

华中科技大学硕士学位论文创新点概述

论文题目：基于贝叶斯蒙特卡罗的 SV 扩展模型比较及在中国资本市场的应用

学科（专业）、研究方向：数量经济学；贝叶斯统计

论文中的主要创新点：

（1）本文介绍和比较了随机波动率的三种传统 MCMC 方法，随机游走的 M-H，独立马尔科夫链 M-H，线性滤波器 FFBS 的估计精度和对中国股市的适应程度。

（2）考虑到均值和带状矩阵特性，使用快速带和稀疏矩阵特殊性算法，更为高效的抽样来拟合这些模型。这个方法适合所有的随机波动率模型，区别与传统的抽样和卡尔曼滤波器。

（3）运用高效抽样的方法对 SV 模型，SV-M 模型进行了贝叶斯分析，并将分析扩展到 SV-AR（2），SV-MA(1)模型和 SV-MG 模型。

（4）本文参考 Kass and Raftery（1995）的贝叶斯因子方法，计算模型边缘似然密度，使用新型的评价准则 marginal likelihoods 来对模型进行比较研究。

（5）基于扩展模型分别研究股市收益波动率与前一期的相关性，股市波动的随机游走性，收益是否会受到当期的波动影响，即波动与收益的关系和股市的波动大部分来自于当前的冲击还是受到过去冲击的持续影响。

硕士学位论文主要研究成果的发表或获奖情况

序号	发表论文题目/ 获奖成果名称	刊物名称/ 获奖部门	刊物/奖励级别	发表论文/ 获奖时间	作者署名名次
1					

华中科技大学

硕士学位论文

基于贝叶斯蒙特卡罗的 SV 扩展模型比较及
在中国资本市场的应用

院 系 _____ XXXX

专业班级 _____ XXXX

姓 名 _____ XXX

学 号 _____ XXX

指导教师 _____ XXX

2018 年 3 月 12 日

学位论文原创性声明

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。

作者签名： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密口，在 年解密后适用本授权书

2、不保密口。

（请在以上相应方框内打“√”）

作者签名： 年 月 日

导师签名： 年 月 日

摘 要

大量研究表明金融时间序列具有高峰厚尾，波动聚集，非线性结构，杠杆效应，波动的长记忆性等特点。基于资本资产定价及风险管理的需要，金融数据波动率的正确估计和预测必须准确描述上述特征。大量的文献研究表明，相比于 ARCH 和 GARCH 模型，随机波动率模型从理论和实践上能更好的刻画和预测金融市场的波动。此外，随机波动率模型在时间的连续性方面能很好的与金融市场中的期权定价，利率期限结构等理论模型相结合。

考虑到国内外学者已有的研究基础，本文主要介绍了关于随机波动率的三种传统贝叶斯蒙特卡罗模拟估计方法：随机游走的 M-H，独立马尔科夫链 M-H，线性滤波器 FFBS。同时引进了一种高效的 MCMC 方法，对 SV 模型，SV-M 模型进行了贝叶斯分析，比较其优劣，并将分析扩展到 SV-AR(2)，SV-MA(1) 模型和 SV-MG 模型，对我国股市中的中证 500 指数收益波动进行分析。

根据实证分析结果，可以得到以下结论：中证 500 指数的对数收益波动率与前一期有很强的相关性，但与两期之前的相关性较弱，未来的收益受到当前的影响，但与过去无关。进一步证明了，我国股市的波动是随机游走过程。其次，中证 500 指数的收益会受到当期的波动影响，且收益与波动呈现负相关，当波动越大，收益越小。最后，中证 500 指数对数收益率的波动受到前期的扰动小，股市的波动大部分来自于当前的冲击。并通过新型的评价准则 marginal likelihoods，对随机波动率族模型簇对比研究，发现高斯混合随机波动率模型能更好的刻画我国中证 500 指数波动。

关键词：随机波动率；MCMC；线性滤波器；马尔可夫链

Abstract

A large number of studies have shown that financial time series have the characteristics of high-thick tail, fluctuating aggregation, nonlinear structure, leverage effect, long memory and so on. Based on the needs of capital asset pricing and risk management, the correct estimation and forecast of financial data volatility must accurately describe the above characteristics. A large number of literature studies have shown that compared with the ARCH and GARCH models, the stochastic volatility model can better describe and predict the volatility of financial markets in theory and practice. In addition, the stochastic volatility model can be combined with the theoretical models of option pricing and term structure of interest rates in the financial market in terms of time continuity.

Considering the existing research foundations of domestic and foreign scholars, this paper mainly introduces three traditional Bayesian Monte Carlo simulation methods for estimating stochastic volatility: random walk MH, independent Markov chain MH, linear filter FFBS . At the same time, an efficient MCMC method was introduced. Bayesian analysis was performed on SV model and SV-M model to compare their advantages and disadvantages, and the analysis was extended to SV-AR(2) and SV-MA(1) models. The SV-MG model analyzes the fluctuation of the returns of China Securities 500 Index in China's stock market.

According to the empirical analysis results, the following conclusions can be drawn: The volatility of the logarithmic returns of the CSI 500 has a strong correlation with the previous period, but its correlation with the previous two periods is weak, and future returns are affected by the current situation. But it has nothing to do with the past. It is further proved that the volatility in China's stock market is a random walk. Second, the earnings of the CSI 500 index will be affected by the fluctuations of the current period, and the income and volatility will be negatively correlated. When the fluctuations are greater, the returns will be smaller. Finally, the volatility of the CSI 500 logarithmic rate of return was less affected by the previous period, and most of the stock market volatility came from the current impact. And through the new evaluation criteria marginal likelihoods, a comparative study of clusters of stochastic volatility family models shows that the Gaussian mixture stochastic volatility model can better describe the volatility of the China Securities 500 Index.

Key words: stochastic volatility; MCMC ; linear filter; Markov chain

目录

摘 要	4
Abstract	5
1. 导论	7
1.1. 研究背景及意义.....	7
1.2. 相关研究及文献综述.....	9
1.2.1 国外文献综述	10
1.2.2 国内文献综述	11
1.3. 研究思路.....	13
1.4 本文创新	14
2. 标准 SV 模型理论分析	14
2.1. Gibbs 抽样和 Metropolis-Hastings 抽样	14
2.1.1 Gibbs 抽样	14
2.1.2 Metropolis-Hastings 抽样:	15
2.2. 标准 SV 的 MCMC 估计.....	16
2.3. 随机游走与 Metropolis-Hastings	18
2.4. 独立马尔科夫链与 Metropolis-Hastings	19
2.5. 线性滤波器抽样方法.....	20
3. SV 模型的扩展与更新	25
3.1. 高斯近似分布与 Newton-Raphson 的 SV-AR(1)	25
3.2. SV-AR(2) 模型的贝叶斯估计.....	29
3.3. SV-M 模型的贝叶斯估计	31
3.4. SV-MA(1) 模型的贝叶斯估计.....	31
3.5. SV-MG 模型的贝叶斯估计	33
4. 中证 500 指数波动率的实证研究.....	34
4.1. 样本及数据选取.....	34
4.2. 样本数据特征.....	35
4.3. 标准 SV 的参数估计及其比较.....	37
4.4. 扩展 SV 的参数估计.....	41
5. 主要结论与展望	45
参考文献	47
致 谢	49
附录	50

1. 导论

1.1. 研究背景及意义

金融资产的波动衡量了所有因素最终导致金融市场的风险大小,也是市场风险最有效和最直接的量化工具。市场参与者的投资决策,相关企业的负债大小等等都受到金融波动的影响。金融资产的波动也是金融学中的基础内容,涉及到金融学理论的方方面面。如 APM 模型及均值-方差模型, CAPM 理论、B-S 期权定价。

在早期,学者对金融资产风险的估计主要是利用马克维茨的均值-方差模型,通过金融资产收益率的方差或者标准差来衡量。之后,大量的研究者发现,金融资产的价格在任意时刻是服从对数正态分布,且是随机游走的。基于此,对于金融风险的衡量在金融资产收益标准差模型的基础上,对金融资产求对数收益率。由于金融的波动时随机的,价格的形成过程也是随机的,仅仅只用标准差来衡量金融资产风险的大小是有所欠缺的,金融资产是服从随机游走的,随机性对金融资产有巨大的影响。将金融资产波动的时变和随机性引入到风险的估计中,一些学者假设金融资产的波动是异方差的和持续性的,提出了条件异方差模型。另一些学者考虑到价格的随机性,假设金融资产的波动也是服从随机游走的,提出了 SV 模型。

波动性一直是经济或金融时间序列的特征之一。在证券市场的发展过程中,股票价格形成机制的理论研究一直是受到极大的关注,并带动了证券市场的其他理论的研究与发展。例如市场均衡理论,期权定价理论,市场有效性理论,资产定价理论等。从宏观的系统上看,金融资产的价格形成是一个复杂的非线性系统,由于金融市场的高度复杂性,使得股票的价格难以分析,描述和预测。大量实证研究表明金融经济中的时间序列呈现出新的特点,高峰厚尾,波动的聚集性,非线性结构,杠杆效应,波动的长记忆性。对于波动率的正确估计和预测,捕获上述特征是必不可少的。为了描述上述波动特征,许多统计模型被相继提出,恩格尔(1982)提出了自回归条件异方差(ARCH)模型, Bollerslev(1986)提出广义自回归条件异方差(GARCH)模型,泰勒(1986)提出了随机波动(SV)模型。ARCH 模型, GARCH 模型和 SV 模型三者的主要区别是前者假设条件方差是一个关

于过去的波动性和观测回报平方和的确定性函数，而在后者的资产收益率的条件波动率的对数服从潜在固定的一阶自回归（AR（1））与单正态新息过程。在这两种波动模型中，资产回报的新息通常被假定为遵循标准正态分布或学生 t 分布。根据经验，SV 模型是首选，相对于预测金融资产收益率的波动率建立 ARCH / GARCH 模型，SV 模型在估算收益率波动中比 ARCH / GARCH 模型表现出更低的持续性。

随着我国经济的快速发展和信息的透明化，我国股票市场体系逐步完善。在时间的推移的过程中，金融资产的时间序列样本规模持续增大。研究表明，大多数金融资产的波动都是非线性的特征。非线性时间序列的分析方法已经成为研究我国金融资产波动的主流之一。我们将非平稳和非线性的理论方法结合与创新，精确的分析和预测我国金融资产时间序列波动的趋势。将非对称性、非线性、非平稳三者结合起来，利用随机波动率模型簇对我国金融资产的波动进行扩展式研究，将是一个具有十分重要意义的研究。

其次，我国股市的起步晚，发展迅速。在发展过程中隐藏着诸多矛盾。股市呈现出强烈的波动性和不稳定性。内部和外部的各种不确定冲击都深刻影响着国内股市。可能是由于国内金融市场缺乏整体性，金融监管不到位，金融监管机构对于金融风险的管控严重经验不足。研究金融市场的风险，对金融监管具有借鉴意义。为监管提供经验，有利于制度的建设和金融市场秩序的维护。本文扩展随机波动率模型为市场风险定量测度提供有力工具，为金融市场良好持续的发展提供基础。

最后，金融研究领域的不断扩大，经济和金融学家们一直研究和创新各种各样的金融资产波动率模型。通过大量的波动率模型研究得知，在金融资产持续波动且幅度较大时，隐含着更高的风险酬金。金融资产的价格受到收益率波动的影响，如果能够准确的预测金融资产未来的波动，对金融资产的定价和套利有深远的意义。

本文正是基于上述基础，通过 SV 扩展模型研究以下问题：1. 股市收益波动率的随机游走性和自相关性，2. 当期的波动对收益的影响，即波动与收益的关系 3. 股市的波动大部分来自于当前的冲击，还是受到过去冲击的持续影响。模型通过 PYTHON 3.6 编程对参数进行估计。

1.2. 相关研究及文献综述

随机波动率模型的研究在 20 世纪 80 年代拉开序幕, Tylor (1986) 年提出标准的随机波动率模型, 扩展传统的 ARCH 波动率研究模型。在模型中, 假设金融资产服从离散时间分布。Hull 和 white (1987) 在 Tylor 基础上, 假设金融资产服从连续时间分布。在研究中, 令波动率服从期权定价的 Ornstein-Uhlenbeck 随机过程。相比于 B-S 期权定价模型, 随机波动率表现的更稳健。

金融领域的波动大多数是持续和隐性的, 在随机波动率模型中, 波动是一个持续变化随机过程, 并将波动率作为潜在变量, 因此, 许多学者认为随机波动率模型更适合描述和刻画金融资产的波动游走。但是也正因为波动率作为潜在变量, 是不可观测的, 模型的估计需要大量的抽样和计算。受早期计量经济建模和人工计算的限制, 相对 ARCH 和 GARCH 模型, 随机波动率模型发展缓慢。到 20 世纪末期, 随着计算机和智能算法的发展, 抽样模拟的方法得到推广, 在随机波动率模型方面也有许多优秀的学术成果。大部分的实证表明, 在金融资产的长期波动预测中, 随机波动率模型相对 ARCH 簇模型能更好的描述金融数据特征。在风险管理和资产定价方面极具前景。

随机波动率模型存在不可观测变量, 模型结构复杂, 参数估计一直是模型的首要问题, 也是难点问题。早期主要有两种估计方法, 广义矩估计和拟最大似然估计。广义矩估计的假设条件宽松, 实现过程容易, 计算简便等, 成为了随机波动率模型参数估计的主要方法。主要是利用模型的遍历性和平稳性, 构建矩条件, 通过矩条件估计模型参数。该方法首次是 Hasen(1982)提出的, Melino 和 Turnbull (1990) 将其运用到随机波动率的模型估计中。Aandersen 和 Sorensen (1997) 理论推导中证明, 广义矩估计在随机波动率模型中, 样本矩收敛于无条件期望特别缓慢, 估计的精度和稳健型差, 只有在大样本下才是最优的。广义矩估计有较差的有限样本性质, 学者发展和改进了广义矩估计方法。模拟矩阵方法和有效矩阵方法相继被提出, 用于估计时间连续的随机波动率模型。特别是在利率期限结构等连续金融波动预测中有广泛的应用。

拟最大似然估计方法是建立在随机波动率模型的状态空间形式上的, 在假定状态方程的干扰项是正态分布的情况下, 通过变形, 可以转化为高斯线性空间状态模型, 其本质思想是基于极大似然估计。Ruiz (1994) 利用标准滤波方法得到模型

的对数似然函数。拟最大似然估计方法在估计随机波动率模型参数时，理论证明参数的估计是具有渐进正态性，也是一致估计。此估计方法多应用于多元 SV 模型和长记忆 SV 模型的参数估计，估计过程相对简单。但在有限样本的估计中，参数估计的误差大，精度低。对此 Fuller (1996) 在模型中引入补偿参数，消除内窗值对估计的影响。苏卫东和张世英 (2002) 将拟最大似然估计和遗传算法相结合，克服了算法在局部取的最优解。

1.2.1 国外文献综述

近年来随着抽样模拟估计的发展，贝叶斯蒙特卡罗模拟也迅速的到广泛应用。Jacquier, Polson 和 Rossi (1994) 首次将贝叶斯分析用于模型估计。在模型的估计中，使用贝叶斯蒙特卡罗模拟对参数进行后验分布抽样，将非线性滤子考虑到模型的整体估计中，作为贝叶斯处理的主要对象，证明模拟得到波动率的有效估计是可行的。但是在抽样模拟过程中，需要以观测值为条件的联合分布函数。考虑到可观测性，现实中这种方法是不可行的，将联合分布抽样分解成一系列的单变量分布抽样，每次抽样都以其它所有参数已知为条件。在这个分解的过程中，可能存在单个变量的分布不是标准的分布形态，不能通过正常的抽样方式取得。Jacquier, Polson 和 Rossi 为非标准分布的变量，提出了反复抽样的 Metropolis 接受拒绝独立链。并通过大量的模拟和仿真实验研究，在有限的样本估计下，MCMC 方法的参数估计精度和稳健型都要优于 GMM 和 QML 方法。

Shephard 和 Chib (1998) 在模型似然函数的问题上，提供了一个统一，切实可行的方法。将隐性马尔可夫链引入到蒙特卡罗模拟抽样中。对于随机波动率模型的潜在变量，通过马尔可夫链建立一个反复赋权的过程，使用渐进的方式抵消掉抽样过程中的噪音。同时，在基于模拟的过程上，提出了模型误差检验，有效滤波等。最后，用实际数据模拟，与其它的抽样方法比较，结果显示这种方法是最优的。

在方法不断创新的过程中，随机波动率模型的扩展形式被相继提出。Jacquier (1994) 在标准随机波动率模型和联合分布抽样的基础上，假设扰动项是服从厚尾 T 分布的，而不是正态分布。Koopman, S. J 和 Uspensky (2003) 在研究风险与预期收益的关系时，将波动引入到收益方程中，提出了 SV-M 模型，很好的

分析了收益与金融市场波动的关系。为了描述金融市场波动的长记忆性过程, Breidt, Crato 和 Lima (1994) 将随机波动率模型和 ARFIMA 模型结合, 研究和创新了一系列长记忆随机波动率模型。David Raggi (2004) 在用道琼斯综合指数进行实证分析时, 发现收益率是一种突变带跳的结构, 随后用贝叶斯蒙特卡罗模拟分析了带跳的仿射随机波动率模型, 发现在标准的随机波动率模型中, 扩散项过高的估计了波动率。

从计算的角度来看, ARCH 类型和 SV 模型的主要区别在于后者需要更多的计算负担。由于波动率在 SV 模型中非线性地进入, 他们的估计需要先进的数值方法, 例如 Ghysels (1996), Andersen (1999 年) 和 Ozturk、Richard (2015 年) 等。在多元分析的背景下, 多元 GARCH 模型已经有许多成功的应用。Harvey (1994) 提出了第一个多变量 SV (MSV) 潜在因子模型, 然后由 Shephard (1996), Pitt 和 Shephard (1999), Jacquier (1999 年) 发展, 除了捕捉简约性之外, MSV 还与套利定价理论有直接联系。基于包含均值方程的方法, 这些模型可以分为加法和乘法。Harvey 等人提出的第一个多变量因子模型是相加的, 在这些模型中, 收益率被分解为两个附加成分。第一个是较低维度, 它反映了所有资产共有的因素, 而另一个因素是特定于资产的特质因素。Shephard (1996), Pitt、Shephard (1999) 以及 Aguilar、West (2000) 扩展了这些模型。但是这些模型存在干扰项的协方差函数对角性假设太强的缺陷。Yu、Meyer (2006) 将时变波动率和与多元随机因素波动率模型相结合。此外, Philipov 和 Glickman (2006) 还开发了一个高维因子 MSV 模型, 其中因子协方差矩阵由 Wishart 方法可的。Nakajima (2016) 采用动态多变量因子 SV 模型, 使用蒙特卡洛 (MC) 近似估计, 通过高效重要性抽样 (EIS) 产生似然函数, 计算最大似然估计。高效重要性抽样是由 Richard 和 Zhang (2007) 引入的一种后向顺序 IS 程序。

1.2.2 国内文献综述

在 SV 模型与 ARCH 模型簇的比较方面: 余素红和张世英 (2002) 针对金融时间序列的两大特性, 即“高峰厚尾”和“平方序列微弱而持续的自相关”, 对比了 SVAR (1) 模型和 GARCH (1, 1) 模型。发现不论是从数学理论还是从上海股市的实证结果上, SVAR (1) 模型优于 GARCH (1, 1) 模型。王春峰和蒋祥林等 (2003) 基于马尔可夫

链蒙特卡罗 (MCMC) 动态模拟了中国股票市场的波动性, 并通过残存, 拟合优度, SIC 准则等综合比较了随机波动率模型和 ARCH 模型, 实证结果表明, 随机波动率模型能更好的刻画中国股票市场的异方差和波动性的序列相关性。刘凤芹和吴喜之 (2004), 根据对称和非对称两类评价准则, 比较了标准的 SV 模型、非对称 SV 模型和 GARCH 模型, GARCH 模型对深圳股市的预测效果最弱, 非对称 SV 模型次子, 标准的 SV 模型效果最好。

在随机波动率模型簇及其应用上: 田秋荣和栾长福 (2009) 在标准随机波动率模型的基础上, 假设误差项服从 t 分布、GED 分布, 利用 SV-N, SV- t 和 SV-GED 模型对中国的深指和上证指数数据进行仿真分析。实证结果表明, 深指和上证指数都具有强烈的波动持续性, 并且深指的波动水平更高, 风险大。SV-GED 模型相对其他模型能更好的刻画中国股市风险。吴萌和徐全智 (2005) 根据市场有效理论, 在随机游走过程上, 建立了厚尾 SV 模型, 长记忆 SV 模型和杠杆效应 SV 模型研究沪市日收益率。发现沪市有杠杆效应, 长记忆性和波动性持续性。孙有发和邓飞其 (2006) 在基于反馈理论上, 结合标准的随机波动率模型, 将反馈带跳的波动率模型运用到证券定价中。实证结果显示, 该模型能很好的模拟股市价格行为, 同时也能刻画出股市价格受到重大冲击, 股市价格的自我增强和乘数效应。周彦和张世英等 (2007) 发现带跳的随机波动率模型对股价波动的描述不是很理想, 将跳跃因子引入连续时间的随机波动率模型中, 对 1996-1997 和 2002-2004 年的中国股市进行实证分析, 发现我国股市 2002-2004 年的波动和跳跃较之前有所减少。刘伟 (2007) 使用随机波动率和 Black-Scholes 模型, 对中国股票波动进行了假设检验, 否定了股票指数服从 LEVY 过程。黄超 (2010) 利用 MCMC 抽样算法, 卡尔曼滤波和高斯平滑估计了跳跃厚尾的随机波动率模型, 并将其与标准随机波动率模型和厚尾波动率模型进行比较。发现在刻画中美股市波动时跳跃厚尾的随机波动率模型更好。刘鑫和孟昭为 (2011) 估计一元随机波动率时, 基于单步移动和多步移动抽样提出一个高效的 MCMC 算法。扩展到多元的随机波动率模型, 提出了带交叉杠杆随机波动率模型。吴鑫育和马超群等 (2012) 针对金融时间序列的有偏和厚尾特征, 提出了非对称 SV 模型。非对称 SV 模型具有尺寸效应和杠杆效应。在非对称 SV 模型引入偏广义误差分布来刻画股票市场波动。模型估计方面将有效重要性抽样和极大似然估计相结合。在对上证指数收益波动分析中, SV-LS 模

型表现最优，证明了我国沪市具有很强的波动持续性和杠杆效应。朱喜安和马兴祥（2012）使用 Gibbs 抽样算法 DIC 准则，研究了带解释变量杠杆 SV 模型。以 2010 年 4 月 16 日至 2012 年 2 月 17 日的上证指数为样本，进行实证分析，发现一周内，周一，周二和周五存在周内效应。在周一和周二有较高的风险和较低的收益，周五有较高的收益和较低的风险。苏理云和彭相武等（2016）根据蒙特卡罗模拟方法估计厚尾模型，然后结合 EM 算法估计混合正态参数，最后使用卡尔曼滤波对样本进行预测。研究表明在高斯混合状态空间下，AMF 相对 KF, EKF 更为有效的刻画了股市收益序列的尾部波动。吴鑫育和李心丹等（2018）将丰富日内高频信息的测度引入标准的随机波动率模型，同时考虑到误差修正和波动率的长记忆、非对称性，构建了双因子非对称波动率模型。在基于连续粒子滤波算法，用极大似然估计方法估计量模型模型参数。并且该模型能够很好的捕捉到股市的动态特征。

上述国内外的文献表明，相比于 ARCH 和 GARCH 模型，SV 模型从理论和实践上能更好的刻画和预测证券市场的波动。此外，SV 模型在时间的连续性方面能很好的与证券市场中的期权定价理论，利率期限结构理论等相联系起来。本文将使用高效的 MCMC 算法估计 SV 模型簇，并引入模型 SV-AR(2) 中，SV-M 模型和 SV-MA(1) 模型，分别研究股市收益波动率与前一期的相关性，股市波动的随机游走性，收益是否会受到当期的波动影响，即波动与收益的关系和股市的波动大部分来自于当前的冲击还是受到过去冲击的持续影响。

1.3. 研究思路

在研究和分析了国内外学者已有的成果上，本学位论文主要介绍了关于随机波动率的三种传统 MCMC 方法，随机游走的 M-H，独立马尔科夫链 M-H，线性滤波器 FFBS。同时引进了一种高效的 MCMC 方法，基于此，对 SV 模型，SV-M 模型进行了贝叶斯分析。同时，也引进了 SV-AR(2)，SV-MA(1) 模型和 SV-MG 模型对我国股市进行分析。全文分为五个部分，各个部分的研究内容安排如下：

第一章：本章主要论述随机波动率模型研究的理论，背景和意义；介绍了随机波动率模型的一些估计方法和当前研究现状。最后介绍本文的研究内容安排。

第二章：介绍了蒙特卡洛模拟抽样的主要方法，Gibbs 抽样与 M-H 抽样。然后

介绍了随机游走的 M-H，独立马尔科夫链 M-H，线性滤波器 FFBS。

第三章：引进了 J.Chan 的 MCMC 方法，并对传统的 SV 模型，SV-AR (2)，SV-M 模型进行了贝叶斯估计。同时，引进了 SV-MA(1), SV-MG 模型，进行了详细的贝叶斯分析。

第四章：本章先说明采用中证 500 指数的收盘价作为样本数据，然后基于前文的不同贝叶斯估计方法进行实证分析，比较了不同的方法下的估计结果。然后用 SV-AR(2)模型，SV-M 模型，SV-MA(1)模型和 SV-MG 模型进行了贝叶斯实证分析，用边缘最大似然进行了模型的比较。

总结：对全文工作的总结和对未来工作的展望。

1.4 本文创新

本文的主要创新点作如下归纳：

(1) 比较了随机波动率的三种传统 MCMC 方法，随机游走的 M-H，独立马尔科夫链 M-H，线性滤波器 FFBS 的估计精度和对中国股市的适应程度。

(2) 考虑到均值和带状矩阵特性，使用快速带和稀疏矩阵特殊性算法，更为高效的抽样来拟合这些模型。这个方法适合所有的随机波动率模型，区别与传统的抽样和卡尔曼滤波器。

(3) 本文参考 Kass and Raftery (1995)的贝叶斯因子方法，计算模型边缘似然密度，使用新型的评价准则 marginal likelihoods 来对模型进行比较研究。

2. 标准 SV 模型理论分析

2.1. Gibbs 抽样和 Metropolis-Hastings 抽样

2.1.1 Gibbs 抽样

在贝叶斯蒙特卡罗模拟抽样中，Gibbs 抽样方法是运用最为广泛的。Gibbs 抽样的详细步骤如下：

首先需要确定待估参数和潜在变量的初始， $\theta = (\theta_1, \theta_2 \cdots, \theta_m)$ ，其中

$\theta_1, \theta_2, \dots, \theta_m$ 指代模型中待估参数， $X = (X_1, X_2, \dots, X_m)$ 为状态变量， $y = (y_1, y_2, \dots, y_t)$ 为观测样本序列。

(1) θ_1^1 从 $p(\theta_1 | \theta_2^0, \theta_3^0, \dots, \theta_m^0, X^0, y)$ 密度分布中抽取

θ_2^1 从 $p(\theta_2 | \theta_1^1, \theta_3^0, \dots, \theta_m^0, X^0, y)$ 密度分布中抽取

\vdots

θ_m^1 从 $p(\theta_m | \theta_1^1, \theta_3^0, \dots, \theta_{m-1}^0, X^0, y)$ 密度分布中抽取

(2) X_1^1 从 $p(X_1 | \theta^1, X_{-X_1}, y)$ 密度分布中抽取

\vdots

X_t^1 从 $p(X_t | \theta^1, X_{-X_t}, y)$ 密度分布中抽取

(3) 重复(1)，(2)步骤N次直至马氏链平稳，剔除前M次不平稳状态。

在未知抽样值 θ_j^i 和已知其余变量，通过之前算法，在基于 θ^t 密度分布决定 θ^{t+1} 密度分布基础上，可以得到一个马尔可夫链 $\{\theta^t, X^t\}$ ，抽样完成后取均值为 θ 和 X_t 的估计值：

$$\theta = \frac{1}{N - M} \sum_{i=1}^{N-M} \theta^i, \quad X_t = \frac{1}{N - M} \sum_{i=1}^{N-M} X_t^i$$

Gibbs 抽样有一个重要的缺陷，要求抽样的密度分布是一个标准的分布函数，例如：正态分布，T 分布，Gamma 分布，Beta 分布等等。

2.1.2 Metropolis-Hastings 抽样：

通过前文的 Gibbs 抽样方法可知，只能应用于标准的密度分布函数。但在实际生活中，有许多参数是非标准分布的，限制了 Gibbs 抽样的使用。为此，Metropolis-Hastings 抽样方法被提出用于解决这类非标准分布的抽样。

Metropolis-Hastings 抽样步骤如下：

(1) 初始值 X_0 产生于初始分布

(2) W 产生于建议分布 $g(* | X_t)$

(3) U 产生于均匀分布 $U(0, 1)$

若 $\alpha(X_t, Y) \geq U$ ，则接受 W 并令 $X_{t+1} = M$ ，否则 $X_{t+1} = X_t$

(4) 增加 t ，返回到 (2)

其中 $\pi(X)$ 是参数后验分布， $\alpha(X_t, Y) = \min \left(1, \frac{\pi(M) g(X_t | Y)}{\pi(X_t) g(Y | X_t)} \right)$ 。

2.2. 标准 SV 的 MCMC 估计

Taylor (1986) 在假设金融资产服从离散时间分布，提出标准的随机波动率模型，解释了金融资产波动的自回归。模型如下：

$$\begin{aligned} y_t &= e^{h_t/2} \varepsilon_t \\ h_t &= \mu + \phi h_{t-1} + \tau \eta_t \end{aligned}$$

ε_t 和 η_t 都是标准状态分布， $\varepsilon_t, \eta_t \sim N(0, 1)$ 。对于所有的 h 都有 $E(\varepsilon_t \eta_{t+h}) = 0$ ，所有的 $l \neq 0$ ，我们都有 $E(\varepsilon_t \varepsilon_{t+l}) = E(\eta_t \eta_{t+l}) = 0$ 。 τ^2 ：是对数随机波动率。要求 $|\phi| < 1$ ，保证 h_t 是一个平稳过程。

$y^n = (y_1, \dots, y_n)'$, $h^n = (h_1, \dots, h_n)'$ ，同时定义 $h_{a:b} = (h_a, \dots, h_b)'$ 。

参数的先验分布是进行贝叶斯分析的前提，在先验分布的选取时，应遵循合理性原则和计算方便原则。共轭先验分布的形式是先验分布的首要选择关于不确定的初始波动率 $h_0 \sim N(m_0, C_0)$ ，令 $\theta = (\mu, \phi)'$ 且 $\varsigma = \tau^{-2}$ ， (θ, τ^2) 的先验分布为正态倒伽马分布，所以 $(\theta, \tau^2) \sim NIG(\theta_0, V_0, v_0, s_0^2)$ ，得到 $\theta | \varsigma \sim N(\theta_0, \varsigma^{-1} V_0)$ 和 $\varsigma \sim G(\frac{V_0}{2}, \frac{V_0 s_0^2}{2})$ 。利用先验分布，推导出各个参数的满条件后验分布：

(1) (θ, τ^2) 的后验分布：

在 $h_{0:n}$ 的条件下， (θ, τ^2) 的后验分布依旧是正态倒伽马分：

$$(\theta, \tau^2 | y^n, h_{0:n}) \sim NIG(\theta_1, V_1, v_1, s_1^2)$$

对于 $\pi(\theta, \varsigma)$ 可知 $\pi(\theta, \varsigma) = p(\theta | \varsigma) p(\varsigma) f(h | \theta, \varsigma)$ 。对

于 (θ, ς) 的先验分布，易得：

$$\begin{aligned} p(\theta, \varsigma) &\propto p(\theta \mid \varsigma) p(\varsigma) \\ &\propto \varsigma^{d/2} \exp\left\{-\frac{\varsigma}{2}(\theta - \theta_0)' V_0^{-1}(\theta - \theta_0)\right\} \varsigma^{(v_0/2)-1} \exp\left\{-\frac{\varsigma}{2} v_0 s_0^2\right\} \quad (1) \\ &\propto \phi^{[(v_0+d)/2]-1} \exp\left\{-\frac{\phi}{2}[v_0 s_0^2 + (\beta - \theta_0)' V_0^{-1}(\beta - \theta_0)]\right\} \end{aligned}$$

h 的条件分布为：

$$\begin{aligned} l(\theta, \varsigma) &= f(h \mid \theta, \varsigma) \\ &\propto \varsigma^{-n/2} \exp\left\{-\frac{\varsigma}{2}(h - X\theta)'(h - X\theta)\right\} \quad (2) \\ &= \varsigma^{-n/2} \exp\left\{-\frac{\varsigma}{2}[Q(\theta) + S_e]\right\} \end{aligned}$$

其中 $X=(1_n, h_{0:n-1})$ ，根据标准的线性回归模型：

$$Q(\theta) = (\theta - \hat{\theta})' X' X (\theta - \hat{\theta}), \quad \hat{\theta} = (X' X)^{-1} X' h, \quad S_e = (h - X\hat{\theta})' (h - X\hat{\theta})$$

结合公式 (1) 和 (2)，得到：

$$\begin{aligned} \pi(\theta, \varsigma) &= l(\theta, \varsigma) p(\theta, \varsigma) \\ &\propto \varsigma^{\frac{v_0+n+d}{2}-1} \exp\left\{-\frac{\varsigma}{2}[v_0 s_0^2 + S_e + (\theta - \theta_0)' V_0^{-1}(\theta - \theta_0) + Q(\theta)]\right\} \\ &\propto \varsigma^{\frac{d}{2}} \exp\left\{-\frac{\varsigma}{2}(\theta - \theta_1)' V_1^{-1}(\theta - \theta_1)\right\} \varsigma^{\frac{v_1}{2}-1} \exp\left\{-\frac{\varsigma}{2} v_1 s_1^2\right\} \end{aligned}$$

整理可得：

$$\begin{aligned} v_1 &= v_0 + n \\ V_1^{-1} &= V_0^{-1} + X' X \\ V_1^{-1} \theta_1 &= V_0^{-1} \theta_0 + X' h_{1:n} \\ v_1 s_1^2 &= v_0 s_0^2 + (y - X \theta_1)' (y - X \theta_1) + (\theta_1 - \theta_0)' V_0^{-1} (\theta_1 - \theta_0) \end{aligned}$$

(2) h_0 的后验分布：

已知 $h_0 \sim N(m_0, C_0)$ ，易得 $h_1 \mid h_0 \sim N(\mu + \phi h_0, \tau^2)$ ，根据联合正态分布

的边缘分布，我们可得 $h_0 \mid h_1 \sim N(m_1, C_1)$ ：

$$\begin{aligned} C_1^{-1} m_1 &= C_0^{-1} m_0 + \phi \tau^{-2} (h_1 - \mu) \\ C_1^{-1} &= C_0^{-1} + \phi^2 \tau^{-2} \end{aligned}$$

(3) h_t 的条件先验分布

在给定 h_{t-1} , θ , τ^2 的情况下, 对于 $t = 1, \dots, n-1$, 我们得到:

$$\begin{pmatrix} h_t \\ h_{t+1} \end{pmatrix} \sim N \left\{ \begin{pmatrix} \mu + \phi h_{t-1} \\ (1 + \phi)\mu + \phi^2 h_{t-1} \end{pmatrix}, \tau^2 \begin{pmatrix} 1 & \phi \\ \phi & (1 + \phi^2) \end{pmatrix} \right\}$$

根据联合正态分布可知, $E(h_t | h_{t-1}, h_{t+1}, \theta, \tau^2)$ 和 $V(h_t | h_{t-1}, h_{t+1}, \theta, \tau^2)$, 分别为

$$\begin{aligned} \mu_t &= \left(\frac{1 - \phi}{1 + \phi^2} \right) \mu + \left(\frac{\phi}{1 + \phi^2} \right) (h_{t-1} + h_{t+1}) \\ \nu^2 &= \tau^2 (1 + \phi^2)^{-1} \end{aligned}$$

因此:

$$\begin{aligned} (h_t | h_{t-1}, h_{t+1}, \theta, \tau^2) &\sim N(\mu_t, \nu^2) \\ (h_n | h_{n-1}, \theta, \tau^2) &\sim N(\mu_n, \tau^2) \end{aligned} \quad t = 1, \dots, n-1$$

对于 $t = n$, $\mu_n = \mu + \phi h_{n-1}$ 。

对于模型里的潜在变量, 我们使用单步 MCMC 的方法进行抽样。定义 $h_{-t} = (h_{0:t-1}, h_{t+1:n})$, 根据贝叶斯定理, 得到潜在变量的条件后验分布:

$$p(h_t | h_{-t}, y^n, \theta, \tau^2) = p(h_t | h_{t-1}, \theta, \tau^2) p(h_t | h_{t+1}, \theta, \tau^2) p(y_t | h_t)$$

从上述分布观测可知, 该分布不是标准分布, 需要使用 M-H 算法来抽样

2.3. 随机游走与 Metropolis-Hastings

如果对一个过程没有太多的信息, 在欧式空间上, 可以令目标分布迭代服从一个随机游走过程。而其中的随机干扰项可以设为一球面对称的建议分布。人工会选择性的控制建议转移函数的变化范围, 一般选择为球面高斯分布。对此处的状态变量, 假设为高斯分布 $N(0, \sigma_r^2)$, 其中人为控制方差项 σ_r^2 。

对于 $t = 1, \dots, n-1$, $\nu_t^2 = \nu^2$, $t = n$ 时 $\nu_n^2 = \tau^2$ 。根据潜在变量 h_t 的条件后验分布和 h_t 的条件先验分布:

$$\begin{aligned}
p(h_t \mid h_{-t}, y^n, \theta, \tau^2) &= p(h_t \mid h_{t-1}, \theta, \tau^2) p(h_t \mid h_{t+1}, \theta, \tau^2) p(y_t \mid h_t) \\
&= p(h_t \mid h_{t-1}, h_{t+1}, \theta, \tau^2) p(y_t \mid h_t) \\
&= f_N(h_t; \mu_t, \nu_t^2) f_N(y_t; 0, e^{h_t})
\end{aligned}$$

对于 $t = 1, \dots, n$ ，随机游走 Metropolis 抽样算法的步骤如下：

1. 已知的当前的状态： $h_t^{(j)}$
2. 在 1 的情况下，从建议的正态分布 $N(h_t^{(j)} \mid \nu_h^2)$ 抽取 h_t^*
3. 计算接受概率

$$\alpha = \min \left(1, \frac{f_N(h_t^*; \mu_t, \nu_t^2) f_N(y_t; 0, e^{h_t^*})}{f_N(h_t^{(j)}; \mu_t, \nu_t^2) f_N(y_t; 0, e^{h_t^{(j)}})} \right)$$

4. 更新后的状态：

$$h_t^{j+1} = \begin{cases} h_t^* & w.p. \quad \alpha \\ h_t^j & w.p. \quad 1 - \alpha \end{cases}, \circ$$

综合上述后验分布，我们得到采样的步骤如下：

1. $p(h \mid y, \mu, \phi, \tau^2)$

$$p(h_0 \mid h_1, \mu, \phi, \tau^2)$$

$$p(h_t \mid h_{-t}, \mu, \phi, \tau^2)$$
2. $p(\mu \mid y, h, \phi, \tau^2) = p(\mu \mid h, \phi, \tau^2)$
3. $p(\phi \mid y, \mu, h, \tau^2) = p(\phi \mid h, \mu, \tau^2)$
4. $p(\tau^2 \mid y, \mu, h, \phi) = p(\tau^2 \mid \mu, h, \phi)$

2.4. 独立马尔科夫链与 Metropolis-Hastings

已知 h_t 的满条件分布：

$$p(h_t \mid h_{-t}, y^n, \theta, \tau^2) = f_N(h_t; \mu_t, \nu_t^2) f_N(y_t; 0, e^{h_t})$$

Kim, Shephard 和 Chib (1998) 发现其对数似然函数可以表示为：

$$\log p(y_t | h_t) = \text{const} - \frac{1}{2} h_t - \frac{y_t^2}{2} \exp(-h_t)$$

在 μ_t 处泰勒展开 $\exp(-h_t)$ 得到:

$$\begin{aligned} \log p(y_t | h_t) &\approx \text{const} - \frac{1}{2} h_t - \frac{y_t^2}{2} (e^{-\mu_t} - (h_t - \mu_t) e^{-\mu_t}) \\ g(h_t) &= \exp \left\{ -\frac{1}{2} h_t (1 - y_t^2 e^{-\mu_t}) \right\} \end{aligned}$$

对于 $t = 1, \dots, n-1$, $v_t^2 = v^2$, $t = n$ 时 $v_n^2 = \tau^2$ 。结合 $f_N(h_t; \mu_t, v_t^2)$

和 $g(h_t)$, 可以得到下面的建议分布:

$$\begin{aligned} q(h_t | h_{-t}, y^n, \theta, \tau^2) &= f_N(h_t; \mu_t, v_t^2) g(h_t) \\ &= N(h_t; \tilde{\mu}_t, v_t^2) \end{aligned} \quad t = 1, \dots, n$$

其中 $\tilde{\mu}_t = \mu_t + 0.5 v_t^2 (y_t^2 e^{-\mu_t} - 1)$, 由此可见, 建议分布是相处独立的, 抽样过程也相互独立。在抽样接受和拒绝时, 目标分布于建议分布相比得到重点抽样权, 然后再计算接受概率。

对于 $t = 1, \dots, n$, 独立的 Metropolis 抽样算法的步骤如下:

1. 已知的当前的状态: $h_t^{(j)}$
2. 在 1 的情况下, 从正态分布 $N(\tilde{\mu}_t, v_t^2)$ 抽取 h_t^*
3. 计算接受概率

$$\alpha = \min \left(1, \frac{f_N(h_t^*; \mu_t, v_t^2) f_N(y_t; 0, e^{h_t^*})}{f_N(h_t^{(j)}; \mu_t, v_t^2) f_N(y_t; 0, e^{h_t^{(j)}})} \times \frac{f_N(h_t^{(j)}; \tilde{\mu}_t, v_t^2)}{f_N(h_t^*; \tilde{\mu}_t, v_t^2)} \right)$$

4. 更新后的状态:

$$h_t^{j+1} = \begin{cases} h_t^* & w.p. \quad \alpha \\ h_t^j & w.p. \quad 1 - \alpha \end{cases}$$

2.5. 线性滤波器抽样方法

随机波动率模型假设潜在的状态变量是来自同一个总体, 有相似的特征。基于上述假设, 将全部状态变量作为一个总体进行联合抽样, 从而提高估计效率。

但是随机波动率模型是一个非线性的模型，在进行潜在变量联合抽样之前要先将模型转换成空间线性模型。即对观测方程的两端平方后，取对数， $y_t^* = \log y_t^2$ 和

$$\varsigma_t = \log \varepsilon_t^2$$

此刻 SV-AR (1) 变成了可观察残差为非正态分布的 DLM 模型：

$$\begin{aligned} y_t^* &= h_t + \varsigma_t \\ h_t &= \mu + \phi h_{t-1} + \tau \eta_t \end{aligned} \quad \eta_t \sim N(0, 1)$$

模型能转化成线性高斯的状态空间形式，则可利用“向前滤波，向后抽样”方法完成状态向量的联合抽样。模型的观测方程经过对数变换之后， ς_t 服从自

由度为 1 的对数卡方分布 $\log \chi_1^2$ ，模拟计算可以得到： $E(\varsigma_t) = -1.27$

$$V(\varsigma_t) = \frac{\pi^2}{2} = 4.935$$

这里用正态分布 $N(\alpha, \sigma^2)$ 近似 ς_t ， $z_t = y_t^* - \alpha$ ， $\alpha = -1.27$ ， $\sigma^2 = \frac{\pi^2}{2}$ ，

替换后方程为：

$$\begin{aligned} z_t &= h_t + \sigma v_t \\ h_t &= \mu + \phi h_{t-1} + \tau \eta_t \end{aligned}$$

这是一个简单的 DLM 模型， $v_t, \eta_t \sim N(0, 1)$ ，在此基础上，我们使用向前滤波，向后平滑（FFBS）的方法完成潜在变量的抽样。我们将上述方程写成标准的动态线性模型：

$$\text{观测方程: } y_t = F_t' \beta_t + \varepsilon_t \quad \varepsilon_t \sim N(0, V)$$

$$\text{状态方程: } \beta_t = G_t \beta_{t-1} + \omega_t \quad \omega_t \sim N(0, W)$$

$$\text{初始先验: } (\beta \mid D_0) \sim N[m_0, C_0]$$

相对应的：

$$y_t = z_t, \quad \beta_t = (u_t, h_t)', \quad V = \sigma^2, \quad W = \tau^2 F_t = (0, 1)', \quad G_t = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

观测和状态误差序列是独立的，并且是相互独立的，且与初始先验独立。给

定 $t = 0$ 的先验信息 D_0 时,任一时刻 t 的有效信息集合为 $D_t = \{y_t, D_{t-1}\}$, 这里 y_t 是序列在 t 时刻的观察值, 为了给出潜在变量, 给出下列递推修正:

$$p(\beta_t | y^t) \Rightarrow p(\beta_{t+1} | y^t) \Rightarrow p(y_{t+1} | \beta_t) \Rightarrow p(\beta_{t+1} | y^{t+1})$$

1. 向前滤波过程

1.1 $t-1$ 时刻的后验分布: 对于均值 m_{t-1} 和方差矩阵 C_{t-1} 有:

$$\beta_{t-1} | y^{t-1} \sim N(m_{t-1}, C_{t-1})$$

1.2 t 时刻的先验分布:

由 (1) 和状态方程可知:

$$a_t = E(\beta_t | D_{t-1}) = E(G_t \beta_{t-1} + \omega_t | D_{t-1}) = G_t \beta_{t-1}$$

$$R_{t-1} = \text{var}(\beta_t | D_{t-1}) = \text{var}(G_t \beta_{t-1} + \omega_t | D_{t-1})$$

$$= G_t C_{t-1} G_t' + W_t$$

得到 $\beta_t | y^{t-1} \sim N(a_t, R_{t-1})$

1.3 一步预测分布:

结合 (2) 和状态方程, 得到关于 y_t 和 β_t 关于 D_{t-1} 的条件下的联合正态分布, 其协方差为:

$$\text{cov}(y_t, \beta_t | D_{t-1}) = \text{cov}(F_t' \beta_t + v_t, \beta_t | D_{t-1})$$

$$= F_t' \text{var}(\beta_t | D_{t-1}) = F_t' R_t = A_t' Q_t$$

其中 $A_t = R_t F_t' / Q_t$, 另外 y_t 的均值 $f = F_t' a_t$ 和方差 $Q_t = F_t' R_t F_t + V_t$ 。

可有观测方程和 (2) 得到, 因此有:

$$\begin{pmatrix} y_t \\ \beta_t \end{pmatrix} | D_{t-1} \sim N \left[\begin{pmatrix} f_t \\ a_t \end{pmatrix}, \begin{pmatrix} Q_t & Q_t A_t' \\ A_t Q_t & R_t \end{pmatrix} \right]$$

整理得到: $y_t | D_{t-1} \sim N(f_t, Q_t)$

1.4 t 时刻的后验分布:

由贝叶斯定理有:

$$p(\beta_t | D_t) \propto f(y_t | \beta_t, D_{t-1}) p(\beta_t | D_{t-1})$$

而由观测方程直接有：

$$f(y_t | \beta_t, D_{t-1}) \propto \exp \left\{ - (y_t - F_t' \beta_t)^2 / (2V_t) \right\}$$

由 (2) 有：

$$p(\beta_t | D_{t-1}) \propto \exp \left\{ - (\beta_t - a_t)' R_t^{-1} (\beta_t - a_t) / 2 \right\}$$

结合上述两式，两端取自然对数，得：

$$\begin{aligned} -2 \ln p(\beta_t | D_t) &= (\beta_t - a_t)' R_t^{-1} (\beta_t - a_t) + (y_t - F_t' \beta_t)^2 V_t^{-1} + \text{const} \\ &= \beta_t' (R_t^{-1} + F_t F_t' V_t^{-1}) \beta_t - 2 \beta_t' (R_t^{-1} a_t + F_t y_t V_t^{-1}) + \text{const}_2 \end{aligned}$$

其中 const_1 和 const_2 都是与 β_t 无关常数：

因为 $(R_t^{-1} + F_t F_t' V_t^{-1}) C_t = I$ ，所以 $C_t^{-1} = (R_t^{-1} + F_t F_t' V_t^{-1})$ ，

又因为 $C_t^{-1} m_t = (R_t^{-1} a_t + F_t y_t V_t^{-1})$

$$\begin{aligned} \text{故得 } -2 \ln p(\beta_t | D_t) &= \beta_t' C_t^{-1} \beta_t - 2 \beta_t' C_t^{-1} m_t + \text{const}_2 \\ &= (\beta_t - m_t)' C_t^{-1} (\beta_t - m_t) + \text{const}_2 \end{aligned}$$

整理后： $p(\beta_t | D_t) \propto \exp \left\{ - (\beta_t - m_t)' C_t^{-1} (\beta_t - m_t) / 2 \right\}$

易得： $\beta_t | y^t \sim N(m_t, C_t)$

在证明中，我们遇到一些等式，归纳如下：

$$\begin{aligned} m_t &= a_t + A_t e_t \\ C_t &= R_t - A_t A_t' Q_t \\ A_t &= R_t F_t / Q_t \\ e_t &= y_t - f_t \end{aligned}$$

2. 向后平滑：

在动态线性模型中，通常的平滑分布如下：

$$\begin{aligned} \pi(\beta | y^n) &= p(\beta_n | y^n) \prod_{t=1}^{n-1} p(\beta_t | \beta_{t+1}, \dots, \beta_n, y^n) \\ &= p(\beta_n | y^n) \prod_{t=1}^{n-1} p(\beta_t | \beta_{t+1}, y^n) \end{aligned}$$

结合已知的 $(\beta_1, \dots, \beta_{t-1})$ ，整理得到：

$$\pi(\beta_t, \dots, \beta_n \mid y^n) = p(\beta_n \mid y^n) \prod_{k=1}^{n-1} p(\beta_k \mid \beta_{k+1}, y^t) \quad t = 1, \dots, n-1$$

$$\pi(\beta_t, \beta_{t+1} \mid y^n) = p(\beta_{t+1} \mid y^n) p(\beta_t \mid \beta_{t+1}, y^n) \quad t = 1, \dots, n-1$$

上述方程对于得到 $\beta_t \mid V, W, y^n$ 的边缘分布提供了一个简单的递归形式，在连续获得 $\beta_t \mid y^t$ 的更新分布后，从时间的反方向上获得 $\beta_t \mid y^n$ ，从而获得连续分布 $\beta_t \mid y^n$ ：

$$\beta_t \mid V, W, y^n \sim N(m_t^n, C_t^n)$$

$$m_t^n = m_t + C_t G_{t+1}' R_{t+1}^{-1} (m_{t+1}^n - a_{t+1})$$

$$C_t^n = C_t - C_t G_{t+1}' R_{t+1}^{-1} (R_{t+1} - C_{t+1}^n) R_{t+1}^{-1} G_{t+1} C_t$$

结果表明， β_t 的满条件分布是正态分布，分解的分布可从更新方程中得到。

在明确地定义条件后，由贝叶斯定理给出

$$\begin{aligned} p(\beta_t \mid \beta_{t+1}, \sigma^2, W, y^t) &\propto p(\beta_{t+1} \mid \beta_t, \sigma^2, W, y^t) p(\beta_t \mid \sigma^2, W, y^t) \\ &\propto f_N(\beta_{t+1}; G_{t+1} \beta_t, W_{t+1}) f_N(\beta_t; m_t, C_t) \end{aligned}$$

右边的第一个式子是似然比函数，第二个式子是先验分布，从这个角度看，观测 β_{t+1} 是从一个标准的线性回归方程得到，其中回归的设计矩阵为 G_t ，参数 β_t 的先验分布通过更新方程得到。依据标准的线性回归，很容易得到：

$$(\beta_t \mid \beta_{t+1}, \sigma^2, W, y^t) \sim N[(G_t' W^{-1} G_t + C_t^{-1})^{-1} (G_t' W^{-1} \beta_{t+1} + C_t^{-1} m_t), (G_t' W^{-1} G_t + C_t^{-1})^{-1}]$$

对于潜在变量 β 的抽样方案如下：

1. 从更新分布 $\beta_t \mid y^t \sim N(m_t, C_t)$ 中抽取 β_n
2. 从平滑后的分布 $(\beta_t \mid \beta_{t+1}, \sigma^2, W, y^t)$ 中抽取 β_t
3. 将 t 减小到 $t - 1$ 并返回步骤 2，直到 $t=1$

3. SV 模型的扩展与更新

3.1. 高斯近似分布与 Newton-Raphson 的 SV-AR(1)

在本节中，考虑到均值和带状矩阵特性，使用了更为高效的算法来拟合这些模型。模型表述如下（ y_t 表示为原始数值的对数形式）：

$$\begin{aligned} y_t &= \mu + \varepsilon_t^y & \varepsilon_t^y &\sim N(0, e^{h_t}) \\ h_t &= \mu_h + \phi_h(h_{t-1} - \mu_h) + \varepsilon_t^h & \varepsilon_t^h &\sim N(0, \omega_h^2) \end{aligned}$$

对数潜在波动率 h_t 服从一个 $|\phi_h| < 1$ 的稳定 AR(1) 过程，初始化的 h_1 为：

$$h_1 \sim N(\mu_h, \omega_h^2 / (1 - \phi_h^2))$$

对于均值的 SV 模型，我们假设 μ 和 ϕ_h 的先验分布如下：

$$\begin{aligned} \mu &\sim N(\mu_0, V_\mu), & \mu_h &\sim N(\mu_{h0}, V_{\mu_h}) \\ \phi_h &\sim N(\phi_{h0}, V_{\phi_h}) \mathbf{1}(|\phi_h| < 1), & \omega_h^2 &\sim IG(V_h, S_h) \end{aligned}$$

在这一部分中，本文借鉴一个适合所有的随机波动率模型的通用方法，由于观测方程关于潜在变量 h 不是一个线性的，在数据和其他模型参数条件下，估算 h 的联合分布成了估算模型中最困难的步骤，这个方法的一个主要特点是建立在快速带和稀疏矩阵算法上，而不是使用传统的卡尔曼滤波器。

更具体地说，这个方法是利用了问题的特殊结构，即 h 的对数条件密度的 Hessian 矩阵是一个带矩阵，它只包含少量的非零元素，沿着一条狭窄的对角线带。这一特性对于开发高效的采样算法非常重要。通过下列顺序抽样得到后验描述：

1. $p(h \mid y, \mu, \mu_h, \phi_h, \omega_h^2)$
2. $p(\mu \mid y, h, \mu_h, \phi_h, \omega_h^2) = p(\mu \mid y, h)$
3. $p(\mu_h \mid y, \mu, h, \phi_h, \omega_h^2) = p(\mu_h \mid h, \phi_h, \omega_h^2)$
4. $p(\omega_h^2 \mid y, \mu, h, \mu_h, \phi_h) = p(\omega_h^2 \mid h, \mu_h, \phi_h)$
5. $p(\phi_h \mid y, \mu, h, \mu_h, \omega_h^2) = p(\phi_h \mid h, \mu_h, \omega_h^2)$

第 1 步:

抽样中, 联合概率密度分布函数 $p(h \mid y, \mu, \mu_h, \phi_h, \omega_h^2)$ 是高维和非标准分布。对于标准的随机波动率模型, 这一步可以使用 Kim, Shepherd, and Chib (1998) 的混合拟合分布采样来完成。但是这种方法对于受到模型的限制, 不能推广应用于一些扩展的随机波动率模型。这里我们介绍和使用 Chan and Jeliazkov (2009) 的方法。

根据 Chan (2014) 的方法, 我们首先得到 $p(h \mid y, \mu, \mu_h, \phi_h, \omega_h^2)$ 的高斯近似分布, 接着使用这个近似分布作为 Metropolis-Hastings 算法的建议密度函数, 潜在变量可以通过精密采样得到 (Chan and Jeliazkov (2009))。

为了用高斯密度函数近似 $p(h \mid y, \mu, \mu_h, \phi_h, \omega_h^2)$, 由贝叶斯定理可知:

$$p(h \mid y, \mu, \mu_h, \phi_h, \omega_h^2) \propto p(y \mid \mu, h) p(h \mid \mu_h, \phi_h, \omega_h^2)$$

通过上式, 我们首先推导 $p(y \mid \mu, h)$ 和 $p(h \mid \mu_h, \phi_h, \omega_h^2)$ 密度分布的显示表达式, 可以证明后者的密度函数是高斯的 (Chan, 2014), 定义 H_{ϕ_h} 为如下的三角矩阵:

$$H_{\phi_h} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi_h & 1 & 0 & \dots & 0 \\ 0 & -\phi_h & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -\phi_h & 1 \end{pmatrix}$$

将潜在变量等式改写成 $H_{\phi_h} h = H_{\phi_h} \mu + \varepsilon_t^h$, 可以得到:

$$(h \mid \mu_h, \phi_h, \omega_h^2) \sim N(\sigma_h, (H_{\phi_h}^{-1} \sum_h^{-1} H_{\phi_h})^{-1})$$

式中 $\sum_h = \text{diag}(\omega_h^2 / (1 - \phi_h^2), \omega_h^2, \dots, \omega_h^2)$, 并有 $\sigma_h = H_{\phi_h}^{-1} \tilde{\sigma}_h$, 其中:

$$\tilde{\sigma}_h = (\mu_h, (1 - \phi_h)\mu_h, \dots, (1 - \phi_h)\mu_h)'$$

上述整理后, 可以得到对数密度分布函数如下:

$$\log p(h \mid \mu_h, \phi_h, \omega_h^2) = -\frac{1}{2} (h' H_{\phi_h}' \sum_h^{-1} H_{\phi_h} h - 2h' H_{\phi_h}' \sum_h^{-1} H_{\phi_h} \sigma_h) + c_1$$

式中 c_1 为常数, 且独立于对数波动率 h 。

下一步，我们通过 h 的高斯分布对 $p(y \mid \mu, h)$ 密度函数近似。为此，在给定
点 $\tilde{h} = (\tilde{h}_1, \dots, \tilde{h}_T)' \in \mathfrak{R}^T$ 处对函数 $p(y \mid \mu, h) = \sum_{t=1}^T \log(y_t \mid \mu, h_t)$ 进行二阶
泰勒展开，得到：

$$\begin{aligned} \log p(y \mid \mu, h) &\approx \log(y \mid \mu, \tilde{h}) + (h - \tilde{h})' f - \frac{1}{2} (h - \tilde{h})' G (h - \tilde{h}) \\ &= -\frac{1}{2} (h' G h - 2 h' (f + G \tilde{h})) + c_2 \end{aligned}$$

同上式中 c_2 为常数，且独立于对数波动率 h 和 $f = (f_1, \dots, f_T)'$ ，
 $G = \text{diag}(G_1, \dots, G_T)$ 。

$$\text{对此有： } f_t = \frac{\partial}{\partial h_t} \log p(y_t \mid \mu, h_t) \Big|_{h_t = \tilde{h}_t} \text{ 和 } G_t = \frac{\partial^2}{\partial h_t^2} \log p(y_t \mid \mu, h_t) \Big|_{h_t = \tilde{h}_t}。$$

从上可知， G 是对数密度函数在 \tilde{h} 处估计的负海塞矩阵。对于标准的随机波
动率模型， G 是一个对角线矩阵。特别的，当给定 μ 和 h_t 时， y_t 的对数概率密度
分布如下：

$$\log p(y_t \mid \mu, h) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} h_t - \frac{1}{2} e^{-h_t} (y_t - \mu)^2$$

对上式求一阶和二阶导数，易得：

$$\frac{\partial}{\partial h_t} \log p(y_t \mid \mu, h_t) = -\frac{1}{2} + \frac{1}{2} e^{-h_t} (y_t - \mu)^2$$

$$\frac{\partial^2}{\partial h_t^2} \log p(y_t \mid \mu, h_t) = -\frac{1}{2} e^{-h_t} (y_t - \mu)^2$$

现在，结合上述 $p(y \mid \mu, h)$ 和 $p(h \mid \mu_h, \phi_h, \omega_h^2)$ ，我们可以得到：

$$\begin{aligned} p(h \mid y, \mu, \mu_h, \phi_h, \omega_h^2) &= \log p(y \mid \mu, h) + \log p(h \mid \mu_h, \phi_h, \omega_h^2) + c_3 \\ &\approx -\frac{1}{2} (h' K_h h - 2 h' k_h) + c_4 \end{aligned} \quad (3)$$

c_3 和 c_4 是独立于波动率 h 的常数，式中：

$$K_h = H_{\phi_h}' \sum_h^{-1} H_{\phi_h} + G \quad k_h = f + G \tilde{h} + H_{\phi_h}' \sum_h^{-1} H_{\phi_h} \sigma_h$$

实际上表达式 3 是正态分布 $N(\hat{h}, K_h^{-1})$ 的对数密度函数的核函数, $\hat{h} = K_h^{-1} k_h$ 。因此, 我们可以用均值为向量 \hat{h} , 精确的方差矩阵 K_h 的高斯分布来近似 $p(h | y, \mu, \mu_h, \phi_h, \omega_h^2)$ 。其中也请注意, K_h 是一个带状矩阵, 其非零元素只出现在主对角线和对角线上, 且 K_h 主要分布在主对角线以下, 接而, \hat{h} 可以通过线性系统 $K_h x = k_h$ 快速求解。我们可以用 Chan and Jeliazkov (2009) 的精确采样方法在 $N(\hat{h}, K_h^{-1})$ 进行抽样。这个高斯分布被用作为 Metropolis-Hastings 算法中的接受和拒绝的建议分布。

最后, 对于二阶泰勒展开式中的给定点 \tilde{h} , 是通过 Newton-Raphson 的方法得到, 具体细节, 请参考 (Kroese, Taimre, and Botev, 2011)。首先, 我们由 (3) 得到对数密度分布在 $h = \tilde{h}$ 处的负海塞矩阵 K_h 和在该点的梯度 $-K_h \tilde{h} + k_h$ 。因此, 我们实施 Newton-Raphson 的方法如下: 初始化 $\tilde{h} = h^{(l)}$, 常数化向量 $h^{(l)}$ 。对于 $l = 1, 2, \dots$ 在估算出 K_h 和 k_h 下, $\tilde{h} = h^{(l)}$, 计算:

$$h^{(l+1)} = h^{(l)} + K_h^{-1}(-K_h h^{(l)} + k_h) = K_h^{-1} k_h$$

重复这个过程直到达到某种收敛标准。第 2, 3, 4 步是根据线性回归以及联合分布的边缘分布, 相对应的满条件分布是标准分布。

第 2 步:

$$(\mu | y, h) \sim N(\hat{\mu}, D_\mu)$$

$$\text{其中 } D_\mu^{-1} = V_\mu^{-1} + \sum_{t=1}^T e^{-h_t} \text{ 和 } \hat{\mu} = D_\mu^{-1}(V_\mu^{-1} \mu_0 + \sum_{t=1}^T e^{-h_t} y_t)$$

第 3 步:

$$(\mu_h | h, \phi_h, \omega_h^2) \sim N(\hat{\mu}_h, D_{\mu_h})$$

$$\text{其中 } D_{\mu_h}^{-1} = V_{\mu_h}^{-1} + X_{\mu_h}' \sum_h^{-1} X_{\mu_h} \text{ 和 } \hat{\mu}_h = D_{\mu_h}^{-1}(V_{\mu_h}^{-1} \mu_{h0} + X_{\mu_h}' \sum_h^{-1} Z_{\mu_h})$$

$$\text{等 式 中 的 } X_{\mu_h} = (1, 1 - \phi_h, \dots, 1 - \phi_h)',$$

$$Z_{\mu_h} = (h_1, h_2 - \phi_h h_1, \dots, h_T - \phi_h h_{T-1})' \text{ 和 } \sum_h = \text{diag}(\omega_h^2 / (1 - \phi_h^2), \omega_h^2, \dots, \omega_h^2)$$

第 4 步:

$$(\omega_h^2 \mid h, \mu_h, \phi_h) \sim IG(V_h + T / 2, \tilde{S}_h)$$

$$\text{其中 } \tilde{S}_h = S_h + ((1 - \phi_h^2) (h_1 - \mu_h)^2 + \sum_{t=2}^T (h_t - \mu_h - \phi_h (h_{t-1} - \mu_h))^2) / 2$$

第 5 步:

结合相对应的波动率方程和参数的先验分布, 由贝叶斯定理可知:

$$p(\phi_h \mid h, \mu_h, \omega_h^2) \propto p(\phi_h) g(\phi_h) e^{-\frac{1}{2\omega_h^2} \sum_{t=2}^T (h_t - \mu_h - \phi_h (h_{t-1} - \mu_h))^2}$$

$$\text{其中 } g(\phi_h) = (1 - \phi_h^2)^{1/2} \exp(-\frac{1}{2\omega_h^2} (1 - \phi_h^2) (h_1 - \mu_h)^2), \quad p(\phi_h) \text{ 已经由先验分}$$

布给出。可以验证, $p(\phi_h \mid h, \mu_h, \omega_h^2)$ 是一个非标准分布, 基于上述, 使用独

立的 Metropolis-Hastings 算法, 在建议分布 $\phi_h \sim N(\hat{\phi}_h, D_{\phi_h}) \mathbf{1}(|\phi_h| < 1)$ 下抽样,

得到 ϕ_h , 表达式中, $D_{\phi_h} = (V_{\phi_h}^{-1} + X_{\phi_h}' X_{\phi_h} / \omega_h^2)^{-1}$, $\hat{\phi}_h = D_{\phi_h} (V_{\phi_h}^{-1} \phi_{h0} + X_{\phi_h}' z_{\phi_h} / \omega_h^2)$,

$$\text{其中 } X_{\phi_h} = (h_1 - \mu_h, \dots, h_{T-1} - \mu_h)', \quad z_{\phi_h} = (h_2 - \mu_h, \dots, h_T - \mu_h)'$$

所以, 在得到当前 ϕ_h 的抽样后, 建议样本 ϕ_h^* 以概率 $\min \left\{ 1, \frac{g(\phi_h^*)}{g(\phi_h)} \right\}$ 接受, 否

则, 马尔科夫链停留在当前的状态 ϕ_h 。

3.2. SV-AR(2) 模型的贝叶斯估计

相比于 SV-AR (1) 中, 对数波动率 h_t 服从 AR (1) 过程, SV-AR (2) 则服从 AR (2) 过程:

$$\begin{aligned} y_t &= \mu + \varepsilon_t^y & \varepsilon_t^y &\sim N(0, e^{h_t}) \\ h_t &= \mu_h + \phi_h (h_{t-1} - \mu_h) + \rho_h (h_{t-2} - \mu_h) + \varepsilon_t^h & \varepsilon_t^h &\sim N(0, \omega_h^2) \end{aligned}$$

我们假设关于 (ϕ_h, ρ_h) 的特征多项式的根, 都是在单位园外, 即 AR (2) 是一个平稳的过程, 进一步, 我们初始化其过程:

$$h_1, h_2 \sim N\left(\mu_h, \frac{(1 - \rho_h)\omega_h^2}{(1 + \rho_h)((1 - \rho_h)^2 - \phi_h^2)}\right)$$

对于转移方程的变体 AR (2) 的估计, 我们只需要在标准随机波动率的主要算法上进行一些细小的调整即可。定义:

$$H_{\theta_h} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ -\rho_h & -\phi_h & 1 & 0 & \dots & 0 \\ 0 & -\rho_h & -\phi_h & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -\rho_h & -\phi_h & 1 \end{bmatrix}$$

我们将关于 h_t 的状态方程改写成:

$$H_{\theta_h} h = \tilde{\gamma}_h + \varepsilon^h \quad \varepsilon^h \sim N(0, P_h)$$

其中 $\varepsilon^h = (\varepsilon_1^h, \dots, \varepsilon_T^h)'$, $\tilde{\gamma}_h = (\mu_h, \mu_h, (1 - \phi_h - \rho_h)\mu_h, \dots, (1 - \phi_h - \rho_h)\mu_h)'$ 。

P_h 是一个对角矩阵, 矩阵中的前两个元素是:

$$(1 - \rho_h)\omega_h^2(1 + \rho_h)^{-1}((1 - \rho_h)^2 - \phi_h^2)^{-1}$$

剩余 T-2 个元素是 ω_h^2 :

$$P_h = \text{diag}((1 - \rho_h)\omega_h^2(1 + \rho_h)^{-1}((1 - \rho_h)^2 - \phi_h^2)^{-1}, (1 - \rho_h)\omega_h^2(1 + \rho_h)^{-1}((1 - \rho_h)^2 - \phi_h^2)^{-1}, \dots, \omega_h^2)$$

类似 AR (1) 过程, 有 $(h \mid \mu_h, \phi_h, \rho_h, \omega_h^2) \sim N(\gamma_h, (H_{\theta_h}' P_h^{-1} H_{\theta_h})^{-1})$, 其中

$\gamma_h = H_{\theta_h}^{-1} \tilde{\gamma}_h$ 。可以得到对数似然函数:

$$\log p(h \mid \mu_h, \phi_h, \rho_h, \omega_h^2) = -\frac{1}{2} (h' H_{\theta_h}' P_h^{-1} H_{\theta_h} h - 2h' H_{\theta_h}' P_h^{-1} H_{\theta_h} \gamma_h) + c_5 \quad (4)$$

其中 c_5 是独立于 h 的常数。因此, 我们只用将 (3) 式替换成 (4) 式, 其余的步骤可以由标准的 随机波动率模型实现。

3.3. SV-M 模型的贝叶斯估计

关于SV的另一种变体，SV-M (Koopman and Hol Us-pensky (2002)用来研究波动率的反馈情况。在SV-M模型下随机波动率进入观察方程：

$$\begin{aligned} y_t &= \mu + \alpha e^{h_t} + \varepsilon_t^y & \varepsilon_t^y &\sim N(0, e^{h_t}) \\ h_t &= \mu_h + \phi_h(h_{t-1} - \mu_h) + \varepsilon_t^h & \varepsilon_t^h &\sim N(0, \omega_h^2) \end{aligned}$$

估算SV-M，我们只需要在原来的主算法上，进行两个细节的调整即可，首先，关于条件似然函数在 h_t 的一阶导数和二阶导数变成：

$$\begin{aligned} \frac{\partial}{\partial h_t} \log p(y_t | \mu, h_t) &= -\frac{1}{2} - \frac{1}{2} \alpha^2 e^{h_t} + \frac{1}{2} e^{-h_t} (y_t - \mu)^2 \\ \frac{\partial^2}{\partial h_t^2} \log p(y_t | \mu, h_t) &= -\frac{1}{2} \alpha^2 e^{h_t} - \frac{1}{2} e^{-h_t} (y_t - \mu)^2 \end{aligned}$$

所以，对数波动率 h 可以依照之前算法进行抽样。其次，我们将抽样过程的第2步，从密度分布 $p(\mu, \alpha | y, h, \mu_h, \phi_h, \omega_h^2) = p(\mu, \alpha | y, h)$ 中联合抽样 (μ, α) ，更新成 $\beta = (\mu, \alpha)'$ ， $V_\beta = \text{diag}(V_\mu, V_\alpha)$ ， $\beta_0 = (\mu_0, \alpha_0)'$ ，

$$X_\beta = \begin{pmatrix} 1 & e^{h_1} \\ \vdots & \vdots \\ 1 & e^{h_T} \end{pmatrix}$$

依据标准的分布的结果，我们的到 $(\mu, \alpha | y, h) \sim N(\hat{\beta}, D_\beta)$ ，其中

$$D_\beta^{-1} = V_\beta^{-1} + X_\beta' \sum_y^{-1} X_\beta \quad \text{和} \quad \hat{\beta} = D_\beta (V_\beta^{-1} \beta_0 + X_\beta' \sum_y^{-1} Y)$$

式中 $\sum_y = \text{diag}(e^{h_1}, \dots, e^{h_T})$ 。

3.4. SV-MA(1)模型的贝叶斯估计

在这个版本的随机波动率中，我们假设误差项是移动平均过程。在此处，只讨论一阶移动平均过程。

$$\begin{aligned}
y_t &= \mu + \varepsilon_t^y & u_t &\sim N(0, e^{h_t}) \\
\varepsilon_t^y &= u_t + \psi u_{t-1} & \varepsilon_t^h &\sim N(0, \omega_h^2) \\
h_t &= \mu_h + \phi_h(h_{t-1} - \mu_h) + \varepsilon_t^h
\end{aligned}$$

假设, $u_0 = 0$ 且满足可逆条件 $|\psi| < 1$, 上述随机波动率模型称之为 SV-MA

(1) 模型。对于观测方程中的 MA(1) 误差, 需要对主算法进行一些修改。首先, 我们对数据进行下列变换后, 可以借用之前方法, 对 h 进行抽样。定义:

$$H_\psi = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ \psi & 1 & 0 & 0 & \dots & 0 \\ 0 & \psi & 1 & 0 & \dots & 0 \\ 0 & 0 & \psi & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \psi & 1 \end{pmatrix}$$

我们可以将移动平均方程更改为:

$$\varepsilon_t^y = H_\psi u \quad u \sim N(0, \sum_y)$$

然后, 我们将数据 y 通过 $\tilde{y} = H_\psi^{-1}(y - \mu \mathbf{1})$ 形式进行变换, 其中 $\mathbf{1}$ 是 $T \times 1$ 的列向量, 所以 $(\tilde{y} \mid \psi, \mu, h) \sim N(0, \sum_y)$, 我们可以通过变化后的 \tilde{y} , 像之前一样对 h 进行抽样。

下一步, 对于 μ 的抽样, 我们改写观察方程, 可以的到 $y = \mu + H_\psi u$, 根据线性回归的特性, 我们有 $(y \mid \psi, \mu, h) \sim N(\mu, \Omega_y)$, 其中 $\Omega_y = H_\psi \sum_y H_\psi'$, 并且是一个带状矩阵。

同理 $(\mu \mid \psi, y, h) \sim N(\hat{\mu}, D_\mu)$ 是标准正态分布, 其中的参数

$$D_\mu^{-1} = V_\mu^{-1} + \mathbf{1}' \Omega_y^{-1} \mathbf{1} \text{ 和 } \hat{\mu} = (V_\mu^{-1} \mu_0 + \mathbf{1}' \Omega_y^{-1} y) \quad (\text{Chan (2013)})。$$

最后, 关于 ψ 的抽样, 由于 ψ 是对于给定 y, μ, h 是独立的, 在给定 $p(\psi)$ 下, 我们有对数似然函数:

$$\begin{aligned}
\log p(\psi \mid y, \mu, h) &\propto \log p(\psi) + \log p(y \mid \psi, \mu, h) \\
&\propto \log p(\psi) - \frac{1}{2} \tilde{y}' \Omega_y^{-1} \tilde{y}
\end{aligned}$$

当 ψ 的维度较高时，我们求得 $\hat{\psi}$ 和负海塞矩阵 K_{ψ}^{-1} 使似然函数 $\log p(\psi \mid y, \mu, h)$ 最大，然后使均值为 $\hat{\psi}$ ，方差为 K_{ψ}^{-1} 的高斯分布为建议分布，使用独立的Metropolis-Hastings的算法抽样。

3.5. SV-MG 模型的贝叶斯估计

在此之前的 FFBS 是基于近似正态的多步抽样方法，这种抽样方法中的分布是一种单一近似分布。Kim (1998) 采用了七组不同均值和方差的正态分布，通过赋予的不同的权重将正态分布与目标分布逼近即：

$$\log \chi_1^2 \sim \sum_{i=1}^7 \pi_i N(\mu_i, \omega_i^2)$$

根据 Kim (1998) 的研究结果，表 1 给出了各个正态分布的均值、方差和权重值。混合后的近似分布也显示在图中，用红色虚线代表其密度函数。容易看出，此时的密度函数与对数卡方分布已经相当接近，预期将提高算法的估计精度。
(Kim 1998)

表 1 拟合正态

i	π_i	μ_i	ω_i^2
1	0.0073	11.4004	5.79596
2	0.10556	5.24321	2.61369
3	0.00002	9.83726	5.1795
4	0.04395	1.50746	0.16735
5	0.34001	0.65098	0.64009
6	0.24566	0.52478	0.34023
7	0.2575	2.35859	1.26261

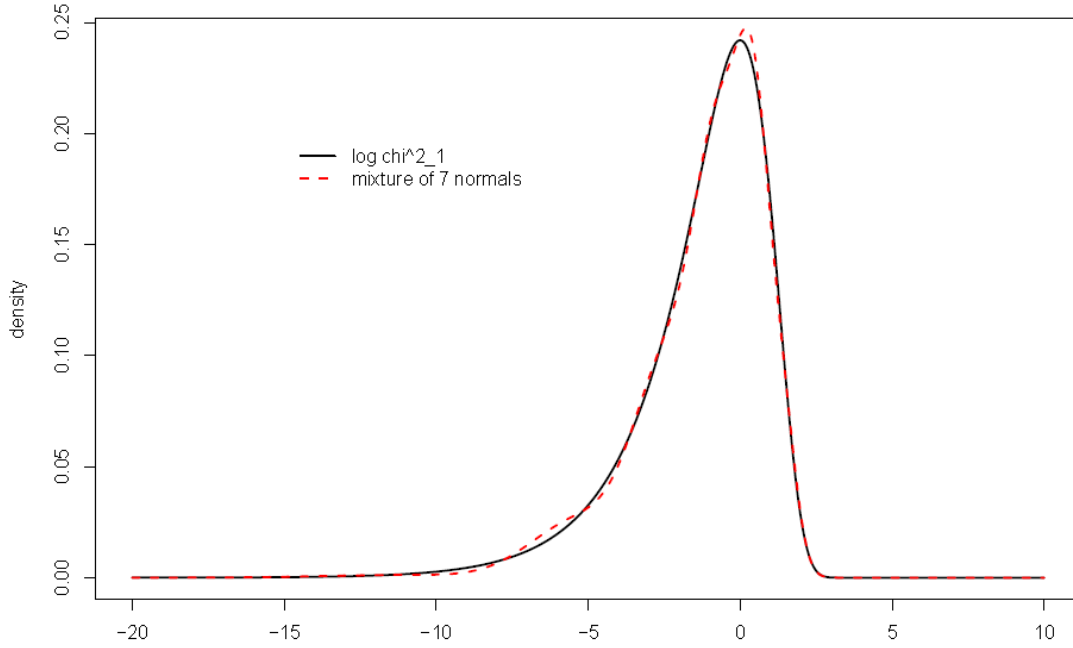


图 1 拟合分布图

为了完成基于混合正态近似的联合抽样方法，定义 Z_1, \dots, Z_n 为潜在指示变量， $Z_t \in \{1, \dots, 7\}$ ， $\Pr(Z_t = i) = \pi_i$ $i = 1, \dots, 7$

因此，在条件 Z 的情况下，将 y_t 变换成 $\log y_t^2$ ：

$$\begin{aligned}\log y_t^2 &= h_t + \log \varepsilon_t^2 \\ h_t &= \mu + \phi h_{t-1} + \tau_\eta \eta_t\end{aligned}$$

将上述的方程等式重写成标准的正态线性模型：

$$\begin{aligned}\log y_t^2 &= h_t + v_t & v_t &\sim N(\mu_{z_t}, \omega_{z_t}^2) \\ h_t &= \mu + \phi h_{t-1} + w_t & w_t &\sim N(0, \tau_\eta^2)\end{aligned}$$

此时可利用上小节所示，基于 FFBS 正态近似的联合抽样方法完成状态变量的抽样过程。

4. 中证 500 指数波动率的实证研究

4. 1. 样本及数据选取

我国的股票市场主体都是个人投资者，风险意识淡薄，承受压力有限，一遇

到某些突发消息，股市便大起大落，非理性行为比较普遍，市场波动特征被严重扭曲。中国股票在 2011 年反弹后，进入了震荡行情。特别是最近 2015 年来 A 股的最大特征就是震荡市，大切换。由于整体经济承压较大，市场风险偏好降低，震荡的区间非常狭窄，指数只有 10% 的波动空间。基于上述，将实证考察的时间范围定在 2016 年 1 月 2017 年 4 月份的 300 个交易日。大盘股市值较大，波动相对较小，我们选择中证 500 指数进行研究。

根据公式 $R_t = \ln(P_t / P_{t-1}) * 100$ 计算出日对数收益率，此处的 P_t 为股票指数第 t 个交易日的收盘价。本文全部数据均来源万德数据库。



图 2 对数波动率

图 1 描述了中证 500 指数 300 天来日收益率的时间序列图，从图中可看出，在 2016 年上半年股票市场的波动剧烈，下半年后趋于平缓。2016 年初，熔断新政实施，A 股开市两天，市场发生了四次熔断，其中两次触发 5% 熔断阈值，两次触发 7% 熔断阈值，呈现断崖式暴跌，导致两次交易提前休市，引起市场恐慌。在此之后，政府主力维稳，市场行情放缓。

4.2. 样本数据特征

关于数据的特征描述主要有三个方面：集中趋势、离散程度和分布形态。由于本文需要通过收益率的分布形态来进行先验假设，更加关注收益率的分布形态。对此本文采用偏度，峰度和标准差等相关指标对数据的特征进行描述。

(1) 标准差。收益率的离散程度主要通过标准差反映出来，同时标准差也是

衡量风险大小的主要指标，计算公式为：

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \mu)^2}$$

(2) 偏度。分布的对称性可以由偏度反映出来。若偏度等于 0 时，则分布是中心对称的。当偏度小于 0 时，分布是左偏。反之，偏度大于 0 时，数据的分布是右偏的。偏度一般通过数据的三阶矩计算，为：

$$S(x) = E[(X - \mu_x)^3 / \sigma^3]$$

本文研究的是对数收益率，在收益率分布的偏度大于 0 时，小于平均收益的天数偏多。反之，大于平均收益的天数偏多。

(3) 峰度。数据的分布形状或者特性一般由峰度体现，正态分布的峰度为 3，以此为标准，当峰度大于 3，表明该分布相对于正态分布具有尖峰厚尾的特性；反之，若峰度小于 3，具有低峰薄尾的特征；当峰度值极大时，可能存在许多偏离均值的异常值。其峰度是由数据四阶矩来计算：

$$k(x) = E[(X - \mu_x)^4 / \sigma^4]$$

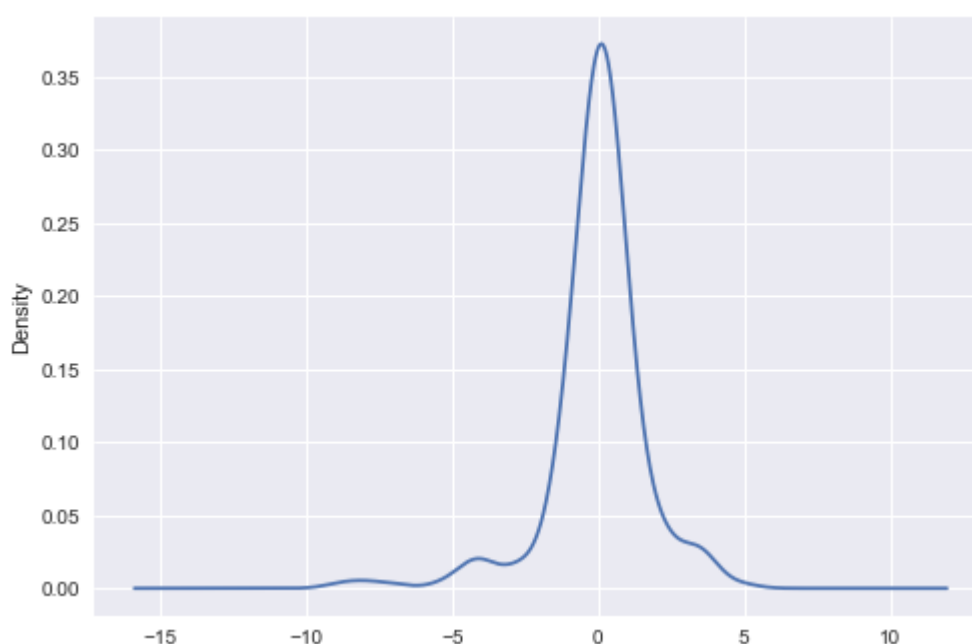


图 3：日对数收益率的分布图

通过 python 软件对中证 500 指数日对数收益率进行统计分析，均值为 -0.0288%，且标准差为 1.70122，呈现出中证 500 指数日对数收益率低于 0，但接近于 0，风险相对较大，与我国股票市场贴近，符合日对数收益率的分布图。

中证 500 指数日对数收益率的偏度为-1.5094，收益率分布的均值在左侧，分布左偏，是呈对称性的，具有偏斜特征。通过偏度分析可知，中证 500 指数日对数收益率小于平均收益的天数偏多。其峰度为 6.6692，比正态分布的峰值要大。这明显的表现出中证 500 指数日对数收益率的分布具有明显的尖峰厚尾的特征。由于中证 500 指数日对数收益率不服从正态分布，所以必须寻找合适的模型来描述厚尾现象。而 SV 模型可以有力的刻画这些特征。

4.3. 标准 SV 的参数估计及其比较

参数先验分布的设置是贝叶斯统计分析的前提条件。根据 hereafter 的观点，参数的先验分布分别为：

$$\mu \sim N(0, 100), \phi \sim N(0, 100), \tau^2 \sim IG(10 / 2, 0.28125/2), h_0 \sim N(0, 100)$$

首先在对模型的参数进行估计之前，现对每个参数进行 1000 此的迭代，以消除初始值的影响，并保证序列的收敛。图 1，2，3 分别是随机游走 M-H，独立 M-H，FFBS 三种估计的参数迭代路径图，从图中可以看出，1000 次后迭代后每个参数的迭代路径围着均值波动得频繁且幅度比较小，我们认为此时序列已经达到收敛，然后舍弃原来迭代，取后面 3000 次的参数均值为估计值。本文所有模型的实现均由 python 3.6 编程实现。

通过参数抽样的轨迹图 4，5，6，可以发现参数 μ 和 ϕ 马尔可夫链在后期处于平稳状态，收敛速度快。参数 τ^2 带跳，收敛速度缓慢。观察图 1，2，3 的三种方法下参数样本的自相关图，可以看到参数 μ 和 ϕ 的马尔科夫链基本处于平稳，而参数 τ^2 的收敛速度缓慢。观察右侧的参数样本相关图，可以看到参数 μ 的相关性较低，参数 ϕ 在随机游走 M-H 算法下，自相关系数很高，没有衰减的迹象。独立 M-H 和 FFBS 算法下，参数 ϕ 的样本自相关程度快速递减，FFBS 算法自相关函数的递减速度稍快。参数 τ^2 在三种不同方法估计下，自相关程度递减缓慢，相对于其它方法，FFBS 算法的递减速度最快。

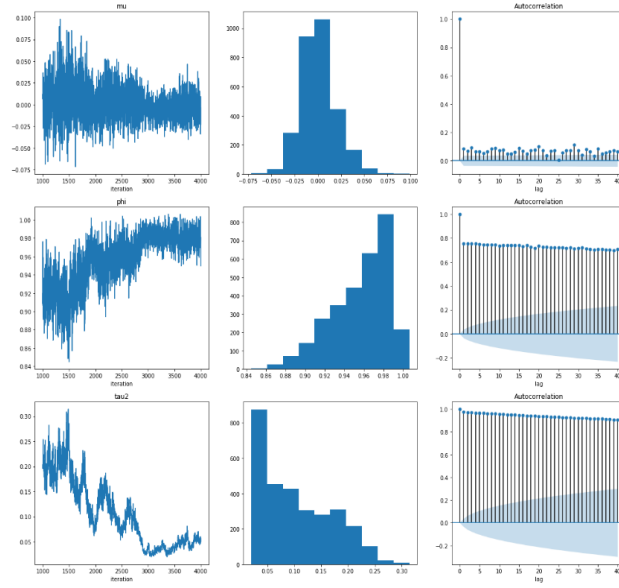


图 4 随机游走 M-H 参数结果

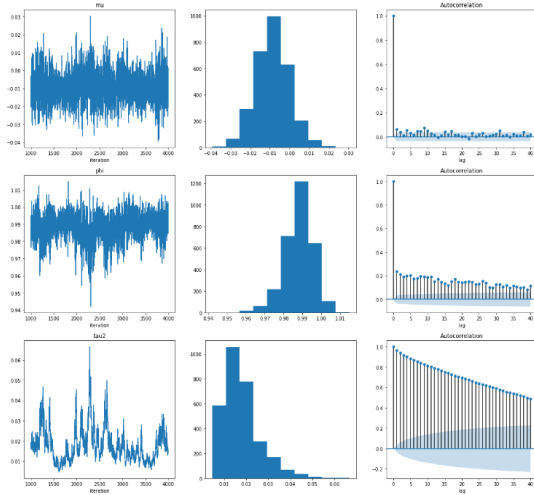


图 5 独立 M-H 参数结果

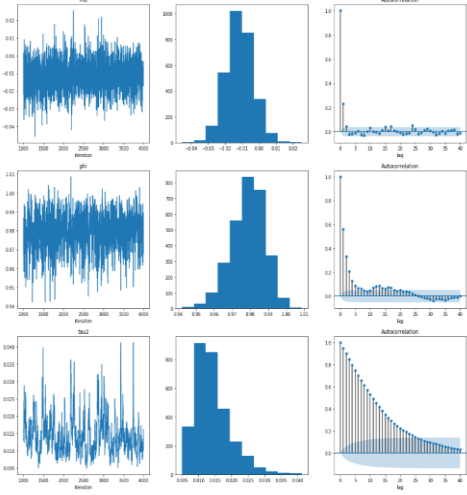


图 6 卡尔曼滤波参数结果

对于参数 μ ，卡尔曼滤波和独立 M-H 算法下的核密度估计较为接近，随机游走 M-H 算法下呈现厚尾特性。参数 ϕ 的核密度在随机游走 M-H 和独立 M-H 抽样下都是左偏，相对于随机游走的 M-H，独立 M-H 的核密度函数峰度更大。参数 τ^2 核密度在随机游走 M-H 抽样下右偏，峰度小，在独立 M-H 抽样和 FFBS 抽样下都是呈现尖峰特性，FFBS 抽样下的峰度更大。综上分析，说明随机游走 M-H 抽样下参数的马尔科夫链尚未收敛，用于中证 500 指数收益率波动估计会产生一定偏差。

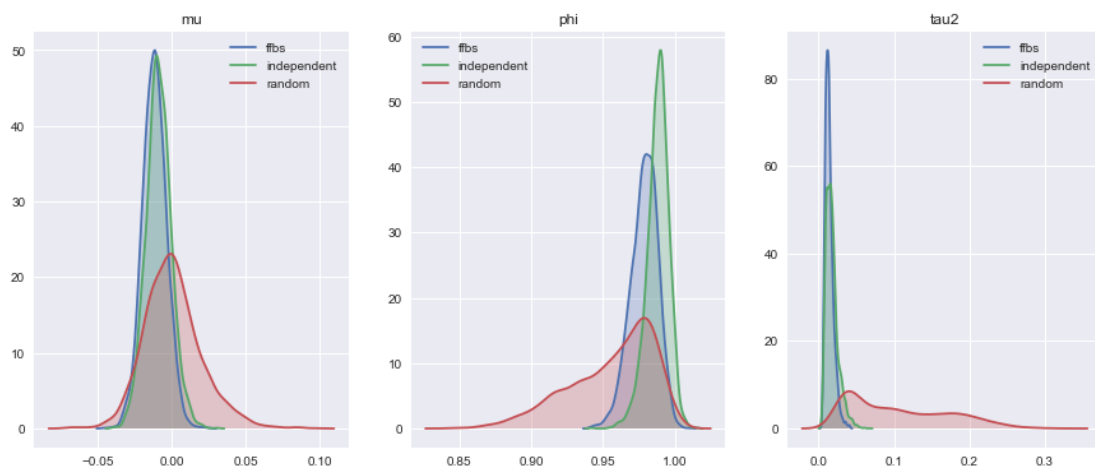


图 7 三种抽样方法下参数分布图

根据随机游走 M-H，独立 M-H，FFBS 三种算法，利用蒙特卡罗贝叶斯抽样模拟，得出参数估计表 2，给出了中证 500 指数在标准 SV 模型下参数的后验均值，标准差，5%和 95%分位数的贝叶斯估计值。

通过对表 2 的详细对比分析，我们可以总结出以下结论：

(1) 标准 SV 模型模拟的中证 500 指数的波动水平参数 μ 的贝叶斯估计值，在随机游走 M-H，独立 M-H，FFBS 三种算法下分别为 0.000844，-0.008695 和 -0.011478，90%的后验置信区间为 $[-0.028108, 0.036069]$ ， $[-0.022296, -0.022296]$ ， $[-0.024457, 0.001774]$ 。可以发现，波动水平参数 μ 的后验均值在三种算法下都接近 0，说明中证 500 指数对数收益率是围绕着 0 波动的。其中 FFBS 的置信区间最小，后验密度峰度大，精确度相对于其它两种算法更高。

(2) 中证 500 指数的波动持续性参数 ϕ 的贝叶斯估计值，分别为 0.956208，0.988173 和 0.978667，90%的后验置信区间为 $[0.900636, 0.992019]$ ， $[0.974495, 0.999687]$ ， $[0.962185, 0.992942]$ 。三种算法下中证 500 指数的波动持续性参数 ϕ 都在 0.95 以上，显示出中证 500 指数的波动具有极强的持续性，过去的波动很大程度上影响未来的走势，也验证了股票市场追涨杀跌的情绪。而独立 M-H 算法下，估计出的中证 500 指数波动持续性最强，且置信区间最小，相对精确度高。

(3) 股市波动的扰动水平通常采用标准差 τ^2 来度量，分别为 0.104580，0.017740 和 0.014228，90%的后验置信区间为 $[0.027360, 0.221595]$ ， $[0.007954, 0.017740]$ ， $[0.014228, 0.017740]$ 。三种算法下中证 500 指数的扰动水平参数 τ^2 都在 0.1 以下，显示出中证 500 指数的波动具有极强的持续性，过去的波动很大程度上影响未来的走势，也验证了股票市场追涨杀跌的情绪。而独立 M-H 算法下，估计出的中证 500 指数波动持续性最强，且置信区间最小，相对精确度高。

0.034174], [0.007424, 0.024809]。显然，随机游走 M-H 算法下，波动的扰动大，精度低，中证 500 指数波动噪音大于另两种算法。

从结果来看，独立 M-H 和 FFBS 算法下的参数比较接近，而随机游走 M-H 算法与其它两种算法差别较大。从标准差上分析，随机游走 M-H 算法的所有参数的标准差大于其它算法，进一步凸显了其算法的不稳定性，参数的马尔科夫链不收敛。

表 2 三种抽样方法下参数结果

不同抽样	参数	均值	标准差	5%分位数	95%分位数
随机游走 M-H	μ	0.000844	0.019525	-0.028108	0.036069
	ϕ	0.956208	0.029325	0.900636	0.992019
	τ^2	0.104580	0.065525	0.027360	0.221595
独立 M-H	μ	-0.008695	0.008501	-0.022296	0.005224
	ϕ	0.988173	0.007738	0.974495	0.999687
	τ^2	0.017740	0.008236	0.007954	0.034174
FFBS	μ	-0.011478	0.008113	-0.024457	0.001774
	ϕ	0.978667	0.009435	0.962185	0.992942
	τ^2	0.014228	0.005508	0.007424	0.024809

同样，我们也获得了模型潜在变量的后验估计结果，计算了每个时刻的后验均值，绘出三种算法的状态变量估计图，见下图。观测可知，2015 年 1 月到 2017 年四月，中证 500 指数收益率的波动率持续下降。FFBS 算法下波动率的估算结果是平滑减小，独立和随机游走算法下的波动率是波动式下降。



图 8 三种抽样方法下波动率时序图

4. 4. 扩展 SV 的参数估计

根据 hereafter 的观点，SV 扩展模型参数的先验分布假设如下：

SA-AR(2) 中参数 μ ， μ_h 和 ω_h^2 的先验分布见上文， ϕ_h 更改为 $\theta_h = (\phi_h, \rho_h)'$ ，并假设 $\theta_h \sim N(\theta_{h_0}, V_{\theta_h})1(\theta_h \in A)$ ，式中 $\theta_{h_0} = (0.97, 0)'$ ， $V_{\theta_h} = \text{diag}(0.1^2, 1)$ ， $A \in R^2$ 是 θ_h 特征多项式的根在单位圆之外的集合。SV-M 中参数 $\alpha \sim N(0, 100)$ ，SV-MA(1) 中参数 $\psi \sim N(0, 1)1(|\psi| < 1)$ 。

在模型抽样估计中，参数抽样 4000 次，为保证参数的收敛稳健性，取后 3000 次抽样值进行后验估计。参数估计结果见表 3；

表 3 扩展 SV 参数结果

	SV	SV_AR(2)	SV-M	SV-MA(1)	SV-MG
μ	0.05838	0.06985	0.10827	0.05873	0.04675
μ_h	0.68924	0.56038	0.66514	0.66404	0.53333
ϕ_h	0.98573	0.9918	0.98714	0.98382	0.99304
ω_h^2	0.05129	0.10225	0.05458	0.04034	0.12386
ρ_h	-	-0.06824	-	-	-

α	-	-	-0.04852	-	-
ψ	-	-	-	-0.0388	-

(1) 对于中证 500 指数收益率 μ 在, SV-M 模型下的模拟后验均值为 0.10827%, 要比其他的 4 个模型大上很多。其他模型的收益率参数 μ 的后验均值在 0.05%左右, 体现出中证 500 指数日收益率接近于 0。股票市场是随机游走的, 因此日收益率的均值理论上接近于 0, SV-M 模型的估计存在偏差。

(2) 波动水平 μ_h 在五个模型中都在 0.6 左右, 其中 SV-AR(2)和 SV-MG 模型在 0.5 左右。这说明在 SV, SV-M, SV-MA (1) 模型下, 中证 500 指数体现出更加强烈的波动性。而 SV-AR(2)和 SV-MG 模型的参数 μ_h 的后验均值相对小, 中证 500 指数波动性弱。

(3) 对于股市波动的持续性参数 ϕ_h 的后验均值均在 0.98 左右, 最大的是 SV-MG 模型, 参数估计值为 0.99304, 最小的是 SV-MA (1) 模型, 参数估计为 0.98382。体现出中证 500 指数波动的持续性强, 其中 SV-MG 模型能更好的刻画中证 500 指数对数收益率序列高的波动持续性, 同时也反映出我国股市波动集聚性的特征。

(4) 中证 500 指数波动的扰动, 可以通过方差表现出来。这里我们使用模型的精度参数 ω_h^2 来衡量。通常情况下, 波动持续性参数 ϕ_h 值越小, 波动的方差就大, 即 ω_h^2 越小, 从侧面说明了波动过程是难以预测的。在上述模型参数 ω_h^2 的贝叶斯估计中, SV-MA(1)模型的精度参数 ω_h^2 最小, 同时参数 ϕ_h 的估计值也最小。因此对中证 500 指数收益率的随机模拟效果并不是很好。然而, SV-MG 模型的精度参数 ω_h^2 最大, 参数 ϕ_h 的估计值也最最大。相对于其他模型, SV-MG 模型能更好的模拟中证 500 指数收益率的波动。

(5) 模型 SV-AR(2) 中, 二阶自相关强度的参数 ρ_h 为-0.06824, 表明对数收益波动率与前期有很强的相关性, 但与两期之前的相关性较弱, 即未来的收益受到当前的影响, 但与过去无关。进一步证明了, 我国股市的波动是随机游走过程。SV-M 模型中的反馈参数 α 后验均值为-0.04852, 表明了中证 500 指数的收益会受到当期的波动影响, 且收益与波动呈现负相关, 当波动越大, 收益越小。SV-MA (1) 模型中的滞后扰动参数 ψ 为-0.0388, 说明波动受前期之后的扰动小, 股市的波动大部分来自于当前的冲击。

对于模型的比较, 本文参考 Kass and Raftery (1995) 的贝叶斯因子方法。对于模型 M_k , 计算边缘似然密度:

$$p(y | M_k) = \int p(y | \theta_k, M_k) p(\theta_k | M_k) d\theta_k$$

θ_k 为模型的具体参数, 边缘似然密度可以被解释为在实际观测数据 y 中对 M_k 模型下数据的密度预测。因此, 如果观察到的数据在模型之中, 那么相关的边缘似然密度将越大 (Joshua C.C. Chan, 2015)。

表 4 扩展 SV 参数结果

ML	-509.2	-509.7	-513.9	-551.3	-508.2
	(0.02)	(0.08)	(0.04)	(0.01)	(0.01)

结合模型的复杂程度和拟合程度, 计算最大似然性。SV, SV-AR (2), SV-M, SV-MA (1), SV-MG 的边缘似然密度分别为-509.2, -509.7, -513.9, -551.3, -508.2。其中相比较而言, SV-MG 模型的 ML 值最大, SV-MG 更能解释我国的股票市场波动性, 与前文结论相呼应。同时, 通过图 7 扩展 SV 参数分布图, 可以看出, SV-MG 模型的参数后验密度更为集中, 估计的精度更高。

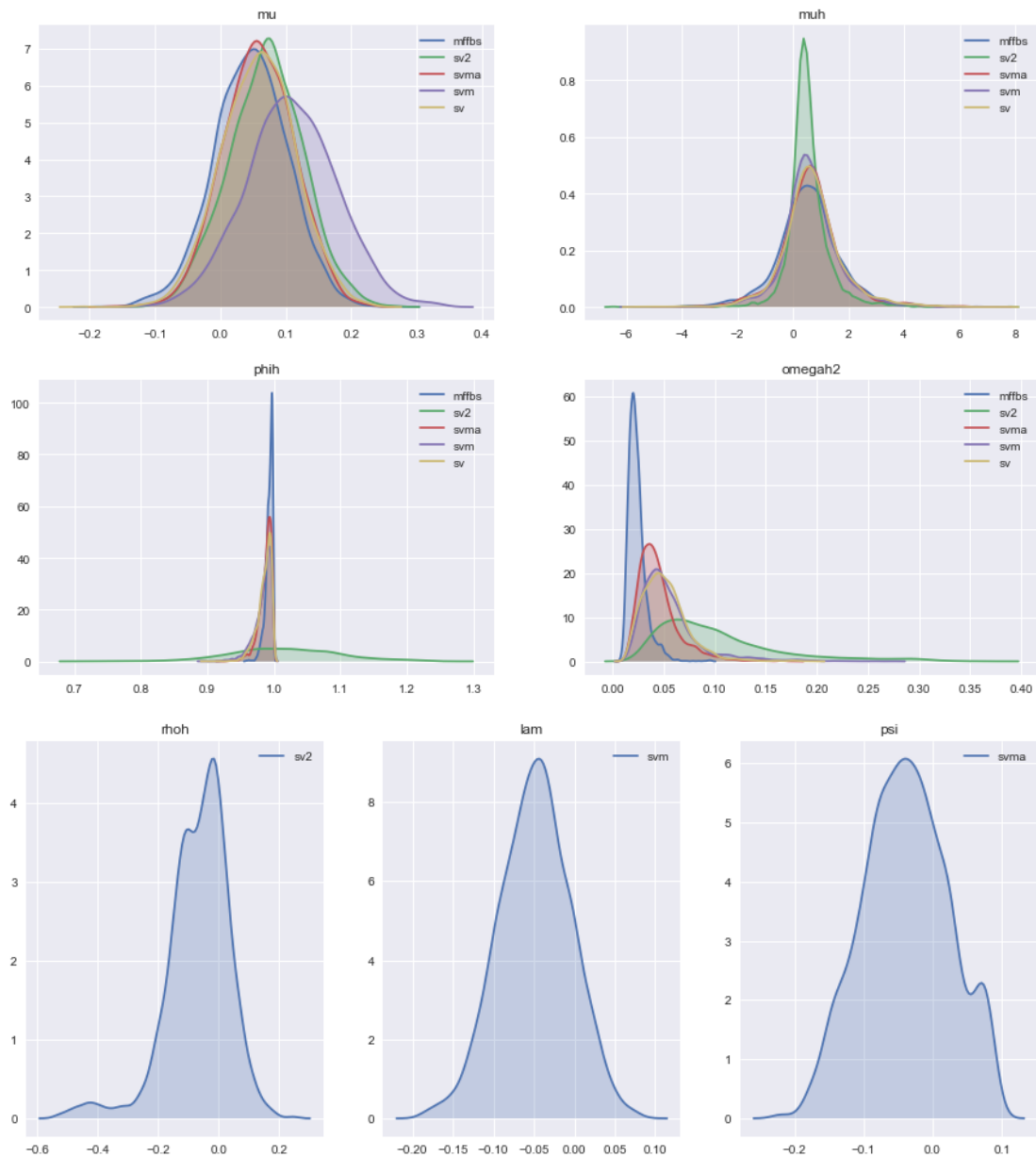


图 7 扩展 SV 参数分布图

根据估计结果，绘出不同模型下的状态变量估计图，见图 8。观测可知，2015 年 1 月到 2017 年四月，中证 500 指数收益率的波动率持续下降。SV-AR (2)模型中，由于需要两个前定数据，第一个起点值的无法估计，下图第一个起点是随机生成的结果。通过下图，发现不同扩展模型下的整体趋势基本相似，在每个波动时点的波动大小不同，SV-MG 模型的平滑减小。

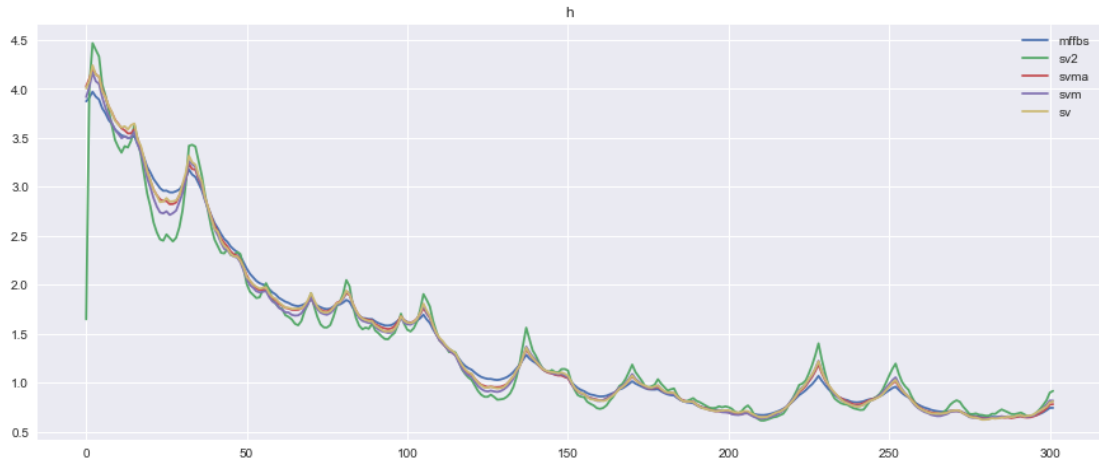


图 8 扩展 SV 波动率时序图

5. 主要结论与展望

金融经济中的时间序列呈现出高峰厚尾，波动的聚集性，非线性结构，杠杆效应，波动的长记忆性等特性，本文主要使用 SV 模型对这些特性进行捕捉。介绍了关于随机波动率的三种传统 MCMC 方法，随机游走的 M-H，独立马尔科夫链 M-H，线性滤波器 FFBS。同时引进了一种高效的 MCMC 方法，基于此，对 SV 模型，SV-M 模型进行了贝叶斯分析。在标准 SV 模型，SV-M 模型上，引进 SV-AR (2) 模型，SV-MA(1) 模型和 SV-MG 模型对我国股市中证 500 指数对数收益率进行分析。通过新型的评价准则 marginal likelihoods，对 SV 族模型的对比研究。最后，本文的主要结论如下：

(1) 通过对中证 500 指数的日对数收益率的描述统计量可知，其偏度为 -1.5094，峰度为 6.692，日对数收益率不服从正态分布，而是左偏，出现尖峰厚尾的特性。

(2) 独立 M-H 和 FFBS 算法下的 SV 模型的参数比较接近，而随机游走 M-H 算法与其它两种算法差别较大，呈现出中证 500 指数的日对数收益率波动持续性相对较弱，精度低，置信区间大。

(3) 在 SV 扩展模型分析中，表明中证 500 指数的对数收益波动率与前一期有很强的相关性，但与两期之前的相关性较弱，未来的收益受到当前的影响，但

与过去无关。进一步证明了，我国股市的波动是随机游走过程。其次，中证 500 指数的收益会受到当期的波动影响，且收益与波动呈现负相关，当波动越大，收益越小。最后，中证 500 指数对数收益率的波动受到前期的扰动小，股市的波动大部分来自于当前的冲击。

（4）SV 模型簇的对比研究发现，在刻画动荡的市场环境时，通过赋予各个正态分布相应的权重从而组合成与目标分布逼近的近似形式的联合抽样要好于单个假设分布抽样。SV-MG 模型的 ML 值最大，更能解释我国的股票市场波动性。并且，模型参数的后验密度更为集中，估计的精度更高。

同时本文也存在诸多方面的问题：

（1）参数先验分布的设置是贝叶斯统计分析基础，在模型估计时，使用的先验分布是参考之前研究者的研究成果，而这些参数的先验假设并不一定适用于这些模型。

（2）在模型的参数结果中，我们以后验均值作为模型参数的估计量。通过参数的核密度函数分布图，初步判断存在左偏或者右偏的现象。以后验均值作为模型参数的估计量可能会使估计结果偏大或者偏小。

（3）在使用 SV 模型扩展簇的时候，并没有考虑到一些特殊分布，也没考虑到是否存在跳跃点，或者杠杆效应等。本文进一步将会考虑到随机波动率的结构变化，长记忆性等。

参考文献

- [1]. Taylor S J. Modeling financial time series [M]. John Wiley, Chichester, 1986
- [2]. Hull, J., A. White. The Pricing of Options on Assets with Stochastic Volatilities. The Journal of Finance[J], 1987, 42: 281-300
- [3]. Hansen, Lars Peter. Large sample properties of generalized method of moments estimators. Econometrical[J], 1982, 50: 1029-1054
- [4]. Koopman, S.J. Shephard and J. N. Shephard and J. Doornik. Review of Pack 2.2: statistical algorithms for models in state space. Economics Journal[J], 1999, 2: 113-116
- [5]. Eric Jacquier, E. N.G. Polson, and P. E. Rossi. Bayesian analysis of stochastic volatility models. Journal of Economic Statistics[J], 1994, 12: 371-389
- [6]. Sangjoon Kim, Neil Shephard, Siddhartha Chib. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. Review of Economic Studies[J], 1998, 65: 361-393
- [7]. Melino, A., S.M. Turnbull. Pricing foreign currency option with stochastic volatility. Journal of Econometrics[J], 1990, 45: 239-265
- [8]. Andersen, T.H. Chung, B. Sorensen. Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. Journal of Econometrics[J], 1991, 91: 61-87
- [9]. Berger, J.O., Pericchi, L.R. The intrinsic Bayes factor for model selection and prediction. The Journal of the American Statistical Association[J], 1996, 91: 109-122
- [10]. Carter C.K., R. Kohn. On Gibbs sampling for state space models. Biometrika, 1994, 81, 541-553
- [11]. Ghysels, E., A. C. Harvey, and E. Renault, 'Stochastic Volatility', in G. Maddalena C. R. Rao (eds), Handbook of Statistics, 1996
- [12]. Ozturk, S. S., and J.-F. Richard, 'Stochastic Volatility and Leverage: Application to a Panel of S&P500 Stocks', Finance Research Letters[J], 2015, 12: 67-76.
- [13]. Harvey, A., E. Ruiz, and N. Shephard, 'Multivariate Stochastic Variance Models', The Review of Economic Studies[J], 1994, 61: 247-64.
- [14]. Jacquier, E., Polson, N. G., and P. E. Rossi, 'Stochastic Volatility: Univariate and Multivariate Extensions', 1999, Discussion Paper No. 99s-26, CIRANO
- [15]. Yu, J., and R. Meyer, 'Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison', Econometric Reviews[J], 2006, 25: 361-84.
- [16]. Philipov, A., and M. E. Glickman, 'Multivariate Stochastic Volatility via Wishart Processes', Journal of Business & Economic Statistics[J], 2006, 24: 313-28.
- [17]. Nakajima, J., 'Bayesian Analysis of Multivariate Stochastic Volatility with Skew Return Distribution', Econometric Reviews[J], 2016, 1-23.
- [18]. Richard, J.-F., and W. Zhang, 'Efficient High Dimensional Monte Carlo Importance Sampling', Journal of Econometrics[J], 2007, 141: 1385-411.
- [19]. P. Agnolucci. Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. Energy Economics[J], 2009, 31(2): 316-321.
- [20]. O. J. Blanchard and M. Riggi. Why are the 2000s so different from the 1970s? A structural interpretation of changes in the macroeconomic effects of oil prices. Journal of the European Economic Association[J], 2013, 11(5): 1032-1052.
- [21]. T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. Journal of

Econometrics[J], 1986, 31(3):307–327.

[22]. C. Brooks and M. Prokopczuk. The dynamics of commodity prices. Quantitative Finance[J], 2013,13(4):527–542.

[23]. J. C. C. Chan. Moving average stochastic volatility models with application to inflation forecast. Journal of Econometrics[J], 2013, 176(2):162–172.

[24]. J. C. C. Chan. The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. Journal of Business and Economic Statistics[J], 2015. Forthcoming.

[26]. J. C. C. Chan and E. Eisenstadt. Marginal likelihood estimation with the Cross-Entropy method. Econometric Reviews[J], 2015 ,34(3):256–285.

[27]. J. C. C. Chan and A. L. Grant. Issues in comparing stochastic volatility models using the deviance information criterion. CAMA Working Paper, 2014.

[28]. J. C. C. Chan and I. Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. International Journal of Mathematical Modelling and Numerical Optimization[J], 2009, 1(1):101–120.

[29]. R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica[J], 1982, 50(4):987–1007.

[30]. W. M. Fong and K. H. See. Modelling the conditional volatility of commodity index futures as a regime switching process. Journal of Applied Econometrics[J], 2001, 16(2):133–163.

[31]. 余素红, 张世英. SV 与 GARCH 模型对金融时间序列刻画能力的比较研究[J]. 系统工程, 2002(05):28–33.

[31]. 王春峰, 蒋祥林, 李刚. 基于随机波动性模型的中国股市波动性估计[J]. 管理科学学报, 2003(04):63–72.

[32]. 刘凤芹, 吴喜之. 基于 SV 模型的深圳股市波动的预测[J]. 山西财经大学学报, 2004(04):96–99.

[33]. 田秋荣, 栾长福. 基于 SV 模型的沪深股市风险分析[J]. 科学技术与工程, 2009, 9(10):2834–2837.

[34]. 吴萌, 徐全智. SV 族模型对沪市综合指数的实证分析[J]. 成都大学学报(自然科学版), 2005(01):47–49+54.

[35]. 孙有发, 邓飞其. 带跳及反馈的随机波动证券定价模型[J]. 华南理工大学学报(自然科学版), 2006(06):59–63.

[36]. 周彦, 张世英. 基于 MCMC 方法的连续时间 SV 模型建模研究[J]. 工业工程, 2007(01):83–86.

[37]. 刘鑫, 孟昭为. 带有交叉杠杆的多元随机波动率模型[J]. 重庆理工大学学报(自然科学), 2011, 25(12):111–115.

[38]. 吴鑫育, 马超群, 汪寿阳. 具有有偏厚尾的非对称 SV 模型及其实证研究[J]. 系统工程, 2012, 30(01):61–66.

[39]. 朱喜安. 基于带解释变量杠杆 SV 模型上证指数收益率周内效应及特征分析[A]. 中国数量经济学会. 21 世纪数量经济学 (第 13 卷) [C]. 中国数量经济学会:, 2012:21.

[40]. 苏理云, 彭相武, 王杰, 邱冬阳. 基于状态空间 SV-T-MN 模型的股指波动率预测[J]. 数理统计与管理, 2016, 35(05):929–942.

[41]. 吴鑫育, 李心丹, 马超群. 双因子非对称已实现 SV 模型及其实证研究[J]. 中国管理科学, 2018(02):1–13.

致 谢

附录

标准随机波动率模型高效抽样的 PYTHON 程序:

```
import numpy as np
import pandas as pd
from scipy import stats
from scipy.stats import norm
import matplotlib.pyplot as plt
from scipy import sparse
import statsmodels.tsa.stattools as sts
import statsmodels.api as sm
nloop=4000
burnin=1000
M=1000
data5=pd.read_csv('C:\\Users\\wuboyuan\\Desktop\\data5.csv', header=None)
data5=100*data5.apply(np.log).diff(1)[1:]
y=data5.T.values[0]
T=len(y)
# 先验分布参数
phih0=0.97
vphih=0.1**2
mu0=0
vmu=10
muh0=1
vmuh=10
nuh=5
sh=0.2**2*(nuh-1)
# 初始化马尔可夫链
muh=np.log(np.var(y))
phih=0.98
omegah2=0.2**2
h=muh+np.sqrt(omegah2)*np.random.randn(T,1)
exph=np.exp(h)
# 存储空间
store_theta=np.zeros([nloop-burnin,4])
store_h=np.zeros([nloop - burnin,T])
store_q=np.zeros([nloop - burnin,2])
# 循环外的计算
hphi=np.eye(T)-sparse.coo_matrix((phih*np.ones(T-1), (np.arange(1, T), np.arange(T-1))), shape=(T, T))
newnuh=T/2+nuh
counth=0
countphi=0
```

```

# MCMC
for loop in range(nloop):
    invdmu=1/vmu+sum(1./exph)
    muhat=(invdmu**(-1))*(1/vmu*mu0+sum((np.matrix(y).T)/exph))
    mu=muhat+((np.linalg.cholesky(np.matrix(invdmu)).T)**(-1))*np.random.randn(
    1)
    hish=hphi.T*sparse.coo_matrix((np.hstack((float((1-phih**2)/omegah2)],1/float(omegah2)*np.ones(T-1))), (np.arange(T), np.arange(T))), shape=(T, T))*hphi
    deltah=(hphi**(-1))*np.vstack((muh, np.ones([T-1, 1])*(1-phih)*muh))
    hishdeltah=hish*np.matrix(deltah)
    s2=np.power((np.matrix(y).T-mu), 2)
    errh=1
    ht=h
    while errh>10**(-3):
        expht=np.exp(ht)
        sinvexpht=s2/expht
        fh=-0.5+0.5*sinvexpht
        gh=0.5*sinvexpht
        kh=hish+sparse.coo_matrix((gh.T.tolist()[0], (np.arange(T), np.arange(T))), shape=(T, T))
        newht=kh**(-1)*(fh+np.multiply(gh, ht)+hishdeltah)
        errh=max(abs(newht-ht)).tolist()[0][0]
        ht=newht

    cholhh=np.linalg.cholesky(kh)
    lph=lambda
x:-0.5*(x-deltah).T*hish*(x-deltah)-0.5*sum(x)-0.5*np.exp(-x).T*s2
    logc=lph(ht).tolist()[0][0]+np.log(3)
    flag=0
    while flag==0:
        hc=ht+(cholhh.T)**(-1)*np.random.randn(T, 1)
        alparc=(lph(hc)+0.5*(hc-ht).T*kh*(hc-ht)-logc).tolist()[0][0]
        if alparc>np.log(np.random.uniform(0, 1)):
            flag=1
        alpar=(lph(h)+0.5*(h-ht).T*kh*(h-ht)-logc).tolist()[0][0]
        if alpar<0:
            alpmh=1
        elif alparc<0:
            alpmh=-alpar
        else:
            alpmh=alparc-alpar
    if alpmh>np.log(np.random.uniform(0, 1)) or loop==0:
        h=hc
        exph=np.exp(h)

```

```

        counth=counth+1

errh=np.vstack(((h[0]-muh)*np.sqrt(1-phih**2),h[1:]-float(phih)*h[:-1]-muh*(1-phih)))

newsh=(sh+sum(np.power(errh,2))/2)[0]
omegah2=1/(np.random.gamma(newnuh,1./newsh))

xphi=np.matrix(h[:-1]-muh)
yphi=np.matrix(h[1:]-muh)
dphi=1/(1/vphih+xphi.T*xphi/omegah2)
phihat=dphi*(phih0/vphih+xphi.T*yphi/omegah2)
phic=phihat+np.sqrt(dphi)*np.random.randn(1)
g=lambda
x:-0.5*np.log(omegah2/(1-np.power(x,2)))-0.5*(1-np.power(x,2))/omegah2*(h[0]-muh)**2

if abs(phic).tolist()[0][0]<0.999:
    alpmh=np.exp(g(phic)-g(phih))
    if alpmh.tolist()[0][0]>np.random.uniform(0,1):
        phih=phic
        countphi = countphi+1

hphi=np.eye(T)-sparse.coo_matrix((((phih*np.ones(T-1)).tolist()[0])*np.ones(T-1),
(np.arange(1,T),np.arange(T-1))),shape=(T,T))

dmuh=1/(1/vmuh+((T-1)*(1-phih)**2+(1-phih**2))/omegah2)

muhhat=dmuh*(muh0/vmuh+(1-phih**2)/omegah2*h[0]+(1-phih)/omegah2*sum(h[1:]-float(phih)*h[:-1]))
muh=muhhat+np.sqrt(dmuh)*np.random.randn(1)

if loop>burnin-1:
    i=loop-burnin-1
    store_h[i,:]=np.array(h.T)[0]
    store_theta[i,:]=[float(mu),float(muh),float(phih),float(omegah2)]
    u=(np.matrix(y).T-mu)/np.exp(h/2)
    rtmp=sts.acf(u,nlags=20)
    Q=T*(T+2)/(T-np.arange(21)[1:])*np.matrix(rtmp[1:])**2.T
    rtmp2=sts.acf(np.power(rtmp,2),nlags=20)
    Q2=T*(T+2)/(T-np.arange(21)[1:])*np.matrix(rtmp2[1:])**2.T
    store_q[i,:]=[float(Q),float(Q2)]

#参数模拟图:
plt.figure(figsize=(20,22))
index=np.arange(0,12,3)
title=['mu','muh','phi','tau2']

```

```

for idx in range(len(index)):
    plt.subplot(4, 3, 1+index[idx])
    plt.plot(pd.DataFrame(store_theta)[idx])
    plt.title(title[idx])
    plt.xlabel('iteration')
    plt.subplot(4, 3, 1+index[idx]+1)
    plt.hist(pd.DataFrame(store_theta)[idx])
    ax1=plt.subplot(4, 3, 1+index[idx]+2)
    sm.graphics.tsa.plot_acf(pd.DataFrame(store_theta)[idx], lags=40, ax=ax1)
    plt.xlabel('lag')

```