Bayesian inference on QGARCH model using the adaptive construction scheme

Tetsuya Takaishi
Hiroshima University of Economics
Outline

- Motivation
- Stylized facts of financial data
- Quadratic GARCH model
- Bayesian inference
  - Markov Chain Monte Carlo
  - Metropolis-Hastings method
  - Adaptive construction of proposal density
- Numerical simulations
- Summary
Motivation

In finance volatility is an important value to measure risk.

To forecast volatility we use a model which mimics the property of the volatility and captures the stylized facts of the financial time series.

popular model

GARCH-type model

Quadratic GARCH model

We need to estimate model parameters.

Markov Chain Monte Carlo method based on the Bayesian inference
The performance of the Markov Chain Monte Carlo method depends on how to implement it.

We want to find a method with a good performance.

Good performance $\Rightarrow$ autocorrelation between Monte Carlo data is small

Metropolis-Hastings method $\Rightarrow$ small statistical error

Our method:

We use the Student’s t-distribution for the proposal density.

We construct it adaptively to fit the Monte Carlo data.

We show that our method works well for the QGARCH model and gives small autocorrelation times.
Stylized facts of financial data

- price return

\[ r(t) = \ln(p(t)) - \ln(p(t - \Delta t)) \]

Stylized facts for price returns

Many empirical studies show some properties which cannot be obtained from Gaussian noise

- fat tailed distribution
- volatility clustering
- absence of autocorrelations in return
- long time correlation in absolute return
- etc
Oct. 19 1987

Gopikrishnan et al., cond-mat/9905305

Gaussian

price return

volatility clustering

\( S&P\, 500 \)
Fat-tailed distribution

Normalized S&P500 returns

$P(x) \sim x^{-(1+\alpha)}$

Gopikrishnan et al., cond-mat/9905305
Quadratic GARCH model

GARCH(1,1) model  Bollerslev(1986)

time series \( y_t \)

\[ y_t = \sigma_t \varepsilon_t, \]

\[ \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \]

QGARCH model  Engle(1993)

Sentana(1995)

\[ \sigma_t^2 = \omega + \gamma y_{t-1} + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \]

Asymmetry

\[ \varepsilon_t \sim N(0,1) \]

Gaussian noise with mean 0 and variance 1

\[ \sigma_t : \text{volatility} \]

Our task is to evaluate \( \alpha, \beta, \gamma \) and \( \omega \) from the time series \( y_t \).
Bayesian inference

Bayes’ theorem

\[ \pi(\theta \mid y) = \frac{L(y \mid \theta)\pi(\theta)}{f(y)} \]

\( \theta = \alpha, \beta, \gamma, \omega \)

\( \pi(\theta \mid y) \) : posterior distribution

\( L(y \mid \theta) \) : likelihood function of QGARCH model

\( \pi(\theta) \) : prior distribution

Probability distribution of \( \theta \)

\[ \pi(\theta \mid y) \propto f(y \mid \theta)\pi(\theta) \]

If there is some information on theta, then we use it for \( \pi(\theta) \)

\[ \pi(\theta) = const. \]

Bayes’ theorem tells us the probability distribution of theta
Likelihood function of QGARCH model

\[ L(y \mid \theta = \alpha, \beta, \gamma, \omega) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{y_t^2}{\sigma_t^2}\right) \]

\[ y_t : \text{time series} \]

\[ \sigma_t^2 = \omega + \gamma y_{t-1} + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \]
Markov Chain Monte Carlo

$$\langle \theta \rangle = \frac{1}{Z} \int \theta \pi(\theta \mid y) d\theta$$

Model parameters are evaluated as expectation values.

probability distribution of theta

We use the Markov Chain Monte Carlo method to integrate it numerically.

Z: normalization constant

First, generate theta with probability distribution: $$\pi(\theta \mid y)$$

We obtain a set of theta: $$(\theta_1, \theta_2, \theta_3, \ldots, \theta_n)$$

$$\langle \theta \rangle \approx \frac{1}{n} \sum_{i=1}^{n} \theta_i$$

The generation of theta is performed by MCMC.

$$\delta \langle \theta \rangle \propto \sqrt{\frac{2\tau}{n}}$$

$$\tau : \text{autocorrelation time}$$
**Metropolis-Hastings method**

Consider to generate $\theta$ from $\pi(\theta \mid y)$

1. Draw $\theta' \mid y$ from a proposal density $g(\theta)$

2. Then, accept $\theta'$ with the following acceptance. Otherwise keep the old $\theta$

$$P_{\text{accept}} = \min\left(1, \frac{\pi(\theta' \mid y) \ g(\theta)}{\pi(\theta \mid y) \ g(\theta')}\right)$$

The performance of the method depends on $g(\theta)$. 

---

Metropolis et. al (1953) 
Hastings (1970)
Adaptive construction of proposal density

Posterior distribution of GARCH model

\[ \pi(\theta = \alpha, \beta, \omega \mid y) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi \sigma_t^2}} \exp \left( - \frac{y_t^2}{\sigma_t^2} \right) \]

Distributions are similar with Gaussian.

To cover the tails of the distribution we use multi-dimensional Student’s t-distributions.

Mitsui, Watanabe (2003)
Asai(2006)

The parameters are determined by the maximum likelihood method. They showed a good performance of the method.

Here we determine the parameters of the Student’s t-dist. from the MCMC results.
4-d student-t distribution

\[ P(\theta) = \frac{\Gamma\left(\frac{\nu + 3}{2}\right)}{(\det \Sigma)^{1/2} \Gamma\left(\frac{\nu}{2}\right)} \sqrt{\frac{1}{(\pi \nu)^3}} \left[ 1 + \frac{(\theta - m)^t \Sigma^{-1} (\theta - m)}{\nu} \right]^{-\frac{\nu + 4}{2}} \]

\[ \theta = (\alpha, \beta, \gamma, \omega) \]

We need the following parameters.

Average \( m = (\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\omega}) \)

Covariance matrix
\[ E((\theta - m)(\theta - m)^t) = \frac{\nu \Sigma}{\nu - 2} \]

unknown parameters

We estimate these parameters from a short MCMC run.

The parameters can be updated adaptively during the MCMC simulation.
Numerical Simulations

Input data

\[ y_t = \sigma_t \varepsilon_t, \]
\[ \sigma_t^2 = \omega + \gamma y_{t-1} + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \]

We generated 2000 data with \( \alpha = 0.07 \), \( \beta = 0.8 \), \( \gamma = -0.05 \), \( \omega = 0.1 \).

1. We start with the Metropolis method and discard 3000 updates.
2. We accumulate 1000 data with the Metropolis method and estimate the parameters of the 4d-Student-t distribution.
3. We perform the Metropolis-Hastings method with the 4d-Student-t distribution.
4. We update the parameters of the 4d-Student-t distribution every 1000 updates.
5. We accumulate 100000 data for analysis.
\[ y_t = \sigma_t \varepsilon_t, \quad \alpha = 0.07 \quad \beta = 0.8 \quad \gamma = -0.05 \quad \omega = 0.1. \]

\[ \sigma_t^2 = \omega + \gamma y_{t-1} + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \]
Convergence of the covariance matrix

\[ V = E((\theta - m)(\theta - m)^t) = \frac{\nu \Sigma}{\nu - 2} \]
Acceptance at Metropolis-Hastings algorithm

\[ \nu = 10 \]
Adaptive Metropolis

time history of sampled data

Adaptive

Metropolis
Adaptive

Metropolis

\[ \beta \]

\[ t \]

\[ 10000 \quad 12000 \quad 14000 \]

\[ 0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1.0 \]
Autocorrelation function of $\alpha$
## Numerical Simulations

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>$\omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>0.07</td>
<td>0.8</td>
<td>-0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.071(18)</td>
<td>0.791(56)</td>
<td>-0.046(19)</td>
<td>0.105(35)</td>
</tr>
<tr>
<td>$2\tau$</td>
<td>4.1(13)</td>
<td>10(5)</td>
<td>3.0(4)</td>
<td>11(5)</td>
</tr>
</tbody>
</table>

**Metropolis**

<table>
<thead>
<tr>
<th></th>
<th>0.070(18)</th>
<th>0.794(53)</th>
<th>-0.047(19)</th>
<th>0.103(33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2\tau$</td>
<td>340(100)</td>
<td>840(280)</td>
<td>54(7)</td>
<td>820(290)</td>
</tr>
</tbody>
</table>

Autocorrelation time

$$\tau = \frac{1}{2} + \sum_{t=1}^{\infty} ACF(t)$$

- **Small autocorrelation time**
- **Large autocorrelation time**
Summary

• We used a multi-dimensional student-t distribution as a proposal density for the MH method.

• We estimated the parameters of the student-t distribution using MCMC results and updated the parameters adaptively during the MCMC simulation.

• The autocorrelations between the sampled data are decreased considerably by our method.

• The performance of our method is similar to the method with the maximum likelihood method.

• Our method is considered to be an alternative efficient method to the Bayesian inference of the QGARCH model.
Convergence of the covariance matrix

\[ V = E((\theta - m)(\theta - m)^t) = \frac{v\Sigma}{v - 2} \]
Acceptance at MH, measured every 1000
MC history of $\alpha$
Autocorrelation function

Adaptive

Metropolis
Acceptance at Metropolis-Hastings algorithm
Empirical analysis

US Dollar and Japanese Yen exchange rate
Sampling period: 1 March 2000 to 29 February 2008

\[ r(t) = 100 \times \left[ \ln(p(t)) - \ln(p(t - \Delta t)) \right] \]

(1) Start with Metropolis algorithm
(2) First 3000 data are discarded as thermalization
(3) Every 1000 we update the parameters of the proposal density and do MH update.
(4) Accumulate 100000 data for analysis
## Results

<table>
<thead>
<tr>
<th>Adaptive</th>
<th>α</th>
<th>β</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.03151(4)</td>
<td>0.9403(1)</td>
<td>0.01104(3)</td>
</tr>
<tr>
<td>2τ</td>
<td>2.8(3)</td>
<td>3.8(4)</td>
<td>4.1(5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metropolis</th>
<th>α</th>
<th>β</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0318(5)</td>
<td>0.9391(14)</td>
<td>0.0114(4)</td>
</tr>
<tr>
<td>2τ</td>
<td>400(60)</td>
<td>650(100)</td>
<td>620(80)</td>
</tr>
</tbody>
</table>
Metropolis method

Local update

We want to generate theta with \( \pi(\theta \mid y) \propto \exp(-f(\theta)) \)

- draw \( \theta' = \theta + d(\varepsilon - 0.5) \)
  - uniform random number in [0,1]
- calculate \( dh = f(\theta') - f(\theta) \)
- accept with \( \min(1, \exp(-dh)) \)

detailed balance condition

\[
P(\theta \leftarrow \theta')P(\theta') = P(\theta' \leftarrow \theta)P(\theta)
\]
Student-t (Tsallis) distribution

\[
P(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \sqrt{\frac{1}{\pi(\nu-2)}} \left[1 + \frac{1}{\nu-2} x^2\right]^{-\frac{\nu+1}{2}}
\]

\[
\nu = \frac{3-q}{q-1}
\]

Tsallis

\[
\frac{\Gamma\left(\frac{1}{q-1}\right)}{\Gamma\left(\frac{3-q}{2(q-1)}\right)} \sqrt{\frac{q-1}{\pi(5-3q)}} \left[1 + \frac{q-1}{5-3q} x^2\right]^{-\frac{1}{q-1}}
\]

\[\{\text{q} \to 1\} \text{ Gaussian distribution}\]
probability distribution of GARCH return

\[ \alpha = 0.1, \beta = 0.88, \omega = 1 - \alpha - \beta \]

Student’s t-distribution with \( \nu \approx 6 \)

q-Gaussian \( q=1.287 \)
(a) Price returns

Exponential decay, $\tau_{ck} = 4$ min

Noise level

(b) Absolute value of price returns

$-0.3$
Fig. 3.4  Autocorrelation of returns and absolute returns for the (a) Shanghai index and (b) NYSE as a function of the time-lag $\Delta t$. 
(a) Cumulative distribution

- Lévy Regime
  - $\alpha \approx 1.7$
  - $\alpha = 2$
  - $\alpha \approx 3$

- Positive tail
- Negative tail

Normalized S&P500 returns