Journal of Statistics: Advances in Theory and Applications
Volume 12, Number 1, 2014, Pages 27-37

A BAYESIAN APPROACH TO VOLATILITY MODELS

ANUPAM DUTTA

Department of Mathematics and Statistics
University of Vaasa
Wolffintie 34
Vaasa-65200
Finland
e-mail: adutta@uwasa.fi

Abstract

The objective of this paper is to investigate the properties of GARCH (1,1) model and to perform inference using a Bayesian approach. In doing so, the Markov Chain Monte Carlo (MCMC) approach is used for estimating the parameters of GARCH (1,1) and the *t*-GARCH (1,1) models. We examine the U.S.- Japan and the U.S.-U.K. exchange rate series. The empirical analysis reveals that the MCMC approach is found to be effective for each return series.

1. Introduction

Stochastic volatility (SV) models are useful as they are employed to estimate the value of market risk. These models are also used for pricing financial derivatives. There exists a large body of research on volatility models. Engle [3], for example, proposes the autoregressive conditional heteroskedastic (ARCH) model for modelling financial time series exhibiting time-varying volatility clustering. Bollerslev [2] later extends Engle's original work by developing a technique that allows the

2010 Mathematics Subject Classification: 97K50, 97M30.
Keywords and phrases: stochastic volatility, Bayesian approach, Markov Chain Monte Carlo (MCMC).
Received August 19, 2014

conditional variance to be an autoregressive moving average (ARMA) process. This extended process is known as generalized autoregressive conditional heteroskedasticity model or GARCH model. Since then multiple extensions of this GARCH model have been proposed. The exponential GARCH (EGARCH) model of Nelson [11] and the GJR model of Glosten et al. [6], which are two popular extensions of the GARCH model, make an important improvement over the symmetric GARCH model by presenting the asymmetric response of volatility to positive and negative returns.

However, although GARCH-type models are usually estimated using the classical maximum likelihood technique, the Bayesian approach also offers an attractive alternative. It enables small sample results, probabilistic statements on nonlinear functions of the model parameters, selection and combination of non-nested models. Due to these numerous advantages, the study of GARCH-type models from a Bayesian view point can be considered very promising.

In the present paper, an attempt has been made to study the GARCH (1,1) model and to perform inference using a Bayesian approach. The Bayesian computational method used for making inference about the SV model parameters is the Markov Chain Monte Carlo (MCMC) approach. The empirical analysis shows that the application of this MCMC approach improves the inference.

The rest of the paper is organized as follows: Section 2 explains GARCH (1,1) model. MCMC procedure is described in Section 3. Section 4 summarizes the data and their properties. Results are discussed in Section 5. Section 6 concludes.

2. GARCH Models

The basic model able to represent non-correlated series with excess kurtosis and autocorrelated squares, proposed by Engle [3], is given by

$$y_t = \varepsilon_t \sigma_t, \tag{1}$$

where ε_t is an i.i.d process with mean zero and variance 1 and σ_t is the volatility that evolves over time.

The volatility, σ_t^2, in the basic ARCH (1) model is defined as

$$\sigma_t^2 = w + \alpha y_{t-1}^2, \tag{2}$$

where $w > 0$ and $\alpha \geq 0$ for σ_t^2 to be positive.

The ARCH (1) model can easily be extended to the ARCH (q) model as follows:

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i y_{t-i}^2. \tag{3}$$

However, early applications of ARCH models needed many lags to adequately represent the dynamic evolution of the conditional variances. In some applications, q could be even 50. To avoid computational problems when estimating such a large number of parameters, the parameters were restricted in an *ad hoc* manner. For example, Engle [4] assumed that $\alpha_i = \frac{\alpha(q+1-i)}{q(q+1)}$.

Later, Bollerslev [2] implemented the same kind of restriction used to approximate the infinite polynomial of the Wald representation by the ratio of two finite polynomials, usually of very low orders. As a result, he proposed the GARCH (p, q) model given by

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2. \tag{4}$$

Then, the GARCH (1,1) model is simply given by

$$\sigma_t^2 = w + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{5}$$

where $w > 0, \alpha \geq 0,$ and $\beta \geq 0$ to guarantee that σ_t^2 is positive.

Usually a GARCH (1,1) model with only three parameters in the conditional variance equation is adequate to obtain a good model fit for financial time series. Indeed, Hansen and Lunde [8], provided compelling evidence that it is difficult to find a volatility model that outperforms the simple GARCH (1,1). This led us to the estimation of GARCH (1,1) models for the present study.

However, the likelihood for GARCH (1,1) can be written as

$$L = \prod_{t=1}^{n} (2\sigma_t^2)^{-1/2} \exp\left(-\frac{y_t^2}{2\sigma_t^2}\right),$$

where $y = (y_1, \ldots, y_n)$. Thus the log-likelihood is

$$\log L = -\frac{1}{2}\sum_{t=1}^{n}\left[\log \sigma_t^2 + \frac{y_t^2}{\sigma_t^2}\right].$$

3. Markov Chain Monte Carlo (MCMC) Methods

The Bayesian approach begins with a prior distribution for an unknown and unobservable parameter, say θ, where θ is treated as a random variable with a distribution over the parameter space. At this stage, before we see any data, say y, the prior distribution reflects our degree of belief about. Having seen the data, our degree of belief can be updated by using Bayesian calculus. Our prior is now updated in to the posterior distribution. A Bayesian model is a probability model that consists of a likelihood function and a prior distribution. In the Bayesian calculation, the posterior distribution of given y is

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}.$$

All inference procedures like moment calculation, estimation, and decision making are based on this posterior distribution. Inference in the Bayesian approach often requires advanced Bayesian computation, and here we focus on Markov Chain Monte Carlo (MCMC) sampling.

The aim of MCMC methods is to construct a Markov chain, which is a sequence of (possibly vector) random variables, $\{\theta_t, t = 1, 2, \ldots, k\}$ with k-dimensional vector of variables. These variables are to be generated according to the model where the next state θ_{t+1} is sampled from some one-step ahead conditional distribution, $p(\theta_{t+1}|\theta_t)$, which depends only on the current state of the chain, θ_t. Such a Markov chain, under regularity conditions, will have an equilibrium or stationary distribution, and it is this distribution that plays a central role in MCMC. The Markov chain method can be used to sample from any (joint) posterior distribution, which defines the equilibrium distribution of the Markov chain. The key aspect in the Bayesian inference setting is to define precisely the form of the one-step ahead conditional distribution so that the equilibrium distribution is required Bayesian posterior distribution.

The MCMC approach, however, is the main inference technique implemented in this paper. As described in Bernardo and Smith [1], the basic Bayesian MCMC procedure is the following:

• Construct a Markov chain on the parameter space Θ, which is straightforward to sample from, and whose equilibrium distribution is $p(\theta|y)$.

• Run the Markov chain sampling process for a long time.

• The expected values with respect to $p(\theta|y)$ of functions $b(\theta)$ of interest are estimated by using the simulated values from the chain. This refers to the Monte Carlo integration, where samples from a (posterior) probability distribution $p(\theta|y)$ are used to estimate the (posterior) expectation for general functions $b(\theta)$ with respect to that distribution, i.e.,

$$E_{\theta|y}[b(\theta)] = \int b(\theta)p(\theta|y)d\theta.$$

The expectation $E_{\theta|y}[b(\theta)]$ is approximated by drawing sample $\{\theta^{(1)}, \ldots, \theta^{(n)}\}$ from the distribution for some suitably large N, and taking the average. That is,

$$E_{\theta|y}[b(\theta)] \approx \frac{1}{N}\sum_{i=1}^{N} b(\theta^{(i)}).$$

For example, the expectation, $E_{\theta|y}[b(\theta)]$ with $b(\theta) = \theta$ can be approximated from the sample using the sample mean of

$$E_{\theta|y}[\theta] \approx \frac{1}{N}\sum_{i=1}^{N} \theta^{(i)}.$$

● In order to implement this strategy, algorithms for constructing the Markov chain with specified equilibrium distribution are required. Fortunately, several proscribed Markov chain construction schemes exist and we outline the two most commonly-used methods. The first method is the Gibbs sampler introduced by Geman and Geman [5]. The second method is the Metropolis-Hasting (MH) originally developed by Metropolis et al. [9] and further generalized by Hastings [7]. These two samplers are simple to implement and are effective in practice when used for Bayesian inference.

4. Data

The study utilizes the U.S.-Japan and U.S.-U.K. daily exchange rate series. The time period for each series ranges from January, 2000 to January, 2012. We choose exchange rate data because they typically exhibit high degree of volatility clustering. We consider daily returns of the exchange rate as daily returns exhibit stronger degree of short-term volatility clustering than intra-day data and are less noisy. Moreover, the weekend effect is less important for currency data than for stock return data. Therefore, the need for using weekly data to avoid such minor problems is less important.

However, the return series is calculated as follows:

$$y_t = 100 \times \ln \frac{P_t}{P_{t-1}},$$

where P_t denotes the observed daily price at time t and y_t is the corresponding daily return.

Table 1 displays the main empirical properties of these two return series. Inspecting these properties indicates that both series show almost zero means and excess kurtosis (always above 3) for the normal distribution value.

Table 1. Summary statistics of the returns series

Series	N	Mean	St. Dev.	Skewness	Kurtosis
U.S.-Japan	3050	0.00969	0.66400	0.355	3.531
U.S.-U.K.	3050	– 0.00142	0.62302	– 0.048	3.905

This table presents the descriptive statistics of both returns series.
St. Dev. indicates the standard deviation of individual series.

5. Results and Discussions

5.1. Bayesian inference for GARCH (1,1) model

For the GARCH (1,1) model defined in (5), we consider the following priors:

$$\log \alpha_0 \sim N(0, \sigma^2_{\alpha_0}),$$

and

$$\alpha_1, \beta_1 \sim \text{Dirichlet}(\delta_1, \delta_2, \delta_3),$$

where we choose $\sigma^2_{\alpha_0} = 5, \delta_1, \delta_2, \delta_3 = 1.$

The Dirichlet prior is the most commonly used prior for parameters restricted to a simplex region and can be tailored to respect genuine prior beliefs. Here, we considered the uniform prior. However, a prior model favoring large $\alpha_1 + \beta_1$ (and near non-stationarity) can also be employed.

Hence the posterior distribution for the GARCH (1,1) model can be written as

$$p(\alpha_0, \alpha_1, \beta_1 | Y) = \prod_{t=1}^{n} (2\sigma_t^2)^{-1/2} \exp\left(-\frac{y_t^2}{2\sigma_t^2}\right) \times \alpha_0^{-1} \exp\left(-\frac{(\log \alpha_0)^2}{2\sigma_{\alpha_0}^2}\right)$$

$$\times\, \alpha_1^{\delta_1 - 1} \beta_1^{\delta_2 - 1} (1 - \alpha_1 - \beta_1)^{\delta_3 - 1}.$$

A similar Bayesian approach has been implemented by Nakatsuma [10]. He considers normal prior for all the parameters. But such a prior may not be appropriate since (α_1, β_1) are restricted to lie between 0 and 1.

5.1.1. Posterior output analysis summary

We now study the posterior statistics for each data series, which have been obtained by MCMC implementation. Table 2 displays posterior summaries for the parameters in the GARCH (1,1) model for each series considered in this study. To explain the stability and persistence of GARCH (1,1) model, the sum of α_1 and β_1 needs to be examined. It is evident from Table 2 that each series estimates the values of $\alpha_1 + \beta_1$ to be significantly close to one. And this is very usual in practice. In addition, the estimated values of α_1 are close to one and that of β_1 are close to zero. Thus, there exists considerable persistence in volatility, moving towards non-stationarity.

Table 2. Posterior statistics of GARCH (1,1) model

Series		U.S.-Japan	U.S.-U.K.
α_1	Mean	0.9247	0.9012
	Median	0.9259	0.9036
	Std.	0.0102	0.0091
β_1	Mean	0.0732	0.0956
	Median	0.0739	0.0961
	Std.	0.0089	0.0097
$\alpha_1 + \beta_1$	Mean	0.9979	0.9968
	Median	0.9992	0.9999
	Std.	0.0095	0.0118

This table presents the posterior summaries for GARCH (1,1) model with Std. implying the standard deviation.

5.2. Bayesian inference for student-*t* GARCH (1,1) model

5.2.1. The student-*t* GARCH model

The student-*t* GARCH (1,1) model can be formulated as follows:

$$y_t = \varepsilon_t \sigma_t^2, \quad \varepsilon_t \sim N(0, k\lambda_t),$$

where

$$\lambda_t \sim IGamma(\frac{v}{2}, \frac{v}{2}),$$

and $0 < \alpha_1 + \beta_1 < 1$ to guarantee stationarity. Moreover, the parameters λ_t, $t = 1, \ldots, n$ modify the model so that

$$y_t | \sigma_t^2 \sim t(0, k\sigma_t^2, v),$$

with $v > 0$ representing the degree of freedom and k being a constant. Also $v > 2$ guarantees the finiteness of the conditional variance of y_t. If we set $k = \frac{v-2}{v}$, then the conditional variance of y_t becomes σ_t^2.

However, setting $\lambda_t = 1 \; \forall t$ yields the original GARCH (1,1) model. Finally, for finite kurtosis, the restriction is $\upsilon > 4$.

5.2.2. Posterior output analysis summary

The posterior summaries for the t-GARCH (1,1) model parameters with $\upsilon = 5$ are shown in Table 3. Results reveal that the posterior mean and median values of $(\alpha_1 + \beta_1)$ are markedly less than 1. This indicates that there is less persistence than indicated in the GARCH (1,1). In contrast, each of the given series produces the values of $(\alpha_1 + \beta_1)$ to be close to 1 indicating significant persistence in volatility.

Table 3. Posterior statistics of t-GARCH (1,1) model

	Series	U.S.-Japan	U.S.-U.K.
α_1	Mean	0.9101	0.8909
	Median	0.9115	0.8917
	Std.	0.0993	0.0157
β_1	Mean	0.0844	0.0633
	Median	0.0849	0.0640
	Std.	0.0274	0.0094
$\alpha_1 + \beta_1$	Mean	0.9945	0.9542
	Median	0.9987	0.9574
	Std.	0.0192	0.0096

This table presents the posterior summaries for t-GARCH (1,1) model with Std. implying the standard deviation.

6. Conclusion

This paper presents a modest attempt to investigate the properties of GARCH (1,1) model and to perform inference using a Bayesian approach. To do so, the Markov Chain Monte Carlo (MCMC) approach is employed. The study, however, inspects two important exchange rate series which are U.S.-Japan and U.S.-U.K.. The MCMC approach used for the estimation of the GARCH (1,1) and the t-GARCH (1,1) models is found to be effective with each of the exchange rate series under study. While comparing the findings using the same MCMC sampling scheme for the two GARCH models, there is clear evidence that the persistence parameters $(\alpha_1 + \beta_1)$ are close to 1, and thus are closer to nonstationarity in the GARCH (1,1) model than those in the t-GARCH (1,1) model. We further note that the posterior median values of $(\alpha_1 + \beta_1)$ are estimated to be very close to 1 for each series considered in this study.

References

[1] J. M. Bernardo and A. F. M. Smith, Bayesian Theory, John Wiley and Sons, 1974.

[2] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 73 (1986), 151-184.

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

[4] R. F. Engle, Estimates of the variance of U.S. inflation based on the ARCH model, Journal of Money, Credit and Banking 15 (1983), 286-301.

[5] S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (1984), 721-741.

[6] L. R. Glosten, R. Jagannathan and D. E. Runkle, On the relation between the expected value and the volatility of the nominal excess return on stocks, The Journal of Finance 48 (1993), 1779-1801.

[7] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970), 97-109.

[8] P. R. Hansen and A. Lunde, A forecast comparison of volatility models: Does anything beat a GARCH (1,1) model ?, Journal of Applied Econometrics 20 (2004), 873-889.

[9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, Equation of state calculations by fast computing machines, Journal of Chemical Physics 21 (1953), 1087-1092.

[10] T. Nakatsuma, Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach, Journal of Econometrics 95 (2000), 57-69.

[11] D. B. Nelson, Conditional heteroskedasticity in asset returns: A new approach, Econometrica 59 (1991), 347-370.

■