

doi: 10.11830/ISSN.1000-5013.201702024



采用 MCMC 方法的上海股市 随机波动模型

赵慧琴¹, 刘金山²

(1. 广东财经大学 华商学院, 广东 广州 511300;
2. 华南农业大学 数学与信息学院, 广东 广州 510642)

摘要: 采用贝叶斯统计中的马尔科夫链-蒙特卡罗 (MCMC) 方法对上海股市的随机波动性进行研究, 基于 Gibbs 抽样的 MCMC 数值计算过程, 对上海股市的随机波动率模型 (SV) 进行参数估计, 并在 WinBUGS 软件中实现. 根据信息判别准则 (DIC), 对比拟合的 SV-N, SV-T, SV-MT 模型参数, 结果表明: SV-T 模型最能反映上海股市波动具有尖峰厚尾的特性, 可进一步用于预测样本外的波动率结果.

关键词: 随机波动率模型; 马尔科夫链-蒙特卡罗方法; 股市波动; 贝叶斯分析; 上海股市

中图分类号: O 212 **文献标志码:** A **文章编号:** 1000-5013(2017)02-0262-04

Stochastic Volatility Modeling of Shanghai Stock Exchange Using MCMC Method

ZHAO Huiqin¹, LIU Jinshan²

(1. Huashang College, Guangdong University of Business Studies, Guangzhou 511300, China;
2. College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China)

Abstract: One method is by Markov chain Monte Carlo (MCMC) bias statistics method. In this paper, we study the stochastic volatility of Shanghai Stock market, and estimate the parameters of the stochastic volatility model (SV) of Shanghai Stock market based on the MCMC sampling, and implement the Gibbs software in the WinBUGS software. By comparing the parameters of SV-N, SV-T, SV-MT model, and according to discriminative information criterion, we find the SV-T model is the best model in China reflecting the fluctuation of the stock market of Shanghai which has peak thick tail characteristics, this model can also be used to step out of sample forecasting volatility results.

Keywords: stochastic volatility models; Markov chain Monte Carlo method; stochastic volatility; Bayesian analysis; Shanghai Stock

经济或金融时间序列均存在着普遍的波动性现象, 随机波动率模型 (SV) 可以模拟和预测波动状况. 近年来, 随机波动率模型在我国得到了不断的发展, 研究者们提出一种最常用的扩展模型, 即标准模型 (SV-N 模型)^[1-3]. SV 族模型较难找到精确的似然函数, 需要模拟建立完全的似然函数进行参数估计, 如马尔科夫链-蒙特卡罗 (MCMC) 方法. 国内已有一些相关研究, 但较少将 SV 族的 SV-MT, SV-T 模型与 SV-N 模型进行比较分析^[4-7]. 因此, 本文基于贝叶斯统计中的 MCMC 方法, 结合 Gibbs 抽样, 对 SV 族的模型参数进行估计.

收稿日期: 2017-02-14
通信作者: 赵慧琴 (1982-), 女, 讲师, 博士, 主要从事经济统计和概率统计的研究. E-mail: zhq_6285144@163.com.
基金项目: 广东省青年创新人才类重点课程建设项目 (2014WQNCX177); 广东省质量工程经管综合实验教学中心建设项目 (2013 年度); 广东省质量工程统计学专业实验教学示范中心建设项目 (2014 年度)

1 SV 族模型的统计结构

1.1 SV-N 模型

标准的 SV 模型形式^[8]为

$$y_t = \exp(\theta_t/2)\epsilon_t, \quad \epsilon_t \sim N(0,1), \tag{1}$$

$$\theta_t = \mu + \phi(\theta_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0,\sigma^2). \tag{2}$$

式(1),(2)中: y_t 为 t 时刻的标准化对数收益率; θ_t 为对数波动,服从一个持续性参数为 ϕ 的高斯自回归 AR(1)过程;误差项 ϵ_t 与 η_t 是不相关的,都服从均值 0,方差 σ^2 的正态分布;对于 $|\phi| < 1$,SV 模型的协方差是平稳的.

潜在波动 θ_t 的分布为

$$\theta_t \sim N(\mu, \tilde{\sigma}^2), \quad \tilde{\sigma}^2 = \sigma^2/(1 - \phi^2). \tag{3}$$

$y_t | \theta_t$ 的分布为

$$y_t | \theta_t \sim N(0, \exp \theta_t), \quad t = 1, 2, \cdots, n. \tag{4}$$

若令 $\phi = 2\phi_1 - 1, \tau = \sigma^2$, 且 $y_{1:n} = (y_1, y_2, \cdots, y_n)'$, $\theta_{1:n} = (\theta_1, \theta_2, \cdots, \theta_n)'$, 根据 Kim 等^[9]提出的方法, 假定模型参数 $\phi, \tau, \mu, \theta_1$ 的先验分布为

$$\phi \sim B(20, 1.5), \quad \tau \sim \text{Ga}(2.5, 0.025), \quad \mu \sim N(0, 100), \quad \theta_1 \sim N(\mu, \sigma^2).$$

$\mu, \phi, \tau, \theta_0$ 的先验分布联合密度函数为

$$\pi(\mu, \phi, \tau, \theta_{1:n}) = \pi(\mu)\pi(\phi)\pi(\tau)\pi(\theta_0 | \mu, \phi, \tau) \times \prod_{t=1}^n \pi(\theta_t | \mu, \phi, \tau). \tag{5}$$

根据贝叶斯定理,后验分布的参数为

$$\pi(\mu, \phi, \tau, \theta_{0:n} | y_{1:n}) \propto L(\mu, \phi, \tau, \theta_{0:n})\pi(\mu, \phi, \tau, \theta_{0:n}). \tag{6}$$

式(6)中: $L(\mu, \phi, \tau, \theta_{0:n})$ 为 SV 模型的似然函数.

1.2 SV-T 模型

SV-T 模型表示一个混合的正态分布模型,其形式为

$$y_t = \exp(\theta_t/2)\epsilon_t, \quad \epsilon_t \sim t_\nu, \tag{7}$$

$$\theta_t = \mu + \phi(\theta_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0,\sigma^2). \tag{8}$$

式(7),(8)中: ϵ_t 服从自由度为 $\omega > 2$ 的 t 分布.

若假定 λ_t 服从倒伽马分布,是一个潜在的变量,则有

$$y_t = \exp(\theta_t/2) \sqrt{\lambda_t} \epsilon_t, \quad \epsilon_t \sim N(0,1), \tag{9}$$

$$\theta_t = \mu + \phi(\theta_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0,\sigma^2). \tag{10}$$

$$\lambda_t \sim \text{IG}((\omega/2) \cdot (\omega/2)). \tag{11}$$

取参数 ω 的先验分布 $\omega \sim \chi^2(8)$, 其他参数的先验分布同 SV-N 模型^[10]. ω 和 θ_t 的后验条件分布为

$$\pi(\omega | \phi, \mu, \tau, \theta_{0:n}; y_{1:n}) \propto \omega^{3-n/2} \exp(-\frac{\omega}{2}) \cdot \left(\frac{\Gamma((\omega+1)/2)}{\Gamma(\omega/2)}\right)^n \prod_{t=1}^n \left(1 + \frac{y_t \exp(-\theta_t)}{\omega}\right)^{-(\omega+1)/2}. \tag{12}$$

1.3 SV-MT 模型

SV-MT 模型形式为

$$y_t = d \cdot \exp \theta_t + \exp(\theta_t/2)\epsilon_t, \quad \epsilon_t \sim t_\omega, \tag{13}$$

$$\theta_t = \mu + \phi(\theta_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0,\sigma^2). \tag{14}$$

式(13),(14)中: ϵ_t 为服从自由度为 $\omega > 2$ 的 t 分布.

在模型中, d 的先验分布近似服从正态分布,其他参数的先验分布与 SV-T 模型参数相同. μ, τ, ϕ 的后验条件分布也与 SV-T 模型参数相同,即

$$\pi(d | \phi, \mu, \tau, \omega, \theta_{0:n}; y_{1:n}) \propto \exp(-\frac{d^2}{2}) \prod_{t=1}^n \left(1 + \frac{(y_t - d \cdot \exp \theta_t)^2}{\omega \cdot \exp \theta_t}\right). \tag{15}$$

1.4 模型的比较

基于贝叶斯方法,采用信息判别准则(DIC)^[11-12]对 SV 模型进行比较,具体形式为

$$D=-2\log P(y|\psi)+2P(y).$$

(16)

DIC 准则可以很好地比较复杂的统计模型,即

$$\text{DIC}=\overline{D}+p_D=2E_{\psi|y}(D)-D(E_{\psi|y}(\psi)).$$

(17)

式(16),(17)中: $P(y|\psi)$ 为似然函数; \overline{D} 表示模型拟合数据的优劣,其值越大,则模型拟合数据越劣质;
 $p_D=E_{\psi|y}(D)-D(E_{\psi|y}(\psi))=\overline{D}-D(\bar{\psi})$,用于衡量模型的复杂程度,其值越大,则说明模型越复杂.

2 上海股市 SV 族模型参数的贝叶斯估计

在 SV 族模型中,其参数用 MCMC 方法中的 Gibbs 抽样进行估计.数据源于新浪财经网 2013 年 9 月 2 日至 2014 年 8 月 29 日的上海股市的收盘价格,共 242 个数据.在进行模型参数的贝叶斯估计前,用 WinBUGS 软件对模型参数进行 10 000 次预迭代,再进行 20 002 次迭代,以保证参数的收敛性.3 种模型参数的贝叶斯估计,如表 1 所示.由表 1 可知:波动水平参数 μ 在 SV-N 模型中的绝对值较大,说明在 SV-N 模型中上证综指的波动性体现得比较强烈;3 个模型中的波动持续性参数 ϕ 的估计值相差不大,都在 0.98 以上,说明 3 个模型都能刻画上证综指收益率序列的波动持续性,也说明我国股市具有波动集聚性;在 SV-T 和 SV-MT 模型中加入自由度参数 ω ,其值分别为 16.260,16.480,体现了上海股票尖峰厚尾的特性^[13],而 SV-N 模型并未体现这一特性.SV-T 模型参数的后验核密度估计,如图 1 所示.

表 1 SV 族模型参数的贝叶斯估计

Tab.1 Bayesian estimation of SV model parameters

模型	参数	均值	方差	2.50%分位数	97.50%分位数
SV-N	μ	-0.138 20	0.110 90	-0.353 80	0.068 85
	ϕ	0.983 50	0.018 08	0.926 10	0.998 80
	τ	203.90	70.40	90.27	380.40
SV-T	μ	-0.007 027	0.110 600	-0.292 200	0.138 700
	ϕ	0.988 100	0.009 251	0.962 900	0.998 700
	τ	85.12	30.41	40.77	152.10
	ω	16.260	4.751	8.811	26.930
SV-MT	μ	-0.010 75	0.097 80	-0.298 30	0.084 08
	ϕ	0.991 800	0.005 625	0.977 800	0.999 100
	τ	110.30	44.83	46.60	213.60
	ω	16.480	4.854	8.893	27.640
	d	0.001 393	0.093 000	-0.179 900	0.182 200

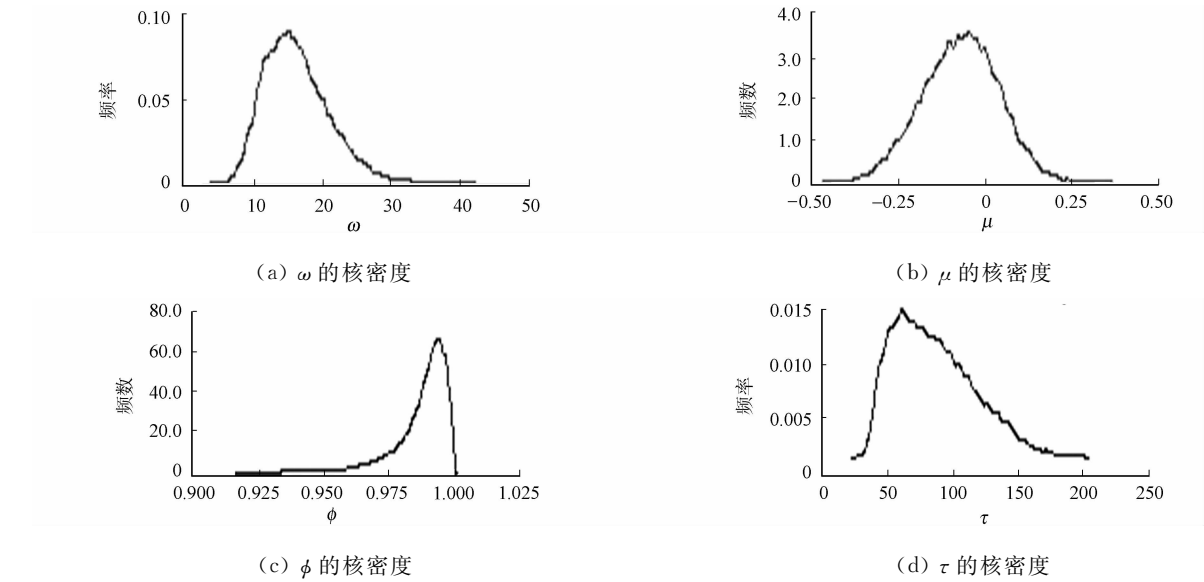


图 1 SV-T 模型参数的后验核密度估计图

Fig.1 Posterior density function for SV-T model

各模型的 DIC 值,如表 2 所示.由表 2 可知:SV-T 模型的 \bar{D} 值最小,而 SV-N 模型的 \bar{D} 值最大,表明 3 个模型中,模拟上海综指的数据最好是 SV-T 模型,最差是 SV-N 模型;对于 P_D 值的模拟,SV-T 模型最大,而 SV-MT 最小.说明对于上海综指的模拟,SV-MT 模型相对于其他模型来说是最简单的模型,而 SV-T 模型是最复杂的;对于综合了模型的复杂度和拟合优劣的 DIC 值,SV-T 模型的值最小,其值为 389.676,而 SV-N 模型的值最大,其值为 398.868,这说明在模拟上海股市波动特性时,SV-T 模型要优于 SV-N 模型和 SV-MT 模型.

3 结束语

对 SV-T,SV-N,SV-MT 模型的参数估计、后验核密度估计和 DIC 值进行比较,可以得出在模拟我国上海股市的波动性中,SV-T 模型比 SV-N,SV-MT 模型更优,更能反映上海股市的尖峰厚尾的特性,并且证明了上海股市具有很强的波动持续性.利用 SV-T 模型,可进一步用于预测样本外的波动率结果.对于国外已有的杠杆 SV 模型^[14]、长记忆 SV 模型^[15-16]研究,将是今后进一步研究的方向.

参考文献:

[1] NAKAJIMA J, OMORI Y. Leverage, heavy-tails and correlated jumps in stochastic volatility models[J]. Computational Statistics and Data Analysis, 2009, 53(6): 2535-2553.

[2] BROTO C, RUIZ E. Estimation methods for stochastic volatility models: A survey[J]. Journal of Economic Surveys, 2004, 18(5): 613-649.

[3] WANG J J J, CHAN J S K, BORIS C S T. Stochastic volatility models with leverage and heavy-tailed distributions: A Bayesian approach using scale mixtures[J]. Computational Statistics and Data Analysis, 2011, 55(1): 852-862.

[4] 王泽锋,史代敏.基于 DIC 准则的 ASV 模型和 SV 模型的实证比较[J].数量经济技术经济研究,2007,24(5):134-141.

[5] 顾锋娟,岑仲迪.基于 GADCH 类模型和 SV 类模型的沪深两市波动性研究[J].数学的实践与认识,2011,41(1):14-22.

[6] 朱慧明,李峰,杨锦明.基于 MCMC 模拟的贝叶斯厚尾金融随机波动模型分析[J].运筹与管理,2007,16(4):111-115.

[7] 何永涛,王飞.随机波动模型贝叶斯估计效率研究[J].统计与信息论坛,2015,30(4):7-14.

[8] TAYLOR S J. Modeling stochastic volatility[J]. Mathematical Finance, 1994, 4(2): 183-204.

[9] KIM S, SHEPHARD N, CHIB S. Stochastic volatility: Likelihood inference and comparison with ARCH models[J]. Review of Economic Studies, 1998, 65(3): 361-393.

[10] MEYER R, YU Jun. BUGS for a Bayesian analysis of stochastic volatility models[J]. Econometrics Journal, 2000, 3(2): 198-215.

[11] SPIEGELHALTER D J, BEST N, CARLIN P B. Bayesian measurements of model flexibility and fit (with discussion)[J]. Journal of Royal Statistical Society, 2002, 64(4): 583-639.

[12] 蔡晓薇. MCMC 方法下最优 Copula 的估计及选取[J].统计与信息论坛,2011,26(10):33-38.

[13] 邱世斌,陈燕武.马尔可夫链在前沿资产组合中的应用[J].华侨大学学报(自然科学版),2009,30(2):221-224.

[14] LSHIHARA T, OMORI Y, ASAI M. Matrix exponential stochastic volatility with cross leverage[J]. Computational Statistics and Data Analysis, 2016, 100(8): 331-350.

[15] SO M K P, KWOK S W Y. A multivariate long memory stochastic volatility model[J]. Physica A, 2006, 362(2): 450-464.

[16] HO H C. Sameple quantile analysis for long-memory stochastic volatility models[J]. Journal of Econometicle, 2015, 189(2): 360-370.

(责任编辑:钱筠 英文审校:黄心中)