Estimating GARCH Modeling Using Metropolis-Hastings Method in R

Min Wang1,2, Yunshun Wu1*

1School of Mathematical Sciences, Guizhou Normal University, Guiyang, China
2School of Mathematical Sciences, Xiamen University, Xiamen, China
Email: *wjmpum@gznu.edu.cn

Abstract
This paper mainly talks about a popular approach of volatility of a GARCH-type model in R, while the disturbances are independent and have identical Student-t distribution. It uses the Metropolis-Hastings method to perform the computations and gives the programs in details in R.

Keywords
Student’s t Distribution, Degree of Freedom, GARCH t Model, R, Metropolis-Hastings Method

1. Introduction
The R-project is an open statistics programming software. Modelling and forecasting volatility or, in other word, the covariance structure of asset returns, is important. A central property of economic time series, being common to many financial time series, is that their volatility varies over time. Describing the volatility of an asset is a key issue in financial economics. Returns were modelled as independent and identically distributed over time. The most popular class of models for time-varying volatility is represented by GARCH type models [1]. GARCH models are commonly used for describing, estimating and predicting the dynamics of financial returns. Recent surveys of the existing GARCH models literature can be found in Davidson [2] and Rombouts et al. [3]. In contrast to Engle’s [4] ARCH model, a double autoregressive (DAR) model, which is a special case of the ARMA-ARCH models in Weiss [5] and an example of weak ARMA models in Francq and Zakoian [6] [7], is also increasingly concerned about researchers. Under the assumption of the disturbance following a normal distribution, Ling [8] considered the structure and the maximum likelihood es-
In addition, stochastic volatility (SV) models have enjoyed great popularity in analyzing financial data in the last couple of decades [9].

In this paper, an R package named bayesGARCH [10] was mentioned to contrast with our procedure. This package makes use of the priors based on the work of Nakatsuma [11], consists of Metropolis-Hastings (MH) algorithm [12] a proper algorithm to sample the posterior distribution.

The remainder of this paper is organized as follows. In Section 2 we describe the GARCH (1, 1)-t process and introduce how to estimate the parameters by using the Metropolis-Hastings method. Through an example, we contrast our model with the bayesGARCH package in Section 3. Section 4 provides the R procedure to execute our model in details. An empirical example is reported in Section 5, and Section 6 concludes this paper.

2. Univariate GARCH-t Model

Let $Y_t$ denote an asset return. The general structure of an asset return series modeled by a GARCH-type model can be written as Audronė Virbickaitė et al. 2014 [13]:

$$Y_t = \mu_t + a_t = \mu_t + h_t \epsilon_t.$$  

In general, $\epsilon$ is a Normal variable. Without loss information, we set $\mu_t = 0$. To capture the fat tail so prominent a Student-t distribution is used for conditional density. The model is, GARCH-t, is

$$Y_t \mid \mathcal{F}_{t-1} \sim S_{\psi_{t}}(0, h_t^2), \quad h_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + \beta h_{t-1}^2. \quad (1)$$

where $S_{\psi_{t}}(0, h_t^2)$ is a Student-t distribution with mean 0 and $\psi_t$ is the degree of freedom parameter. $h_t$ is the conditional variance given $\mathcal{F}_{t-1}$ in the GARCH (1, 1) model. $\alpha_0 > 0, \quad \alpha_1, \beta \geq 0$ are restrictions for positive variance, and $\alpha + \beta < 1$ for the covariance stationarity. Then the posterior density function could be:

$$f(y_t \mid \mathcal{F}_{t-1}) = \frac{\Gamma\left(\frac{\psi_{t}+1}{2}\right)}{\frac{1}{\pi^2 \Gamma\left(\frac{1}{2}\psi_{t}\right)}} \left[\frac{(\psi_t - 2) h_t^2}{2} + 1 + \frac{y_t^2}{(\psi_t - 2) h_t^2}\right]^{\frac{1}{2}(\psi_{t}+1)} \cdot \frac{1}{\psi_t \cdot \Gamma\left(\frac{\psi_t}{2}\right)} \left[1 + \frac{\psi_t}{2}\right]^\frac{\psi_t}{2}. \quad (2)$$

It is indicated that the $\epsilon$ obey the Student-t distribution in (1). The following, we will show how we get (2), the density distribution of $\epsilon$ is

$$p(\epsilon) = \frac{\Gamma\left(\frac{\psi_{t}+1}{2}\right)}{\sqrt{\psi_t} \cdot \Gamma\left(\frac{\psi_t}{2}\right)} \left(\frac{\psi_t}{2}\right)^\frac{\psi_t-1}{2}.$$  

Through the equation $Y_t = h_t \cdot \epsilon_t$, the distribution of $Y_t$ is

$$P(Y_t \leq y_t) = P(h_t \cdot \epsilon_t \leq y_t) = P(\epsilon_t \leq \frac{y_t}{h_t^2}) = F_{\epsilon_t}\left(\frac{y_t}{h_t^2}\right).$$

So the density is
Because, \( Var(\epsilon_t) = \frac{\psi}{\psi - 2} \), to ensure the conditional variance of \( Y_t \) to be \( h_t^2 \), we replace \( "h_t" \) with \( \frac{\psi - 2}{\psi} h_t \), then the (3) changed into (2).

Given a dataset \( Y_t = (y_{1t}, y_{2t}, \cdots, y_{Tt}) \), model parameter \( \Gamma = (\alpha_0, \alpha, \beta, \psi) \) and the posterior density is

\[
p(\Gamma | Y_T) \propto p(\Gamma) \prod_{t=1}^{T} f(y_t | \psi, \ldots, \psi^t)
\]

According to Jensen [14], the priors are independent and identically distributed (IID) as \( N(0, 100) \) with the following restrictions \( \alpha_0 > 0, \alpha_i \geq 0, \beta \geq 0 \) to impose identification and \( \psi > 2 \). Since the variance is non-negative, the positive constraints on coefficients are reasonable. The restriction on the degrees of freedom parameter \( \psi \) ensures the conditional variance to be finite and the restrictions on the GARCH parameters \( \alpha_0, \alpha_i \) and \( \beta \) guarantee its positivity.

One of the most popular MCMC algorithm used in estimating GARCH model parameters, is the Metropolis-Hastings (MH) method. We emphasize the fact that only positivity constraints are implemented in the MH algorithm; no stationarity conditions are imposed in the simulation procedure. We employ a MH sampler. Given the current value \( \Gamma \) of the chain, the proposal \( \Gamma' \) is sampled from

\[
h(\Gamma') \sim \begin{cases} N(\Gamma, V) & \text{with probability } p \\ N(\Gamma, 100V') & \text{with probability } 1 - p \end{cases}
\]

where \( V \) is the inverse Hessian matrix of \( \log(p(\Gamma | Y_T)) \). The accepted probability equal to \( \min\left\{ \frac{p(\Gamma | Y_T)}{p(\Gamma' | Y_T)} \right\} \), and \( p = 0.9 \) empirically. After the test sample collecting the new \( \{\Gamma^{(i)}\}_{i=1}^{N} \), the predictive density is

\[
p(y_{t+1} | Y_t) = \frac{1}{N} \sum_{i=1}^{N} f(y_{t+1} | 0, h_{t+1}^{(i)}, \psi^{(i)}).
\]

3. The Priors and bayesGARCH Package

In this section, an R package named bayesGARCH was applied to deal with the
GRACH model, then we’ll describe the difference of the priors between our model and the package. The package bayesGARCH provides functions for the Bayesian estimation of the parsimonious and effective GARCH (1, 1) model with Student-t innovations. The priors in the bayesGARCH package distributions on $\alpha_0$ is a bivariate truncated Normal distribution; $\alpha_1$ is a univariate truncated Normal distribution; $\beta$ is a translated Exponential distribution. The estimation procedure is fully automatic and thus avoids the tedious task of tuning an MCMC sampling algorithm. It is obviously there are some differences set of priors between our model and the package, so we couldn’t cite the package directly. Next, we will show the estimation results to demonstration that the different priors could inference the estimation result.

3.1. Example

We set the initial value of $\Gamma$ is $(1, 0.2, 0.4, 5)$. 800 simulation values are obtained by simulation in $R$.

```r
R > T = 800
R > y = c(rep(0,T))
R > y[1] = 0.5; h[1] = 0.5
R > alpha0 = 1; alpha1 = 0.2; beta = 0.4; psi = 5
R > for(i in 2:T)
+ {  
+h[i] = alpha0 + alpha1 * y[i-1]^2 + beta * h[i-1]
+y[i] = rt(1, psi)
+y[i] = sqrt(h[i]) * y[i]
}
```

Then we use bayesGARCH package to estimate $\Gamma$, we remain the number of MCMC chains and the length of each MCMC chain to be the default value.

```r
MCMC <- bayesGARCH(y)
smpl <- formSmpl(MCMC, lbi = 50)
summary(smpl)
```

We receive the estimation of $\Gamma$ is $(1.2630, 0.2367, 0.4829, 11.3285)$. Obviously, there’s a big difference estimation of $\psi$ and the estimation of other parameters is equally unsatisfactory.

3.2. The Priors

The difference sets of the priors see Table 1. The GARCH (1, 1) model with Student-t innovations in package bayesGARCH is written in Geweke [15].
Table 1. The difference sets of priors.

<table>
<thead>
<tr>
<th></th>
<th>our GARCH-t</th>
<th>bayesGARCH package</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>N (0, 100)</td>
<td>a bivariate truncated Normal distribution</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>N (0, 100)</td>
<td>a bivariate truncated Normal distribution</td>
</tr>
<tr>
<td>$\beta$</td>
<td>N (0, 100)</td>
<td>a univariate truncated Normal distribution</td>
</tr>
<tr>
<td>$\psi$</td>
<td>U (2, 100)</td>
<td>a translated Exponential distribution</td>
</tr>
</tbody>
</table>

Note. This table shows the priors set in our model are all independent each other and more typical.

4. Conducting GARCH-t in R

The estimate results in our model

In this section, we design and developed the R program. We will show the whole procedure of how we conduct GARCH-t in R. Our results is (1.0103, 0.1838, 0.4694, 9.2587) is closer to the true value, and we can see the procedure in Figure 1.

The R program

We set the all necessary initial values to be zero, the number of MCMC chain to be 10,000, meanwhile $\Gamma$ is a 10,000 times 4 matrix. We leave out the R programs for these simple settings, giving only the necessary parts of the program.

```
R > for(i in 2:N)
R > { 
R > u = runif(1,0,1)
R > gammaA = rnorm(1,gamma[i-1,1],0.1*v)*(u <= 0.9)
R > +rnorm(1,gamma[i-1,1],5*v)*(u > 0.9)
R > for(j in 2:T)
R > { 
R > hh[j] = gammaA^2 + gamma[i-1,2]^2*y[j-1]^2 + gamma[i-1,3]^2*hh[j-1]
R > oldhh[j] = gamma[i-1,1]^2 + gamma[i-1,2]^2*y[j-1]^2
R > + gamma[i-1,3]^2*oldhh[j-1]
R > prop = dt(y/sqrt(hh),gamma[i-1,4])/sqrt(hh)
R > old = dt(y/sqrt(oldhh),gamma[i-1,4])/sqrt(oldhh)
R > ratio = sum(log(prop) - log(old))
R > ratio = exp(ratio)
R > gamma[i,1] = gamma[i-1,1] + (gammaA - gamma[i-1,1])*(runif(1) < ratio)
```

The above procedure could give the estimate value of $\alpha_0$, and the estimation of the other rest of the parameters is very like it, we won’t go back to that.
5. Empirical Application

This section will use Metropolis-Hastings method by the actual financial data to fit the model. The using data is the DAX(Ibis) index of German in European stock market, a total of 1860 data. In order to make the data smooth, we need to do some processing to the original data. Let $x_t$ denote the logarithm and $y_t$ denote the order difference of $x_t$, i.e. $y_t = x_t - x_{t-1}$, see Figure 2.

![Figure 1](image1.png)

**Figure 1.** MCMC results of GARCH (1, 1)-t.

![Figure 2](image2.png)

**Figure 2.** We can see the data in figures (a) and (b) shows a growing trend, while data in figure (c) is stable through the numeric transform.
We apply the R program in the section 5 to estimate the DAX data, achieve the result of $\Gamma$ is $(0.0028, 0.0600, 0.0676, 7.7485)$. That is, the equation of variance is $h_t^2 = 0.0028 + 0.0600 \cdot Y_{t-1}^2 + 0.676 \cdot h_{t-1}^2$. Through compute the Ljung-Box test statistic for examining the null hypothesis of independence in ARCH model, we get the $\chi^2$ squared 0.0023 with p-value 0.9621. It obvious to show our GARCH (1, 1)-t model is sufficient.

6. Conclusion

In this article, the R program to estimate GARCH-t model has been developed. The parameters’ distribution has been modeled using Gaussian model with the most common setting. The results we achieved in each of our experiments with either simulation study or real data application, are quite encouraging.

Foundation

This research was partially supported by the PhD research startup foundation of Guizhou Normal University (Grant No. GZNUD[2017]27) & Science and Technology Foundation of Guizhou Province (LKS[2013]5 & LKS[2012]11), China. & Teaching Project of Guizhou Normal University in 2016: Contract No. [2016] XJ No. 09.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

https://doi.org/10.1016/0304-4076(86)90063-1

https://doi.org/10.1198/073500103288619359

https://doi.org/10.1016/j.ijforecast.2013.07.006

https://doi.org/10.2307/1912773

https://doi.org/10.1017/S0266466600011397

https://doi.org/10.1016/S0378-3758(97)00139-0


