota, 1\}$.

The weighted LR function $L_{sn}(\theta)$ is modified as

$$\tilde{L}_{sn}(\theta) = -\frac{1}{2n}\sum_{t=1}^{n} w_t\Big[\log \tilde{h}_t(\theta) + \frac{\tilde{\varepsilon}_t^2(\gamma)}{\tilde{h}_t(\theta)}\Big].$$

Let $\widehat{\theta}_{sn} = argmax_{\Theta}\tilde{L}_{sn}(\theta)$.

Since the weight w_t is determined by $\{y_t\}$ itself,

$\widehat{\theta}_{sn}$ is called the self-weighted QMLE.

Theorem 3.1 *Suppose that Assumptions 2.1-2.4 and 3.1-3.2 hold. Then,*

$$
\begin{aligned}
(i) &\quad \widehat{\theta}_{sn} \longrightarrow_p \theta_0, \\
(ii) &\quad \sqrt{n}(\widehat{\theta}_{sn} - \theta_0) \longrightarrow_{\mathcal{L}} N(0, \Sigma_0^{-1}\Omega_0\Sigma_0^{-1}),
\end{aligned}
$$

if $E\eta_t^4 < \infty$ *and* $J > 0$

where $\Sigma_0 = E[w_t U_t(\theta_0) U_t'(\theta_0)]$,

$\Omega_0 = E[w_t^2 U_t(\theta_0) J U_t'(\theta_0)]$,

$J = \begin{pmatrix} 1 & \kappa_3 \\ \kappa_3 & \kappa \end{pmatrix}$,

$\kappa_3 = E\eta_t^3/\sqrt{2}$,

$\kappa = (E\eta_t^4 - 1)/2$ *and*

$$
U_t(\theta) = [\frac{1}{\sqrt{h_t}} \frac{\partial \varepsilon_t(\gamma)}{\partial \theta}, \frac{1}{\sqrt{2}h_t} \frac{\partial h_t(\theta)}{\partial \theta}].
$$

To use the result, we need to select a weight w_t. Many weights satisfy Assumptions 3.1-3.2.

When $\iota = 1/2$ (i.e., $E|\varepsilon_t| < \infty$), as in Ling (2005), one natural weight is

$$w_t = (\max\{1, C^{-1}\sum_{k=1}^{\infty} \frac{1}{k^9}|y_{t-k}|I\{|y_{t-k}| > C\}\})^{-4},$$

for some $C > 0$.

It has a connection to Huber's robust estimator for the regression model.

When $q = s = 0$ (AR-ARCH model), for any $\iota > 0$, the weight can be selected as

$$w_t = (\max\{1, C^{-1}\sum_{k=1}^{p+r} \frac{1}{k^9}|y_{t-k}|I\{|y_{t-k}| > C\}\})^{-4}.$$

When $\iota \in (0, 1/2)$ (i.e., $E|\varepsilon_t| = \infty$) and $q > 0$ or $s > 0$, the weight needs to be modified as follows:

$$w_t = (\max\{1, C^{-1} \sum_{k=1}^{\infty} \frac{1}{k^{1+8/\iota}} |y_{t-k}| I\{|y_{t-k}| > C\}\})^{-4}.$$

$\iota =$?

Simulation method to verify the condition (2.7) for a possible ι.

Another way is to use the Hill estimator to estimate the tail index of $\{y_t\}$

The constant ι can be any value less than the tail index of $\{y_t\}$.

$C = ?$

when $E\|U_t(\theta_0)\|^2 < \infty$, $\Sigma_0^{-1}\Omega_0\Sigma_0^{-1} \to$ the asymptotic covariance matrix of the QMLE as $C \to \infty$.

When $E\|U_t(\theta_0)\|^2 = \infty$,

$$\|\Sigma_0^{-1}\Omega_0\Sigma_0^{-1}\| \leq \lambda\|\Sigma_0^{-1}\| \to 0 \text{ as } C \to \infty,$$

for some $\lambda > 0$. That is, the asymptotic variance of the self-weighted QMLE can be as small as we want only if C is large enough.