Bayesian inference for regime switching stochastic volatility model with fat-tails and correlated errors

Tomohiro Ando

Graduate School of Business Administration Keio University, 2-1-1 Hiyoshi-Honcho, Kohoku-ku, Yokohama-shi, Kanagawa, 223-8523, Japan.
andoh@kbs.keio.ac.jp

Summary. The stochastic volatility model provides a successful tool in accounting for well-documented time-varying behavior in the volatility of financial time series. The purpose of the paper is to extend the regime-switching stochastic volatility (RSSV) model to handle with both heavy-tailed and correlated errors. A Bayesian approach with Markov Chain Monte Carlo is employed for estimating model parameters. To evaluate the goodness of the estimated model, the deviance information criterion (DIC) is employed. Fitting the proposed model to Nikkei 225 index returns, we find that the proposed model is better than the standard RSSV model in the sense that it gives the smaller DIC score than that of the standard RSSV model.

Key words: heavy-tail, leverage effect, regime-switching

1 Introduction

Modelling of the financial asset return evolutions has recently attracted much attention in the financial econometrics research. The stochastic volatility (SV) model [Tay82] provides a successful tool in accounting for well-documented time-varying behavior in the volatility of financial time series. Despite its popularity, it has been observed that the standard SV model is too restrictive for many financial time series. To refine the standard SV model, many researches have been conducted; including the model with a leverage effect [MY00], with a jump component [BS01], with a heavy-tailed distribution of asset returns [JNR04], with a nonlinear conditional volatility innovation [And06] and so on. Comprehensive reviews of the SV models can be found in the work of [She05].

Recently, [SLL98] generalized the standard SV model by allowing a regime-switching property, which is governed by a first-order Markov process. An advantage of the regime-switching stochastic volatility (RSSV) model is that can take into account the structural shift in financial time series (See also [HS94]). The RSSV model can be extended in two natural ways: (a) with a leverage effect, and (b) with a heavy-tailed distribution of the asset returns. The purpose of the paper is to extend the RSSV model to handle with both heavy-tailed and correlated errors.
A Bayesian approach with Markov Chain Monte Carlo (MCMC) is employed for estimating model parameters. It allows us to estimate the proposed model easily because the model parameter estimation can be done without evaluating the likelihood function. To evaluate the goodness of the estimated model, a deviance information criterion (DIC; [SBCv02]) is employed. An advantage of DIC is that only requires the already produced MCMC posterior samples in practical computation.

The paper is organized as follows. Section 2 proposes the extended RSSV model that treats both heavy-tailed and correlated errors. Section 3 presents the details of Bayesian MCMC estimation procedure. Section 4 reviews DIC for model evaluation. Section 5 applies the proposed model to Nikkei 225 index returns. Section 6 gives concluding remarks.

2 Regime switching stochastic volatility model with heavy-tails and correlated errors

In this section, we extend the standard RSSV model [SLL98,KS03] by introducing both heavy-tailed and correlated errors. Consider an asset return process $Y_n = (y_1, ..., y_n)'$ described below:

$$
\begin{align*}
    y_t &= \exp(h_t/2)\sqrt{\lambda_t} \varepsilon_t, \\
    h_t &= \mu_{s_t} + \phi(h_{t-1} - \mu_{s_{t-1}}) + \tau v_t, \\
    \mu_{s_t} &= \mu + \eta(s_t), \quad \eta(s_t) > 0, \quad s_t = \{1, 2, ..., S\}, \\
    P(s_t = j | s_{t-1} = i) &= p_{ij}.
\end{align*}
$$

The measurement error $\sqrt{\lambda_t} \varepsilon_t$ is assumed to follow the heavy-tailed student-$t$ distribution with mean zero and unknown degrees of freedom by using a Gamma scale mixture of normal distributions $\varepsilon_t \sim N(0, 1)$ and $\lambda_t^{-1} \sim \text{Gamma}(\nu/2, \nu/2)$. The scaling factor $\exp(\mu_{s_t}/2)$ specifies the amount of the model volatility on day $t$, $\tau$ determines the volatility of log-volatilities and $\phi$ measures the autocorrelation. The proposed model also incorporates the leverage effect by allowing the correlation $\rho$ between $u_t$ and $v_{t+1}$:

$$
\begin{pmatrix}
    \varepsilon_t \\
    s_{t+1}
\end{pmatrix}
\sim N(0, \Sigma)
\quad \text{and} \quad
\Sigma = \begin{pmatrix}
1 & \rho \tau \\
\rho \tau & \tau^2
\end{pmatrix}.
$$

Negative correlations can produce the leverage effect in which the negative shock to $y_t$ increases of the volatility $h_{t+1}$. The underlying state $s_t$ takes one of $S$ states $s_t \in \{1, 2, ..., S\}$, which is assumed to be governed by an $S$-state Markov process [Ham89]. The transition probability parameter $p_{ij}$ represents the transition probability that the next state will be state $j$, given that the current state is $i$. Therefore, $\mu_{s_t} = \mu + \eta(s_t)$ is a function of the underlying state $s_t$, which follows an $S$-state Markov process. As an identification condition, we require that $\eta(1) = 0$ and $\eta(s_t)$ is the monotone increasing function, i.e., $\eta(i) > \eta(j)$ if $i > j$. In this paper, we assume there are two regimes $S = 2$, say, a low-volatility state ($s_t = 1$) and a high-volatility state ($s_t = 2$).
3 Bayesian inference via Markov Chain Monte Carlo

A critical difference between the SV type models and ARCH [Eng82] type models is the difficulty level of the likelihood computation. Since ARCH type models specify the volatility of the current return as a non-stochastic, we can easily evaluate the likelihood as a stochastic process and the likelihood function depends on high-dimensional integrals:

$$L(Y_n|\theta) = \prod_{t=1}^{n} f(y_t|F_{t-1}, \theta) = \prod_{t=1}^{n} \int f(y_t|h_t, \theta)f(h_t|F_{t-1}, \theta)dh_t,$$

where $F_{t-1}$ denotes the history of the information sequence up to time $t-1$, $\theta$ is unknown parameter vector. For the proposed model, $f(y_t|h_t, \theta)$ is a normal distribution with mean $\theta$ and standard deviation $\exp(h_t/2)\sqrt{\lambda}$, $	heta = (\mu, \tau, \phi, \nu, \eta(2), p_{11}, p_{22})$ and the conditional density function $f(h_t|F_{t-1}, \theta)$ is expressed as follows:

$$f(h_t|F_{t-1}, \theta) = \sum_{i=1}^{2} \sum_{j=1}^{2} f(h_t, s_t = i, s_{t-1} = j|F_{t-1}, \theta)$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij} P(s_{t-1} = j|F_{t-1}) f(h_t|s_t = i, s_{t-1} = j, F_{t-1}, \theta),$$

where the filtering probability may be decomposed by Bayes’ Rule as

$$P(s_t = i|F_t) = \sum_{j=1}^{2} P(s_t = i, s_{t-1} = j|F_t)$$

$$= \sum_{j=1}^{2} \left[ \frac{f(y_t, s_t = i, s_{t-1} = j|F_{t-1}, \theta)}{\sum_{k=1}^{2} \sum_{l=1}^{2} f(y_t, s_l = k, s_{l-1} = l|F_{t-1}, \theta)} \right]$$

Unfortunately, the density $f(h_t|s_t = i, s_{t-1} = j, F_{t-1}, \theta)$ in (3) cannot not be expressed in the closed form. Furthermore, the variance structure of $f(y_t|h_t, \theta)$ includes the volatility $h_t$ and this makes an obstacle to obtain analytical evaluation. Thus it is not straightforward to construct the likelihood function and to perform the maximum likelihood method.

Even though the integration in the likelihood function (2) cannot be evaluated analytically, Bayesian approach via MCMC method allows us to estimate various types of SV models easily (See for e.g., [JPR94], [KSC98], [CNS02]). In this approach, the latent volatility vectors $H_n = (h_1, ..., h_n)'$ and $A_n = (A_1, ..., A_n)'$ and the regime vector $S_n = (s_1, ..., s_n)'$ in the proposed model is considered as model parameter. The estimation of the proposed model (1) therefore involves estimating three latent variables, $H_n$, $A_n$ and $S_n$ in addition to the model parameters $\theta$. Enlarging with $\omega = (H_n, A_n, S_n, \theta)'$ as an element, the inference is done by producing a sample from the posterior distribution

$$\pi(\omega|Y_n) \propto f(Y_n|H_n, A_n, \theta)f(H_n|A_n, S_n, \theta)f(S_n|\theta)f(\theta)$$
\[
\begin{align*}
\pi(h_t | \omega_{-h_t}, Y_n) &\propto \begin{cases} 
  f(y_t | h_t, \lambda_t, \theta) f(h_t | s_t, \lambda_t, \theta), & (t = 1), \\
  f(h_{t+1} | h_t, s_{t+1}, \lambda_t, \theta) f(h_t | h_{t-1}, s_t, s_{t-1}, \lambda_{t-1}, \theta) f(y_t | h_t, \lambda_t, \theta), & (t \neq 1, n), \\
  f(h_n | h_{n-1}, s_n, s_{n-1}, \lambda_{n-1}, \theta) f(y_n | h_n, \lambda_n, \theta), & (t = n),
\end{cases}
\end{align*}
\]
where \( \omega_{-h_t} \) denotes the rest of the \( \omega \) vector other than \( h_t \). Considering the following decomposition
\[
f(S_n | Y_n, H_n) = f(s_n | Y_n, H_n) \prod_{t=1}^{n-1} f(s_t | S^{t+1}, Y_n, H_n), \quad S^t = (s_t, \ldots, s_n),'
\]
the underlying state \( f(S_n | Y_n, H_n) \) is then generated by adopting the multi-move sampler [CK94, SLL98]. Sampling of \( \{ \tau, \phi, \mu, \nu, \eta(2) \} \) can be performed by using the following conditional posterior density functions:
\[
\begin{align*}
  f(\tau | \omega_{-\tau}, Y_n) &\propto f(H_n | A_n, S_n, \theta) f(\tau), \quad f(\phi | \omega_{-\phi}, Y_n) \propto f(H_n | A_n, S_n, \theta) f(\phi), \\
  f(\mu | \omega_{-\mu}, Y_n) &\propto f(H_n | A_n, S_n, \theta) f(\mu), \quad f(\rho | \omega_{-\rho}, Y_n) \propto f(H_n | A_n, S_n, \theta) f(\rho), \\
  f(\nu | \omega_{-\nu}, Y_n) &\propto f(A_n | \lambda_n, \theta) f(\nu), \quad f(\eta(2) | \omega_{-\eta(2)}, Y_n) \propto f(H_n | A_n, S_n, \theta) f(\eta(2)).
\end{align*}
\]
Finally, the conditional posterior density functions of transition probabilities $p_{11}$ and $p_{22}$ are given by $f(p_{11} | \omega_{11} - \pi_{11}, Y_n) \sim \text{Be}(0.5 + n_{11}, 0.5 + n_{12})$ and $f(p_{22} | \omega_{22} - \pi_{22}, Y_n) \sim \text{Be}(0.5 + n_{22}, 0.5 + n_{21})$, respectively. Here $n_{ij}$ is the number of transitions from state $i$ to $j$, which can be easily counted for given $S_n$. (See for e.g., [KN99]). The outcomes from MCMC algorithm can be regarded as a sample from the posterior density function after a burn-in period. For more details on MCMC method, we refer to [CL96].

4 Deviance information criterion for model diagnosis

As a Bayesian model diagnosis criterion, we employ the deviance information criterion [SBCv02]:

$$
\text{DIC} = -2 \int \log L(Y_n | \theta) \pi(\theta | Y_n) + P_D,
$$

where the first term is the posterior mean of the log-likelihood and the second term is the effective number of parameters, which is defined as the difference between the posterior mean of the likelihood and the likelihood evaluated at the posterior mean of the parameter: $P_D = 2[\log L(Y_n | \bar{\theta}) - \int \log L(y_n | \theta) \pi(\theta | Y_n)]$, where $\bar{\theta}$ is the posterior mean. The model with smaller value of DIC is favored. [BMY04] applied the DIC to assess several different SV models. As shown in equation (2), the likelihood function has no analytical form as it is marginalized over the latent volatilities. We therefore approximated the likelihood function by using the particle filtering method [Kit96].

5 An empirical study

This section analyzes the daily returns of Nikkei 225 index from January 25, 2001 to February 17, 2005 on which the market was open leading to a set of 1,000 samples. The returns $y_t$ are defined as the differences in the logarithm of the daily closing value of Nikkei 225 index $y_t = \log(x_t) - \log(x_{t-1})$, where $x_t$ is the closing price on day $t$. Figure 1 (a) shows the transformed log-difference of Nikkei 225 index. The vertical axis is the differences in the logarithm of the daily closing value of Nikkei 225 index and the horizontal axis is the time. It may be seen from Figure 1 (a) that the transformed Nikkei 225 index returns are exhibiting time-varying volatility. The basic statistics the mean, standard deviation, skewness and kurtosis are given as $-0.0002$, $-0.0002$, $-0.01157$ and $4.2597$ respectively. Considering that the kurtosis is above three and there are some outliers in Figure 1 (a), the transformed Nikkei 225 index returns would follow a heavy-tailed distribution.

In our application, the total number of MCMC iterations is chosen to be 1,100,000 in which the first 100,000 iterations are discarded as a burn-in period sample. Due to higher posterior correlations amongst the parameters and thus slower convergence of the MCMC sampling algorithm, we stored every 1,000 th iteration after a burn-in period. Similar MCMC sampling scheme was also implemented by [BMY04]. To improve the efficiency of the Metropolis-Hastings (MH) algorithm, we approximated the posterior density by using Laplace approximation. This density is used for a proposal density in MH sampling.
Using 1,000 draws for each of the parameters, we calculated the posterior means, the standard errors and the 95% confidence intervals, respectively. The 95% confidence intervals are estimated using the 2.5th and 97.5th percentiles of the posterior samples. These quantities are reported in Table 1. We also assessed the convergence of MCMC simulation by calculating the convergence diagnostic (CD) test statistics [Gew92]. Since the absolute values of CD test statistics are less than 1.96 (at significance level 5%), we can judge that MCMC sampling works well. As shown in Table 1, the upper limits of the 95% posterior credibility interval of $\rho$ were below 0, which suggests some evidence of negative correlation between $\epsilon_t$ and $v_t$. Similar result is also reported by [Wat99], where Tokyo stock price index (TOPIX) is an-
Bayesian inference for regime switching stochastic volatility model

Table 1. Comparison of the parameter estimates for the daily returns of Nikkei 225 index. The posterior means, the standard errors (SD), the 95% confidence intervals, and the convergence diagnostic (CD) statistics are calculated.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>95% conf. interval</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td>τ</td>
<td>0.1718</td>
<td>0.0319</td>
<td>[0.1238 0.2261]</td>
<td>-0.03</td>
</tr>
<tr>
<td>φ</td>
<td>0.4436</td>
<td>0.0047</td>
<td>[0.4366 0.4502]</td>
<td>0.06</td>
</tr>
<tr>
<td>µ</td>
<td>-5.0446</td>
<td>0.0353</td>
<td>[-5.0965 -4.9823]</td>
<td>-0.04</td>
</tr>
<tr>
<td>ν</td>
<td>13.0001</td>
<td>0.0498</td>
<td>[12.9185 13.0807]</td>
<td>-0.05</td>
</tr>
<tr>
<td>ρ</td>
<td>-0.4930</td>
<td>0.0486</td>
<td>[-0.5789 -0.4154]</td>
<td>-1.22</td>
</tr>
<tr>
<td>p_{11}</td>
<td>0.9961</td>
<td>0.0069</td>
<td>[0.9866 0.9999]</td>
<td>1.88</td>
</tr>
<tr>
<td>p_{22}</td>
<td>0.9963</td>
<td>0.0078</td>
<td>[0.9873 0.9999]</td>
<td>1.02</td>
</tr>
<tr>
<td>η_2</td>
<td>0.4482</td>
<td>0.0333</td>
<td>[0.3911 0.5016]</td>
<td>1.38</td>
</tr>
</tbody>
</table>

The transition parameters \( p_{11} \) and \( p_{22} \) have posterior means close to unity, implying that both states 1 and 2 tend to persist for a long time.

Figures 1 (b) and (c) plot the posterior mean of the volatility \( \exp(h_t/2) \sqrt{\lambda_t} \) and the posterior mean of underlying state \( s_t \). These figures indicate that the Nikkei 225 index yields tend to belong to a low-volatility regime. In fact, as the time goes, the transformed Nikkei 225 index returns seem to be distributed within a relatively small range. We also fitted the standard RSSV model [SLL98]. The DIC score (4) of the proposed model is -5465.733, which is smaller than that of the standard RSSV model -5460.720. It indicates that the proposed model is more adequate to describe Nikkei 225 index. We therefore conclude that the simultaneous consideration of the leverage effect and the heavy-tail distribution of asset returns are important.

6 Summary and conclusions

This paper considers a generalization of the regime-switching stochastic volatility model [SLL98] by introducing both heavy-tailed and correlated errors. To carry out the model parameter inference, a Bayesian Markov chain Monte Carlo method is employed. The goodness of the fitted model is then evaluated by the deviance information criterion [SBCv02]. Fitting the proposed model to Nikkei 225 index returns, we find that the proposed model fit Nikkei 225 index data better than the standard RSSV model in the sense that it gives the smaller DIC score than that of the standard RSSV model. The proposed model can be extended to multivariate stochastic volatility type model [PS99]. We would like to discuss this extension in a future paper.

References


Biometrika, 81, 541–553. (1994)

[CNS02] Chib, S., Nardari, F., Shephard, N.: Markov Chain Monte Carlo Methods
(2002)


[Gew92] Geweke, J.: Evaluating the accuracy of sampling-based approaches to cal-
culating posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P.,
Smith, A.F.M. (ed), Bayesian Statistics, 4. Oxford University Press, UK
(1992)

[Ham89] Hamilton, J.D.: A new approach to the economic analysis of nonstationary


Volatility Models (with discussion). Journal of Business and Economic

Volatility Models with Fat-tails and Correlated Errors, Journal of Econo-
metrics, 122, 185–212 (2004)

[KS03] Kalimipallia, M., Susmelb, R.: Regime-switching stochastic volatility and
(2004)

[Kit96] Kitagawa, G.: Monte Carlo filter and smoother for Gaussian nonlinear
state space models. Journal of Computational and Graphical Statistics,
5, 1–25 (1996)


393 (1998)


[PS99] Pitt, M., Shephard, N.: Time varying covariances: a factor stochastic
volatility approach (with discussion). In: Bernardo, J., Berger, J.O.,
Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics, 6. Oxford Uni-

[She05] Shephard, N.: Stochastic Volatility: Selected Readings. Oxford University
Press (2005)

[SLL98] So, M., Lam, K., Li, W.: A stochastic volatility model with Markov switch-
Bayesian inference for regime switching stochastic volatility model

