

中文摘要

连续时间模型在金融领域中具有广泛的应用。随着国际国内金融市场的迅速发展,金融市场的波动也日益加剧,风险不断增大。对于金融市场上波动的特性及其内在机制和经济含义的深入分析,已经成为经济学研究中的一个重要方向。基于此,本文以贝叶斯原理为工具分析了资产收益的连续时间模型。本文的主要工作和创新点如下:

1) 运用“马尔可夫链蒙特卡罗”模拟(MCMC)方法进行参数估计,这种方法可以有效的处理高维参数及高维隐含变量的估计问题。使用C++语言开发了包含隐含变量的连续时间模型估计的基于MH算法的MCMC方法,并用该方法估计了双指数跳跃扩散模型。

2) 研究了连续时间资产收益变结构模型。给出了连续时间BS变结构模型和连续时间随机波动变结构模型,提出了应用MCMC方法的连续时间变结构模型的单一变结构点的定位方法,并提出了连续时间多变结构点模型的变结构点定位方法;该方法在确定变结构点位置的同时,又能估计相应的模型参数。用该方法对上海股市综合指数的收益序列进行了变结构分析,理论与实证结果表明该方法是有效且可行的。

3) 研究了抛物线跳跃扩散模型。首先将跳跃因子引入到抛物线扩散模型中;接着,用基于Milstein的方法获得参数后验分布的离散密度,使用MCMC方法来估计抛物线跳跃扩散模型,最后用所得的估计值模拟了资产收益序列。通过上面的结果,不但发现MCMC方法较其他方法,如ML方法,更适合含有隐含变量的模型估计;还说明了抛物线跳跃扩散模型(HJD)能够很好的拟合资产收益的经验特征,如有偏、尖峰厚尾以及泰勒效应,即HJD模型更能反映资产收益的序列特征。

4) 应用似然比检验和贝叶斯因子分别比较了连续时间常规模型与变结构模型、抛物线扩散模型与抛物线跳跃扩散模型的优劣,并从实证的角度利用两个模型预测的准确性来进行模型比较。

本论文是国家自然科学基金资助项目《基于连续时间模型的贝叶斯分析》(No:70301006)和《多变量矩序列长期均衡关系及动态金融风险规避策略研究》(No:70471050)的组成部分。

关键词: 连续时间模型 贝叶斯分析 马尔可夫链蒙特卡罗 变结构 抛物线跳跃扩散模型 随机波动 MH算法

ABSTRACT

The continuous-time models can be applied a lot when studying finance. In the financial market, the volatility and the risk is becoming larger and larger along with the development of the financial market at home and abroad. Now studying the economic meaning, inherent mechanism and characteristic of the financial volatility are important. Bayes theory was used to analyze Continuous-time models. The key points and main achievements of this work are listed as follows:

1) The parameters of Continuous-time models were estimated using Markov chain Monte Carlo method (MCMC). The method is based on Markov-chain Monte Carlo methodology and applies to a wide class of models including systems with unobservable state variables and nonlinearities. In the paper, The method is applied to the estimation of parameters in Double Exponential Jump Diffusion Model (DEJD).

2) The Continuous-time structural change models were studied. The method for detecting and locating single structural change point using MCMC method is proposed. The method for detecting and locating multiple structural change points of continuous time models is also put forward. The method suggested is proved to be effective and feasible by theoretical analysis as well as the structural change analysis the distribution of the return series of composite index of Shanghai stock markets.

3) The Hyperbolic Jump-Diffusion model was put forward. Based on Bibby and Sørensen's Hyperbolic Diffusion model (HYD), we model the log-price as a deterministic linear trend and jumps plus a diffusion process with drift zero and with a diffusion coefficient (volatility) which depends in a particular way on the instantaneous asset price. We name the model Hyperbolic Jump-Diffusion model (HJD). It is shown that the model possesses a number of properties encountered in empirical studies of asset prices. The model is rather successfully fitted to different price index data sets. We proposed a MCMC method to estimation parameters and implied variables, the MCMC method based on the Milstein scheme.

4) The likelihood ratio test based on stochastic simulation was used to compare standard BS model with structural change BS model. With this method we compare the goodness of fit between HYD and HJD models.

The research is sponsored by National Natural Science Foundation of China: Bayes Analysis of Continuous-time Models (No. 70301006) and Research on Long Run Equilibrium in Multivariate Moments Series and Avoiding Tactics of Dynamic Financial Risk (No. 70471050).

KEY WORDS: Continuous-time Models, Bayes Analysis, Markov Chain Monte Carlo Method (MCMC), Structural Change, Hyperbolic Jump Diffusion Model, Stochastic Volatility, Metropolis-Hasting Algorithm

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的设计工作和取得的设计结果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写的研究成果，也不包含为获得天津大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本设计所做得任何贡献已在论文中做了明确说明并表示谢意。

学位论文作者签名：胡静华

签字日期：2006 年 8 月 28

日

学位论文版权使用授权书

本学位论文作者完全了解天津大学有关保留、使用学位论文的规定。特授权天津大学可以将学位论文的全部或部分内 容编入有关数据库进行检索，并采用并影印、缩印或扫描复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名：胡静华

导师签名：黄

签字日期：2006 年 8 月 28 日

签字日期：2006 年 8

月 28 日

第一章 绪 论

本章主要介绍论文选题的经济及金融背景与方法论背景，阐述了相关理论与建模方法的国内外研究现状，指出了存在的问题，给出本文选题的理论意义与实际意义。最后，介绍了本文研究结构安排与主要创新工作。

1.1 研究背景

1.1.1 经济及金融系统的不稳定性与脆弱性

经济及金融系统是一个复杂系统，综观其发展历程，呈现出较强的不稳定性。无论是现代资本主义市场经济、社会主义计划经济还是社会主义市场经济，经济及金融的波动性就从来没有停止过。

自20世纪70年代以来，由于布雷顿森林体系的崩溃导致国际货币体系的瓦解，以及20世纪70年代末美联储利率体制的调整，即以货币总量管理代替利率管理的目标，造成了世界经济环境的剧烈动荡。个人、企业以及金融机构投资的风险也空前加大。此后，全球范围内一些大的金融波动就层出不穷：1982年爆发了拉美国家债务危机、1994年底发生了墨西哥金融危机、1997年7月发生了东南亚金融危机、1999年1月和2002年7月分别由巴西和乌拉圭金融动荡引起的拉美金融危机等，这些金融波动无不伴随着汇率动荡、货币贬值、股市爆跌、公司破产、银行倒闭等现象。

在这样的背景下，一方面各种规避风险的措施与工具（如金融衍生产品）应运而生，这促进了新兴的经济与金融理论的诞生与发展；另一方面，人们迫切需要了解经济及金融波动的原因及其规律性。

多年来，为揭示经济及金融波动的本质，国际学术界对经济及金融系统的运行规律进行了不懈地探索。然而，传统的经济计量学由于其本身的缺陷，不可能

为这一问题提供有力的分析工具。正是在这一深刻的社会经济背景下，现代经济计量学应运而生。现代经济计量学方法论的发展，为波动性的动态建模分析提供了坚实的方法论基础。

1.1.2 经济及金融风险的时变性与传染性

在对大量的经济、金融时间序列数据的分析中，人们发现经济变量的波动性（或不确定性）并非固定不变，而是随时间变化的，即具有时变性。在对波动的时变性进一步研究中，人们发现波动的时变性又表现出了明显的持续性，即当前波动会持续地作用于未来波动的变化过程。然而，尽管人们已经认识到大量的关于资本收益的时间序列表现出了明显的波动聚集特征，但只是到了20世纪的80年代，人们才开始真正研究基于资本收益的二阶矩和高阶矩的动态建模问题。

另外，在经济全球化、一体化的浪潮中，由于国际间信息流、技术流、资金流等的流动性，世界各国经济、金融系统从最初孤立分散系统整合为在子系统间存在较强耦合作用的世界经济大系统。这既增加各国经济之间的联系、促进经济发展，也为风险在世界范围内的传播创造了机会，加大了全球金融市场之间的相互影响，导致了各个市场之间波动的互动效应，金融风险在不同市场之间传导、放大，使得全球金融市场的波动性和风险不断加大。

1.1.3 经济及金融风险的规避与金融资产定价

在经济及金融风险的规避方面，已经有相应的预警机制、景气分析等手段。现在重点介绍在金融领域风险的规避与定价。

早在1952年，Markowitz^[1]建立了资产组合的风险模型，第一次把数理工具引入金融研究，从而能够进行定量的检验和预测。在Markowitz工作的基础上，Sharpe (1964)^[2]、Litner (1965)^[3]、Mossin (1966)^[4]各自独立地研究了任一证券组合收益率与某个共同因子的关系，从而导出资本资产定价模型(Capital Asset Pricing Model, CAPM)。

20世纪70年代，随着金融创新的不断进行，金融衍生产品的定价成为理论研究的重点。Black和Scholes (1973)^[5]建立了期权定价模型。期权定价理论是金融理论研究的又一大突破，并迅速被运用于金融实践。此后，Ross (1976)^[6]

又建立了套利定价理论(Arbitrage Pricing Theory, APT), 研究了多时期证券市场的均衡定价, 非对称信息下的金融市场等问题。

随着金融风险规避策略和金融资产定价研究的深入, 新的金融衍生品(期货、期权、互换等)也在不断被创造出来, 与此同时“金融工程”诞生并迅速发展。这些理论与工具满足了不同投资者的收益-风险偏好, 为投资者进行组合投资、防范金融风险提供了可能性。

1.1.4 方法论背景

连续时间模型在金融领域的应用最早可以追溯到上个世纪60年代末70年代初, 由Merton (1973)^[7]最早提出, 他在讨论消费和投资组合问题时将连续时间模型应用到金融领域中。30多年以来, 连续时间模型已经成为金融经济研究中的一个重要组成部分。在金融的许多领域如资产定价、衍生产品定价、期限结构理论和投资组合选择等中, 连续时间模型被证明是最具有吸引力的方法。同时, 计量经济学的快速发展也使得连续时间方法的检验有章可循。

除了Merton关于连续时间方法的专著^[7]以外, 还有许多学者的出色的研究成果也被应用到金融经济领域。其中有Bhattacharya(1989)^[8]和Constantinides(1989)^[9]、Harrison (1985)^[10]、Malliaris (1982)^[11]、Ingersoll (1987)^[12]、Dothan (1990)^[13]、Duffie (1988, 1996)^{[14][15]}以及Karatzas和Shreve (1988, 1998)^{[16][17]}, 他们不仅将1969年至1990年这一阶段连续时间方法的发展过程作了详尽的介绍, 同时还给出了这一领域大量的参考文献。有关连续时间资产收益模型的进展情况可参见第二章。

1.2 研究现状

1.2.1 连续时间金融模型的研究现状

连续时间方法研究的重要突破主要集中在1969~1980年间。Black、Scholes和Merton开创了连续时间方法应用于金融领域的先河, 他们的期权定价模型毫无疑问是最有影响力的贡献。在他们的文章发表后, 数以百计的有关衍生产品定价的文章出现了。其中包括以各种不同资产作为标的资产的期权、远期合约、期

货合约、掉期等等。1981年以后，随着研究的进一步深入，使得连续时间方法的理论与金融市场中的实际情况逐渐趋于一致。同时，金融连续时间模型的估计方法在这一时期也成为了研究的主要方向。模型能够得到有效的参数估计归功于计量经济学的不断进步。在1990~1999年间，定价方面的研究集中在对衍生证券组合的定价问题上。此外，许多研究者致力于完善模型来缓解理论与实际不相符的矛盾。

1981年至1999年这一阶段的研究方向可以归纳为如下几个方面：

1) 在完善的市场中，动态随机最优控制问题和静态状态空间表示的同构关系的确立。Cox和Huang (1989)^{[18][19]}、Karatzas、Lehoczky和Shreve (1987)的论文介绍的方法对于解决投资组合的选择问题和资产定价非常有帮助，而这些问题对于投资者来说往往具有一定的约束条件。他们的文章提供了在约束条件下用于消费和投资组合的清晰的解决方案。

2) Duffie和Huang (1985)^[20]对部分有价证券的连续交易的研究。

3) 对于连续时间金融模型的已知理论和一些经验的矛盾的缓解。这些矛盾包括：股权溢价的反周期变动；资产定价中可预测的股票收益；衍生产品定价的波动微笑和有偏问题；信用风险理论中，违约风险变动与无风险利率变动之间存在的负相关关系；期限结构理论中，已有的单因素期限结构模型不能解释收益曲线形状的大幅变化。国内和国际的投资组合选择问题对理论研究者提出了又一挑战。研究者试图解决股权溢价问题的努力促进了具有更多参数的连续时间模型的发展。

4) 连续时间理论中加入了市场冲突 (frictions) 的概念，有助于这一理论更好地解释一些固有的事实。与此有关的内容包括税费、交易费、不完全市场、市场进入限制以及信息不对称等。包括Back (1992, 1993)^{[21][22]}、Breman和Xia (1999)^[23]、Detemple (1986)^[24]、He和Pearson (1991)^[25]以及Veronesi (1999)^[26]在内的许多学者已经开始对这一领域做更深入的研究。

5) 模型对于来自市场的数据的拟合优劣。如何使资产定价模型更好地拟合市场数据，从而具有足够的吸引力被应用到实际中呢？这一方面的内容包括无风险的期限结构模型和对受信用风险影响的证券进行定价的简化型方法。例如，无套利的期限结构模型能够拟合市场上的利率和波动的期限结构^[27]。

6) 由Brennan和Schwartz (1985)^[28]以及McDonale和Siegel (1986)^[29]最初提出的实物期权理论在这一阶段也得到了发展。

1.2.2 连续时间模型的估计研究现状

随着连续时间金融模型的发展,相应的连续时间金融模型的估计问题成为研究的热门。这一领域中多数惊人的贡献都来自于计量经济学的理论。比如有关矩的理论以及用于估计连续时间金融模型的参数和非参数方法。这一阶段出现的新的估计方法主要有: Duffie和Singleton的模拟矩SMM (simulated method of moments) 方法^[30]、Aït-Sahalia的极大似然MLE (maximum likelihood estimation) 方法^[31]、Singleton (1999)^[32]对仿射过程基于特征函数的估计方法等等。

尽管近年来有关连续时间模型的估计也取得了很大的进展,但其一般的估计方法仍旧是高度专业化的,对于未经特殊训练的人而言,也是很难实施的,一些学者也提出过一些替代的估计程序,见Lo (1988)^[33]、Hansen与Scheinkman (1995)^[34]、Aït-Sahalia (1996、2002)^[35,32]、Gallant与Tauchen (1996)^[36]、Stanton (1997)^[37]、Bandi与Phillips (1998, 2000)^[38,39]、Chacko与Viceira (1999)^[40]以及Singleton (1999)。而最近发展起来的抽样估计技术对这一问题的解决将会起着重要的作用,见Bollerslev和Zhou (2000)^[41]。

“马尔可夫链蒙特卡罗” (Markov chain Monte Carlo) 模型方法是数理统计领域近 10 年来非常热门的研究领域,这种方法可以有效地处理高维参数及高维隐含变量的估计问题。已经有一些质量很好的文章运用 MCMC 方法研究一些计量经济模型,效果非常好。应用 MCMC 方法通常需要较强的贝叶斯 (Bayes) 分析基础,从而可以针对问题提出高效率的算法;需要较强编程技巧以解决高质量的计算问题。目前存在的 MCMC 算法在解决前面所阐述问题上显得不够完善。

国内在金融计量经济学领域开始使用MCMC (Markov Chain Monte Carlo) 方法主要在 2000 年以后,王春峰等 (2000)^[42]针对现有VaR 计算中主流方法的缺陷,创新性地提出了一种基于马尔科夫链蒙特卡罗 (MCMC) 模拟的VaR 计算方法,以克服传统Monte Carlo 模拟的高维、静态性缺陷,从而提高估算精度。朱崇军 (2002)^[43]用MCMC 样本确定的缺失数据的后验分布收敛到精确分布的问题,并给出了几种度量形式下的收敛性。王春峰等 (2003)^[44]通过基于马尔可夫链蒙特卡罗(MCMC) 模拟的贝叶斯分析方法,较好地估计了随机波动性模

型中的参数与波动性序列。林静等（2005）^[45]运用基于Gibbs抽样的马尔科夫链蒙特卡洛方法动态模拟出参数后验分布的马尔科夫链，在回归参数的先验分布为多元正态分布时，给出随机截尾条件下，回归参数在指数回归模型中的贝叶斯估计。孟利锋（2003）^[46,47]、许启发和张世英（2005）^[48]使用MCMC方法估计了厚尾SV模型和Box—Cox—SV模型，黄大海和郑丕谔（2004）^[49]使用基于MCMC方法比较了两类波动模型差异。

在连续时间领域使用MCMC方法进行模型估计的研究成果在近3年才出现，吴振翔和缪柏其（2004）^[50]用MCMC方法来估计上证指数收益率线性扩散方程中的有关参数，同时采用Monte Carlo方法给出下一交易日收益率的分布，并用实际数据检验了其有效性。胡素华、张世英和张彤（2006a）^[51]对使用基于MH算法的MCMC方法估计了连续时间模型进行了总结，非线性模型和带跳跃的非线性模型。胡素华、张世英和张彤（2006b）^[52]使用MCMC方法估计了非线性连续时间模型——正态逆高斯模型（NIG）的参数；胡素华、张世英和张彤（2006）^[53,54]用MCMC方法估计了带跳跃的连续时间模型。周彦^[55,56]使用MCMC方法估计了随机波动模型和带有跳跃的随机波动模型的参数。

1.3 问题的提出与研究意义

金融市场的发展离不开相关理论的支持和指导，同时金融市场的发展又对金融理论提出了创新要求。随着研究的不断深入，量化分析成为现代金融理论发展的主要特点和趋势，概率统计、随机分析、非线性系统理论、现代控制理论、市场均衡与非均衡理论、人工智能等学科和理论广泛应用于金融研究之中，使之呈现出科学化、精细化的特征。在近些年的研究中，金融工程学和计量经济学获得了长足发展，出现了许多热点问题和研究领域，特别是资产价格连续时间建模理论、模型估计方法及其在金融风险管理中的应用研究取得了长足的进展。然而，还有许多问题有待于进一步解决。

1.3.1 问题的提出

连续时间方法与无套利均衡分析在资产定价理论中一直处于核心地位。同时，很多至今仍有影响的理论与模型都是建立在对资产价格过程作出一系列的假

设基础之上，例如，Black-Scholes期权定价模型就是如此（见本文第二章）；而这些假设有着太大的局限性，且与现实有着很大的差距。为了解决这一问题，近年来很多学者提出了更为接近现实的连续时间过程，假设其中的波动是时变的，这方面的例子主要有Hull和White（1987）与Heston（1993）的随机波动期权定价公式，Duffie和Kan（1996）、Dai和Singleton（2000）的期限结构指数仿射模型。

一般认为金融产品价格（例如股票价格、股市指数、利率等）服从一个扩散过程(diffusion process)，但是越来越多的实证研究显示汇率和权益市场(currency and equity markets)经常受到外在因素的冲击(external shock)，从而导致金融产品市场价格的“跳跃”(jump)。所以很有必要将这种跳跃引入扩散过程的模型中，目前已经有一些文章讨论这一问题，但是问题的难点在于如何估计跳跃的时间和幅度。现有文献的参数估计方法没能完善地解决这一问题。

美国股市经历了1987年10月的“股灾”。亚洲资本市场也经历了1998年的区域金融危机。那么一个很自然的问题是：连续时间的金融模型是否存在结构变化(structural change)问题。变结构在统计领域探讨得很久，而连续时间金融模型的变结构问题的难点在于：参数向量的维数高、模型通常包含隐含的不可观测的变量（其维数等同于观测值向量的维数）。传统的参数方法需要对变结构模型加上过多的假设条件，其效果往往不尽如人意。

针对前面提出的几个问题，本文具体从以下方面展开研究：

1) 对金融产品价格建立模型时将跳跃过程引入抛物线扩散模型，同时就增加了待估计参数的维数，模型的估计运用贝叶斯原理和MCMC原理，根据MCMC原理在参数和隐含变量的后验分布中抽取后验样本，从而得到参数和隐含变量的后验估计。

2) 关于连续时间金融模型的变结构问题，我们研究出一种基于MH算法的MCMC抽样算法，该算法在抽取参数向量的后验样本的同时，抽取一个能反映变结构时间的向量。目前还没有一个有效的算法来解决如此难的问题，这项研究非常具有挑战意义。

3) 研究如何对上述新模型进行参数估计，在本项目中，我们将采用贝叶斯方法和MCMC原理来估计参数。通过蒙特卡罗模拟方法验证参数估计方法，针

对我们提出的模型，选择一组合乎实际的参数，通过模型和待估参数生成一组数据。利用我们提出的参数估计方法，估计该模型的参数，如果估计方法正确，参数的估计值和真实值应该很接近，这样在一定的误差范围内，我们可以验证估计方法的正确性。

4) 模型选择：在理论上应用一些贝叶斯模型的选择方法进行比较，例如似然比检验、贝叶斯因子以及 DIC 准则；从实证分析的角度看，利用我们的模型和现有的模型进行预测，并利用预测的准确性进行比较。

5) 连续时间资产收益模型在金融工程中的应用，主要研究模型在衍生资产定价中的应用以及模型的实证分析，我们利用纽约股市成分指数、道琼斯工业指数、纳斯达克和标准普尔氏尔等市场指数进行实证分析。同时也会分析亚洲的金融市场，如日经 225、香港恒生以及我国的上海和深州的市场指数。

1.3.2 研究意义

选择“基于连续时间模型的贝叶斯分析”作为研究题目，其思想来源于张彤副教授主持的国家自然科学基金项目《基于连续时间模型的贝叶斯分析》(No:70301006)和张世英教授主持的国家自然科学基金项目《多变量矩序列长期均衡关系及动态金融风险规避策略研究》(No: 70471050)。

连续时间金融模型是近 20 年金融计量分析中比较活跃的研究领域，主要是为金融产品价格及其波动性建立模型，从而为估计和预测金融产品价格的波动提供基础，并将其应用到衍生金融产品及证券组合的定价。

连续时间模型研究的极其重要的一个原因是，它能更容易地刻画资产收益的波动特性，能够直接研究无套利基础上的资产定价与套期保值问题。因此这方面的进一步研究步仅有着重要的理论价值，而且有着很大的现实意义。本选题的理论意义在于为金融产品价格建立更实用模型并进行估计和预测；实践意义在于：对股票市场的指数、证券组合以及金融产品价格进行实证分析，并应用于期权定价。我国股票市场已经有 10 多年的发展历史，可以利用历史数据为我国股票市场建立模型，为投资者提供有效的分析工具。

1.4 本文结构安排与主要创新

1.4.1 本文结构安排

本文共六章，具体安排如下：

第一章为绪论。主要介绍论文选题的经济及金融背景与方法论背景，阐述了相关理论与建模方法的国内外研究现状，指出了存在的问题，给出本文选题的理论意义与实际意义。最后，介绍了本文研究结构安排与主要创新工作。

第二章为资产价格连续时间资产收益模型及估计方法评述。本章首先讨论了连续时间线性金融模型的进展情况，这些线性模型通过对 BS 模型添加随机波动因子或者跳跃因子从而体现不同的资产收益分布经验特征。接着总结了连续时间非线性模型的进展，相对于线性模型来说，由于非线性模型的复杂性导致这一类模型的进展相对缓慢，主要有两大类线性模型 CEV 模型和广义抛物线扩散模型。在连续时间模型不断发展的同时，模型的估计方法也在不断改进，本章的最后讨论了连续时间模型的估计方法，并总结了各种估计方法的优缺点和适用范围。

第三章为连续时间资产收益模型的 MCMC 估计。“马尔可夫链蒙特卡罗” (Markov chain Monte Carlo) 模型方法是数理统计领域近 10 年来非常热门的研究领域，这种方法可以有效地处理高维参数及高维隐含变量的估计问题。本章首先介绍了有关贝叶斯统计方法，依据贝叶斯原理，参数的后验分布正比于模型的似然函数与参数的先验分布的乘积；之后讨论了如何采用 Euler 或 Milstein 方法对连续时间模型进行离散近似从而获得模型的近似似然函数；接着给出了 MCMC 算法和取样的收敛性检验；最后通过对双指数跳跃扩散模型的估计来验证 MCMC 估计方法的有效性。

第四章为连续时间资产收益模型的变结构分析。本文提出了应用 MCMC 方法的连续时间变结构模型的单一变结构点的定位方法，并用仿真试验验证了该方法的有效性。接着提出了连续时间多变结构点模型的变结构点的定位方法；该方法在确定变结构点位置的同时，又能估计相应的模型参数。用该方法对上海股市综合指数的收益序列进行了变结构分析，理论与实证结果表明该方法是有有效且可行的。

本章首先考虑了连续时间最简单形式的 BS 模型的变结构问题，同时将该方

法推广到连续时间随机波动模型,并给出了连续时间随机波动变结构模型中变结构点位置参数、隐含波动变量以及模型参数的联合后验分布,从而可以参照连续时间 BS 变结构模型的等变结构点的定位方法对随机波动变结构模型的变结构点进行定位。

第五章为资产价格的抛物线跳跃扩散模型研究。本文以 Bibby 与 Sørensen (1997) 抛物线扩散模型为基础,用一个线性趋势和泊松跳跃过程加上一个扩散过程为资产价格对数建模,该扩散过程是一个漂移项为 0、扩散系数依赖与瞬时的资产价格,该模型称之为抛物线跳跃扩散模型。通过模型成功的拟合了几个不同价格指数数据集合,表明抛物线跳跃扩散模型能够很好的体现资产收益分布的检验特征。在本章中,开发了基于 Milstein 的 MCMC 算法,成功的估计了模型的参数及隐含跳跃变量,并验证了所开发算法的有效性。

第六章总结和展望。

1.4.2 本文主要创新工作

论文的主要创新性工作可以概况如下:

1) 开发了基于 MH 算法的“马尔可夫链蒙特卡罗”模拟 (MCMC) 方法,这种方法可以有效的处理含有高维参数及高维隐含变量模型的估计问题。目前国内外的文献主要采用的是基于 Gibbs 算法的 MCMC 方法,该方法的优点是可以直接使用 Winbugs 软件,简单方便;缺点是只能从标准分布中直接抽样,这里的标准分布包括包括正态分布、t 分布、beta 分布、gamma 分布、二项分布以及 Dirichlet 分布。而基于 MH 算法的 MCMC 方法恰好可以解决这一问题,并用该方法估计了由 Kou (2002) 提出的双指数跳跃扩散模型,从实证结果看,该方法具有非常好的准确性和可靠性,其应用前景非常乐观。

2) 研究了连续时间资产收益变结构模型。提出了连续时间 BS 变结构模型和连续时间随机波动变结构模型,同时提出了应用 MCMC 方法的连续时间变结构模型的单一变结构点的定位方法以及连续时间多变结构点模型的变结构点定位方法;该方法在确定变结构点位置的同时,又能估计相应的模型参数。用该方法对上海股市综合指数的收益序列进行了变结构分析,理论与实证结果表明该方法是非常有效且可行的。

3) 提出了抛物线跳跃扩散模型,在国内外的文献中尚未见到这类模型的研

究。首先将跳跃因子引入到抛物线扩散模型中；接着用基于 Milstein 的方法获得参数后验分布的离散密度以及隐含跳跃因子的后验密度，使用 MCMC 方法来估计抛物线跳跃扩散模型；最后用所得的估计值模拟了资产收益序列。通过上面的结果，不但发现基于 MH 算法的 MCMC 方法较其他方法，如 ML 方法，更适合还有隐含变量的模型估计；还说明了抛物线跳跃扩散模型 (HJD) 能够很好的拟合资产收益的经验特征，如有偏、尖峰厚尾以及泰勒效应，即 HJD 模型更能反映资产收益序列的统计特征。

4) 应用似然比检验和贝叶斯因子分别比较了连续时间常规模型与变结构模型、抛物线扩散模型与抛物线跳跃扩散模型的优劣，并从实证的角度利用两个模型预测的准确性来进行模型比较。

第二章 连续时间资产收益模型及估计方法评述

连续时间金融模型最早由Merton (1969) [57]在 60 年代末提出,最初应用于消费和投资组合的动态随机规划中。在 30 多年的发展过程中,连续时间方法在金融工程的期权定价[58,59]、衍生资产定价[60,61]、期限结构理论[62]和资产组合选择[1,63,64]等研究方面得到了广泛的应用。

许多资产及衍生资产定价理论都是基于有价证券价格的扩散模型,但是一些模型在实际的应用中没有取得理想的效果,主要是这些连续时间资产收益模型不能获得期权价格的解析解。最早的Black—Scholes期权定价公式是建立在严格假设基础之上,但是由于这些严格假设都偏离了实际情况,因而存在一定的偏差。将实际期权价格偏离Black—Scholes期权价格称之为“波动微笑”,如果能够对资产的价格建立合理模型将会解决“波动微笑”的问题。另外由于资产的价格对波动非常敏感,且隐含资产收益分布具有有偏和尖峰的特征,但是我们已经知道跳跃和随机波动可以解释资产收益的这些特征[65,66]。但是,资产收益的这些特征都是在众多模型中分开讨论的,同时用离散观测值对连续模型进行参数估计的技术一直是建模的障碍。

2.1 资产收益连续时间模型——线性模型

自Merton (1973)和Black—Scholes (1973)年用布朗运动和正态分布研究资产收益和期权定价以来,众多的研究者通过对Merton和Black—Scholes研究成果的扩展来描述资产收益分布的经验特征和波动微笑问题。如Merton (1976) [67]的常跳跃扩散模型, Hull和White (1987) [68]以及Heston (1993) [69]的随机波动模型, Bates (1996) [70]和Adersen等 (2002) [71]的随机波动跳跃模型 (SVJ), Duffie、Pan和Singleton (2000) [72]的收益和波动同时存在跳跃模型 (SVCJ和SVIJ), Eraker、Johannes和Polson (2003) [73]也考虑了的波动跳跃的模型, Eraker (2004) [74]扩展了Duffie等的波动跳跃模型。这些模型都是从波动的建模、收益跳跃和波

动跳跃这几个方面对BS模型进行扩展的，下面就依次介绍这几种扩展模型。

Merton (1973)、Black 和 Scholes (1973) 提出了资产收益的模型 (BS 模型) 如下：

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (2-1)$$

其中， S_t 为资产价格， W_t 为标准布朗运动， μ, σ 为参数。

资产收益的模型是在严格的假设下提出的，这些假设包括：首先，股票价格服从对数正态分布，股票收益的方差是常数；其次，股票不会分红；第三，买卖股票和期权没有交易费用等等。但是这些理想情况与现实情况严重背离，自他们之后，众多的研究者对该模型进行了扩展，主要是从跳跃、随机波动以及收益波动与资产价格的关系三个方面进行了扩展。

2.1.1 收益跳跃模型

Merton (1976) 扩展了 BS 模型，将具有常强度跳跃模型引入其中，跳跃规模是标准正态独立同分布变量：

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t + \kappa_t dq_t \quad (2-2)$$

其中 q_t 是强度为 λ 的泊松过程； $\kappa_t \sim iidN(\bar{\kappa}, \delta^2)$ 为每次跳跃的规模，其中 $\bar{\kappa}$ 和 δ^2 分别为跳跃规模的均值和方差；其它变量和参数同上。跳跃的引入可以体现尖峰和波动微笑的特征，但是由于假定跳跃规模是独立同分布的正态分布，所以并不能很好的体现收益分布的有偏性。

2.1.2 随机波动模型

在 BS 模型中，资产收益的均值和方差都是常数，但是现实情况并不是这样。Hull 和 White (1987) 使用了随机波动模型，之后 Heston (1993) 给出了在随机波动模型下，期权价格的封闭形式解。Hull 和 White (1987) 给出的对数方差随机波动模型为：

$$\begin{aligned}\frac{dS_t}{S_t} &= (\mu + cV_t)dt + \sqrt{V_t}dW_{1,t} \\ d \ln V_t &= (\alpha - \beta \ln V_t)dt + \eta \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)\end{aligned}\quad (2-3)$$

其中 W_1 和 W_2 是相互独立的标准布朗运动, V_t 为波动过程, 参数 α, β, η 体现了收益分布中的高峰特征; 参数 ρ 体现了资产收益的非对称性, 即杠杠效应, $\rho < 0$ 说明收益分布左偏, $\rho > 0$ 收益分布右偏; 参数 η 体现收益分布的厚尾特征。另外, 依据对波动建模的不同, 随机波动模型还可以采用平方根的形式建模:

$$\begin{aligned}\frac{dS_t}{S_t} &= (\mu + cV_t)dt + \sqrt{V_t}dW_{1,t} \\ dV_t &= (\alpha - \beta V_t)dt + \eta \sqrt{V_t} \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)\end{aligned}\quad (2-4)$$

各参数的含义同平方根随机波动模型。对数方差随机波动模型在形式上更加接近于标准离散时间随机波动模型, 因而可以直接与离散时间模型进行比较; 但是对数方差形式在衍生资产定价数值计算方面没有平方根方差形式便利 (Melino 和 Turnbull 1990^[75]、Benzoni 1998^[76])。

随机波动体现了资产收益过程中的高峰特征、收益的非对称 (杠杠效应) 以及厚尾特征。但是 SV 模型的峰度随着取样频率的增加而减少, 而这正与资产收益的经验分布特征恰好相反, 而加入跳跃就能解决这一问。

2.1.3 随机波动—收益跳跃模型 (SVJ)

仅仅含有跳跃或随机波动的模型只能从某个方面来反映资产收益分布的部分特征, Bates (1996) 将收益的跳跃和随机波动结合, 提出了 SVJ 模型:

$$\begin{aligned}\frac{dS_t}{S_t} &= (\mu + cV_t - \lambda \bar{\kappa})dt + \sqrt{V_t}dW_{1,t} + \kappa_t dq_t \\ dV_t &= (\alpha - \beta V_t)dt + \eta \sqrt{V_t} \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)\end{aligned}\quad (2-5)$$

或

$$\begin{aligned}\frac{dS_t}{S_t} &= (\mu + cV_t - \lambda \bar{\kappa})dt + \sqrt{V_t}dW_{1,t} + \kappa_t dq_t \\ d \ln V_t &= (\alpha - \beta \ln V_t)dt + \eta \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)\end{aligned}\quad (2-6)$$

其中, q_t 是服从强度为 λ 的泊松过程, $\kappa_t \sim iidN(\bar{\kappa}, \delta^2)$ 为每次跳跃的规模,

其中 $\bar{\kappa}$ 和 δ^2 分别为跳跃规模的均值和方差。该模型体现了有偏和高峰特征，更重要的是，该模型改变了 SV 模型只能在存在极高波动风险时才能解释波动微笑的缺点，在 SVJ 模型中，只要加入合适的跳跃参数就可以解决波动微笑，跳跃规模方差的增加可以增大收益分布的两个尾部（左尾和右尾），同时跳跃规模均值小于 0，说明左尾较右尾更厚，反之则相反。

由于假定跳跃强度 λ 固定不变不能解释股票价格波动的集聚性，Adersen、Benzoni 和 Lund 于 2001 也提出了 SVJ 模型，但是其假定跳跃强度为 $\lambda(t)$ 且 $\lambda(t)$ 是瞬时方差的仿射函数，且跳跃规模 κ_t 服从对数正态分布，即

$$\begin{aligned} \lambda(t) &= \lambda_0 + \lambda_1 V_t \\ \ln(1 + \kappa_t) &\sim N(\ln(1 + \bar{\kappa}) - 0.5\delta^2, \delta^2) \end{aligned} \quad (2-7)$$

其中， $\bar{\kappa}$ 为跳跃规模的均值， δ^2 为跳跃规模对数的方差。

2.1.4 随机波动—收益和波动跳跃模型

Duffie、Pan 及 Singleton (2000) 提出了收益和波动具有跳跃的两个模型，一个是收益和波动跳跃同时发生、且两者跳跃规模大小相关的模型 (SVCJ)；另外一个为收益和波动跳跃独立发生且规模也独立 (SVIJ)。Eraker、Johannes 和 Polson (2003) 考虑了波动和收益的跳跃效应，相应的收益和波动跳跃模型如下

$$\begin{aligned} \frac{dS_t}{S_t} &= (\mu + cV_t - \lambda^* \bar{\kappa}^*) dt + \sqrt{V_t} dW_{1,t} + \kappa_t^* dq_t^* \\ dV_t &= (\alpha - \beta V_t) dt + \eta \sqrt{V_t} \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right) + \kappa_t^* dq_t^* \end{aligned} \quad (2-8)$$

其中 W_1 和 W_2 是相互独立的标准布朗运动； q_t^*, q_t^* 是跳跃强度分别为 λ^*, λ^* 的泊松过程，且与 W_1 和 W_2 不相关； κ_t^*, κ_t^* 是收益和波动的跳跃规模， $\bar{\kappa}^*$ 为收益跳跃的均值；假定参数和初始条件能够保证方程 (2-8) 的解的存在。

SVIJ 模型的收益和波动跳跃分别以强度为 λ^*, λ^* 的泊松过程随机达到，其中，跳跃 $\kappa_t^* \sim \exp(\mu_s)$ ，收益跳跃 $\kappa_t^* \sim N(\mu_s, \sigma_s^2)$ 。SVCJ 模型的跳跃具有相同的到达时间，即 $q_t^* = q_t^* = q_t$ ，相应的跳跃规模为 $\kappa_t^* \sim \exp(\mu_s)$ 、 $\kappa_t^* | \kappa_t^* \sim N(\mu_s + \rho_s \kappa_t^*, \sigma_s^2)$ 。

Eraker (2004) 还提出了状态非独立、相关跳跃随机波动模型 (SVSCJ)，该

模型除了满足 SVCJ 的模型条件外，还满足跳跃频率依赖于波动，即：

$$\lambda^s = \lambda^v = \lambda_0 + \lambda_1 V_t \quad (2-9)$$

2.1.5 线性模型的比较

上述的几种线性连续时间资产收益模型分别加入随机波动因子或者跳跃因子对 BS 模型进行了扩展，不同因子的加入也就能体现不同的资产收益分布经验特征，具体结果见表 2.1。

表 2.1：连续时间线性模型特征

模型名称	模型形式	特征
收益跳跃模型	$dS_t/S_t = \mu dt + \sigma dW_t + \kappa dq_t$	尖峰厚尾、波动微笑
随机波动模型	$dS_t/S_t = (\mu + cV_t)dt + \sqrt{V_t}dW_{1,t}$ $d \ln V_t = (\alpha - \beta \ln V_t)dt + \eta \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)$	有偏性、尖峰厚尾
随机波动-收益跳跃模型	$dS_t/S_t = (\mu + cV_t - \lambda \bar{\kappa})dt + \sqrt{V_t}dW_{1,t} + \kappa dq_t$ $dV_t = (\alpha - \beta V_t)dt + \eta \sqrt{V_t} \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)$	有偏性、尖峰厚尾、波动微笑
随机波动-收益和波动跳跃模型	$dS_t/S_t = (\mu + cV_t - \lambda^s \bar{\kappa}^s)dt + \sqrt{V_t}dW_{1,t} + \kappa^s dq_t^s$ $dV_t = (\alpha - \beta V_t)dt + \eta \sqrt{V_t} \left(\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right) + \kappa^v dq_t^v$	有偏性、尖峰厚尾、波动微笑

2.2 资产收益连续时间模型——非线性模型

在线性模型发展的同时，资产收益的非线性模型也在发展，但是由于非线性模型在数值计算上的困难，直到 Eberlein 和 Keller (1995)^[77]用正态倒数高斯

(NIG) 为股票价格建模以来，研究人员才开始用非线性连续时间模型来描述资产收益的分布，同时非线性模型得到了快速的发展。Barndorff-Nielsen 和 Shephard (2001)^[78]，Schoutens (2003)^[79]，Barndorff-Nielsen 和 Shephard (2005)^[80]对 NIG 模型进行了扩展和应用。与此同时，股票价格的广义抛物线扩散模型也取得了发展，主要成果有 Rydberg (1997)^[81]，Bibby 和 Sørensen (1997)^[82]，T.H. Rydberg (1999)^[83]，Cont 和 Tankov (2004)^[84]，Emmer 和 Kläupelberg (2004)^[85]等。

2.2.1 CEV 模型

Cox和Ross (1976)^[86]和Cox (1996)^[87]认为收益的波动受资产价格的影响，但是收益的波动相对于资产价格的弹性不变，因而采用下式建模：

$$dS_t = \mu S_t dt + \sigma S_t^\gamma dW_t \quad (2-10)$$

该模型为方差常弹性模型，记为 CEV；其中 $\gamma(0 < \gamma < 1)$ 是方差弹性（其实当 $\gamma=1$ 时，该模型就是 BS 模型，所以在此不考虑 $\gamma=1$ 的情况），CEV 模型较 BS 模型的优点就是解决了收益分布非对数正态的假定以及体现了波动微笑的特征。对于长期期权，期权价格要么是 0 要么是无穷大。另外 CEV 模型没有尖峰特征，更准确地说，CEV 模型中的收益分布有比正态分布更瘦的右尾，因而在期权定价中，隐含波动是执行价格的单调函数。

依据漂移项和扩散项的函数形式不同，CEV 模型又可以具体化为以下几类模型。

2.2.1.1 Vasicek 模型

Vasicek (1977)^[88]提出了该模型，模型形式如下：

$$dS_t = (\alpha + \beta S_t) dt + \sigma dW_t \quad (2-11)$$

该模型常常用于债券期权、债券期货以及其他衍生证券的定价。

2.2.1.2 CIR 模型

Cox、Ingersoll和Ross (1985)^[89]在用广义均衡资产定价模型来研究利率的期限结构中提出了CIR模型，CIR模型是 $\gamma = \frac{1}{2}$ 时的CEV模型，CIR模型形式如下：

$$dS_t = \xi(\theta - S_t) dt + \sigma \sqrt{S_t} dW_t \quad (2-12)$$

CIR 模型最早是用于分析单因子广义均衡期限结构模型，在该模型中条件波动的变化是正比于资产价格的。

2.2.1.3 广义 Chan 模型

Chan et al (1992)^[90]提出了广义Chan模型，形式如下：

$$dS_t = \xi(\theta - S_t) dt + \sigma_0 (S_t^\gamma + \sigma_1) dW_t \quad (2-13)$$

该模型还有另外一种形式：

$$dS_t = \xi(\theta - S_t)dt + \sigma_0(e^{\gamma S_t + \gamma t} + \sigma_1)dW_t \quad (2-14)$$

这两个模型形式仅仅是扩散项的形式不同。

前面两个模型的参数估计较第三个模型简单，因而常常用于利率的衍生资产定价中，另外 CEV 类模型通常用于描述利率的动态变化。

2.2.2 广义抛物线类扩散模型

基于资产收益分布的经验特征，T.H. Rydberg (1999) 以 B-S 模型为基础，通过扩散系数（波动）拟合不含确定线性趋势的价格对数的波动部分，而这种扩散系数是依赖于过程的状态，即股票价格。

经验结果表明，股票价格的对数是一个近似不相关的增量过程，但是增量不是独立的。为股票价格而建立的模型是：

$$S_t = \exp(\kappa t + X_t) \quad (2-15)$$

其中

$$X_t = X_0 + \int_0^t v(X_s) dW_s \quad (2-16)$$

通过伊藤公式，有

$$dS_t = S_t \left\{ \left[\kappa + \frac{1}{2} v^2(\log S_t, -\kappa t) \right] dt + v(\log S_t, -\kappa t) dW_t \right\} \quad (2-17)$$

其中

$$v^2(x) = \sigma^2 \exp(-\beta(x-\mu)) \left(\frac{\left(\sqrt{\delta^2 + (x-\mu)^2} \right)^{\lambda-1}}{K_{\lambda-\nu/2}(\alpha \sqrt{\delta^2 + (x-\mu)^2})} \right) \quad (2-18)$$

其中 $K_\lambda(x)$ 为修正的第三种巴塞尔函数，各参数满足 $\lambda, \mu, \beta \in R, \delta, \alpha \geq 0$ 并且有：

$$\begin{aligned} & \text{当 } \lambda > 0 \text{ 时, } \delta \geq 0, \alpha > 0, \alpha^2 > \beta^2 \\ & \text{当 } \lambda = 0 \text{ 时, } \delta > 0, \alpha > 0, \alpha^2 > \beta^2 \\ & \text{当 } \lambda < 0 \text{ 时, } \delta > 0, \alpha \geq 0, \alpha^2 \geq \beta^2 \end{aligned} \quad (2-19)$$

α 是倾斜度参数， β 是对称参数，若 $\beta = 0$ 则密度是关于均值对称的， δ 是标度参数， μ 是位置参数。

通过选择不同的 λ 值，分布就有不同的具体形式。当 $\lambda=1$ 时，该模型就转化为 Bibby 和 Sørensen (1997) 的抛物线扩散模型，同时函数 $v^2(x)$ 就简化为：

$$v^2(x) = \sigma^2 \exp\left[\alpha\sqrt{\delta^2 + (x-\mu)^2} - \beta(x-\mu)\right] \quad (2-20)$$

当 $\lambda = -1/2$ 时，就变成倒数逆高斯扩散模型， $v^2(x)$ 为：

$$v^2(x) = \sigma^2 \exp(-\beta(x-\mu)) \frac{\sqrt{\delta^2 + (x-\mu)^2}}{K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)} \quad (2-21)$$

当函数 $v^2(\bullet)$ 为一常数时，该模型就是基本的 BS 模型。

2.2.3 非线性模型的比较

上述简单的介绍了 CEV 模型和广义抛物线扩散模型，其中 CEV 模型和 BS 模型都是广义抛物线扩散模型的特殊情况，表 2.2 给出了这两种模型在描述资产收益分布经验特征的情况。

表 2.2: 连续时间非线性模型特征

模型名称	模型形式	特征
CEV 模型	$dS_t = \mu S_t dt + \sigma S_t^\lambda dW_t$	有偏性、波动微笑
广义抛物线扩散模型	$dS_t = S_t \left\{ \left[\kappa + \frac{1}{2} v^2(\log S_t - \kappa t) \right] dt + v(\log S_t - \kappa t) dW_t \right\}$ $v^2(x) = \sigma^2 \exp(-\beta(x-\mu)) \frac{\left(\sqrt{\delta^2 + (x-\mu)^2} \right)^{\lambda-1}}{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}$	有偏性、尖峰厚尾、波动微笑、泰勒效应

2.3 连续时间模型估计方法

在模型的参数估计方面，如何从离散取样数据中得到连续时间模型有效推断是参数估计的最大难题，主要存在于：关于离散转移密度的封闭形式不能获得，特别是存在不可观测和序列相关状态变量情况下，后面一种情况一般出现在随机

波动模型中。虽然原则上借助数值计算可以得到最大似然估计 (Lo1988)，但是如果隐含变量必须从似然函数中积分出来，对计算机的需求将是巨大的。

过去几十年中连续时间领域中出现的最显著的成果是连续时间模型的估计，这方面最早的贡献是由Nelson (1989, 1990, 1991) ^[91,92,93]在研究GARCH模型和扩散模型关系中得出的。Nelson发现GARCH模型可以看成是连续时间随机波动模型的近似，因而可以用GARCH模型的最大似然估计值 (MLE) 作为连续时间随机波动模型的有效近似。

总的来说，连续时间模型的估计方法可以概括为以下几类：最大似然估计、广义矩估计、Duffie和Singleton (1993)的模拟矩估计、Gallant和Tauchen (1996)的有效矩估计、Ait Sahalia (1996)等的非参数方法、Duffie, Pan和Singleton (2000)以及Liu (1997) ^[94]开发了一种利用经验特征函数 (ECF empirical characteristic functions) 的估计方法以及贝叶斯方法。

2.3.1 最大似然估计 (MLE)

对于连续时间扩散模型，过程的条件密度可以用 Fokker—Planck 等式获得。考虑下面的扩散过程。

$$dY_t = \mu(Y_t, \theta) dt + \sigma(Y_t, \theta) dW_t \quad (2-22)$$

其中 θ 是待估计的参数向量， Y_t 是状态变量。该过程的条件密度 f 由 Fokker—Planck 等式确定，Fokker—Planck 等式为：

$$\frac{\partial f}{\partial t} = -\frac{\partial[\mu f]}{\partial Y} + \frac{1}{2} \frac{\partial^2[\sigma^2 f]}{\partial Y^2} \quad (2-23)$$

给 Fokker—Planck 等式加上一个初始条件就可以求得过程的条件密度。在有些时候，单变量的扩散过程可能没有稳定的分布，在运用 MLE 方法以前，必须确保满足近似正规化条件。在利率应用中，可以获得封闭形式的正确条件密度，包括利率期限结构的 CIR (1985b) 模型，CIR 是一个单变量平方根扩散、具有几何布朗运动的 BS 模型，以及含有 OU 过程的 Vasicek (1977) 模型。在这些情况下，可以应用 MLE 方法来估计扩散模型的参数。

Pearson与Sun (1994) ^[95]用最大似然估计方法估计了CIR模型。在研究利率的变化中，ML方法不适合那些条件密度不具有封闭形式的情况。原则上来说，

似然函数可以用蒙特卡罗模拟来估计，由于需要对每一个条件变量和参数进行模拟，所以在计算上存在问题。虽然Lo (1988) 表明如何用数值方法来完成ML过程，但是这种方法需要进行大量的计算，这是由于在每一个循环中需要解一个偏微分方程。

在最近的研究结果中，当状态变量 Y_t 的观测值是离散情况时，Aït Sahalia (2002) ^[31]用下式作为似然函数的对数近似：

$$L = \sum_{t=0}^n \ln p_Y(\Delta, Y_{t+1} | Y_t; \theta) \quad (2-24)$$

其中 $p_Y(\Delta, Y_{t+1} | Y_t; \theta)$ 为状态变量 Y_t 的转移密度。

Aït Sahalia (2002, 1999) ^[31,96]表明ML方法可以在不能获得条件密度封闭形式下也可以适用。Aït Sahalia使用单变量扩散过程的未知但真实的转移密度函数来作为条件密度的近似。使用转移密度的Hermite扩展，可以获得似然函数显性的近似。但是这种方法在多变量的金融计量经济学连续时间模型中不太适合。

2.3.2 广义矩估计 (GMM)

对于一些多变量扩散模型和单变量扩散的数量限制，我们可以得到条件密度的解析特征。对于许多扩散过程而言，很难并且有时不可能求到Fokker—Planck等式的显性形式，此时，MLE方法就不具有可行性。因此在这种情况下，Hansen与Scheinkman (1995) 运用GMM方法来估计模型的参数。利用离散观测值获得连续时间模型的矩条件，这种方法的关键是使用无限小生成器来描述连续时间马尔可夫过程的特征，同时提出可以用这个生成器来构造稳定马尔可夫过程的矩条件。但是GMM方法不太适合模型中含有隐含观测变量的情况，如随机波动模型。GMM方法具体步骤如下：

首先依据结构模型获得参数的解析矩条件：

$$E(g(Y, \theta)) = 0 \quad (2-25)$$

其次从状态变量的离散观测值构造样本矩 $m(Y, \theta)$ 。这样参数 θ 的GMM估计值 $\hat{\theta}$ 为：

$$\hat{\theta} = \arg \min_{\theta} m(Y, \theta)' \hat{V}^{-1} m(Y, \theta) \quad (2-26)$$

其中 \hat{V} 为一对称正定矩阵，或称为权重函数， \hat{V} 依概率收敛于 V ， V 满足：

$$\lim \hat{V} \xrightarrow{p} V = E \left[g(Y, \theta) g(Y, \theta)' \right] \quad (2-27)$$

其中 \xrightarrow{p} 表示依概率收敛。

GMM估计是大样本估计，也就是说在大样本下的估计是渐进有效的，但是在小样本下却是无效的^[97]。此外，GMM方法很难对同时具有随机波动和跳跃因子的模型进行估计。

2.3.3 模拟矩估计 (SMM)

Duffie 与 Singleton (1993) 提出了模拟矩估计方法，并将该方法应用于时间齐次马尔可夫过程中。当可观测的状态变量（如资产价格、股票指数等）是关于人工变量和未知参数的非解析函数时，GMM方法就很难对这些未知参数进行估计，而SMM方法通过对GMM方法进行扩展从而适合该种情况。在使用SMM时，需要解决了以下几个重要的问题：首先，模拟需要初始条件，且模拟时没有稳定分布的密度函数可以直接取样；另外，模拟的状态变量通过参数向量依赖于模型的结构和模拟的数据。SMM方法特别适合那些似然函数不能获得封闭形式解的情形，SMM方法的具体步骤如下：

首先依据结构模型获得观测变量的转移密度函数 H ：

$$Y_{t+1} = H(Y_t, \theta) \quad t=1, 2, \dots, T \quad (2-28)$$

并令 $Z_t = (Y_t, Y_{t-1}, \dots, Y_{t-l+1})$ ($l \leq t$)， θ 的估计值就依赖于向量矩 $f_t^* = f(Z_t, \theta)$ ，其中 f 是一个可观测函数。

其次给定参数值 θ' ，依据转移函数模拟过程 Y'

$$Y'_{t+1} = H(Y'_t, \theta') \quad (2-29)$$

相应的有 $Z'_t = (Y'_t, Y'_{t-1}, \dots, Y'_{t-l+1})$ 和 $f'_t = f(Z'_t, \theta')$ 。模拟 Y' 的次数为 s ，当 $T \rightarrow \infty$ 时，必有 $Ts \rightarrow \infty$ 。对于任意一个参数向量 θ' 有：

$$G_T(\theta) = \frac{1}{T} \sum_{t=1}^T f_t^* - \frac{1}{Ts} \sum_{i=1}^{Ts} f'_i \quad (2-30)$$

这样，模拟矩估计量就是：

$$\hat{\theta} = \arg \min_{\theta} G_T(\theta)' W_T G_T(\theta) \quad (2-31)$$

其中 W_T 为一半正定矩阵，或称为权重函数， W_T 由状态变量观测值 $\{Y_t\}$ 和模拟状态变量 $\{Y_t^s\}$ 决定。

SMM方法通过比较样本矩和模拟矩而获得参数的估计值，计算模拟矩所使用的数据是使用模拟方法所产生的序列。SMM方法虽然避免了模型的线性化以及似然函数的估计等问题，同时SMM方法应用与扩展都较容易，但是估计效果不及最大似然估计方法^[30]。

Broze、Scaillet和Zakoian（1998）^[98]使用SMM方法利用离散数据估计了连续时间模型，Brandt与Santa-Clara（1999）^[99]使用该方法估计了多变量扩散过程。

2.3.4 有效矩估计（EMM）

Gallant与Tauchen（1996，1997，1998）^[36,100,101]提供了另外一种估计方法。他们通过建立一个辅助模型，该辅助模型具有对数密度的衍生形式。该方法的优点是得分（Score）函数具有解析表达式，通过得分函数的结构模型表达式来构成矩条件。这个解析表达式依赖于辅助模型和结构模型的参数，辅助模型的参数用柯西MLE（伪极大似然估计值）来代替，结构模型的参数通过最小GMM标准函数来估计。EMM的具体方法如下：

首先，EMM引入一个辅助模型或“得分”生成器（score generator），该辅助模型以辅助参数向量 η 表示的条件密度的形式给出，形式如下：

$$f(y_t | Y_{t-1}, \eta) \quad (2-32)$$

其中 $Y_t = \{y_t, y_{t-1}, \dots, y_1\}$ 为历史数据信息。

其次采用柯西最大似然估计方法估计辅助模型参数，估计值 $\hat{\eta}_T$ 满足一阶条件：

$$\frac{1}{T} \sum_{t=1}^T \frac{\partial}{\partial \hat{\eta}} \ln f(y_t | Y_{t-1}, \hat{\eta}_T) = \frac{1}{T} \sum_{t=1}^T S_f(y_t, \hat{\eta}_T) = 0 \quad (2-33)$$

其中 $S_f(y_t, \hat{\eta}_T) = (\partial / \partial \hat{\eta}) \ln f(y_t | Y_{t-1}, \hat{\eta}_T)$ 为柯西得分（score）函数（类似于MLE的似然函数）。

接着，设 $P(Y_i|\theta) = \ln f(y_i|Y_{i-1}, \theta)$ 为待估模型在其参数向量 θ 上的对数似然向量，对 $P(Y_i|\theta)$ 微分，并以 $S_f(Y_i, \hat{\eta}_T)$ 为被积函数求积分，得到如下的向量：

$$m(\theta, \hat{\eta}_T) = E[S_f(Y_i, \hat{\eta}_T)] = \int S_f(Y_i, \hat{\eta}_T) dP(Y_i, \theta) \quad (2-34)$$

其含义表示待估模型的伪得分向量与辅助模型的伪得分向量的差别。由于存在隐含变量，所以很难得出式 (2-34) 的具体解析式，一个替代的方法是，利用 Duffie 与 Singleton(1993) 提出的模拟矩方法 (Simulated Method of Moments, SMM)，由待估模型产生一个模拟序列 $\hat{y}_n(\theta)$ ， $n=1, \dots, N$ ，这样，(2-34) 式可写为：

$$m_N(\theta, \hat{\eta}) = \frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial \eta} \ln \{f[\hat{y}_n(\theta) | \hat{Y}_{n-1}(\theta), \eta]\}_{\eta=\hat{\eta}} \quad (2-35)$$

最后，在前两步的基础上，极小化伪得分向量的二次型，得出待估参数 θ 的 EMM 估计量：

$$\hat{\theta} = \arg \min_{\theta} [m_N(\theta, \hat{\eta})' \hat{I}_T^{-1} m_N(\theta, \hat{\eta})] \quad (2-36)$$

这里， \hat{I}_T 表示伪得分向量的近似协方差矩阵，它对应每个得分的权重，由得分向量的外积得出：

$$\hat{I}_T = \frac{1}{T} \sum_{i=1}^T \frac{\partial}{\partial \eta} \ln f(y_i | Y_{i-1}, \eta) \frac{\partial}{\partial \eta} \ln f(y_i | Y_{i-1}, \eta)' \Big|_{\eta=\hat{\eta}} \quad (2-37)$$

综上，EMM 的基本思想是用模型隐含的矩匹配可观测量的转移密度隐含的矩。Gallant 和 Tauchen 的 EMM 由三步组成：预测、估计、再预测^[98]。EMM 方法最显著的特征是：当辅助模型的得分函数收敛于真实模型的得分函数时，该方法与最大似然方法的估计效果一致。理论上，如果可以模拟隐含路径并能够借助矩条件获得用于估计参数的足够信息，模拟方法是可行的。但是该方法的缺点是，它既需要一个辅助模型和一个连续时间模型捕捉数据的所有特征，而且还需要这两个模型的参数之间存在一对一的映射关系，因而实现起来比较繁琐。

Anderson 与 Lund (1997)^[102] 使用 EMM 方法估计了短期利率的随机波动模型，Benzoni (1999)^[103] 使用该方法估计了标的资产服从随机波动模型的期权定价问题。

2.3.5 经验特征函数估计 (ECF)

该领域的工作最早是由Parzen (1962) ^[104]开始的, 而且按处理的是独立同分布过程还是相关平稳的随机过程可分成两类。这里主要讨论非独立同分布的情况, 独立同分布情况是它的特例; 并且假设研究的过程是严平稳和弱相关的, 这样能保证经验特征函数估计量的渐近特性。

假设随机过程 X 的累积分布函数是 $F(x; \theta)$, 依赖于一个 K 维参数向量 θ 。特征函数的定义为:

$$c(r; \theta) \equiv E[\exp(irX)] = \int \exp(irx) dF(x; \theta) \quad (2-38)$$

而经验特征函数是特征函数的样本对应部分, 其定义为:

$$c_n(r) \equiv \frac{1}{n} \sum_{j=1}^n \exp(irX_j) \approx \int \exp(irx) dF_n(x) \quad (2-39)$$

其中 $i = \sqrt{-1}$, $\{X_j\}_{j=1}^n$ 是一个独立同分布序列, $F_n(x)$ 是经验累积分布函数, r 是转换变量。假定 θ 的真值为 θ_0 , 特征函数是 θ 的一个确定性函数, 而经验特征函数仅通过观测值 $\{X_j\}$ 依赖于 θ_0 。在不引起混淆的情况下, 在 $c(r; \theta)$ 中把 θ 省略掉。

由前面的定义知, 特征函数和经验特征函数分别是分布函数和经验分布函数的傅立叶变换, 由于 Fourier—Stieltjes 变换的唯一性, 特征函数包含与分布函数同样多的信息, 且经验特征函数保留样本中的所以信息。同时注意到特征函数仅包含参数, 而经验特征函数仅包含数据。经验特征函数估计方法的一般思想是最小化经验特征函数和特征函数之间距离的各种测度。例如, 通过选择一组变换变量, r_1, \dots, r_q , 可以最小化下面的距离:

$$\sum_{j=1}^q |c_n(r_j) - c(r_j; \theta)|^2 g(r_j) \quad (2-40)$$

也就是, q 个离散点上的距离。或者通过连续地选择变换变量 r , 这样参数向量 θ 的估计值就是:

$$\hat{\theta} = \min_{\theta \in \Theta} \int |c_n(r) - c(r; \theta)|^2 g(r) dr \quad (2-41)$$

也就是一个区间上的距离。在这两种情况下, $g(\cdot)$ 都是一个权重函数, $g(\cdot)$ 如何确定参见文献[47]。

Singleton (2001) ^[105]、Chacko和Viceira (2003) ^[106]使用经验特征函数方

法估计了扩散过程的参数。

2.3.6 非参数估计 (NPE)

非参数估计方法适合于：不存在一个结构模型能够很好的描述外生变量和内生变量关系的情况。另外，非参数估计方法对数据的要求十分严格，大量高质量的数据是非参数估计效果的关键。幸运的是在金融经济学中，高质量的数据不是问题，而如何选择一个合适的模型却是问题的关键。非参数估计方法具体可以参考文献 Jiang 和 Knight (1997)。

非参数估计的优点在于需要较少的数据生成过程的先验信息。考虑扩散模型

$$dY_t = \mu(Y_t, \theta)dt + \sigma(Y_t, \theta)dW_t \quad (2-42)$$

在一般条件下^{[107][108]}，该扩散过程是一维严马尔可夫过程，具有时不变稳定转移密度。漂移项 (drift) 和扩散项 (diffusion) 可以用下式表示：

$$\begin{aligned} \mu(Y_t, \theta) &= \lim_{\Delta \rightarrow 0} E \left[\frac{Y_{t+\Delta} - Y_t}{\Delta} \middle| Y_t \right] \\ \sigma^2(Y_t, \theta) &= \lim_{\Delta \rightarrow 0} E \left[\frac{(Y_{t+\Delta} - Y_t)^2}{\Delta} \middle| Y_t \right] \end{aligned} \quad (2-43)$$

漂移项表示的是过程 Y_t 随时间的变化，而扩散项表示的是过程 Y_t 关于漂移项的随机波动程度。

使用 Dynkin 算子^{[17][37]}可以获得漂移项和扩散项的一阶以及多阶近似，在此为了简单说明非参数估计方法，仅仅考虑一阶近似。状态变量 Y_t 的观测值集合为 $\{Y_t, t = 0, 1, \dots, n\}$ ，任意两个观测值之间的时间间隔为 Δ ，漂移项和扩散项的一阶近似为：

$$\begin{aligned} \mu(Y_t, \theta) &\approx \frac{1}{\Delta} E[Y_{t+\Delta} - Y_t | Y_t] \\ \sigma^2(Y_t, \theta) &\approx \frac{1}{\Delta} E[(Y_{t+\Delta} - Y_t)^2 | Y_t] \end{aligned} \quad (t=1, 2, \dots, n) \quad (2-44)$$

这样，参数的估计问题就转化为非参数回归问题。一般用 Nadaraya-Watson (NW) 核估计值^[109, 110]来估计 $\hat{\mu}, \hat{\sigma}^2$ ，因而有：

$$\hat{\mu}(y) = \frac{1}{\Delta} \frac{\sum_{i=1}^{n-1} (Y_{i+1} - Y_i) K_h(y - Y_i)}{\sum_{i=1}^{n-1} K_h(y - Y_i)} \quad (2-45)$$

$$\hat{\sigma}^2(y) = \frac{1}{\Delta} \frac{\sum_{i=1}^{n-1} (Y_{i+1} - Y_i)^2 K_h(y - Y_i)}{\sum_{i=1}^{n-1} K_h(y - Y_i)} \quad (2-46)$$

其中 $K_h(u) = K(u/h)/h$, $h = h_n \rightarrow 0$ 是带宽 (bandwidth), $K(\bullet)$ 是标准核, 是一个有界可积函数, 具体形式可以见文献[111,32]。

Pritsker (1998)^[112]使用Vasicek (1977) 的利率模型来评价在有限样本下非参数估计方法中的核密度估计值的效果, Stanton (1997)^[37]用具有持续性的美国利率数据检验了非参数估计的效果。

2.3.7 各种估计方法的比较

连续时间模型的估计方法是过去二十年中连续时间领域中出现的显著成果之一, 正是由于连续时间模型估计方法的进步才促进了连续时间模型的发展。表 2.3 总结了上述的几种连续时间模型估计方法的优缺点和适用范围。

表2.3: 各种估计方法的优缺点

估计方法	优缺点
有效矩估计	通过辅助模型来获得“得分”函数的解析表达式, 但估计的效果依赖辅助模型的选择
模拟矩估计	一致估计, 但仅仅适合于时不变马尔可夫过程
经验特征函数估计	可以获得最大似然估计和矩估计的结果, 但仅仅适用于仿射类扩散 (Affine Diffusion) 模型
非参数估计	适合没有合适模型来描述内生变量、外生变量和参数之间存在的关系, 但估计结果受数据质量的影响以及没有明确的模型形式

2.4 本章小结

连续时间金融模型在过去 20 多年的发展过程中取得了巨大的进步, 本章首先讨论了连续时间线性金融模型的进展情况, 这些线性模型通过对 BS 模型添加随机波动因子或者跳跃因子从而体现不同的资产收益分布经验特征。接着总结了

连续时间非线性模型的进展，相对于线性模型来说，由于非线性模型的复杂性导致这一类模型的进展相对缓慢，主要有两大类线性模型 CEV 模型和广义抛物线扩散模型。在连续时间模型不断发展的同时，模型的各种估计方法也在不断改进，本章的最后讨论了连续时间模型的估计方法，并总结了各种估计方法的优缺点和适用范围，基于 Bayes 原理的 MCMC 估计方法将在下一章中详细讨论。

第三章 连续时间模型参数估计的 MCMC 方法

Pedersen (1995)^[113]首先将参数估计的问题作为缺失值问题,并建立了相应的 MCMC方法,虽然最初这种方法不能适应含有隐含因子的模型参数估计,但是 Eraker (2001)^[114], Jones (1998)^[115]以及Elerian、Chib和Shephard(2001)^[116]扩展了 MCMC方法,为MCMC技术提供了新的进展。Tse、张喜彬和Yu (2003)^[117], 胡素华、张世英和张彤 (2006a, d) 分别用MCMC方法对抛物线扩散和NIG扩散模型进行了估计,孟利峰等 (2004)^[118]也用MCMC方法估计了随机波动模型。实际上,MCMC方法对那些在似然函数中包含高维积分的问题以及模型含有隐含变量的情形特别适用。

MCMC方法具有以下几个优点:首先,MCMC方法可以适用于多变量模型,其可以避免间接推断方法中由于引入辅助模型而产生的参数数量大规模增加的问题;其次,对于那些含有隐含变量的模型,如随机波动模型,在进行参数估计的同时,MCMC方法还可以解决这些隐含变量的滤波和预测问题;第三,MCMC方法从参数的条件后验分布来模拟参数值时不需要精确的参数后验分布,仅仅需要精确后验分布的正比形式,另外在使用MH算法的MCMC方法中,可以直接从各个参数的联合后验分布中对各个参数直接取样;最后,在多变量模型中,各个变量可以非同步取样。

使用MCMC方法进行模型的参数估计,首先要获得参数的后验分布,这样就必须对模型进行离散化的表示,常用的方法是采用Euler方法,这种方法比较简单,但是误差相对较大;另外就是采用Milstein方法,这种方法精度较高,但是计算相对要复杂,尤其是在含有跳跃的模型中。Tse、张喜彬和Yu (2003)采用Milstein方法对抛物线扩散模型进行了估计,并比较了这两种方法的效果。这里首先介绍贝叶斯统计方法,接着讨论连续时间模型的离散方法,最后给出基于Bayes原理的MCMC估计方法并用该方法估计了Kou提出的双指数跳跃扩散模型(Double Exponential Jump Diffusion Model, DEJD),同时用该模型对我国的沪深股市进行了实证分析。

3.1 贝叶斯统计方法

3.1.1 贝叶斯公式的密度函数形式

用随机变量的密度函数形式叙述贝叶斯公式如下：

参数向量 θ 的密度函数在经典统计中记为 $p(X; \theta)$ 或 $p_{\theta}(X)$ ，表示的是在参数空间 $\Theta = \{\theta\}$ 中不同的 θ 对应不同的分布。而在贝叶斯统计中记为： $p(X | \theta)$ ，表示在参数 θ 给定某个值时，总体指标 X 的条件密度分布。根据参数 θ 的先验信息确定其先验分布为 $\pi(\theta)$ 。

依据贝叶斯观点，样本 $X = (x_1, \dots, x_n)$ 的产生分两步。首先，设想从先验分布 $\pi(\theta)$ 产生一个参数向量的实现 θ' ，这一步人们是看不到的，所以用“设想”二字。第二步从总体密度分布 $p(X | \theta')$ 产生一个具体的样本 $X = (x_1, \dots, x_n)$ ，此样本的发生概率与如下的联合密度函数成正比，即

$$p(X | \theta) \propto \prod_{i=1}^n p(x_i | \theta) \quad (3-1)$$

这个联合密度函数综合了总体信息和样本信息，常被称为似然函数，记为 $L(\theta)$ 。在有了样本观测值 $X = (x_1, \dots, x_n)$ 之后，总体和样本所含 θ 的信息都包含在似然函数 $L(\theta)$ 中。

由于 θ' 是设想出来的，它仍然是未知的，是按先验分布 $\pi(\theta)$ 产生的，要把先验信息进行综合，不能只考虑 θ' ，而应对 θ 的一切可能加以考虑，故要用 $\pi(\theta)$ 参与进一步的综合。这样样本 X 和参数 θ 的联合分布 $h(X, \theta) = p(X | \theta)\pi(\theta)$ 把参数的先验信息、总体信息和样本信息这三种可用信息都综合进去了。

我们的任务是对 θ 作统计推断，没有样本信息时，只能根据先验分布对 θ 作推断。在有样本观测值 $X = (x_1, \dots, x_n)$ 之后，应根据 $h(X, \theta)$ 对 θ 作推断。为此，需要把 $h(X, \theta)$ 作如下分解： $h(X, \theta) = \pi(\theta | X)m(X)$ 其中 $m(X)$ 是 X 的边际密度函数，表达式为：

$$m(X) = \int_{\Theta} h(X, \theta) d\theta = \int_{\Theta} p(X | \theta)\pi(\theta) d\theta \quad (3-2)$$

$m(X)$ 与 θ 无关，或者说 $m(X)$ 中不含任何 θ 的信息。因此能用来对 θ 作推断

的仅是条件分布 $\pi(\theta|X)$ ，依贝叶斯公式有：

$$\pi(\theta|X) = \frac{h(X, \theta)}{m(X)} = \frac{h(X, \theta)}{\int_{\Theta} h(X, \theta) d\theta} = \frac{p(X|\theta)\pi(\theta)}{\int_{\Theta} p(X|\theta)\pi(\theta) d\theta} \quad (3-3)$$

这就是贝叶斯公式的密度函数形式。给定样本 X ， θ 的条件分布称为 θ 的后验分布。集中了总体、样本和先验三种信息中有关 θ 的一切信息，又排除了一切与 θ 无关的信息后得到的结果。故基于后验分布 $\pi(\theta|X)$ 对 θ 的统计推断更有效，也是最合理的。

一般来说，先验分布 $\pi(\theta)$ 反映了人们在抽样前对 θ 的认识，后验分布 $\pi(\theta|X)$ 反映了人们在抽样后对 θ 的认识。它们之间的差异是由于样本 X 出现后人们对 θ 认识的一种调整。所以后验分布 $\pi(\theta|X)$ 可看作是人们用总体信息和样本信息（综合称为抽样信息）对先验分布 $\pi(\theta)$ 作调整的结果。

3.1.2 后验分布的计算

在给定样本分布 $p(X|\theta)$ 和先验分布 $\pi(\theta)$ 后，可用贝叶斯公式计算 θ 的后验分布，即：

$$\pi(\theta|X) = [p(X|\theta)\pi(\theta)] / m(X) \quad (3-4)$$

由于 $m(X)$ 不依赖于 θ ，在计算 θ 的后验分布中仅起到一个正则化因子的作用。假如把 $m(X)$ 省略，则贝叶斯公式改写成如下的等价形式：

$$\pi(\theta|X) \propto p(X|\theta)\pi(\theta) \quad (3-5)$$

符号“ \propto ”两边仅差一个不依赖于 θ 的常数因子。(3-5) 式右端虽不是正常的密度函数，但它是后验分布 $\pi(\theta|X)$ 的核，在需要时可以利用适当的方式计算出后验密度，特别是当看出 $p(\theta|X)\pi(\theta)$ 就是某常用分布的核，不用计算 $m(X)$ 就可以很快恢复所缺常数因子。

3.1.3 先验分布的确定

贝叶斯统计中要使用先验信息，而先验信息主要是指经验和历史资料。因此如何用人们的经验和过去的历史资料确定参数向量 θ 的先验分布是贝叶斯学派要研究的问题之一。

在经典统计中，概率是用非负性、正则性和可加性三条公理定义的。概率的确定方法主要是两种。一是古典方法(包括几何方法)，另一种是频率方法。实际中大量使用的是频率方法，所以经典统计的研究对象是大量重复的随机现象，如果不是大量重复的随机现象，我们就不能用频率的方法去确定其有关事件的频率。这无疑就把统计学的应用和研究领域缩小了。譬如，很多经济现象都是不能重复或不能大量重复的随机现象，在这类随机现象中要用频率方法去确定有关事件的概率常常是不可能的或很难实现的。

贝叶斯学派是完全同意概率的公理化定义的，但认为概率也可以用经验确定，这是与人们的实践活动一致的。这就可以使不能重复或不能大量重复的随机现象也可谈及概率。同时也使人们积累的丰富经验得以概括和应用。贝叶斯学派认为：一个事件的概率是人们根据经验对该事件发生可能性所给出的个人信念，这样给出的概率称为主观概率。

主观概率的确定不是随意的，而是要求当事人对所考察的事件有较透彻的了解和丰富的经验，甚至是这一行的专家。并能对周围信息和历史信息进行仔细分析，在这个基础上确定的主观概率就能符合实际。所以应把主观概率与主观臆造区别开来。

自主观概率提出以来，使用的人越来越多，特别在经济领域和决策分析中使用较为广泛，因为在那里遇到的随机现象大多是不能大量重复，无法用频率方法去确定事件概率的。在这个意义上看，主观概率至少是频率方法和古典方法的一种补充，有了主观概率至少使人们在频率观点不适用时也能谈论概率，使用概率与统计方法。主观概率并不反对用频率方法确定概率，但也要看到它的局限性。现实世界中能够在相同条件下进行大量重复的随机现象是不多的，无穷次重复更不可能，在这些场合下，主观概率就是一种补充。

当总体参数 θ 是离散时，即参数空间 Θ 只含有有限个或可数个点时，可对 Θ 中每个点确定一个主观概率。当总体参数 θ 是连续时，即参数空间 Θ 是实数轴或其上某个区间时，要构造一个先验密度 $\pi(\theta)$ 那就有些困难了。当 θ 的先验信息(经验和历史数据)足够多时，可使用以下三种方法：直方图法、选定先验密度函数形式再估计其超参数(根据先验信息选定 θ 的先验密度函数 $\pi(\theta)$ 的形式，例如选择其共轭先验分布；当先验分布中含有未知参数，即超参数时，给出超参数的

估计值,使 $\pi(\theta)$ 最接近先验信息)、定分度法和变分度法(即通过专家咨询获得各种主观概率,然后经过整理加工即可得到累积概率分布曲线)。先验分布的确定并不是本文的研究重点,所以不再做详细的介绍。

早在上个世纪 90 年代, Gelfand和Smith就实现了一种特殊的MCMC方法——Gibbs抽样^[119],该方法可以广泛应用到一大类贝叶斯问题中,也使得贝叶斯分析的应用得到迅速发展。MCMC方法植根于Metropolis和Ulam提出的Metropolis算法,该方法是当时物理学家试图计算复杂的积分时,把它们表达成某个分布的期望,然后通过从该分布抽取样本估计这个期望^[120,121]。Gibbs抽样最初起源于图像处理^[122],这个MCMC方法的利器直到最近才对统计领域产生影响。

MCMC 方法是一种特殊的蒙特卡罗积分模拟方法,它将随机过程中的马尔科夫过程引入到蒙特卡罗模拟中,实现动态模拟(即抽样分布随模拟的进行而改变)。本质上, MCMC 方法是使用马尔科夫链的蒙特卡罗模拟积分方法。

3.2 模型的离散化

目前可以采用两种方法对连续时间模型进行离散:一种是 Euler 方法,其实质就是对描述资产收益的随机微分方程进行 0.5 阶泰勒展开,由于简单性,在众多的文献中经常出现。第二种 Milstein 方法,其实质就是对随机微分方程进行 1.5 阶泰勒展开,显然, Milstein 方法要较 Euler 方法更为精确。本文,主要采用 Milstein 方法。

考虑如下的广义扩散过程:

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \quad (3-6)$$

其中 X_t 是状态变量, W_t 是标准布朗运动, $\mu(\cdot, \cdot), \sigma(\cdot, \cdot)$ 是已知函数, θ 是未知参数向量。接下来分别用 Euler 和 Milstein 方法对模型(3-6)进行离散化,从而获得近似的似然函数。

3.2.1 Euler 方法

Euler方法近似广义扩散过程(3-6)的表达式为:

$$X_{t+\Delta t} = X_t + \mu(X_t, \theta)\Delta t + \sigma(X_t, \theta)\Delta W_t \quad (3-7)$$

其中 $\Delta W_t = \varepsilon_t \sqrt{\Delta t}$, $\varepsilon_t \sim iidN(0,1)$, 假定对于所有参数的先验分布不变并且预先给定, 同时给定状态变量 $X = \{x_t : t = 0, 1, \dots, n\}$ 的 $n+1$ 个观测值, 则关于参数向量 θ 的对数似然函数是:

$$\log p_E(X|\theta) = -\frac{1}{2} \sum_{i=1}^n \log(\sigma(x_i, \theta)^2 \Delta t) - \frac{1}{2} \sum_{i=1}^n \frac{(x_i - x_{i-1} - \mu(x_i, \theta) \Delta t)^2}{\sigma(x_i, \theta)^2 \Delta t} \quad (3-8)$$

其中 $p_E(X|\theta)$ 是基于 Euler 方法的似然函数。

3.2.2 Milstein 方法

对于模型 (3-6) 进行 1.5 阶泰勒展开, 有:

$$X_{t+\Delta t} = X_t + \mu(X_t, \theta) \Delta t + \sigma(X_t, \theta) \Delta W_t + \frac{1}{2} \sigma(X_t, \theta) \frac{\partial \sigma(X_t, \theta)}{\partial X_t} [(\Delta W_t)^2 - \Delta t] \quad (3-9)$$

(3-9) 式也可以写成

$$X_{t+\Delta t} - X_t - \mu(X_t, \theta) \Delta t + g(X_t, \theta) \Delta t = \sigma(X_t, \theta) \sqrt{\Delta t} \varepsilon + g(X_t, \theta) \Delta t \varepsilon^2 \quad (3-10)$$

其中

$$g(X_t, \theta) = \frac{1}{2} \sigma(X_t, \theta) (\partial \sigma(X_t, \theta) / \partial X_t) \quad (3-11)$$

并令

$$a = \sigma(X_t, \theta) \sqrt{\Delta t}, \quad b = g(X_t, \theta) \Delta t \quad (3-12)$$

这样将式 (3-11) 和 (3-12) 代入 (3-10) 式, 有:

$$Y = a\varepsilon + b\varepsilon^2 = b \left[\left(\varepsilon + \frac{a}{2b} \right)^2 - \frac{a^2}{4b^2} \right] \quad (3-13)$$

其中, $Y = X_{t+\Delta t} - X_t - \mu(X_t, \theta) \Delta t + g(X_t, \theta) \Delta t$ 。

令 $Z = \left(\varepsilon + \frac{a}{2b} \right)^2$, 这样 Z 是服从自由度为 1 的非中心 χ^2 分布, 非中心参数

$\lambda = a^2/4b^2$, Elerian (1998) [123] 证明 Z 的密度是:

$$f(z) = \frac{1}{2} \exp \left\{ -\frac{\lambda+z}{2} \right\} \left(\frac{z}{\lambda} \right)^{-1/4} I_{-1/2}(\sqrt{\lambda z}) \quad (3-14)$$

其中

$$I_{-1/2}(\omega) = \sqrt{2} \sum_{j=0}^{\infty} \frac{(2/\omega)^j}{j! \Gamma(j+1/2)} = \sqrt{\frac{2}{\pi\omega}} \cosh(\omega) \quad (3-15)$$

这里 $\cosh(\omega) = (1/2)\{\exp(\omega) + \exp(-\omega)\}$ 是抛物线余弦函数。这样， Y 的密度是：

$$f^*(y) = \frac{1}{b} f\left(\frac{y}{b} + \frac{a^2}{4b^2}\right) \quad (3-16)$$

因而参数向量 θ 基于观测值集合 $X = \{x_t : t = 0, 1, \dots, n\}$ 的对数似然函数为：

$$\log p_M(X|\theta) = \sum_{i=1}^n \left[\log \left\{ f\left(\frac{y_i}{b} + \frac{a^2}{4b^2}\right) \right\} - \log(b) \right] \quad (3-17)$$

其中， $y_t = x_t - x_{t-1} - \mu(x_{t-1}, \theta)\Delta t + g(x_{t-1}, \theta)\Delta t - \kappa q_{t+M}$ 。

按照贝叶斯原理，参数的后验分布是：

$$\pi(\theta|X) = cp(X|\theta)\pi(\theta) \quad (3-18)$$

其中 c 是不依赖于参数向量 θ 的正则化常数， $p(X|\theta)$ 是参数向量 θ 基于观测值集合 $X = \{x_t : t = 0, 1, \dots, n\}$ 的似然函数， $\pi(\theta)$ 是参数向量 θ 的先验密度。

3.3 MCMC 算法

MCMC 算法分为两种，一种是 Gibbs 算法，另一种是 Metropolis-Hasting 算法。在实现一个算法时，即可以用 Gibbs 算法，又可以同时使用 Metropolis-Hasting 算法。

3.3.1 Gibbs 算法

Gibbs 取样最早是 Geman 在图像处理的文献中提出的，是 Metropolis-Hastings 取样的一个特殊情况，其中随机值总是被接受（也就是说备选值的接受概率为 1）。Gibbs 取样器的关键是只考虑一元条件分布，也就是说除了一个变量外，所有的随机变量都赋予固定值的分布。这种条件分布比复杂的联合分布容易模拟，且密度通常只能具有简单的形式（通常是正态， t 分布，或其它常见的先验密度）。因此，该方法是从 n 个一元条件分布依次模拟 n 个随机变量，而不是直接使用完全联合密度（full joint density）来产生一个单独的 n 维向量。

首先考虑一个二元随机变量 (x, y) ，假定希望计算它的两个边际密度 $p(x)$ 和 $p(y)$ 。因为计算条件密度 $p(x|y)$ 或 $p(y|x)$ 要比通过从联合密度 $p(x, y)$ 积分获得边际密度（如 $p(x) = \int p(x, y) dy$ ）容易的多，所以使用 Gibbs 取样器。给定 y 的初始值 y_0 ，通过从条件密度 $p(x|y=y_0)$ 产生一个随机变量获得 x_0 。然后使用 x_0 从基于 x_0 的条件密度 $p(y|x=x_0)$ 抽样产生一个新值 y_1 。取样器按下述过程进行取样：

$$\begin{aligned} x_i &\rightarrow p(x|y=y_i) \\ y_i &\rightarrow p(y|x=x_{i-1}) \end{aligned} \quad (3-19)$$

重复该过程 k 次，产生一个长度为 k 的 Gibbs 序列，其中点 (x_j, y_j) ， $1 \leq j \leq m < k$ 的一个子集作为我们从满联合密度的模拟抽样。通常所有一元分布的一次迭代称为取样器的一次扫描。为了获得所有的 m 个理想取样点（这里取样器上的每一个“点”都是一个二参数向量），链应在足够的燃烧期（burn-in period）之后抽取，以去除初始取样值的影响。这样 Gibbs 序列会收敛到一个独立于起始值的平稳（均衡）分布，这个平稳分布就是我们试图模拟的目标分布^[124]。

当不止两个变量时，该取样器可以扩展。假如联合分布包含 n 个变量 (x_1, \dots, x_n) ，则 Gibbs 抽样按下面的步骤进行：

- Step1: 设定初始状态点 $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ ；
- Step2: 从条件分布密度 $p(x_1 | x_2^{(0)}, \dots, x_n^{(0)})$ 抽取 $x_1^{(1)}$ ；
- Step3: 用 $x_1^{(1)}$ 代替 $x_1^{(0)}$ ，从 $p(x_2 | x_1^{(1)}, \dots, x_n^{(0)})$ 抽取 $x_2^{(1)}$ ；
- Step4: 一直进行，直到从 $p(x_n | x_1^{(1)}, x_2^{(1)}, \dots, x_{n-1}^{(1)})$ 抽取 $x_n^{(1)}$

上述过程完成了第一轮迭代取样。第二轮以 $(x_1^{(1)}, \dots, x_n^{(1)})$ 作为初始值，重复上述迭代取样过程。假如进行了 t 轮迭代，则得到长度为 t 的 Gibbs 序列 $X^{(t)} = (x_1^{(t)}, \dots, x_n^{(t)})$ 。在足够的燃烧期后，Gibbs 序列才能收敛到一个独立于初始值的平稳分布，即我们试图模拟的目标分布。故应该舍弃前 l 轮的迭代值，则去掉燃烧期后的序列 $X^{(l+1)}, \dots, X^{(t)}$ 为马尔可夫链的实现值。

3.3.2 Metropolis-Hasting 算法

应用蒙特卡罗积分的一个问题在于从某个复杂的概率分布密度 $p(\theta)$ 获得样本。MCMC方法最初就是试图解决这个问题，特别是可以追溯到数学物理学家用随机取样试图对非常复杂的函数进行积分，并由此产生了Metropolis-Hastings算法^[125]，关于MH算法的具体实现将在3.4节中详细说明。但成功实现Metropolis-Hastings或其它MCMC方法的关键是产生的马尔科夫链接近稳态的运行数（步数），即燃烧期长度。一般来说燃烧取样时迭代得到的链的前5000到10000元素要舍去，另外也可以用一些收敛检验方法来评估链是否已达到稳态，检验方法具体见文献[126,127,128]。

虽然Gibbs算法可以直接通过Winbugs软件实现，但是Gibbs只能适合后验密度为常见的密度形式，如正态密度等；而对于特殊情况下的密度形式，Gibbs算法就无能为力。反之，MH算法的适用范围就更广，Gibbs算法只是MH算法的一种特例。MH算法与Gibbs算法更相比具有以下几方面的优势：首先在多变量情况下，Gibbs算法只能从各个变量的边际密度中依次对各个变量取样，而MH算法可以直接从各个变量的联合密度中对所有变量同时取样。其次，MH算法可以从复杂的分布密度中取样，而Gibbs算法则不能。虽然MH算法具有这些优点，但是目前尚未有现成的计算软件来实现，只能通过C++语言编程来实现。

3.4 包含隐含变量的MCMC算法

对于 d 维向量扩散过程

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \quad (3-20)$$

如果 X_t 包含不可观测变量，即 $X_t = (Y_t, Z_t)'$ ， Y_t 和 Z_t 分别为 d_1 和 d_2 维，显然 $d = d_1 + d_2$ ；其中 Y_t 为可观测变量，而 Z_t 为不可观测变量。上述扩散过程依Milstein方法的离散形式为：

$$\Delta X_t = \mu(X_t; \theta)\Delta t + \sigma(X_t; \theta)\Delta W_t + \frac{1}{2}\sigma(X_t; \theta)\frac{\partial\sigma(X_t; \theta)}{\partial X_t}\left[(\Delta W_t)^2 - \Delta t\right] \quad (3-21)$$

在这里假设 $\Delta t = 1/m$ ，另外将 $[0, T]$ 分成 $n = mT$ 个相等间隔点，

$0 = t_1 < \dots < t_{n-1} < t_n = T$ 。对于向量 Y_i 来说，需要在每两个观测值之间模拟 $m-1$ 个缺失数据，即当 Y_i 是观测值时， i 是 m 的倍数。也就是说要在两个观测值向量 Y_{i_m} 与 $Y_{i_{m(n)}}$ 之间加入 $d_1 \times (m-1)$ 的模拟数据矩阵，从而形成一个 $d_1 \times n$ 的数据矩阵，令该矩阵为 $\hat{Y}_{d_1 \times n}$ ，其中 \hat{Y}_i 为 $\hat{Y}_{d_1 \times n}$ 的第 i 列。对于不可观测变量向量 Z_i 来说，需要模拟 $d_2 \times n$ 个缺失数据，该数据矩阵为 $\hat{Z}_{d_2 \times n}$ ，其中 \hat{Z}_i 为 $\hat{Z}_{d_2 \times n}$ 的第 i 列。这样由观测数据和模拟数据所形成的矩阵为：

$$\hat{X}_i = \begin{bmatrix} \hat{Y}_i \\ \hat{Z}_i \end{bmatrix} = \begin{bmatrix} \hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n \\ \hat{Z}_1, \hat{Z}_2, \dots, \hat{Z}_n \end{bmatrix} = \begin{bmatrix} Y_{1,1}, Y_{1,2}, \dots, Y_{1,m}, Y_{1,m+1}, \dots, Y_{1,n} \\ Y_{2,1}, Y_{2,2}, \dots, Y_{2,m}, Y_{2,m+1}, \dots, Y_{2,n} \\ \vdots \\ Y_{d_1,1}, Y_{d_1,2}, \dots, Y_{d_1,m}, Y_{d_1,m+1}, \dots, Y_{d_1,n} \\ Z_{1,1}, Z_{1,2}, \dots, Z_{1,m}, Z_{1,m+1}, \dots, Z_{1,n} \\ \vdots \\ Z_{d_2,1}, Z_{d_2,2}, \dots, Z_{d_2,m}, Z_{d_2,m+1}, \dots, Z_{d_2,n} \end{bmatrix} \quad (3-22)$$

在用 MCMC 方法估计该扩散模型时，首先要解决的就是缺失数据 (missing data) 的模拟，之后才是参数的模拟。

假定基于观测值的联合后验密度为：

$$\pi(\hat{X}_i, \theta) \propto p_M(\hat{X}_i | \theta) p(\theta) \quad (3-23)$$

其中 $p(\theta)$ 是参数的先验密度， $p_M(\hat{X}_i | \theta)$ 为基于 Milstein 离散方法所获得的参数似然函数。

3.4.1 缺失数据的条件后验及抽样

由于后验密度的高维性，不太可能一次对所有缺失数据取样。但是使用 Gibbs 取样器^[114]一次获得 $\hat{X}_{d_1 \times n}$ 的一列，从而我们需要知道第 i 列的条件后验，这个后验密度存在下面的比例关系：

$$\pi(\hat{X}_i | \hat{X}_v, \theta) \propto \pi(\hat{X}_i | \hat{X}_{i-1}, \hat{X}_{i+1}, \theta) \quad (3-24)$$

其中 \hat{X}_v 是不含第 i 列的其他数据。其中 $\pi(\hat{X}_i | \hat{X}_{i-1}, \hat{X}_{i+1}, \theta)$ 满足下式：

$$\pi(\hat{X}_i | \hat{X}_{i-1}, \hat{X}_{i+1}, \theta) = p(\hat{X}_i | \hat{X}_{i-1}, \theta) p(\hat{X}_{i+1} | \hat{X}_i, \theta) \quad (3-25)$$

其中 $p(\bullet)$ 为状态变量 X_t 的转移密度，依据前面介绍的 Milstein 方法， $p(\bullet)$ 可由式 (3-16) 进行计算获得。

因此，在 Gibbs 取样的第 h 次循环中，

$$\hat{X}_i^{(h)} \sim \pi\left(\hat{X}_i \mid \hat{X}_{i-1}^{(h)}, \hat{X}_{i+1}^{(h)}, \theta\right) \quad (3-26)$$

显然，当 i 是 m 的倍数时， \hat{Y}_i 为观测数据因而无需模拟，所以只需要模拟 \hat{Z}_i 。

3.4.2 参数的后验密度及取样

由上一小节所获得缺失数据与观测数据，按照 3.2.2 中介绍的 Milstein 方法，很容易获得参数的联合条件后验分布，从而采用 MCMC 的随机游走 MH 算法来模拟参数。

MH 算法思想就是：首先从给定的建议条件密度 $q(\theta'|\theta)$ 随机生成备选 θ' ，这个建议条件密度必须满足一定的性质 (Chib 和 Greenberg 1995)，关于建议密度的具体选择将在下面详细探讨。之后以概率 $T(\theta, \theta')$ 接受 θ' ，其中接受概率 $T(\theta, \theta')$ 如下：

$$T(\theta, \theta') = \min \left\{ 1, \frac{\pi(\theta'|\hat{X})q(\theta|\theta')}{\pi(\theta|\hat{X})q(\theta'|\theta)} \right\} \quad (3-27)$$

其中 $\pi(\theta|\hat{X})$ 是参数的联合条件后验密度。

在使用 MCMC 方法估计时需要注意的是：首先，接受概率与式 (3-18) 中标准化常数“ c ”无关，所以可以用 $p_M(\hat{X}|\theta)\pi(\theta)$ 代替 $\pi(\theta|\hat{X})$ 来计算接受概率，从而式 (3-27) 可以转换成如下形式：

$$T(\theta, \theta') = \min \left\{ 1, \frac{p_M(\hat{X}|\theta')\pi(\theta')q(\theta|\theta')}{p_M(\hat{X}|\theta)\pi(\theta)q(\theta'|\theta)} \right\} \quad (3-28)$$

其次，就是建议密度的选择问题，Chib 和 Greenberg (1995) [129] 总结了一下几种建议密度的形式，具体如下：

第一类是 Metropolis (1953) 提出的，其令 $q(\theta'|\theta) = q_1(\theta' - \theta)$ ，其中 $q_1(\bullet)$ 是一个多变量密度，形式上一般为 t 分布、正态分布或均匀分布。备选样本值 θ' 按

照下式更新：

$$\theta' = \theta + \varepsilon \quad (3-29)$$

其中 ε 是服从 $q_1(\bullet)$ 分布的随机变量。从式 (3-29) 可以看出，备选值 θ' 是当前值 θ 与一个噪声 ε 的和，所以这种情况又称之为随机游走形式。当 $q_1(\bullet)$ 是对称时，即 $q_1(\varepsilon) = q_1(-\varepsilon)$ ，这时接受概率可以简化为：

$$T(\theta, \theta') = \min \left\{ 1, \frac{p_M(\hat{X}|\theta')\pi(\theta')}{p_M(\hat{X}|\theta)\pi(\theta)} \right\} \quad (3-30)$$

第二类是由 Hasting (1970) 提出的，与随机游走形式的建议密度不同，备选值 θ' 独立于当前值 θ ，即 $q(\theta'|\theta) = q_2(\theta')$ ；但与随机游走形式相同的是建议密度函数的形式与第一种情况一样，一般取一个多变量正态分布或者 t 分布。

第三类是由 Chib 和 Greenberg (1994)^[130] 提出的，通过挖掘 $\pi(\theta|\hat{X})$ 的已知形式来选择建议密度。例如，如果 $\pi(\theta|\hat{X})$ 可以写成 $\pi(\theta|\hat{X}) \propto \psi(\theta|\hat{X})h(\theta|\hat{X})$ ，其中 $h(\theta|\hat{X})$ 可以用来直接取样，而 $\psi(\theta|\hat{X})$ 是一致有界的，这样就可以直接将 $h(\theta|\hat{X})$ 作为建议密度函数。在这种情况下，接受概率的计算仅仅与 $\psi(\theta|\hat{X})$ 有关，从而接受概率的计算公式如下：

$$T(\theta, \theta') = \min \left\{ \frac{\psi(\theta')}{\psi(\theta)}, 1 \right\} \quad (3-31)$$

第四种是由 Tierney (1994) 提出的，建议密度 $q(\theta'|\theta)$ 用一个 1 阶向量自回归过程来表示，令 $\theta' = a + B(\theta - a) + z$ 来产生参数向量 θ 的备选值 θ' ，其中 a 是向量， B 是一方阵， z 的密度为 $q(z)$ 。这样建议密度就满足：

$$q(\theta'|\theta) = q(\theta' - a - B(\theta - a)) \quad (3-32)$$

上面给出了几种可选择的建议密度形式，但是在一般情况下都是采用最简单的随机游走形式，即 $q(\theta'|\theta)$ 是关于 $|\theta' - \theta|$ 的密度函数，从而可以简化接受概率 $T(\theta, \theta')$ 的计算，同时又可以很容易生成备选值 θ' 。

另外，建议密度的选择将会影响到基于 MH 算法的效率，建议密度的分散程

度将会从两个方面来影响样本链的性质：一是接受备选值的概率，另外一个马尔可夫链所覆盖的取样空间。如果建议密度函数的分散程度 (Spread or Scale of the Proposal Density) 较大，备选值 θ' 就会大大地偏离现在值 θ ，因而就可能具有很低的接受概率。但是如果分散程度较小，低密度空间将不在取样之中；而这两种情况都说明样本值之间具有很大得自相关性。为了保证基于 MH 算法的 MCMC 方法的取样效率，通过调整建议密度函数的分散程度来获得适当的接受比率。接受比率多大才是合适的呢？下面就简单说明一下当建议密度采用前面两类情况时接受概率大约取值范围。

当建议密度采用随机游走情形时，Roberts、Gelman和Gilks (1994)^[131]认为如果参数的后验密度与建议密度都是正态的，则应该调整建议密度的分散程度 (Spread or Scale) 以至接受概率在一维取样中近似等于 0.45，在多维变量取样中近似于 0.23。Muller (1993)^[132]认为在随机游走链中，接受概率应该在 0.5 左右。

当建议密度采用上述 Hasting (1970) 的独立取样情况时，Chib 与 Geweke 认为重要的是要确保建议密度的尾部分布厚于目标密度的尾部分布^[130]。

3.4.3 MCMC 算法

通过上面的分析，我们分别获得了缺失数据和参数的抽样算法，具体步骤如下：

Step1: 令 $j=1$ ，并对所有参数向量 θ 和隐含变量 X_i 赋初值；

Step2: 对于所有 $i=0,1,\dots,n$ ，采用 Gibbs 算法直接从缺失数据的条件后验密度 (3-26) 中模拟缺失数据；

Step3: 采用随机游走 MH 算法，从参数的联合后验密度中模拟 $\theta^{(j)}$ ，具体步骤如下：

Step 3.1: 给定现在状态 $\theta^{(j-1)}$ ，按照式 (3-29) 生成备选值 θ' ；

Step 3.2: 依据式 (3-27) 计算接受概率 $T(\theta^{(j-1)}, \theta')$ ；

Step 3.3: 以 $T(\theta^{(j-1)}, \theta')$ 的概率接受备选值，即 $\theta^{(j)} = \theta'$ ；以 $1-T(\theta^{(j-1)}, \theta')$ 的概率拒绝备选值，即 $\theta^{(j)} = \theta^{(j-1)}$ ；

Step 4: 令 $j = j + 1$, 若 $j < h$, 返回第 2 步, 其中 h 为取样的总次数; 反之结束算法。

按照上面的步骤可以获得 h 个参数样本向量 $\theta = \{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(h)}\}$, 由不同的 $\theta^{(i)}$ 出发, 由马尔科夫链经过一段时间的迭代后, 可以认为 $\theta = \{\theta^{(h_0)}, \theta^{(h_0+1)}, \dots, \theta^{(h)}\}$ 的边际分布为平稳分布, 此时它收敛。而在收敛出现前的 h_0 次迭代中, 各状态的边际分布还不能认为是 $\pi(\theta|X)$, 因此在实际应用中, 要剔除最初的 h_0 次抽样, 仅仅记录已经收敛后验分布的后 $(h - h_0)$ 次抽样结果, 在本章后面的实证分析中一般取 $h = 60000$ 和 $h_0 = 10000$ 。

另外需要说明的是, 在上面的步骤中, 当模型中不含隐含变量时, Step2 就应该省略。

3.5 参数抽样结果的收敛性分析

在 MCMC 算法中, 模拟样本的收敛性将影响到估计的效果, 下面说明通过计算模拟无效因子来判断模拟的收敛效果, 令模拟参数样本为 $\{\theta^{(i)} : i = 1, 2, \dots, N\}$, 则输出的遍历均值为:

$$\bar{f}_N = \frac{1}{N} \sum_{i=1}^N f(\theta^{(i)}) \quad (3-33)$$

其中函数 $f(\theta^{(i)})$ 是一个实值函数, 一般情况下取

$$f(\theta^{(i)}) = \theta^{(i)} \quad (3-34)$$

依据遍历均值的中心极限定理有:

$$\sqrt{N}(\bar{f}_N - E_\pi(f(\theta))) \xrightarrow{D} N(0, \sigma_f^2) \quad (3-35)$$

其中 $E_\pi(f(\theta))$ 是关于后验密度 $\pi(\theta|x)$ 的期望算子, \xrightarrow{D} 表示的是依分布收敛。为了讨论 $E_\pi(f(\theta))$ 作为遍历均值估计的准确性, 必须要估计方差 σ_f^2 。而估计 σ_f^2 最常用的方法是由 Roberts (1996)^[133] 提出的, 用块均值的方差来作为 σ_f^2 的估计值。

为了使用块均值估计 σ_f^2 , 令 $N = m \times n$, 这里 m 是块的个数, n 是每块中取

样样本的个数，在这里 n 的值要足够大以致块均值 y_k ($k=1,2,\dots,m$) 能够服从独立同分布， y_k 按下式计算：

$$y_k = \frac{1}{n} \sum_{t=(k-1)n+1}^{kn} f(\theta^{(t)}) \quad (3-36)$$

对于 $k=1,2,\dots,m$ ， y_k 收敛于独立同分布 $N\left(E_{\pi}(f(\theta)), \frac{\sigma_f^2}{n}\right)$ ，这样 σ_f^2 的估计值就是：

$$\hat{\sigma}_f^2 = \frac{n}{m-1} \sum_{k=1}^m (y_k - \bar{f}_N)^2 \quad (3-37)$$

同时， \bar{f}_N 的标准误差可以用 $\sqrt{\frac{\hat{\sigma}_f^2}{N}}$ 来估计，这个标准误差也称之为参数估计值的蒙特卡罗标准误差 (Monte Carlo Standard Error, MCSE)。另外还要计算所有取样本的标准方差 $\hat{\sigma}_f$ ：

$$\hat{\sigma}_f = \sqrt{\frac{1}{N-1} \sum_{t=1}^N [f(\theta^{(t)}) - \bar{f}_N]^2} \quad (3-38)$$

Kim *et al* (1998) [134] 提出了用模拟无效因子 (Simulation Inefficiency Factor, SIF) 来评价抽样样本的模拟效果，Meyer 和 Yu (2000) [135] 给出了 SIF 的计算公式：

$$SIF = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_f^2} \quad (3-39)$$

目前 SIF 没有理论上的标准值来判断抽样的效果，SIF 越小，样本的收敛性越好；从实证角度来说，当 SIF 不超过 100 时，抽样样本就具有非常好的收敛效果。

3.6 双指数跳跃扩散模型的 MCMC 估计

3.6.1 双指数跳跃扩散模型及其特征

3.6.1.4 双指数跳跃扩散模型

Kou (2002) 为了能够同时体现资产收益分布的尖峰厚尾和有偏性特征，提出了双指数跳跃扩散模型，该模型如下：

假定资产价格为 $S(t)$ ，是下面随机微分方程的解

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) + d\left(\sum_{i=1}^{N(t)} (V_i - 1)\right) \quad (3-40)$$

其中 $W(t)$ 是标准维纳过程， $N(t)$ 是强度为 λ 的泊松过程，且两者相互独立； $\{V_i\}$ 是独立同分布非负随机变量序列，令 $X_i = \log(V_i)$ ， X_i 服从双指数分布，其密度为：

$$f_{X_i}(x) = \frac{1}{2\eta} e^{-\kappa - x/\eta}, 0 < \eta < 1 \quad (3-41)$$

其中， κ 是 X_i 的均值，同时

$$P(x - \kappa = \xi) = P(x - \kappa = -\xi) = 0.5 \quad (3-42)$$

其中 ξ 是双指数变量， ξ 的均值为 η 、方差为 η^2 。

该模型认为跳跃是市场对外部信息的反映，好消息和坏消息以泊松过程到达，资产价格的跳跃也以泊松过程发生。另外为了简单，同时为了取得各种期权定价问题的解析解，假定漂移项 μ 和波动项 σ 为常数，但是关于波动项 σ 为常数的假设可以推广到随机波动情况。

3.6.1.5 双指数跳跃扩散模型的特征

为了更好的体现双指数跳跃扩散模型的特征，对随机微分方程 (3-40) 采用 Euler 方法离散，可得：

$$\begin{aligned} \frac{\Delta S(t)}{S(t)} &\approx \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma(W(t+\Delta t) - W(t)) + \sum_{i=N(t)+1}^{N(t+\Delta t)} X_i \\ &= \mu\Delta t + \sigma\varepsilon_i\sqrt{\Delta t} + \sum_{i=N(t)+1}^{N(t+\Delta t)} X_i \end{aligned} \quad (3-43)$$

其中 ε_i 是服从标准正态分布的随机过程，泊松过程 $N(t)$ 在区间 $(t, t + \Delta t]$ 有一次跳跃的概率是 $\lambda\Delta t$ ，超过一次的概率是 $o(\Delta t)$ ，因此，只要 Δt 足够小，可以忽略多次跳跃，并有

$$N(t + \Delta t) = \begin{cases} N(t) + 1 & P = \lambda\Delta t \\ N(t) & P = 1 - \lambda\Delta t \end{cases} \quad (3-44)$$

所以有

$$\sum_{i=N(t)+1}^{N(t+\Delta t)} X_i \approx \begin{cases} X_i, & P = \lambda\Delta t \\ 0, & P = 1 - \lambda\Delta t \end{cases} \quad (3-45)$$

总之，当 Δt 很小时，收益可以近似为：

$$\frac{\Delta S(t)}{S(t)} \approx \mu\Delta t + \sigma\varepsilon_t \sqrt{\Delta t} + B_i \cdot X_i \quad (3-46)$$

其中 B_i 服从贝努利分布的随机过程，且 $P(B_i = 1) = \lambda\Delta t, P(B_i = 0) = 1 - \lambda\Delta t$ ，当 $B_i = 1$ 时，说明在时间 $(t, t + \Delta t]$ 内发生了跳跃，反之则没有发生跳跃。 X_i 是收益

的跳跃规模，服从双指数分布。这样收益 $\frac{\Delta S(t)}{S(t)}$ 就近似于正态分布，其均值和方

差为：

$$E\left(\frac{\Delta S(t)}{S(t)}\right) = \mu\Delta t + \kappa\lambda\Delta t + o(\Delta t) \approx \mu\Delta t + \kappa\lambda\Delta t \quad (3-47)$$

$$Var\left(\frac{\Delta S(t)}{S(t)}\right) = \sigma^2\Delta t + 2\eta^2\lambda\Delta t(1 - \lambda\Delta t) + o(\Delta t) \approx \sigma^2\Delta t + 2\eta^2\lambda\Delta t(1 - \lambda\Delta t) \quad (3-48)$$

可以看到这种资产收益分布的一个重要特征就是：相对于正态密度而言，这种分布在均值附近有更高的峰、有两个厚尾，也就是尖峰特征。另外一个需要提到的就是，如果跳跃规模均值 κ 不为 0，则密度是非对称的。实际上，如果 $\kappa < 0$ ，向左偏；反之向右偏。

另外，还发现，资产收益分布的高峰和厚尾特征随着 $|\kappa|, \lambda, \eta$ 的增大而变得更加显著；虽然跳跃规模的对数为正态分布可以体现收益分布的厚尾特征，但是它不可能同时有高峰和厚尾的特征。

3.6.2 双指数跳跃扩散模型的 MCMC 估计

3.6.2.6 跳跃扩散模型的似然函数

由 (3-46) 式可知：

$$\frac{\Delta S(t) - \mu\Delta t - B_i \cdot X_i}{\sigma\sqrt{\Delta t}} \approx \varepsilon_t \sim N(0,1) \quad (3-49)$$

模型参数向量 $\theta = (\mu, \sigma, \kappa, \eta)$ 基于观测值集合的对数似然函数是：

$$\log p(\theta|S) \approx -\frac{n}{2} \log(\sigma^2 \Delta t) - \frac{1}{2} \sum_{i=1}^n \frac{\left(\frac{S_i - S_{i-1} - \mu \Delta t - B_i X_i}{S_i} \right)^2}{\sigma^2 \Delta t} \quad (3-50)$$

3.6.2.7 跳跃扩散模型的 MCMC 方法

马尔可夫链蒙特卡罗方法在众多的统计应用中被证明是一种比传统独立取样更有效的取样方法。在计量经济学和金融学中，MCMC 方法也被广泛使用。依据贝叶斯原理，参数向量 θ 基于数据 S 的条件分布可以通过后验密度 $p(S|\theta)$ 获得，即：

$$\pi(\theta|S) \propto p(S|\theta)\pi(\theta) \quad (3-51)$$

其中 $p(S|\theta)$ 是似然函数， $\pi(\theta)$ 是参数 θ 的先验分布。

在抛物线跳跃扩散模型中，由于引入了跳跃因子，而这些跳跃因子都是隐含变量，从而参数的估计将更加复杂。要解决参数的估计问题，首先要解决隐含变量的估计问题；Eraker (2001) 对含有隐含变量的扩散模型估计提出了相应的解决方案，即在用 MCMC 的 MH 算法从参数的后验分布中抽取每一组参数样本的同时，运用 MCMC 的 Gibbs 算法从隐含变量的后验分布中抽取一组隐含变量。下面是参数向量 θ ，跳跃时间 $B=(B_t, t=1, 2, \dots, n)$ 和跳跃规模 $X=(X_t, t=1, 2, \dots, n)$ 的后验分布：

$$\text{扩散和跳跃参数 } \theta: \pi(\theta|S, B, X) \propto p(S|\theta, B, X)\pi(\theta)$$

$$\text{跳跃时间 } B_t (t=1, 2, \dots, n): \pi(B_t|S_t, S_{t-1}, \theta, X_t) \propto p(S_t|\theta, S_{t-1}, B_t, X_t)\pi(B_t)$$

$$\text{跳跃规模 } X_t (t=1, 2, \dots, n): \pi(X_t|S_t, S_{t-1}, \theta, B_t=1) \propto p(S_t|\theta, S_{t-1}, B_t=1, X_t)\pi(X_t)$$

依据 MCMC 方法从双指数跳跃扩散模型的参数 $\theta=(\mu, \sigma^2, \kappa, \eta^2, \lambda)$ 、跳跃时间 B 和跳跃规模 X 的后验分布中抽样按照 3.4.3 中给出的步骤进行。

在该算法中，要解决的是建议密度 $q(\bullet|\theta)$ 的选择，具体的讨论可以参见 Chib 和 Greenberg (1995)、Gilks et al (1996) [136]。在本文中使用的是随机游走的 MH 算法，即建议密度为 $[-0.5, 0.5]$ 上的均匀分布，参数向量 θ 按照下面的方法生成：

$$\theta' = \theta + \tau \varepsilon \tag{3-52}$$

其中 ε 是在区间上 $[-0.5, 0.5]$ 均匀分布的误差项， τ 是调和参数，通过 τ 的值可以调整误差项的方差，一般来说 τ 值应该使接受概率在 20% 到 30% 之间^[129]。

3.6.3 双指数跳跃扩散模型的实证检验

3.6.3.1 参数的估计结果

数据分别采用上海和深圳股市的周综合指数，即从 1992-5-1 日到 2004-4-30 日共 12 年的上海股市周综合指数和深圳股市周综合指数。

参数先验分布利用文献[65]中介绍的先验分布，各个参数的先验分别假定为： $\mu \sim N(0, 10)$ ， $\sigma^2 \sim IG(5, 0.05)$ ， $\kappa \sim N(0, 10)$ ， $\eta^2 \sim IG(5, 0.05)$ ， $\lambda \sim \beta(2, 40)$ 。实证研究表明，参数的估计值不会随着先验分布的改变而发生明显的变化，这也说明参数的后验分布不会随着参数的先验分布发生显著变化^{[54][117]}。

运用 3.4.3 节介绍的 MCMC 方法，采用上面的数据和参数的先验分布形式，用 MH 算法迭代 60000 次，舍弃掉前面的 10000 个抽样，记录模拟值的后 50000 个抽样样本，所得的参数估计结果如表 3.1 所示，隐含变量 $X = (X_t, t = 1, 2, \dots, n)$ 的结果如图 3.1 和图 3.3。

表 3.1：基于 MCMC 方法的双指数跳跃扩散模型的参数估计结果

数据	参数	均值	置信区间 (95%)	SD	MCSE	SIF	AC
上海股市综合指数	μ	-0.074803	-0.305240, 0.155635	0.117570	0.00117503	5	0.27
	σ^2	0.184264	0.117952, 0.250577	0.033833	0.00064575	18	0.23
	κ	0.027500	-0.069735, 0.061973	0.017589	0.00016755	6	0.26
	η^2	0.004608	0.000514, 0.009165	0.002325	0.00002407	6	0.26
	λ	0.052951	-0.008371, 0.114250	0.031698	0.00038350	8	0.28
深圳股市综合指数	μ	-0.043931	-0.288511, 0.200649	0.124768	0.00144215	7	0.24
	σ^2	0.153357	0.099634, 0.207081	0.027410	0.00055144	20	0.23
	κ	0.016500	-0.014938, 0.047938	0.016040	0.00014850	5	0.24
	η^2	0.004122	0.000536, 0.007707	0.001829	0.00002678	8	0.25
	λ	0.050534	-0.008694, 0.109762	0.030219	0.00038747	8	0.28

在表 3.1 中，SD 是参数抽样样本的标准离差；MCSE 是蒙特卡罗模拟标准误差，通过公式 (3-33) - (3-37) 来计算；SIF 是模拟无效因子，SIF 是用来检验模拟样本的收敛效果的，通过公式 (3-38) - (3-39) 来计算；在计算 MCSE、SIF 时，令 $f(\theta^{(l)}) = \theta^{(l)}$ 、 $m = 50$ 、 $n = 1000$ 。AC 为在 MCMC 算法中的样本接受概率，即备选值接受的总次数与模拟总次数的比率。

从表 3.1 中，跳跃强度为 0.05 以上，也就意味着沪深股市每年发生跳跃的次数为 2 到 3 次（上海综合指数的跳跃规模见图 3.1）。从跳跃规模的均值来看，沪深股市的收益分布都存在有偏性，沪深两市的 κ 都大于 0，所以收益分布右偏。

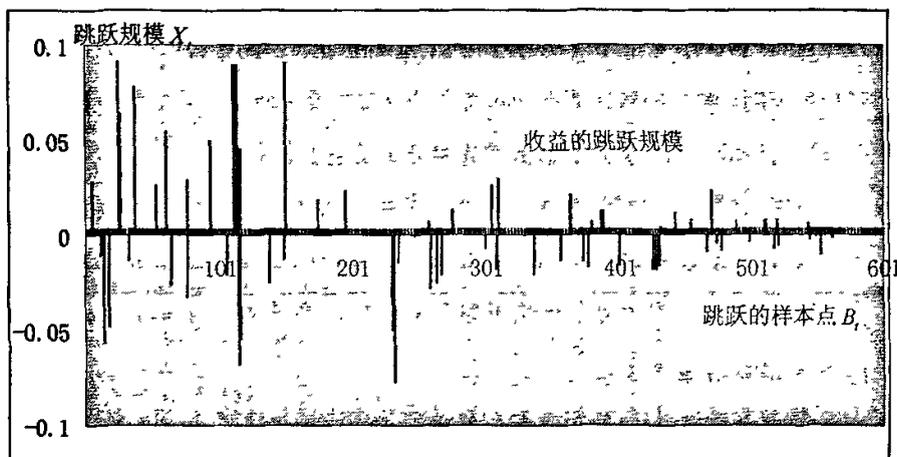


图 3.1：上海综合指数收益的跳跃图

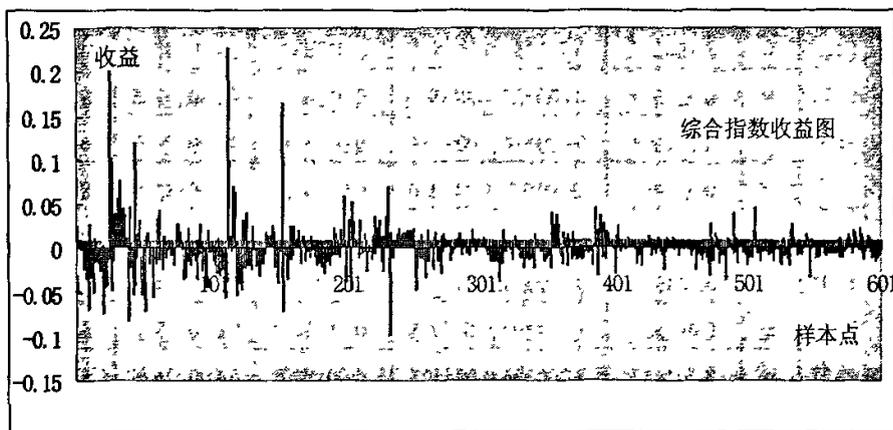


图 3.2：上海交易所综合指数的收益图

通过图 3.1 和图 3.2 的比较，我们可以很好的看出，当上海综合指数收益波动较大时，跳跃明显的发生，且跳跃的规模也较大，反之当收益很平稳时，跳跃很少发生，且跳跃规模都是很小的。上海股市综合指数收益发生跳跃主要集中在以下 2 个区间：1992 年 12 月份和 94 年 8 月份。另外，从 1996 年以后，上海综合指数较为平稳，发生跳跃的次数和规模都较 1996 年以前减少。

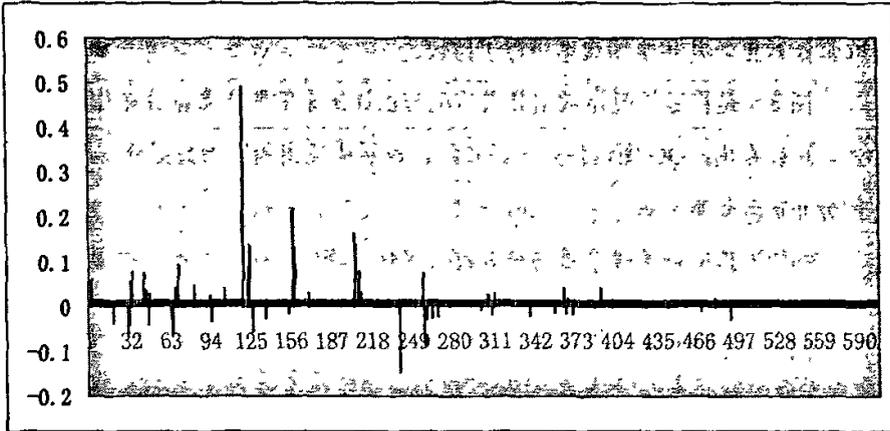


图 3.3：深圳综合指数收益的跳跃图

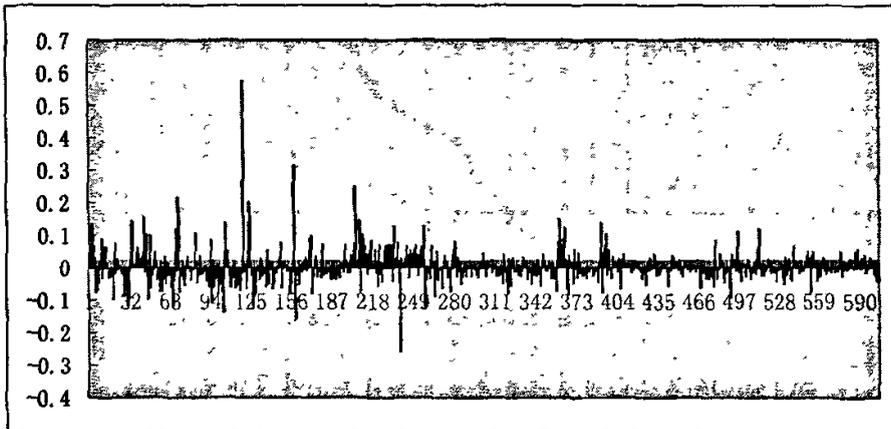


图 3.4：深圳交易所综合指数的收益图

通过比较深圳综合指数收益图 3.4 和收益跳跃图 3.3，也可以看出深圳综合指数收益的跳跃主要发生在 1996 年以前，具体时间也集中在 1994 年的 8 月份和 1992 年的年底。

通过上面的分析，上海和深圳综合指数收益都在 1992 年底和 1994 年 8 月份发生了大规模的跳跃，这也是和当时的政策有关。1992 年 11 月发生了“股票拆

细”，以及 1994 年 8 月 1 日的证监会推出三大利好政策相联系。而在 1996 年以后深圳和上海综合收益指数没有发生大规模的跳跃，这也与 1996 年 12 月 16 日证监会发出通知实行涨跌停板制度相关。

3.6.3.2 结果分析

这里将参数的估计值代入到双指数跳跃扩散模型 (DEJD) 中，模拟出 601 个股指数数据，并和原始数据进行了比较。

图 3.5 是将原始股指数据分布情况与估计出来的股指数据分布作比较的 QQ 图。可以看出，QQ 图近似是一条直线，说明模拟出的股指数据分布与原始数据的分布情况基本一致。

利用模拟所得的股票价格来求收益序列的峰度和偏度，所得结果如下：

$$\begin{aligned} Kurt &= 19.9876 \\ Skew &= 2.5438 \end{aligned} \quad (3-50)$$

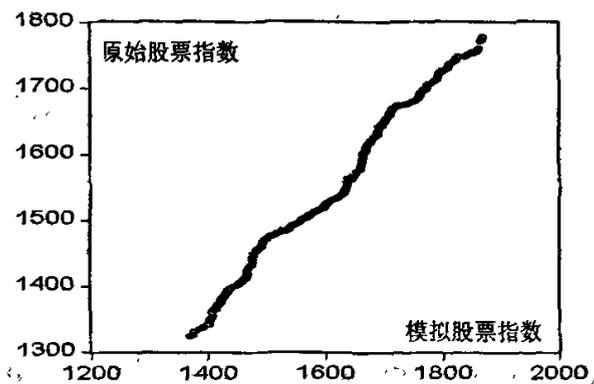


图 3.5: 原始股指数据分布与模拟的股指数据分布的 QQ 图

从上面的结果我们可以看出，模拟收益序列的峰度和偏度远远大于正态分布的峰度和偏度，从而也就证明了双指数跳跃扩散模型能够很好地同时描述资产收益分布的尖峰厚尾和有偏性的特征。另外，在用 MCMC 方法对参数进行估计的同时，还估计了跳跃发生的时间和跳跃的规模，而 Kou (2002) 在估计双指数跳跃扩散模型的时候仅仅估计了模型的参数。所以，模拟结果和估计结果也说明了 MCMC 方法适合估计含有隐含变量的连续时间模型。同时通过该模型所模拟的股指数据和实际数据的分布基本一致，从而也说明了双指数跳跃扩散 (DEJD) 模型可以很好的描述资产收益的分布特征。

3.7 本章小结

MCMC 方法是最近十几年来金融计量经济学领域的热门研究领域，基于贝叶斯原理的 MCMC 估计方法是从正比于参数后验密度的密度函数中取样，从而确定参数的估计值。而其中的关键在于获得参数后验分布的正比形式，而这就需要贝叶斯原理，从模型的似然函数和参数的先验密度中获得参数的后验密度。

MCMC 方法目前有 Gibbs 和 MH 两种算法，目前国内外的绝大部分文献都是使用 Gibbs 算法，直接使用 Winbugs 软件实现 MCMC 方法。本章使用 C++ 语言开发了基于 MH 算法的 MCMC 方法，该方法可以处理参数后验密度为复杂情况下的参数估计问题，还可以解决含有隐含变量的连续时间模型的参数估计以及隐含变量的估计问题。

本章首先介绍了有关贝叶斯统计方法，依据贝叶斯原理，参数的后验分布正比于模型的似然函数与参数的先验分布的乘积；其次讨论了如何采用 Euler 或 Milstein 方法对连续时间模型进行离散化从而获得模型的近似似然函数；接着重点给出了基于 MH 算法的 MCMC 方法以及取样的收敛性检验；最后通过对双指数跳跃扩散模型的估计来验证基于 MH 算法的 MCMC 估计方法的有效性，并对沪深股市进行了实证分析。

第四章 连续时间变结构模型研究

在实证研究中,如何选择合适的模型形式来描述资产收益统计分布特征就显得至关重要,Fama(1965)^[137]和Mandelbrot(1963)^[138]最早指出,资产收益的条件分布具有尖峰厚尾、有偏性以及波动集聚性。而时变条件矩模型恰恰可以描述这些资产收益的经验特征,离散时间模型中主要有Engle(1982)^[139]的ARCH模型和Bollerslev(1986)^[140]的GARCH模型等,在连续时间模型中主要有随机波动模型等。这些模型都一致认为条件矩随着信息集的改变而改变。

另外经验研究结果表明,在不同的时段,资产收益的分布存在着变化,但是常规模型(如ARCH模型、GARCH模型、连续时间的BS模型、随机波动模型等)都不能描述这些特征。在有效市场中,公司的投资和融资计划的公告能够影响企业的期望收益和风险,而这些收益和 risk 的变化会立刻体现在公司股票的价格上;相应的宏观经济信息的公告将会影响所有的证券的收益和风险,当然包括投资组合和指数。既然影响收益和 risk 的信息是随机的在某个时间段出现,这些信息就有序的影响证券收益时间序列的均值和方差参数。Kim和Kon(1996)^[141]指出那些具有有序结构变化的模型在描述资产收益经验特征方面比那些时间独立模型更有效。另外,Lamoureux和Lastrapes(1990)^[142]指出变结构GARCH模型还能刻画资产收益的波动持续性。

关于变结构模型的研究主要分成两大类,一类是已知变结构点的变结构模型的研究;这一类变结构模型的主要是研究模型的稳定性,所采用的方法相对来说也比较简单,Andrews和Fair(1988)^[143]提出了用Wald检验、似然比检验以及拉格朗日乘子检验来检验并比较变结构点前后模型参数的显著性差异来判断是否存在变结构。Ghysels和Hall(1990)^[144]提出了用变结构点前的数据所估计的参数来评价变结构点后数据的矩条件,从而判断是否存在变结构的问题。

另外一类就是未知变结构点的变结构模型的研究,这一类变结构模型的研究首要的问题是确定变结构点的位置。这方面的主要贡献有Bai(1999)^[145]的似然比检验方法、Kim与Kon(1999)^[146]的Bayes分析方法以及Gagliardini(2004)^[147]

的GMM方法。国内的研究成果有柯珂和张世英(2001)^[148]研究了ARCH类变结构问题、黄连洪和张世英(1987)^[149]用GBV方法来确定了线性模型的渐变和突变变结构点的位置、苏卫东和张世英(2003)^[150]用Bayes方法研究了变截距SV模型、徐梅和张世英(2005)^[151]应用DWT(离散小波变换)系数累积平方和的LMSV(长记忆随机波动)模型单一变结构点的检测与基于最大重复离散小波变换(MODWT)系数的变结构点的定位方法。

本章使用基于 Bayes 原理的 MCMC 方法研究了连续时间的变结构 BS 模型, 将变结构点作为一个隐含变量, 从变结构点的后验分布中取样从而确定变结构模型的变结构点。特别指出, 在本章中, 变结构点为一个连续变量, 而不是离散的数据观测点。

4.1 变结构点的贝叶斯诊断及实证分析

4.1.1 变结构点和异常点的联系和区别

异常波动点是指严重偏离股票价格走势的数据点, 按Fox(1972)^[152]的说法, 异常点根据其性质和产生机理, 可以分为两大类: 加性异常点(additive outliers)和革新异常点(innovation outliers)。

加性异常点(AO)是一种孤立的异常点, 往往是由某一时刻的观察或记录误差所造成的, 它们的出现并不波及邻近的观测值。就时间序列而言, 此类异常点是非本质的, 因为它们不涉及到时间序列内在的相关结构。

革新异常点(IO)是一种成片出现的异常点, 在某段时间内一个异常点通过随机序列的相关结构, 波及到与它邻近的一批数据点(通常是它的后继临近点), 从而使这些临近点表现出一定的异常。与AO相比, 此类异常点是更为内在或本质的一类异常点, 它产生的机理也更为复杂。

一般而言, AO对时间序列滞后相关系数的影响是使其偏向于0, 而IO对滞后相关系数的影响是使其偏向于1^[153]。由此可见, 不同类型的异常点可能对估计量产生不同的影响, 因此区别异常点的类型很重要, 而且由于个别异常点的存在可能对模型分析的结果产生举足轻重的作用, 因此在时间序列分析中, 探索异常点并将其与变结构点相区别是非常有意义的工作。

波动变结构点是这样一类点：它们对模型的结构又如此重大的影响，在该点之前与该点之后模型的结构有显著区别。这里模型的结构不仅包括模型的函数形式，还包括模型中的各参数。

4.1.2 变结构点的贝叶斯诊断

波动变结构问题的关键是如何才能有效准确地找到波动的变结构点。一般有两种方法：一是使用经济事件公告的日期；另一种则是使用统计诊断方法。前者虽然简单易行，但在实际中无法确认公告日期到底是哪天，也很难确定到底那一条信息足以改变风险（波动）的结构。而且，在一个非有效的市场上，公告日期与结构变化实际发生的日期不一定一致。因而，统计诊断程序对于发现真正的变结构点是更加有效的。

用统计手段探寻变结构点通常有两种方式：模拟法与贝叶斯法。模拟法适用于波动状态的数目 K 已知情况，其具体做法是选择那些有最大后验概率的点作为变结构点列。然而现实中状态的数目并不是明确知道的，而且，当样本的容量和状态的数目较大时，变结构点的模拟诊断就会在计算上遇到困难，因为对样本点所有可能的部分都必须逐一考察。对于模型结构变化点的贝叶斯检测，黄违洪、张世英（1987）提出了针对线性模型变结构的GBV(General Bayesian Method with Vague Prior Information)检测法，GBV法具有简单、过程直观、通用性强等优点，但其无法判定结构变化点的数目；Kim（1991）^[154]提出一种与GBV法相类似的方法，但他构造了检验线性模型平稳性的贝叶斯显著检验量，以检测单个参数是否发生结构变化，这种程序不仅在计算上更加有效，而且能估计出未知的变结构点数；Kim与Kon（1999）将这种方法直接用于检测方差平稳性。

贝叶斯诊断程序有如下的步骤：

设原始样本容量为 T ，以未知的结构变化点 τ 为界，将原始样本分为两个序列，各序列的容量分别为 $n_1 = \tau$ 和 $n_2 = T - \tau$ 。

第一步：对原始样本进行方差波动的平稳性检验。考察如下的零假设和备择假设：

$$\begin{aligned} H_0 : \rho = \sigma_2^2 / \sigma_1^2 = 1 \\ H_1 : \rho = \sigma_2^2 / \sigma_1^2 \neq 1 \end{aligned} \quad (4-1)$$

其中 σ_1^2 与 σ_2^2 是检验点前后两个样本状态下的方差。传统的贝叶斯检验对先验信息输入高度依赖，当参数的先验信息未知或比较含糊时，Box和Tiao (1973) [155] 提出一种基于无信息先验贝叶斯检验，它建立在最高后验概率密度 (Highest Posterior Density, HPD) 区间基础之上，计算在原假设 H_0 下的无条件概率 p -值，当 p -值小于给定的显著性水平时，拒绝零假设，然后计算以该点划分的两个样本的后验概率，并选择检验样本内有最大后验概率的数据点作为波动的变结构点。

这一检验需要计算 H_0 下的无条件概率 (p -值)。对于容量为 n 的样本，用 HPD 区间法进行检验时，在 H_0 下的无条件 p -值的计算公式 [150] 为

$$p_{\rho=1} = \sum_{\tau} 2 \{ 1 - F_{n_1-1, n_2-1} [\max(\rho', 1/\rho')] \} \pi(\tau|y) \quad (4-2)$$

其中的累加是对所有可能的变结构点 τ 进行， $F_{n_1-1, n_2-1}(\cdot)$ 是自由度为 $(n_1 - 1, n_2 - 1)$ 的 F 分布的概率密度函数，且其中

$$\rho' = (\hat{\sigma}_1^2 / \hat{\sigma}_2^2) \rho \quad (4-3)$$

$\pi(\tau|y)$ 为 τ 的后验概率密度函数，其表达式如下：

$$\pi(\tau|y) \propto \pi(\tau) \tau^{-1/2} (n-\tau)^{-1/2} \Gamma((\tau-1)/2) \Gamma((n-\tau-1)/2) \times S_1^{-(\tau-1)/2} S_2^{-(n-\tau-1)/2} \quad (4-4)$$

其中 $S_1 = \sum_{i=1}^{\tau} (y_i - \bar{y}_1)^2$ ， $\bar{y}_1 = \sum_{i=1}^{\tau} y_i / \tau$ ； $S_2 = \sum_{i=\tau+1}^n (y_i - \bar{y}_2)^2$ ， $\bar{y}_2 = \sum_{i=\tau+1}^n y_i / (n-\tau)$ ， $\Gamma(\cdot)$ 为伽玛函数， $\pi(\tau)$ 是 τ 的先验密度，由于我们并没有关于 τ 的先验信息，因此假定它的先验为均匀分布，也即 $\pi(\tau)$ 在确定 τ 的后验概率时不会产生任何影响。注意到 $F(1)$ 是零假设 H_0 成立时的条件 F 检验统计量，因而方差参数平稳性的无条件 p -值是建立在给定 τ 的 F -检验量基础上的无条件 p -值的加权平均，其权重取决于 τ 的后验密度式 (4-4)。

方差的这种平稳性检验需要在每一个数据点添加到初始样本上时重复进行，直到 p -值小于预定的显著性水平 (H_0 被拒绝) 为止。

第二步：确定波动的变结构点。如果 H_0 被拒绝，就可认为在样本上存在一个波动的变结构点。然后对这一样本计算每一可能的两状态划分的后验概率密度 (见公式 4-4)。如果某一个点所对应的状态划分有着最大的后验概率密度，那么它就是波动的第一个变结构点。对于波动所发现的第一个变结构点就记为 $\hat{\tau}_1$ 。

然后对以 $\hat{\tau}_1 + 1$ 点作为初始点的样本，重复以上两步。并在此样本区间上再

次计算每一数据点的后验概率密度，有着最大后验概率密度分类的点被选作波动的第二个变结构点 $\hat{\tau}_2$ 。继续重复这一程序直到所有的数据点被探寻一遍，这样所有的变结构点被依次找到。

4.2 连续时间资产收益变结构模型

黄违洪和张世英（1987）给出了变结构模型的定义，认为变结构模型就是统计模型的结构发生改变，这种改变包括模型的函数形式、变量的选取和参数、随机变量的统计分布特征不同，但目前在文献上常见的是指在不同的数据取样区间上模型参数发生改变的情况。仅有少量的文献考虑了模型的函数形式发生改变，白昆和张世英（2003）^[156]考虑了SV模型的变结构问题，首先寻找变结构点，之后采用分段建模的思想，在不同的数据取样区间具有不同的模型函数形式。

对于变结构模型来说，最关键的就是寻找模型结构发生变化的时间点，即变结构点。

Black和Scholes（1973）^[157]提出了资产收益的模型（BS模型）如下：

$$dy_t/y_t = \mu dt + \sigma dW_t \quad (4-5)$$

其中， y_t 为资产价格， W_t 为标准布朗运动， μ, σ 为参数。该模型是在严格的假设下提出的，这些假设包括：首先，股票价格服从对数正态分布，股票收益的方差是常数；其次，股票不会分红；第三，买卖股票和期权没有交易费用等等。

本章模型结构变化的角度对基本的BS模型进行扩展，考虑存在 m 个变结构点的连续时间变结构BS模型：

$$\begin{aligned} dy_t/y_t &= \mu_0 dt + \sigma_0 dW_t, & t \in [0, \tau_1] \\ dy_t/y_t &= \mu_1 dt + \sigma_1 dW_t, & t \in (\tau_1, \tau_2] \\ &\vdots \\ dy_t/y_t &= \mu_m dt + \sigma_m dW_t, & t \in (\tau_m, T] \end{aligned} \quad (4-6)$$

其中 y_t 为资产的价格， W_t 为标准布朗运动， $\mu_j, \sigma_j (j=0, 1, \dots, m)$ 为参数， $\tau_j (j=1, \dots, m)$ 为变结构点的位置，且 $0 < \tau_1 < \tau_2 < \dots < \tau_m < T$ 。

变结构BS模型与离散时间变结构模型的差别主要体现在变结构点上，离散时间的变结构点一般都是某一个数据观测点，而在连续时间模型中，变结构点

$\tau_j (j=1, \dots, m)$ 既可能对应于某一个数据观测点，也可能位于两个数据观测点之间，这样在变结构点定位的精度将与数据的取样频率相关，这一点将在下一节的仿真实验中得到印证。

4.3 连续时间单一变结构点的检测和定位

波动变结构分析最为关键的是有效准确地找到波动的变结构点，一般采取两种方法：一是使用经济事件公告的日期；另一种则是使用统计诊断方法。前者简单易用，无需任何数理工具，但在实践中遇到不少的困难。首先，我们很难确定在相关信息发布的很长一段时间区间内，公告日期到底是哪天，也很难确定到底是哪一条信息足以改变资产收益分布的结构。要知道有不少的信息对股票价格的影响太弱而没能改变风险的结构。而且，有许多的信息不是公开发布的，但它们却也有可能对公司的风险结构产生较大的影响。另外采用公告发布时间作为变结构点的另一问题是公告日期是否与结构变化实际发生的经济事件的日期一致，这主要决定于市场是否有效。因此，统计诊断程序相对于公告日期方法能够更加有效的确定变结构点的位置。

对于模型结构变化点的检测，黄违洪和张世英（1987）提出了一种针对线性模型变结构的GBV检测法，GBV法具有计算简单、过程直观、通用性强等优点，但不适合一些复杂形式的模型，如含隐含变量的随机波动模型；GBV法的另一缺陷在于无法判定结构变化点的数目。Kim（1991）提出一种与GBV法相类似的Bayes方法，他构造了检验线性模型平稳性的Bayes显著检验量，以检测到单个参数的非平稳性；Kim与Kon（1999）将这种方法直接用于检测方差平稳性，这种程序不仅在计算上更加有效，而且能估计出未知的变结构点数目。

Perreault等（2000）^[158]使用基于Bayes方法的MCMC方法估计了气象水文时间序列的变结构问题，其基本思想是将未知的变结构点作为一个变结构模型的一个参数，通过变结构模型的似然函数，结合各个参数的先验信息，再利用MCMC的Gibbs算法从各个参数的边际后验分布中获得参数（包括变结构点参数和模型参数）的估计值，从而确定模型的变结构点。

本文利用Perreault等（2000）的思想，首先对模型（4-6）采用Milstein方法

进行离散化，利用变结构模型的转移密度知识，获得模型的似然函数，再利用 MCMC 的 MH 算法从参数的联合后验密度中获得参数的取样样本，从而确定变结构点以及模型的参数。

为了简单起见，首先只考虑仅仅存在单一变结构点的变结构模型，之后将该方法推广到存在多变结构点的变结构模型。

4.3.1 单一变结构点定位的 MCMC 方法

如果最多只存在一个变结构点，这样模型 (4-6) 就可以简化为：

$$\begin{aligned} dy_t/y_t &= \mu_0 dt + \sigma_0 dW_t, & t \in [0, \tau_1] \\ dy_t/y_t &= \mu_1 dt + \sigma_1 dW_t, & t \in (\tau_1, T] \end{aligned} \quad (4-7)$$

假定资产价格 y_t 在时间 $[0, T]$ 上的 $n+1$ 个离散观测值为 $y = (y_0, y_1, \dots, y_n)$ ，这样相邻两个观测点的时间间隔就是 $\Delta = T/n$ 。则通过 Milstein 方法对模型 (4-7) 进行离散化，可得：

$$\begin{aligned} y_{i+1} - y_i - y_i \mu_0 \Delta + \frac{1}{2} \sigma_0^2 y_i \Delta &= \sigma_0 y_i \sqrt{\Delta} \varepsilon_i + \frac{1}{2} \sigma_0^2 y_i \Delta \varepsilon_i^2 & t = 0, 1, \dots, [m_1/T] \\ y_{i+1} - y_i - y_i \mu_1 \Delta + \frac{1}{2} \sigma_1^2 y_i \Delta &= \sigma_1 y_i \sqrt{\Delta} \varepsilon_i + \frac{1}{2} \sigma_1^2 y_i \Delta \varepsilon_i^2 & t = [m_1/T] + 1, \dots, n-1 \end{aligned} \quad (4-8)$$

其中 $[m_1/T]$ 表示的是 m_1/T 的整数部分， $\varepsilon_t (t = 0, 1, \dots, n-1)$ 服从标准正态分布；式 (4-8) 可以转化成如下形式：

$$\begin{aligned} \frac{1}{\sigma_0^2 \Delta} + 2 \left(y_{i+1} - y_i - y_i \mu_0 \Delta + \frac{1}{2} \sigma_0^2 y_i \Delta \right) / \sigma_0^2 y_i \Delta &= \left(\varepsilon_i + \frac{1}{\sigma_0 \sqrt{\Delta}} \right)^2 & t = 0, 1, \dots, [m_1/T] \\ \frac{1}{\sigma_1^2 \Delta} + 2 \left(y_{i+1} - y_i - y_i \mu_1 \Delta + \frac{1}{2} \sigma_1^2 y_i \Delta \right) / \sigma_1^2 y_i \Delta &= \left(\varepsilon_i + \frac{1}{\sigma_1 \sqrt{\Delta}} \right)^2 & t = [m_1/T] + 1, \dots, n-1 \end{aligned} \quad (4-9)$$

根据 3.2.2 中的公式 (3-10) — (3-17) 得到参数 $\theta = \{\mu_0, \mu_1, \sigma_0, \sigma_1, \tau\}$ 基于观测值集合 $y_t (t = 0, 1, 2, \dots, n)$ 对数似然函数为：

$$\begin{aligned} \log p_M(y|\theta) &= \sum_{t=1}^{[m_1/T]} \left[\log \left\{ f \left(\frac{2x_0}{\sigma_0^2 y_t \Delta} + \frac{1}{\sigma_0^2 \Delta} \right) \right\} - \log \left(\frac{\sigma_0^2 y_t \Delta}{2} \right) \right] \\ &+ \sum_{t=[m_1/T]+1}^{n-1} \left[\log \left\{ f \left(\frac{2x_1}{\sigma_1^2 y_t \Delta} + \frac{1}{\sigma_1^2 \Delta} \right) \right\} - \log \left(\frac{\sigma_1^2 y_t \Delta}{2} \right) \right] \end{aligned} \quad (4-10)$$

这样，依据贝叶斯原理，参数的联合后验密度为：

$$p(\theta|y) \propto p_M(\theta|y)p(\theta) \quad (4-11)$$

其中 $p(\theta) = p(\mu_0, \mu_1, \sigma_0, \sigma_1, \tau)$ 为各参数联合先验密度， $p_M(\theta|y)$ 由式 (4-10) 确定。由于各参数的边际后验密度都是正比于式 (4-11) 的，这样通过 MCMC 抽样，从而就可以获得变结构点和模型参数的样本。Perreault 等 (2000) 给出了各参数的先验密度， $\mu_0, \mu_1 \sim N$ ，正态分布， $\sigma_0, \sigma_1 \sim IG$ ，为 Inverted-Gamma 分布， τ 服从均匀分布^[160]。

有了参数的联合后验密度，很容易使用 MH 算法的 MCMC 估计参数的值，这里采用的是随机游走 MH 算法，算法具体步骤如下：

Step1: 赋予初始状态 $\theta^{(0)}$ ；

Step 2: 已知现在状态 $\theta^{(l)}$ ，按照下面的方法生成备选值 θ' ；

$$\theta' = \theta^{(l)} + \tau\varepsilon \quad (4-12)$$

其中 ε 是在区间上 $[-0.5, 0.5]$ 生成的随机数， τ 是调和参数，因此 τ 的选择应该使参数 θ' 被接受的概率在 20% 到 30% 之间^[117]。

Step 3: 由于建议密度采用的是随机游走的形式，因此按照式 (3-28) 来计算接受概率 $T(\theta^{(l)}, \theta')$ ；

Step 4: 以概率 $T(\theta^{(l)}, \theta')$ 接受备选值，即 $\theta^{(l+1)} = \theta'$ ；反之拒绝备选值，即 $\theta^{(l+1)} = \theta^{(l)}$ ；

Step5: 重复前面步骤 2-4，获得 $\{\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots\}$ ，剔除前面的 d 个值，则 $\{\theta^{(d+1)}, \theta^{(d+2)}, \dots\}$ 都具有相同的后验密度 $p(\theta|y)$ 。

这样，从联合后验分布中按照 MH 算法抽取的变结构点参数 τ 的有效样本为 $\tau^{(1)}, \tau^{(2)}, \dots, \tau^{(N)}$ ，因此变结构点的位置就是该样本的均值，即 $\hat{\tau} = \frac{1}{N} \sum_{i=1}^N \tau^{(i)}$ ，若 $\hat{\tau}$ 近似为 0 或 T 则变结构点不存在。另外，我们还可以用 4.3.3 节的似然比对该变结构点进行检验。

4.3.2 单一变结构点定位的仿真试验

变结构 BS 模型在资产收益的均值和方差上存在变化，上一小节讨论了用 MCMC 方法来进行变结构点的定位和变结构模型参数的估计，下面用仿真试验来验证 MCMC 方法的可行性和准确性。假设数据产生于存在唯一变结构点的变结构 BS 模型，通过下面三组试验进行对比分析。各组参数值见表 4.1，第一组是较为常见的情况，第二组和第三组都是与第一组进行对比试验，在不改变参数值的情况下分别改变变结构点的位置和取样频率来与第一组的结果进行对比，从而来验证 MCMC 方法。这三组模拟试验的结果见表 4.1。

表 4.1：仿真试验结果

参数	第一组 ($\Delta=1/52$)		第二组 ($\Delta=1/52$)		第三组 ($\Delta=1/104$)	
	真实值	估计值	真实值	估计值	真实值	估计值
τ	2.5	2.4853077 (0.022012)	3.5	3.507203 (0.027428)	2.5	2.49787 (0.016902)
μ_0	0.15	0.144560 (0.020514)	0.15	0.150194 (0.014407)	0.15	0.147694 (0.020798)
σ_0^2	0.055	0.051589 (0.0000443)	0.055	0.054165 (0.0000321)	0.055	0.052812 (0.0000212)
μ_1	0.30	0.291028 (0.040953)	0.30	0.288608 (0.026136)	0.30	0.291031 (0.0276845)
σ_1^2	0.10	0.102390 (0.000157)	0.10	0.095889 (0.000104)	0.10	0.099607 (0.00005635)

注：括号内的值为参数估计值的方差。

从表 4.1 我们可以看出，在这三组仿真试验中，变结构模型参数的估计值与真实值都非常接近，说明了 MCMC 方法的可行性。从第一组与第二组的对比试验看，变结构点位置的改变不影响 MCMC 方法估计的正确性；第一组与第三组的对比试验可以看出，当取样频率加大时能够改进变结构点位置参数估计的准确性，取样频率增加一倍， τ 的估计值与真实值的差距减少了 0.012。图 4.1 还给出了第一组试验各参数的取样路径，从图 4.1 可以看出，各参数的取样路径具有很好的收敛性。

4.3.3 变结构点的似然比检验

上面已经通过仿真试验验证了 MCMC 方法在进行变结构点定位时的有效性，接下来讨论当数据中可能存在变结构时，标准 BS 模型与变结构模型在拟合数据时的差异。在本小节，以上一节第一组仿真试验的模拟数据为基础，用似然比检验方法来分析变结构 BS 模型与标准 BS 模型在拟合这些数据的优劣，即进行本文中模型 (4-5) 与模型 (4-7) 的拟合优度检验。

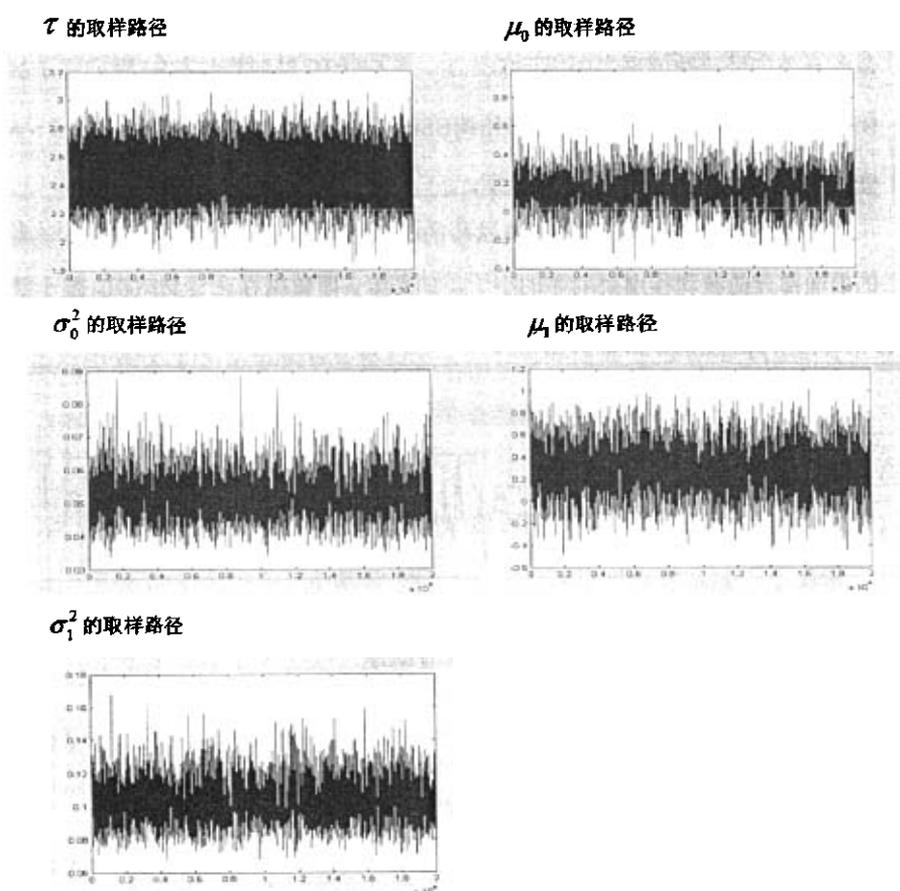


图 4.1：第一组仿真试验的各参数 MCMC 估计值的取样路径

模型选择的零假设和备择假设分别为： H_0 :变结构 BS 模型优于标准 BS 模型， H_1 :标准 BS 模型优于变结构 BS 模型；这样如果零假设为真，似然比检验量为：

$$LR = 2 \ln \lambda = 2(\ln l(\theta_1) - \ln l(\theta_2)) \quad (4-13)$$

其中 $l(\theta_1)$ 和 $l(\theta_2)$ 分别表示变结构 BS 模型和标准 BS 模型的似然函数值, θ_1 和 θ_2 分别为这两个模型的参数向量。

影响似然比检验方法效果的最大问题是关于似然比统计量分布的假设, 在进行单个模型对数据所进行的拟合优度检验时, 只需假定模型的极大似然估计量具有渐近正态分布, 因而可以假定检验统计量服从 χ^2 分布。但是在讨论两个模型的拟合优度对比情况时, 由于有两个模型介入, 检验统计量服从 χ^2 分布的假设就有可能与实际情况产生较大偏差。Rafterty (1996) [159] 给出的似然比检验统计量临界值的经验数值来判断模型的优劣。当 $LR \in (0, 2)$ 时变结构 BS 模型优于标准 BS 模型不显著, 当 $LR \in (2, 6)$ 时变结构 BS 模型显著优于标准 BS 模型, 当 $LR > 6$ 时变结构 BS 模型严格优于标准 BS 模型。

图 4.2 是上一小节第一组仿真试验的模拟数据, 从图 4.2 可以看出, 该数据的前面部分的波动程度要明显的小于后面部分, 即数据存在着变结构。基于该数据的标准 BS 模型的参数估计结果为:

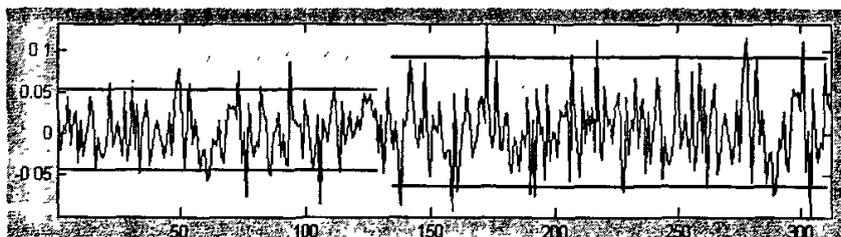


图 4.2: 模拟收益数据

$$\begin{aligned} \mu &= 0.235902 \\ \sigma^2 &= 0.078608 \end{aligned} \quad (4-14)$$

这样, 基于模拟数据的似然比统计量的值为:

$$LR^* = 2 \times (567.475 - 563.368) = 8.214 > 6 \quad (4-15)$$

依据 Rafterty (1996) 给出的标准, 可见变结构 BS 模型严格优于标准 BS 模型, 从而也说明了本文提出的变结构点定位的 MCMC 方法的正确性。

4.3.4 变结构点的随机模拟似然比检验方法

对变结构点是否存在还可以使用一种模型选择的随机模拟似然比检验的方法来进行,通过比较含变结构点模型与不含变结构点模型的优劣来确定是否存在变结构点。可以采用余素红和张世英(2004)^[160]提出的基于随机模拟方法的似然比检验方法对变结构点是否存在进行检验;在这种检验方法中,似然比统计量的分布情况将通过随机模拟得到。

$l(\theta_1)$ 和 $l(\theta_2)$ 分别表示变结构BS模型和标准BS模型的似然函数值,由于本章对变结构定位和参数的估计采用的是MCMC方法,所以在变结构点定位的时候就可以获得精确的似然函数 $l(\theta_1)$ 和 $l(\theta_2)$ 的值,利用随机模拟方法的似然比检验来比较标准BS模型和变结构BS模型的有效性时,步骤如下:

Step1: 对于同一时间序列,分别估计出标准BS模型和变结构BS模型的参数集;

Step2: 零假设和备择假设分别为: H_0 :变结构BS模型优于标准BS模型, H_1 :标准BS模型优于变结构BS模型;

Step3: 利用零假设的变结构BS模型产生 N 列数据 $Y^i, i=1,2,\dots,N$;

Step4: 由模拟的每列数据分别对标准BS模型和变结构BS模型进行估计,并根据式(4-13)计算出对应的 $LR^i, i=1,2,\dots,N$, 从而获得 LR 的样本;

Step5: 对 LR 的样本进行数据处理,拟合得出 LR 的近似分布;

Step6: 有了 LR 统计量的大致分布情况后,将 N 个 LR 样本加上实际观测值 LR^* 共 $N+1$ 个 LR 值由大到小排序,考察实际观测值 LR^* 在所有值中的排序情况,若排序处于极端情况 (LR 显著靠前或显著靠后),说明 LR^* 的值不是来自于该总体,实际情况与零假设前提相悖,应拒绝原假设,若排序情况处于正常的中间状态,则说明实际值 LR 属于该总体,基于零假设情况得到的 LR 统计量的分布正确,所以零假设可以接受。

用仿真试验中第一组的数据做为原始数据,并令 $N=99$,采用上述方法计算 99 个 LR , 加上 LR^* 共有 100 个 LR 样本,这 100 个 LR 样本的分布图见图 4.3。

从 LR 的分布图中可以看出,在零假设成立的情况下, LR 的期望值是正值,大致在 8—10 之间,实际 LR^* 为 8.2,远大于 Raftery (1996) 给出的接受零假设的

标准。另外从图4.3可以直观的看出变结构BS模型优于标准BS模型。这一结论还可以通过检验进一步说明：在100个 LR 值中，观测值 $LR^* = 8.214$ ，由大到小排在第63位，即在零假设下有60%的 LR 模拟值高于实际值 $LR^* = 8.214$ 。可知，只有当排序接近1或100时，该检验结果显著，将拒绝原假设。若将置信度定为95%，则拒绝域是 $[0, 3]$ 或 $[97, 100]$ ，而排序为63的 LR^* 远不在拒绝域内，说明根据样本可以接受零假设。

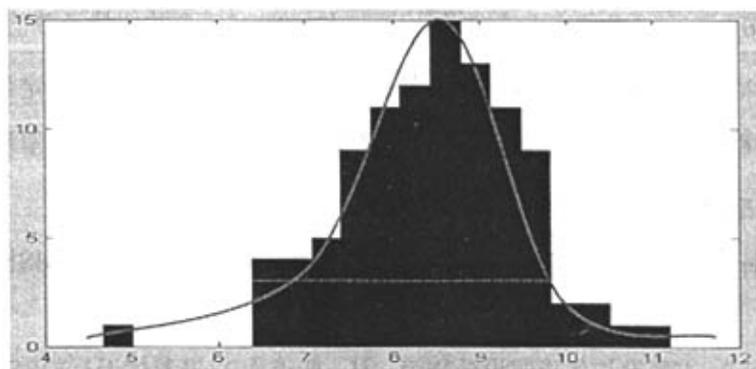


图4.3：在零假设成立的条件下 LR 的分布图

(注：图中直方图表示实际分布，曲线表示的近似拟合分布)

从上面的分析可知，应该接受零假设，即变结构BS模型优于标准BS模型，从而也验证了MCMC方法在确定变结构模型的变结构点位置和参数估计值的有效性。

4.4 连续时间资产收益模型多变结构点的定位

在实际应用中，变结构模型可能存在一个以上的变结构点。Bai (1999)提出了多变结构点模型的似然比检验方法，具体方法是：零假设是存在 $l(l=0,1,2,\dots)$ 个变结构点，备择假设为存在 $l+1$ 个变结构点；在检验变结构点的个数时，以 $l=0$ 开始进行似然比检验，到接受零假设为止，则拒绝零假设的次数就是变结构点的数量，变结构点的位置可以在似然比检验中获得。

Kim 和 Kon (1999)提出了另外一种多变结构点的定位方法，首先定位多变结构模型的第一个变结构点；接着剔除第一个变结构点以前的数据，重新定位剩

余数据的第一个变结构点，直到剩余数据不存在变结构为止，从而确定了所有变结构点的位置。徐梅和张世英（2005）也采用该方法定位变结构长记忆 SV 模型的变结构点。

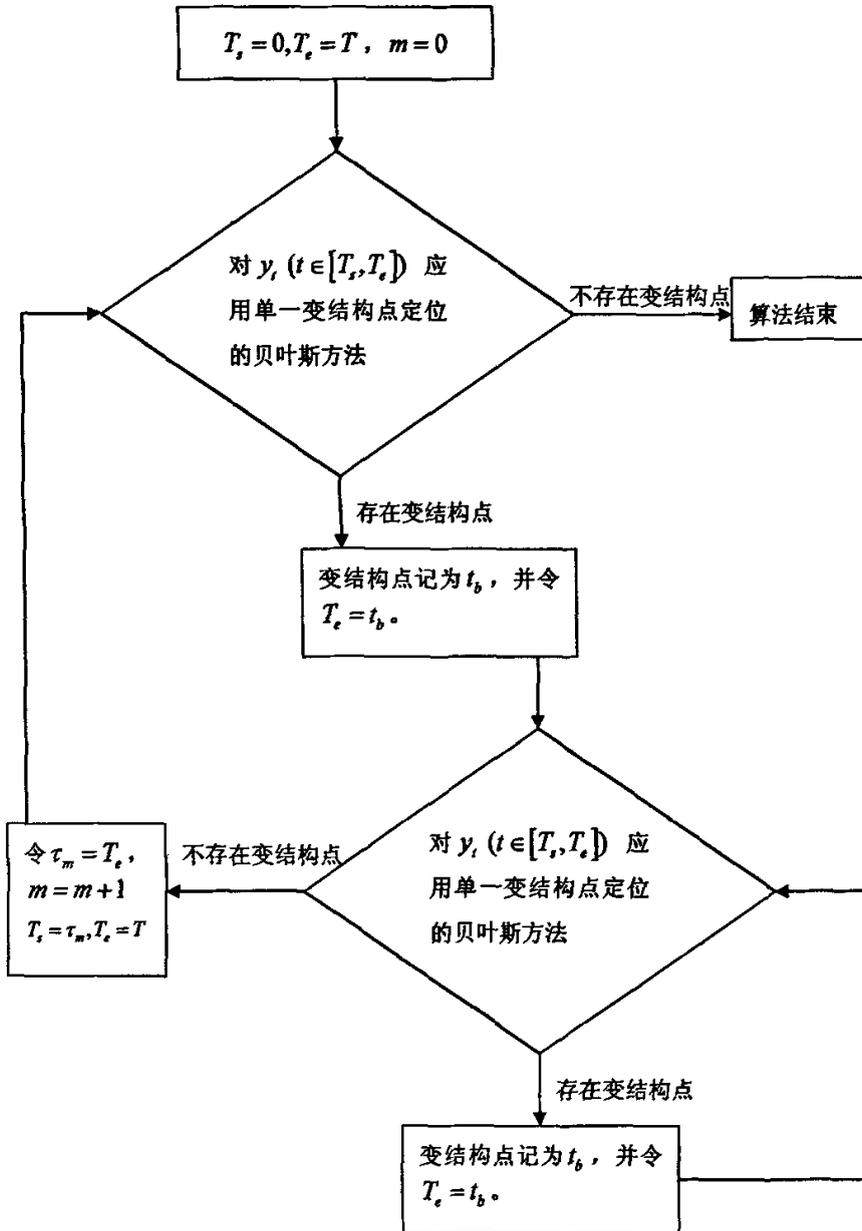


图 4.4: 多变结构点定位算法的流程图

在本节中，借鉴 Kim 和 Kon (1999) 的思想进行 BS 模型的多变结构点的检测与定位。考虑资产的收益 $y_t (t \in [0, T])$ ，对模型 (4-6) 进行多变结构点检测和定位。为了更好说明多变结构点的定位方法，图 4.4 给出了该算法的流程图。

依据图 4.4 的流程图，多变结构点定位的 MCMC 算法的步骤如下：

Step1: 令 $T_1 = 0, T_2 = T$; $m = 0$ ，为变结构点的个数。

Step2: 对 $y_t (t \in [T_1, T_2])$ 应用单一变结构点定位的贝叶斯方法，如果没有检测到变结构点，则算法停止，若 $m = 0$ 说明 $y_t (t \in [0, T])$ 不存在变结构点，若 $m > 0$ ，变结构点个数为 m ，位置分别为 $\tau_i (i = 0, \dots, m-1)$ ；若检测到变结构点，则其变结构点记为 t_b ，令 $T_2 = t_b$ 。

Step3: 对 $y_t (t \in [T_1, T_2])$ 应用单一变结构点的检测与定位的贝叶斯方法，如果不存在变结构点，则令 $\tau_m = T_2$ ， $m = m + 1$ ，并转到 Step5；若存在变结构点，则变结构点记为 t_b 。

Step4: 令 $T_2 = t_b$ ，返回 Step3。

Step5: 令 $T_1 = \tau_m, T_2 = T$ ，返回 Step2。

在上面的算法中，Step2 给出了该算法停止的条件，Step3 和 Step4 步骤的主要功能是寻找 $y_t (t \in [T_1, T_2])$ 的第一个变结构点，之后将第一个变结构点作为起始点、 T 作为终点，返回 Step2。

4.5 连续时间变结构模型在上海股市的应用

本节选用了上海证券交易所综合指数的每日收盘价作为样本数据，对上海股市收益的变结构进行分析。上海综合指数的时间段为 1997/06/04—2002/01/07，共 1112 个数据，日收益序列的样本如图 4.5。

对该收益样本收据采用 4.4 节的多变结构点定位方法，表 4.2 给出了上证指数收益序列的变结构点的位置，表 4.3 给出了不同时间段模型参数的估计值。

从表 4.2 的结果可以看出，上证指数收益存在着三个变结构点。接下来分析表 4.2 的三个变结构点所对应的时间，原因也比较明显。其一是 1998 年 8 月由于东南亚金融危机和国内洪涝灾害的影响，这与上海股市的第一个变结构点的产生不

无直接的关系。其二是1999年6月28日正是香港回归两周年的前夕，也是经历了5月19日的股市狂飙后中央采取响应政策的日子，加之1999年7月1日《证券法》的实施是上海股市波动第二个变结构点产生的原因。第三个变结构点是2000年3月20日，是由于2月份两大利好消息的公布。虽然从变结构的分布特点看，波动结构变化的实际发生点并不是与股市政策一一对应，这是因为市场的非有效性，造成信息的反应是不充分的，存在“反应过度”或“反应不足”的现象。另外，从表3的每个时间段的参数估计结果看，也存在明显的差异。

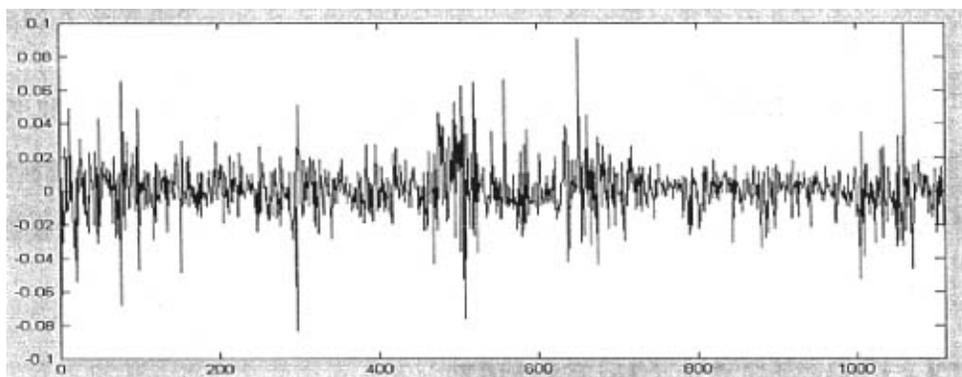


图 4.5：上海综合指数日收益序列

表 4.2：上证指数收益序列的变结构点

$\tau \left(\Delta = \frac{1}{250} \right)$	1.2010231 (0.01183756)	2.0180767 (0.01642308)	2.7056536 (0.01612336)
对应时间	1998/08/21	1999/06/29	2000/03/20

表 4.3：各时间段参数的估计值

时间段	1997/06/04— 1998/08/21	1998/08/21— 1999/06/29	1999/06/29— 2000/03/20	2000/03/20— 2002/01/07
μ	-0.0718411 (0.06127305)	0.3960804 (0.06737929)	0.0461919 (0.07117858)	-0.0149526 (0.02238281)
σ^2	0.0712144 (0.00003342)	0.0505787 (0.00002728)	0.0933567 (0.00011333)	0.0364926 (0.00000625)

4.6 连续时间随机波动模型的变结构分析

在标准 BS 模型中，资产收益的均值和方差都是常数，但是现实情况并不是这样。为了体现了资产收益过程中的高峰特征、收益的非对称（杠杠效应）以及厚尾特征。Hull 和 White (1987) 使用了随机波动模型，之后 Heston (1993) 给出了在随机波动模型下，期权价格的封闭形式解。Heston (1993) 提出的平方根随机波动模型如下：

$$\begin{aligned} d \ln S_t &= \mu dt + \sqrt{V_t} dW_{1,t} \\ dV_t &= \kappa(\eta - V_t) dt + \sqrt{V_t} \sigma dW_{2,t} \end{aligned} \quad (4-16)$$

其中 V_t 为波动过程，参数 α, σ, η 体现了收益分布中的高峰特征； W_1 和 W_2 是标准布朗运动，它们之间的相关系数为 ρ ，参数 ρ 体现了资产收益的非对称性，即杠杠效应， $\rho < 0$ 说明收益分布左偏， $\rho > 0$ 收益分布右偏；参数 η 体现收益分布的厚尾特征。

4.6.1 连续时间变结构随机波动模型

依据前面提出的连续时间变结构 BS 模型，存在 m 个变结构点的连续时间变结构随机波动模型的形式如下：

$$\begin{aligned} d \ln S_t &= \mu_0 dt + \sqrt{V_t} dW_{1,t} \\ dV_t &= \kappa_0 (\eta_0 - V_t) dt + \sqrt{V_t} \sigma_0 dW_{2,t} & t \in (0, \tau_1] \\ \\ d \ln S_t &= \mu_1 dt + \sqrt{V_t} dW_{1,t} \\ dV_t &= \kappa_1 (\eta_1 - V_t) dt + \sqrt{V_t} \sigma_1 dW_{2,t} & t \in (\tau_1, \tau_2] \\ \\ &\vdots \\ d \ln S_t &= \mu_m dt + \sqrt{V_t} dW_{1,t} \\ dV_t &= \kappa_m (\eta_m - V_t) dt + \sqrt{V_t} \sigma_m dW_{2,t} & t \in (\tau_m, T] \end{aligned} \quad (4-17)$$

其中 $\tau_j (j=1, \dots, m)$ 为变结构点的位置，且 $0 < \tau_1 < \tau_2 \dots < \tau_m < T$ ；为了简单起见并假定在任一时间段内 W_1 和 W_2 的相关系数 $\rho=0$ 并在各个时间段保持不变，但是这个假定可以推广到一般情况，即 ρ 在各个期间也可能发生变化。

4.6.2 变结构点位置、模型参数及隐含波动过程的后验密度

依据连续时间变结构 BS 模型的变结构点定位的 MCMC 方法，关键是获得变结构点和参数的后验密度。对于连续时间变结构随机波动模型来说，由于模型中还有隐含过程——随机波动过程，因此获得变结构 SV 模型的变结构点位置、模型参数以及隐含波动过程的后验分布将更加复杂，本节将采用 Euler 方法对 (4-17) 的模型进行离散化。

在确定随机波动变结构点的位置时，首先也是考虑单一变结构点的情形，之后再行多变结构点的定位。假定资产价格 S_t 在时间 $[0, T]$ 上的 n 个离散观测值为 $S = (S_1, S_2, \dots, S_n)$ ，这样相邻两个观测点的时间间隔就是 $\Delta t = T/n$ 。则通过 Euler 方法对模型单一变结构点 SV 模型进行离散化，可得：

$$\begin{aligned} Y_{t+1} &= Y_t + \mu_0 \Delta t + \sqrt{V_t} \Delta W_{1,t} \\ V_{t+1} - V_t &= \kappa_0 (\eta_0 - V_t) \Delta t + \sqrt{V_t} \sigma_0 \Delta W_{2,t} \quad t = 1, 2, \dots, \left[\frac{n\tau_1}{T} \right] \\ Y_{t+1} &= Y_t + \mu_1 \Delta t + \sqrt{V_t} \Delta W_{1,t} \\ V_{t+1} - V_t &= \kappa_1 (\eta_1 - V_t) \Delta t + \sqrt{V_t} \sigma_1 \Delta W_{2,t} \quad t = \left[\frac{n\tau_1}{T} \right] + 1, \dots, n \end{aligned} \quad (4-18)$$

其中 $\left[\frac{n\tau_1}{T} \right]$ 表示的是 $n\tau_1/T$ 的整数部分； $\Delta W_{1,t}, \Delta W_{2,t} (t=1, \dots, n-1) \sim N(0, \sqrt{\Delta t})$ ， N 表示的是正态分布， $Y_t = \ln S_t$ 。令 $\alpha_i = \kappa_i \eta_i$ 和 $\beta_i = -\kappa_i$ ，按照第三章给出的方法，变结构点位置和模型参数基于观测值 $\{S_t, t=1, 2, \dots, n\}$ 和隐含波动过程 $\{V_t, t=1, 2, \dots, n\}$ 的对数似然函数为：

$$\begin{aligned} p_E(\theta|Y, V) &= \sum_{i=1}^{\left[\frac{n}{\Delta t} \right]} \left\{ -\frac{1}{2} \ln(2\pi V_i \sigma_0 \Delta t) - \frac{1}{2V_i \Delta t} \times y_{i,0} \right\} + \\ &\quad \sum_{i=\left[\frac{n}{\Delta t} \right] + 1}^{n-1} \left\{ -\frac{1}{2} \ln(2\pi V_i \sigma_1 \Delta t) - \frac{1}{2V_i \Delta t} \times y_{i,1} \right\} \end{aligned} \quad (4-19)$$

其中， $\theta = \{\tau_1, \alpha_i, \beta_i, \mu_i, \sigma_i^2, (i=0, 1)\}$ ； $y_{i,j} (t=1, 2, \dots, n-1; i=0, 1)$ 满足下式：

$$y_{i,j} = (Y_{t+1} - Y_t - \mu_i \Delta t)^2 + \frac{(V_{t+1} - V_t - \alpha_i^2 \Delta t + \beta_i V_t \Delta t)^2}{\sigma_i^2}, (i=0, 1) \quad (4-20)$$

这样参数的后验密度就是：

$$P(\theta|Y, V) \propto p_E(\theta|Y, V) p(\theta) \quad (4-21)$$

其中 $p(\theta)$ 是参数的先验密度，变结构点的位置参数 τ_i 为均匀分布，周彦 (2006) 其他模型参数的先验密度。

在进行变结构随机波动模型的变结构点定位时，由于变结构点的存在，如何确定隐含波动过程的后验密度是解决问题的关键之一。在变结构点前后，隐含波动过程的后验密度的对数将有所不同，具体形式如下：

$$p(V_{i,t} | Y_i, S, \theta) = \begin{cases} -\frac{1}{2} \ln(2\pi V_i \sigma_0 \Delta t) - \frac{1}{2V_i \Delta t} \times y_{i,t}, t=1, 2, \dots, [\tau_i/\Delta] \\ -\frac{1}{2} \ln(2\pi V_i \sigma_1 \Delta t) - \frac{1}{2V_i \Delta t} \times y_{i,t}, t=[\tau_i/\Delta]+1, \dots, n-1 \end{cases} \quad (4-22)$$

其中 $y_{i,t}$ ($t=1, 2, \dots, n-1; i=0, 1$) 满足 (4-20)。

有了参数的联合后验密度和隐含波动过程的后验密度，从而采用 MH 算法的 MCMC 估计参数的值，单一变结构点随机波动模型的变结构点定位的算法具体步骤如下：

Step1: 赋予初始状态 $\theta^{(0)}$, $\{V_i^{(0)}, t=1, 2, \dots, n\}$;

Step 2: 采用 Hasting (1970) 提出的独立抽样 MH 算法从隐含波动过程的后验密度 (4-22) 中抽取 $\{V_i^{(t+1)}, t=1, 2, \dots, n\}$

Step 3: 给定现在状态 $\theta^{(t)}$ 和隐含波动过程 $\{V_i^{(t+1)}, t=1, 2, \dots, n\}$ ，采用随机游走 MH 算法生成 $\theta^{(t+1)}$;

Step4: 重复前面步骤 2—3，获得 $\{\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots\}$ ，剔除前面的 d 个值，则 $\{\theta^{(d+1)}, \theta^{(d+2)}, \dots\}$ 都具有相同的后验密度 $p(\theta | Y, V)$ 。

这样，从联合后验分布中按照 MH 算法抽取的变结构点参数 τ 的有效样本为 $\tau^{(1)}, \tau^{(2)}, \dots, \tau^{(N)}$ ，因此变结构点的位置就是该样本的均值，即 $\hat{\tau} = \frac{1}{N} \sum_{i=1}^N \tau^{(i)}$ ，若 $\hat{\tau}$ 近似为 0 或 T 则变结构点不存在。

对于按照上述方法确定的变结构随机波动模型的单一变结构点也可以采用 4.3.3 节中的模拟似然比检验方法进行变结构点的检验。

而对于多变结构点随机波动模型的变结构点定位按照 4.4 节中提出的步骤进行，其基本思路就是：首先定位多变结构模型的第一个变结构点；接着剔除第一

个变结构点以前的数据，重新定位剩余数据的第一个变结构点，直到剩余数据不存在变结构为止，从而确定了所有变结构点的位置。关于变结构随机波动模型的多变结构点定位的 MCMC 算法的具体步骤不在赘述。

4.6.3 变结构随机波动模型的实证分析

本节选用了上海证券交易所和深圳证券交易所的综合指数的每日收盘价作为样本数据，对上海股市收益的变结构进行分析。上海股市和深圳股市综合指数的时间段为 2000/01/04—2005/12/30，共 1442 个数据，日收益序列的样本图分别见图 4.6 和图 4.7。

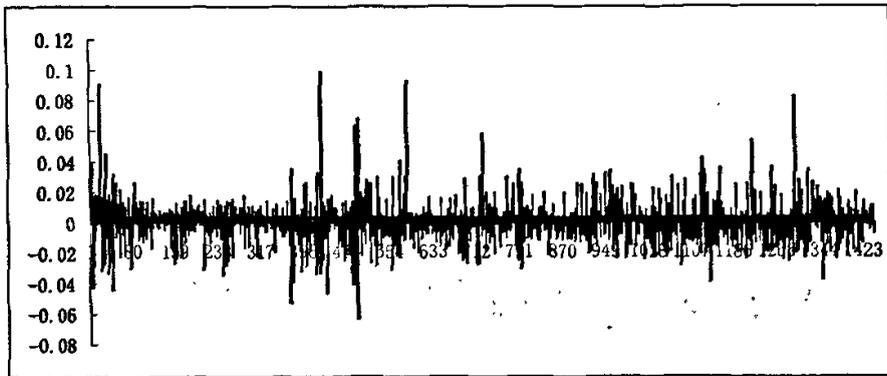


图 4.6：上海综合指数日收益序列

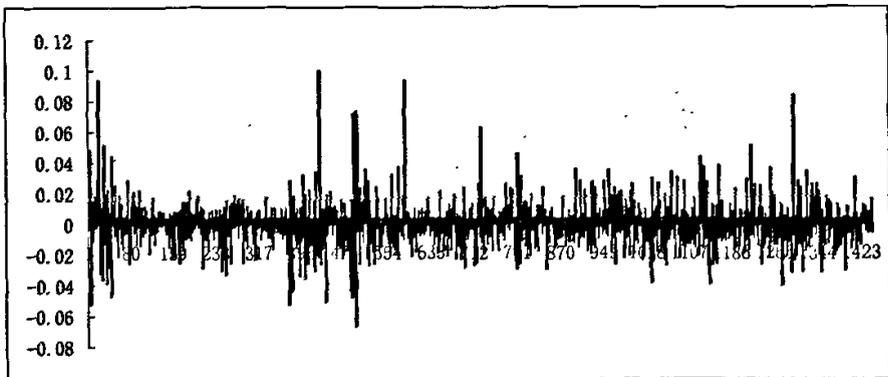


图4.7：深圳综合指数日收益序列

通过对上海股市和深圳股市综合指数收益序列图的直观观察，两大市场的综合指数收益具有较强的一致性，所以在实证分析中仅仅选定上海股市综合指数收益序列作为实证的原始数据来源。

对上海综合指数收益的样本数据采用4.6.2中的单一变结构点定位方法和4.4节的多变结构点定位方法，表4.4给出了上海综合指数收益序列的变结构SV模型的变结构点的位置，表4.4给出了不同时间段模型参数的估计值。图4.8给出了各个变结构点的样本序列图。

表4.4：上海综合指数收益的变结构SV模型的变结构点

变结构点	τ_1	τ_2	τ_3	τ_4
τ 的取值	0.278996 (0.008973)	1.4963 (0.008267)	2.42984 (0.009137)	4.441001 (0.014059)
对应的样本数据点	70	374	607	1110
对应的时间	2000/04/21	2001/07/27	2002/07/22	2004/08/19

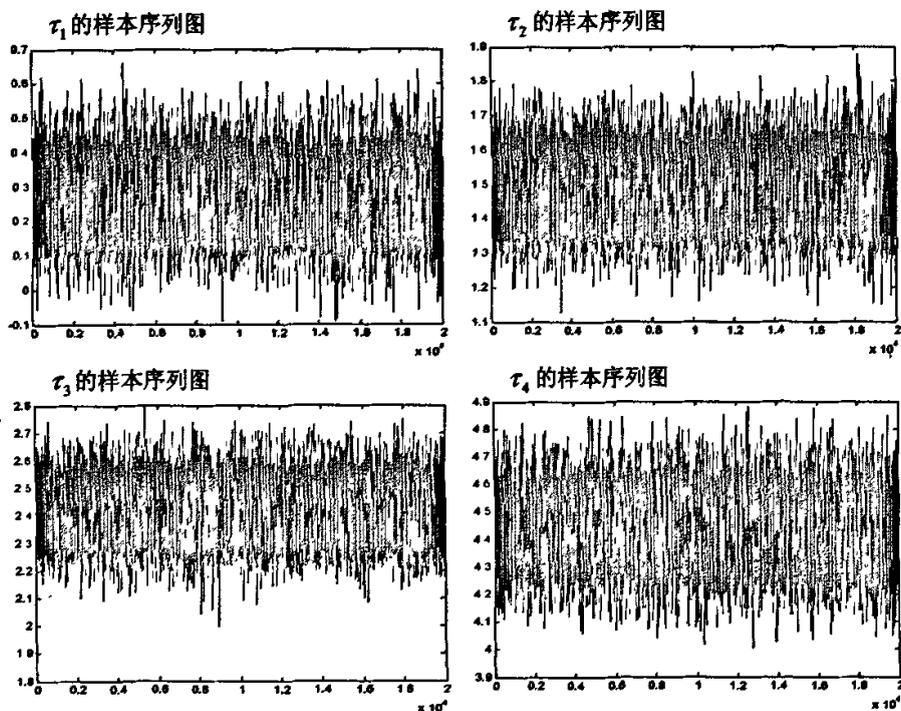


图4.8：变结构点的样本序列图

从表4.4的结果可以看出，上证指数收益在时间段2000/01/04—2005/12/30中存在着各四个变结构点，对应的时间分别为2000/04/21、2001/07/27、2002/07/22和2004/08/19。

接下来分析表 4.4 中的变结构点产生的原因：第一个变结构点是发生在 2000 年 4 月，这是由于同年 2 月份两大利好消息的公布。第二个变结构点产生的原因在于同年 6 月 12 日，国务院发布的《减持国有股筹集社会保障资金暂行办法》，确立了国有股减持主要采取存量发行，原则上采取市场定价方式。第三个变结构发生在 2002 年 7 月，这是由于同年 6 月 24 日，国务院决定停止通过国内证券市场减持国有股，除企业海外发行上市外，对国内上市公司停止执行《减持国有股筹集社会保障资金管理暂行办法》中关于利用证券市场减持国有股的规定，并不再出台具体实施办法，此举引起市场强烈反响。第四个变结构点产生的原因与同年 6 月中小企业板块正式登场、首批八只股票上市，这是落实“国九条”的首项具体措施，中小企业板块出台不仅将使深交所乃至中国证券市场获得新的活力；另外 8 月中国证监会《关于首次公开发行股票试行询价制度若干问题的通知》(征求意见稿)见诸报端是产生变结构点的另外一个影响因素。虽然从变结构的分布特点看，模型结构变化的实际发生点并不是与股市政策一一对应，这是因为市场的非有效性，造成信息的反应是不充分的，存在“反应过度”或“反应不足”的现象。另外，从表 4.5 的每个时间段的参数估计结果看，也存在明显的差异。

表 4.5: 各时间段参数的估计值

时间段	2000/04/21— 2001/07/27	2001/07/27— 2002/07/22	2002/07/22— 2004/08/19	2004/08/19— 2005/12/30
μ	0.102005 (0.019823)	-0.19099 (0.021355)	-0.1056 (0.018952)	-0.09373 (0.020197)
α	0.253672 (0.013919)	0.299182 (0.020017)	0.275892 (0.0141361)	0.284621 (0.018379)
β	-0.01258 (0.014032)	-0.02808 (0.01188)	-0.019934 (0.015859)	-0.027852 (0.013041)
σ^2	0.053734 (0.001653)	0.0886727 (0.002728)	0.062726 (0.001833)	0.0778396 (0.002495)

表 4.5 给出了各个时间段 SV 模型参数的估计值及其标准差。在此需要说明的时，由于在第一个变结构点之前的上海综合指数收益的样本数据个数仅为 70 个，所以将改时间段的数据舍弃。同时图 4.9 给出了 2000/04/21—2001/07/27 这一时间段各参数的取样序列图，其他时间的参数取样序列图省略。从表 4.5 的结

果可以看出，各时间段参数的估计值存在着显著差异。

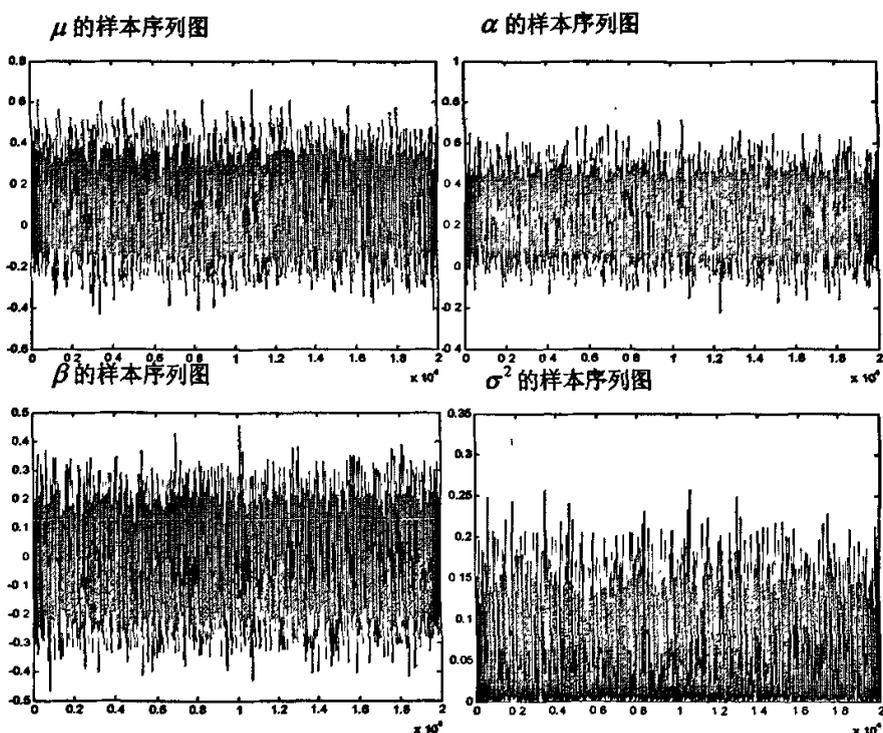


图4.9：2000/04/21—2001/07/27时间段SV模型各参数的样本序列图

4.7 本章小结

目前尚未见到国内外关于连续时间变结构模型研究的报道，主要原因在于变结构点定位以及模型参数估计的难度。此时不仅要解决每个时间段模型参数的估计，同时在连续时间变结构模型中变结构点不再依赖于状态变量的观测值，从而增加了变结构定位的难度。

本章提出了应用 MCMC 方法的连续时间变结构模型的单一变结构点的定位方法，并用仿真试验验证了该方法的有效性。接着提出了连续时间多变结构点模型的变结构点定位方法；该方法在确定变结构点位置的同时，又能估计相应的模型参数。用该方法对上海股市综合指数的收益序列进行了变结构分析，理论与实

证结果表明该方法有效且可行。

同时将该方法从 BS 模型扩展到了连续时间随机波动模型，并给出了各参数和隐含波动过程的后验分布，并给出了相应的 MCMC 算法。讨论了变结构点的检验，同时应用该方法对连续时间随机波动模型进行了实证分析，说明所提方法的可行性。

第五章 资产价格的抛物线跳跃扩散模型

自 Merton 和 Black-Scholes 于 1973 年用布朗运动和正态分布研究资产收益和期权定价以来，众多的研究者通过对 Merton 和 Black-Scholes 研究成果的扩展来更好的描述资产收益分布的经验特征。这些研究成果都是从以下几个方面进行扩展的，一是收益的跳跃，二是随机波动，三是跳跃和随机波动的结合。如 Cox 和 Ross(1976)的 CEV 模型, Merton(1976)的常跳跃扩散模型, Cox、Ingersoll 和 Ross (1985) 的 CIR 模型, Hull 和 White (1987)、Heston (1993) 的平方根随机波动模型, Bibby 和 Sørensen (1997) 的抛物线扩散模型, Bates (1996)^[9] 和 Adersen 等 (2002) 的 SVJ 模型, Duffie、Pan 和 Singleton (2000) 的收益和波动同时存在跳跃模型 (SVCJ 和 SVIJ), Eraker、Johannes 和 Polson (2003) 也考虑了波动跳跃的模型。

在金融学文献中，几何布朗运动被用来描述股票价格运动，虽然在期权定价和其他理论中具有有用的作用，但是几何布朗运动不具有股票收益的一些已知的统计特征，如高峰、波动集聚和长记忆性。为了反映这些特征，有跳跃扩散模型、随机波动模型、随机波动加跳跃模型以及时变列维过程。关于股票收益分布的厚尾特征已经得到了很好的论证，实际上可以用扩散分布 (Eberlein与Keller1995, Kuchler1994^[161]) 来很好的拟合这种分布。另外，股票价格对数的经验研究表明他们的增量具有很小的相关性但不是独立的。增量的平方表明增量之间还是有实质的相关 (Fama1965^[162], Taylor1986^[163])，在这些文章中都表明可以用扩散过程拟合随机过程的边际分布。

Bibby 和 Sørensen (1997) 提出了抛物线扩散过程，利用确定性的线性加上漂移为 0 的扩散过程对价格对数进行建模，扩散系数依赖于瞬时的股票价格，且调整之后的对数价格的分布形式是扩散分布而不是正态分布。通过扩散系数 (波动) 来拟合不含确定线性趋势的价格对数的剩余部分，而这种扩散系数是依赖于过程的状态。并用该模型成功的拟合一些股票数据，同时将抛物线扩散运用到期权定价中。

本章以上述的抛物线扩散模型为基础，用跳跃扩散过程来拟合股票价格对数的线性趋势和跳跃之外的特征，而这种跳跃扩散过程是具有抛物线不变测度的各态历过程。

5.1 抛物线跳跃扩散模型及其性质

5.1.1 抛物线跳跃扩散模型

经验结果表明股票价格的对数是一个近似不相关的增量过程，但是增量不是独立的。为此，股票价格的模型形式是：

$$S_t = \exp\left(\kappa t + X_t + \sum_{i=0}^{q_t} \xi_i\right) \quad (5-1)$$

其中， q_t 是强度为 λ 的泊松过程，即在区间 $[0, t]$ 发生跳跃的次数， $\xi_i \sim iidN(\mu_\xi, \sigma_\xi^2)$, $(i=0, \dots, q_t)$ 为跳跃规模，即在区间 $[0, t]$ 内每次跳跃的大小； X_t 为价格对数除去趋势项和跳跃项之后的稳定过程，其满足：

$$X_t = X_0 + \int_0^t v(X_s) dW_s \quad (5-2)$$

使用伊藤公式，式 (5-1) 可转化为：

$$dS_t = S_t \left\{ \left[\kappa + \frac{1}{2} v^2 \left(\log S_t - \kappa t - \sum_{i=0}^{q_t} \xi_i \right) + \lambda \mu_\xi \right] dt + v \left(\log S_t - \kappa t - \sum_{i=0}^{q_t} \xi_i \right) dW_t + \xi_i dq_t \right\} \quad (5-3)$$

其中 q_t 为泊松过程， W_t 是维纳过程，且两者是相互独立的。

通过选择合适的 $v(\bullet)$ 可以得到不同的模型；最简单的一种就是函数 $v(\bullet)$ 为一个常数，模型 (5-3) 就是 Merton (1976) 的常跳跃扩散模型。本章选择 Bibby 和 Sørensen (1997) 的形式，其 $v(\bullet)$ 满足：

$$v(x) = \sigma \exp\left[\frac{1}{2} \alpha \sqrt{\delta^2 + (x - \mu)^2} - \frac{1}{2} \beta (x - \mu)\right] \quad (5-4)$$

其中参数满足 $\alpha > |\beta| \geq 0$, $\delta > 0$, $\mu \in R$ ； α 为倾斜度参数， μ 为位置参数， δ 是标量参数， β 为对称性参数，当 $\beta=0$ 时，分布是关于 μ 对称的。此时，股票价格的对数除去趋势项和跳跃项后是一种抛物线分布。

5.1.2 抛物线跳跃扩散模型的性质

抛物线跳跃扩散模型的性质可以看成是由两部分构成：抛物线扩散过程与跳跃过程，因此研究抛物线跳跃扩散模型的性质，首先要研究抛物线扩散模型的性质。

5.1.2.1 抛物线扩散模型的性质

为了更好的说明抛物线扩散模型能够体现资产收益的经验特征，用 Euler 近似方法对 5.1.1 中介绍的抛物线扩散模型近似，可得：

$$Y_t \approx \sigma \exp \left[\frac{1}{2} \alpha \sqrt{\delta^2 + \left(\sum_{i=1}^n Y_{t-i} - \mu \right)^2} - \frac{1}{2} \beta \left(\sum_{i=1}^n Y_{t-i} - \mu \right) \right] e_t \quad (5-5)$$

其中， $Y_t = \ln S_t - \ln S_{t-1}$ ， $e_t \sim iidN(0, \Delta t)$ 。首先将式 (5-5) 转换成如下的形式：

$$\begin{aligned} Y_t &= \sigma \exp \left[\frac{1}{2} h_t \right] e_t \\ h_t &= f(Y_{t-1}, Y_{t-2}, \dots) \end{aligned} \quad (5-6)$$

而 Engle^[139]的 ARCH(∞) 模型的形式为：

$$\begin{aligned} Y_t &= \sigma_t e_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{\infty} \alpha_i Y_{t-i}^2 \end{aligned} \quad (5-7)$$

比较式 (5-6) 和 (5-7) 可以发现：抛物线模型就是一种非线性 ARCH(∞) 模型，所以抛物线扩散模型能够体现 ARCH(∞) 模型的特征，其中条件波动的非线性关系还可以体现波动的长记忆特征和波动集聚性。

5.1.2.2 抛物线跳跃扩散模型的性质

上一小节已经讨论了抛物线扩散模型的性质，在 Bibby 和 Sorensen (1997)、Tse、Zhang 和 Yu (2003) 中也详细介绍了抛物线分布能够很好的拟合资产收益的经验特征，如厚尾、有偏、泰勒效应、波动集聚性等。下面主要讨论加入跳跃后，抛物线跳跃扩散模型能否更好的体现这些经验特征。

为了更好的说明抛物线扩散模型能够体现资产收益的经验特征，用 Euler 近似方法对模型 (5-3) 近似，可得：

$$Y_t \approx \sigma \exp \left[\frac{1}{2} \alpha \sqrt{\delta^2 + \left(\sum_{i=1}^{\infty} Y_{t-i} - \mu \right)^2} - \frac{1}{2} \beta \left(\sum_{i=1}^{\infty} Y_{t-i} - \mu \right) \right] e_t + B \xi_t \quad (5-8)$$

其中, $Y_t = \ln S_t - \ln S_{t-1}$, $\xi_t \sim iidN(\mu_\xi, \sigma_\xi^2)$, $e_t \sim iidN(0, \Delta t)$, B 为贝努利随机变量, $B=1$ 的概率为 $\lambda \Delta t$, $B=0$ 的概率为 $1 - \lambda \Delta t$ 。

从式 (5-5) 可得:

$$Y_t \approx Y_t^1 + Y_t^2 \quad (5-9)$$

其中, $Y_t^1 = \sigma \exp \left[\frac{1}{2} \alpha \sqrt{\delta^2 + \left(\sum_{i=1}^{\infty} Y_{t-i} - \mu \right)^2} - \frac{1}{2} \beta \left(\sum_{i=1}^{\infty} Y_{t-i} - \mu \right) \right] e_t$ 和 $Y_t^2 = B \xi_t$ 是正态过程。而 Y_t^1 就是抛物线扩散的 Euler 近似, 所以 Y_t 的分布具有尖峰厚尾等特征。另外, 当跳跃规模的均值 $\mu_\xi \neq 0$ 时, Y_t 的分布不是对称的; 具体来说, 当 $\mu_\xi < 0$ 时, Y_t 的分布是左偏的; 当 $\mu_\xi > 0$ 时, Y_t 的分布是右偏的。

5.2 抛物线跳跃扩散模型的 MCMC 估计

5.2.1 参数的似然函数

运用 MCMC 方法对抛物线扩散跳跃过程进行估计, 首先要获得参数的联合后验分布, 而联合后验分布是以离散的观测值为基础的, 所以首先必须对跳跃抛物线扩散过程进行离散化。对于式 (5-10) 这个广义的跳跃扩散模型来说目前有两种近似离散化方法, 一是 Euler 方法, 另外一种是 Milstein 方法。

由于 Euler 方法其实质就是对式 (5-10) 这个随机微分方程进行 0.5 阶泰勒展开, 而 Milstein 是 1.5 阶泰勒展开, 另外 Tse、Zhang 和 Yu (2003) 已经证明 Milstein 方法较 Euler 方法更精确, 所以本文在对参数进行估计的时候, 使用的是 Milstein 方法。

对于一个广义抛物线跳跃扩散模型:

$$dX_t = \mu(X_t, \theta) dt + \sigma(X_t, \theta) dW_t + \xi_t dq_t \quad (5-10)$$

Milstein 的方法就是 1.5 阶泰勒展开, X_t 的观测值集合为 $X = \{x_t : t = 0, 1, \dots, n\}$, 由于时间 $(t, t + \Delta t]$ 内至多发生一次跳跃, 所以我们可用下

面的公式：

$$X_{t+\Delta t} = X_t + \mu(X_t, \theta)\Delta t + \sigma(X_t, \theta)\Delta W_t + \frac{1}{2}\sigma(X_t, \theta)\frac{\partial\sigma(X_t, \theta)}{\partial X_t}\left[(\Delta W_t)^2 - \Delta t\right] + J_t\xi_t \quad (5-11)$$

来近似广义跳跃扩散过程，而 (5-11) 式可以转换为：

$$X_{t+\Delta t} - X_t - \mu(X_t, \theta)\Delta t + g(X_t, \theta)\Delta t - J_t\xi_t = \sigma(X_t, \theta)\sqrt{\Delta t}\varepsilon + g(X_t, \theta)\Delta t\varepsilon^2 \quad (5-12)$$

其中 $g(X_t, \theta) = \frac{1}{2}\sigma(X_t, \theta)(\partial\sigma(X_t, \theta)/\partial X_t)$ ，令

$$a = \sigma(X_t, \theta)\sqrt{\Delta t}, \quad b = g(X_t, \theta)\Delta t \quad (5-13)$$

依据 3.2.2 中的公式 (3-10) - (3-18) 可得到对数似然函数为：

$$\log p_M(\theta|X) = \sum_{t=1}^n \left[\log \left\{ f\left(\frac{y_t}{b} + \frac{a^2}{4b^2}\right) \right\} - \log(b) \right] \quad (5-14)$$

其中 $p_M(\theta|X)$ 是基于 Milstein 方法的似然函数。

5.2.2 跳跃扩散模型的 MCMC 方法

马尔可夫链蒙特卡罗方法在众多的统计应用中被证明是一种比传统独立取样更有效的取样方法。在计量经济学和金融学中，MCMC 方法也被广泛使用。依据贝叶斯原理，参数向量 θ 基于数据 X 的条件分布可以通过后验密度 $p(\theta|X)$ 获得，即：

$$\pi(\theta|X) = cp(X|\theta)\pi(\theta) \quad (5-15)$$

其中 $p(X|\theta)$ 是似然函数， $\pi(\theta)$ 是参数 θ 的先验密度。

在抛物线跳跃扩散模型中，由于引入了跳跃因子，而这些跳跃因子都是隐含变量，从而参数的估计将更加复杂。要解决参数的估计问题，首先要解决隐含变量的估计问题；Eraker (2001) 对含有隐含变量的扩散模型估计提出了相应的解决方案，即在用 MCMC 的 MH 算法从参数的后验分布中抽取每一组参数样本的同时，运用 MCMC 的 Gibbs 算法从隐含变量的后验分布中抽取一组隐含变量。下面是参数向量 θ ，跳跃时间 $J = (J_t, t=1, 2, \dots, n)$ 和跳跃规模 $\xi = (\xi_t, t=1, 2, \dots, n)$ 的后验密度：

$$\text{扩散和跳跃参数 } \theta: \pi(\theta|X, J, \xi) \propto p(X|\theta, J, \xi)\pi(\theta)$$

跳跃时间 J_t ($t=1,2,\dots,n$): $\pi(J_t|X_t, X_{t-1}, \theta, \xi_t) \propto p(X_t|\theta, X_{t-1}, J_t, \xi_t) \pi(J_t)$

跳跃规模 ξ_t ($t=1,2,\dots,n$): $\pi(\xi_t|X_t, X_{t-1}, \theta, J_t=1) \propto p(X_t|\theta, X_{t-1}, J_t=1, \xi_t) \pi(\xi_t)$

依据 MCMC 方法从双指数跳跃扩散模型的参数 $\theta = (\mu, \sigma^2, \kappa, \eta^2, \lambda)$ 、跳跃时间 B 和跳跃规模 X 的后验分布中抽样按照 3.3.3 中给出的步骤进行。

在该算法中，建议密度 $q(\bullet|\theta)$ 的采用随机游走的形式，即建议密度为 $[-0.5, 0.5]$ 上的均匀分布，参数向量 θ 按照下面的方法生成：

$$\theta' = \theta + \tau \varepsilon \quad (5-16)$$

其中 ε 是在区间上 $[-0.5, 0.5]$ 均匀分布的误差项， τ 是调和参数，通过 τ 的值可以调整误差项的方差，一般来说 τ 值应该使接受概率在 20% 到 30% 之间。

关于参数抽样样本的收敛性问题可以参见第三章中的详细介绍，本章采用还是由 Kim (1998) 提出的模拟无效因子 (SIF) 来评价抽样样本的模拟效果，Meyer 和 Yu (2000) 给出了 SIF 的计算公式。

5.3 数据及实证

本节中，使用随机游走 MH 算法来估计离散化的抛物线跳跃扩散模型，数据分别采用 MSCI 全球指数与 NYSE 指数，这些数据序列是从 1990-01-01 至 2000-12-31 的周观测数据。

联合先验信息是 $\pi(\theta)$ ，基于 Milstein 方法的似然函数为 $p_M(\theta|X)$ ，则参数的联合后验密度为 $\pi(\theta|x) \propto \pi(\theta) p_M(\theta|X)$ 。实证研究表明，参数的估计值不会随着先验密度的改变而发生明显的变化，这也说明参数的后验密度不会随着参数的先验分布发生变化，具体讨论可以参见文献[54]和[117]。

按照随机游走 MH 算法对 HJD 模型进行参数估计，每个参数共产生 60000 个的样本，剔除前面的 10000 个模拟样本，记录的模拟样本为后面的 50000 个。其中使用数据集 MSCI 全球指数的样本抽样值曲线见图 5.1，样本的自相关系数见图 5.2，跳跃规模见图 5.3；各种不同数据集的参数估计结果为表 5.1 所示。

表 5.1: 特殊先验分布的 HJD 模型的 MCMC 结果 (Milstein 离散)

数据	参数	均值	置信区间	SD	MCSE	SIF	AC
MSCI 全球 指数	κ	-0.0801223	-0.130178, -0.0300669	0.0255385	0.00099422	76	0.25
	α	1.45727	0.976772, 1.93777	0.245153	0.00524926	23	0.24
	δ^2	0.0218031	-0.0363314, 0.0799377	0.029661	0.00099962	57	0.31
	μ	6.33956	4.51187, 8.16726	0.932494	0.0186435	20	0.22
	β	0.372285	0.103081, 0.641489	0.137349	0.0031961	27	0.24
	σ^2	0.00754131	0.0046393, 0.0104433	0.0014806	4.07373e-005	38	0.23
	λ	0.023701	-0.0230086, 0.0704105	0.0238314	0.000244879	6	0.29
	μ_{ξ}	0.0493188	-1.87695, 5.14385	0.982793	0.0099683	6	0.29
	σ_{ξ}^2	0.00493034	-0.00013768, 0.00999836	0.0025857	3.16784e-005	8	0.27
MSCI 欧洲 指数	κ	-0.022641	-0.0728653, 0.0275833	0.0256247	0.00104568	83	0.25
	α	1.43025	0.904425, 1.95608	0.26828	0.0065107	29	0.30
	δ^2	0.0441354	-0.0706715, 0.158942	0.0585749	0.0024969	90	0.25
	μ	6.19588	4.40463, 7.98713	0.913904	0.0182337	19	0.25
	β	0.281772	-0.13976, 0.703305	0.215068	0.00717191	56	0.27
	σ^2	0.0100177	0.00584066, 0.0141948	0.00213116	6.99E-05	54	0.23
	λ	0.0238783	-0.0232411, 0.0709978	0.0240405	0.000231919	6	0.29
	μ_{ξ}	0.0206013	-1.89848, 1.93968	0.979123	0.00881288	5	0.29
	σ_{ξ}^2	0.00491937	-1.97E-05, 0.00985841	0.00251992	3.11e-05	8	0.27
NYS E指 数	κ	-0.0369071	-0.0830144, 0.0092001	0.0235241	0.000810165	59	0.29
	α	1.52733	1.01546, 2.0392	0.261158	0.00545198	22	0.22
	δ^2	0.0141633	-0.0296882, 0.0580148	0.0223732	0.00068173	46	0.30
	μ	5.61715	3.99388, 7.24042	0.8282	0.0165839	20	0.23
	β	0.277976	-0.000870043, 0.556821	0.142268	0.00329863	27	0.20
	σ^2	0.00785827	0.00485826, 0.0108583	0.00153062	3.946e-005	23	0.24
	λ	0.0239834	-0.0227865, 0.0707533	0.0238622	0.000261007	6	0.28
	μ_{ξ}	0.0668608	-1.87103, 2.00475	0.988719	0.0094878	5	0.29
	σ_{ξ}^2	0.00487207	-8.111e-005, 0.00982525	0.00252713	2.9394e-005	7	0.22

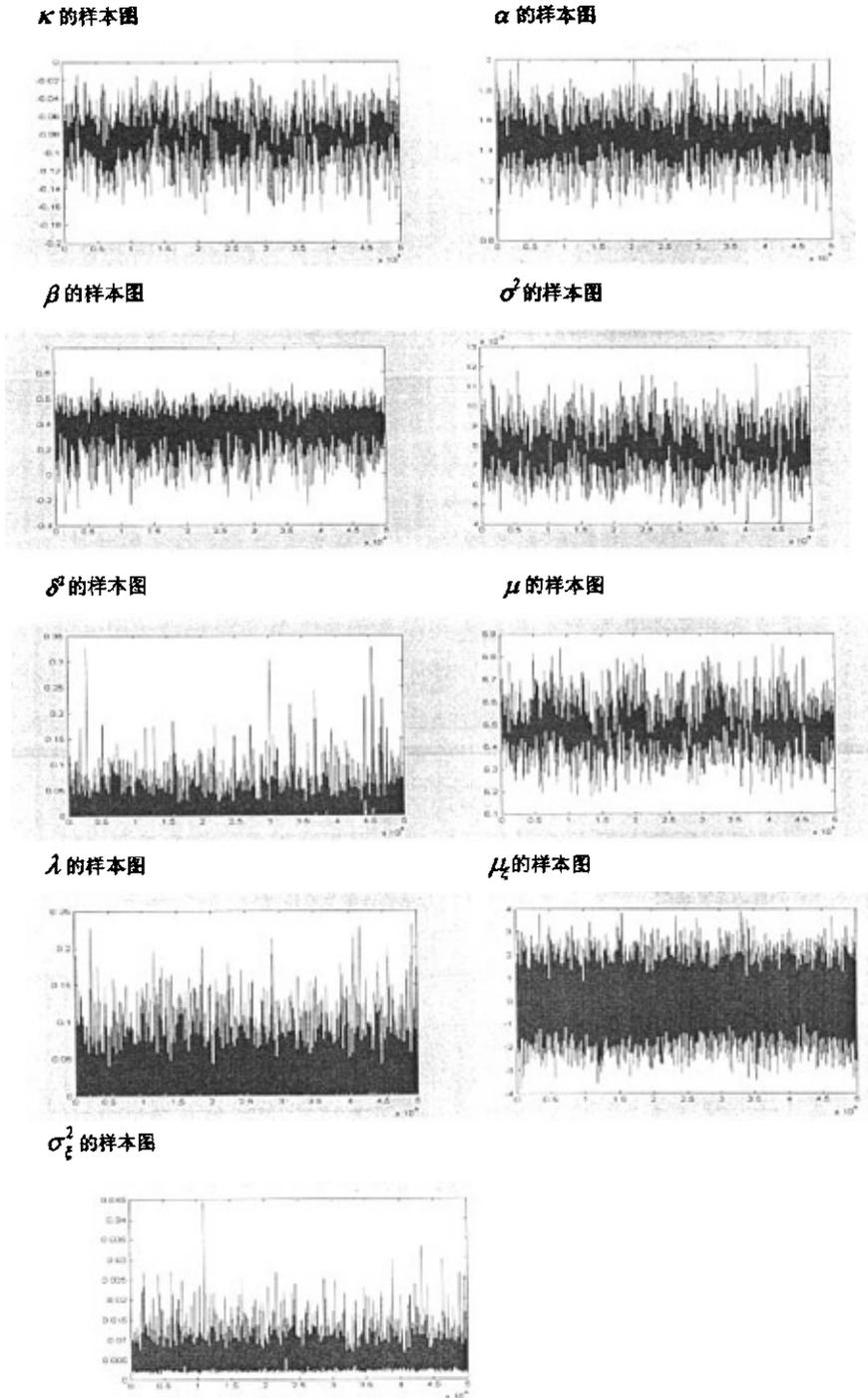


图 5.1: 各参数的样本图 (基于 MSCI 全球指数与 Milstein 离散方法)

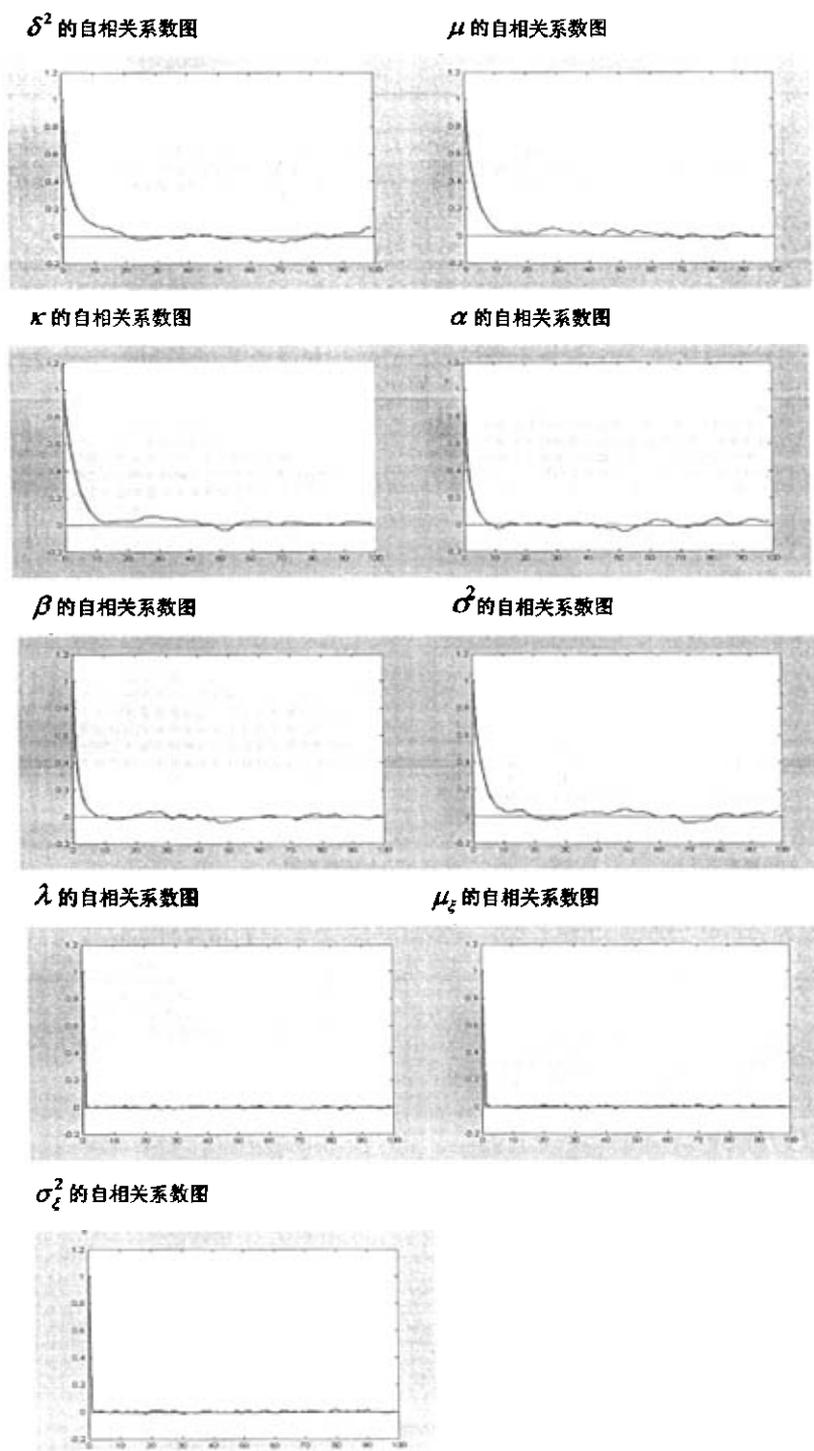


图 5.2: 各参数的自相关系数图 (基于 MSCI 全球指数与 Milstein 离散方法)

表 5.1 总结了个参数的遍历平均 (Mean)、标准离差 (SD)、95%的贝叶斯置信区间 (CI)、蒙特卡罗标准误差 (MCSE) 和 SIF；贝叶斯置信区间说明了各参数估计的有效性。

其中, SD 是参数抽样样本的标准离差; MCSE 是蒙特卡罗模拟标准误差, 通过公式 (3-33) - (3-37) 来计算; SIF 是模拟无效因子, SIF 是用来检验模拟样本的收敛效果的, 通过公式 (3-38) - (3-39) 来计算; 在计算 MCSE、SIF 时, 令 $f(\theta^{(i)}) = \theta^{(i)}$ 、 $m = 50$ 、 $n = 1000$ 。AC 为在 MCMC 算法中的样本接受概率。

利用文献[117]和[61]中介绍的有关参数先验密度的结果, 本文中参数的先验密度采用他们的假定, 各参数的先验分布分别为: $\kappa \sim N(0,10)$, $\alpha \sim \Gamma(1,20)$, $\delta^2 \sim \Gamma(0.05,20)$, $\mu \sim N(5,10)$, $\beta \sim U(-\alpha,\alpha)$, $\sigma^2 \sim IG(5,0.05)$, $\lambda \sim \beta(2,40)$, $\mu_{\xi} \sim N(0,10)$, $\sigma_{\xi}^2 \sim IG(5,20)$ 。

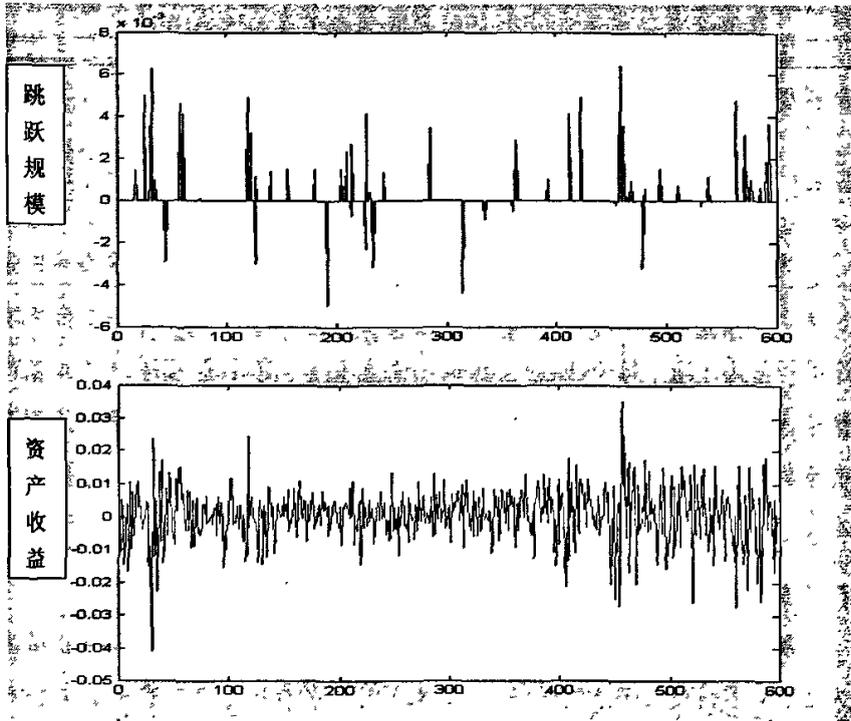


图 5.3: 跳跃规模与资产收益 (基于 MSCI 全球指数与 Milstein 离散方法)

从表 5.1 可以看出, 对于对称参数 β 来说, MSCI 欧洲指数和 NYSE 指数的 β

为是弱对称因为参数 β 不是显著不等于 0 的；但是他们的估计值都是大于 0，说明收益分布向右倾斜的，从而说明了资产收益分布的弱有偏性这一经验特征。另外，对于倾斜度参数 α 而言，三个指数的 α 都是显著不等于 0，且他们的值都在 1.5 左右，说明这些指数收益分布都存在着明显尖峰厚尾的特征。

接下来，我们来看一下设计跳跃的三个参数。首先跳跃强度 λ 都在 0.023 与 0.024 之间，由于采用的是周观测数据，从而收益跳跃发生的次数大约每年 1 到 2 次。跳跃规模的均值，MSCI 全球指数最大、NYSE 指数最小，跳跃规模的方差三种指数都有相同的后验均值。

至于隐含变量跳跃时间和跳跃规模可以见图 5.3，当跳跃发生时资产收益具有明显的波动，从而说明 HJD 模型很好地拟合了资产收益的分布。

5.4 抛物线跳跃扩散 (HJD) 模型的经验特征

首先通过 HJD 模型来模拟资产价格，并说明该模型所模拟的资产收益分布所体现的经验特征，同时，加入跳跃因子以后，说明 HJD 模型相对于抛物线扩散模型有哪些改进。

5.4.3 HJD 模型的经验特征

为了证明 HJD 的后验特征，利用前面在 MSCI 全球指数中所获得的参数抽样集合，将每个参数的抽样集合分解成 1000 组参数抽样值的样本集合（连续 50 个抽样值为一组，每个参数共 50000 个抽样值），使用每一组样本的均值作为参数值，模拟获得 2000 个收益序列，计算每一个收益序列的偏度和峰度以及自相关系数。表 5.2 列示了偏度、峰度和绝对收益的幂分别为 1、1.5、2 时自相关系数的均值；最后一行说明了峰度大于 3、偏度小于 0 等的比例以及 $|r|$ 的幂分别为 1、1.5、2 时自相关系数的均值中幂为 1 时最大的概率。

从表 5.2 我们可以看出，运用 HJD 模型模拟出来的资产收益序列的峰度均值为 22.53、且峰度大于 2 的概率为 0.998，这很好的体现了分布的高峰厚尾的特征。模拟的资产收益序列偏度系数为 0.182，说明跳跃因子的加入使资产收益的分布具有非对称性。另外，从收益的自相关系数来看， $|r|$ 具有微弱而持续的自相关性；

同时, 收益的绝对值在幂等于 1 时, 自相关系数较幂为 1.5 和 2 时大 (幂等于 1 时的自相关系数最大的概率都在 65% 以上), 所以通过 HJD 模型模拟出来的收益序列体现了泰勒效应。

表 5.2: HJD 模型的统计特征

	峰度	偏度	$ r , r ^{1.5}, r^2$ 的自相关系数		
			滞后 100	滞后 200	滞后 300
均值	22.53	0.182	0.145/0.103/0.072	0.095/0.073/0.048	0.055/0.043/0.025
概率	0.998	0.753	0.762	0.728	0.694

综上, 通过 HJD 模型的模拟实验, 发现 HJD 模型很好的体现了众多文献中提及的有关资产收益序列的经验特征: 尖峰厚尾、收益的持续弱自相关性以及泰勒效应等经验特征。另外, HJD 模型除了可以用来为资产价格建模, 还可以为研究期权定价、利率的期限结构以及其它衍生资产定价等方面提供方便。

5.4.4 HJD 模型与抛物线扩散 (HYD) 模型比较

对抛物线跳跃扩散模型与抛物线扩散模型进行比较的时候, 从理论上来说可以应用一些贝叶斯模型的选择方法进行比较, 例如似然比检验、贝叶斯因子以及 DIC 准则; 从实证分析的角度看, 可以利用这两种模型拟合数据的效果进行比较, 即比较这两种模型的残差平方和的大小, 较小的残差平方和说明对数据的拟合效果更好。本文在比较 HJD 与 HYD 模型的时候, 就是首先从实证的角度来比较这两种模型。

首先用这两个模型的拟合 MSCI 全球指数数据时的残差来进行比较, 在对这两个模型进行离散求残差时, 采用的是 Milstein 方法, 这两个模型的残差平方和分别为:

$$\begin{aligned} ESS_{HYD} &= 28385.2 \\ ESS_{HJD} &= 25877.6 \end{aligned} \tag{5-17}$$

可以看出 $ESS_{HJD} < ESS_{HYD}$, 也就是说 HJD 模型比 HYD 模型更好的拟合数据, 实际上 HJD 模型更能反映资产收益序列特征。

另外还可以使用似然比检验方法来判断这两个模型的优劣。具体方法可以

参见第四章中连续时间变结构 BS 模型与标准 BS 模型的似然比检验。

模型选择的零假设和备择假设分别为： H_0 ：抛物线跳跃扩散模型优于抛物线扩散模型， H_1 ：抛物线扩散模型优于抛物线跳跃扩散模型；这样如果零假设为真，似然比检验量为：

$$LR = 2 \ln \lambda = 2(\ln l(\theta_1) - \ln l(\theta_2)) \quad (5-18)$$

其中 $l(\theta_1)$ 和 $l(\theta_2)$ 分别表示抛物线跳跃扩散模型和抛物线扩散模型的似然函数值， θ_1 和 θ_2 分别为这两个模型的参数向量。

所采用的原始数据为1990-01-01至2000-12-31的MSCI全球指数，依据前面的估计结果，在零假设成立的条件下，这两个模型的似然比检验统计量的值为：

$$LR^* = 15.6541 \quad (5-19)$$

接下来令 $N = 99$ ，采用4.3.3中的方法计算99个 LR ，加上 LR^* 共有100个 LR 样本，这100个 LR 样本的分布图见图5.4。

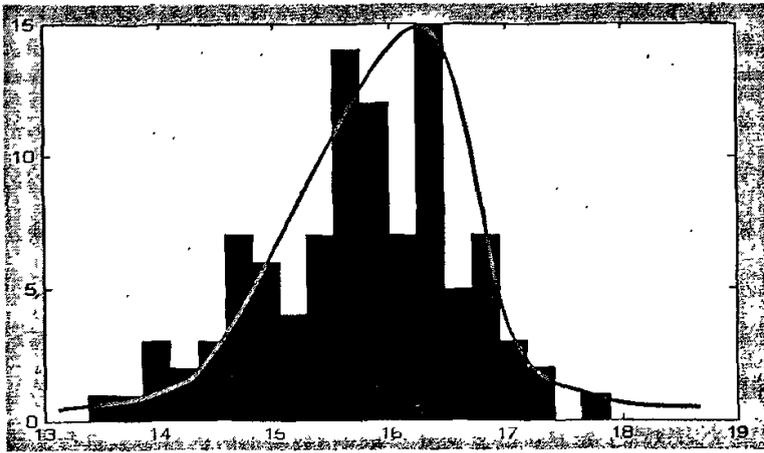


图5.4：在零假设成立的条件下 LR 的分布图

(注：图中直方图表示实际分布，曲线表示的近似拟合分布)

从 LR 的分布图中可以看出，在零假设成立的情况下， LR 的期望值是正值，大致在15—17之间，实际 LR^* 为15.6541，也远远大于Raftery (1996) 给出的接受零假设的标准值6。另外从图5.4可以直观的看出抛物线跳跃扩散模型优于抛物线扩散模型。这一结论还可以通过检验进一步说明：在100个 LR 值中，观测值 $LR^* = 15.6541$ ，由大到小排在第56位，即在零假设下有56%的 LR 模拟值高于实

际值。由 $LR^* = 15.6541$ 可知，只有当排序接近1或100时，该检验结果显著，将拒绝原假设。若将置信度定为95%，则拒绝域是 $[0, 3]$ 或 $[97, 100]$ ，而排序为56的 LR^* 远不在拒绝域内，说明根据样本可以接受零假设。通过模拟似然比检验的结果可以看出，抛物线跳跃扩散模型的拟合优度优于抛物线扩散模型。

5.5 正态逆高斯扩散模型的 MCMC 估计

5.5.1 正态逆高斯扩散模型

正态逆高斯模型 (Normal Inverse Gaussian Diffusion, NIG) 分布和抛物线分布都是广义抛物线分布的两种特例，Rydberg (1997) 指出：NIG 扩散过程较抛物线扩散过程具有更厚的尾部，从而就更有利于描述股票价格。

假定股票价格满足下面的过程：

$$S_t = \exp(\kappa t + X_t) \quad (5-20)$$

其中 κ 是股票价格对数随时间变化的系数， S_t 是股票价格， X_t 是股票价格对数除去趋势项之后的状态变量，其满足

$$X_t = X_0 + \int_0^t v(X_s) dW_s \quad (5-21)$$

通过伊藤公式，有

$$dS_t = S_t \left\{ \left[\kappa + \frac{1}{2} v^2(\log S_t - \kappa t) \right] dt + v(\log S_t - \kappa t) dW_t \right\} \quad (5-22)$$

注意到，如果函数 $v(\bullet)$ 是一个常数，关于 S_t 的模型是一个在 B-S 公式中使用的几何布朗运动。过程 X_t 是一个时变维纳过程，这样式 (5-22) 就是一个简单的 B-S 模型广义形式。关于 $\log S_t$ 的漂移项是一个时间的线性函数 κt 的假定不是必须的，做这样的假定是为了简单和尽可能的不改变几何布朗运动模型。在本文中，要讨论的就是当 $\log S_t$ 有一个确定性漂移时的情况。

通过选择合适的 $v(x)$ 可以得到几个有趣的模型，但是在此仅仅考虑一个特殊而常见的情况，并推广到一般情况。经验研究表明股票收益 $\log S_{t+\Delta} - \log S_t$ 的分布不是正态的，分布的尾部趋向于对数线性的。若选择式 (5-23) 的形式，就是 Bibby 和 Sørensen (1997) 的抛物线扩散过程。

$$v(x) = \sigma \exp \left[\frac{1}{2} \alpha \sqrt{\delta^2 + (x - \mu)^2} - \frac{1}{2} \beta (x - \mu) \right] \quad (5-23)$$

在本文中，选择 $v(x)$ 满足下式：

$$v(x) = \sigma \exp \left(-\frac{1}{2} \beta (x - \mu) \right) \sqrt{\frac{g_\delta(x - \mu)}{K_1(\alpha g_\delta(x - \mu))}} \quad (5-24)$$

其中 $g_\delta(x) = \sqrt{\delta^2 + x^2}$ ， $\alpha, \beta, \delta, \mu$ 是分布参数，他们满足 $\alpha > |\beta| \geq 0$ ， $\delta > 0$ 。 δ 是标度参数， μ 是位置参数， β 决定了对称性， α 决定了分布的倾斜度。

在用MCMC方法估计NIG模型时，首先总结一下在众多金融文献中关于资产收益序列的经验特征，定义 r_t 为股票收益，Ryden (1998) [164] 总结了 r_t 的下列动态性质：

- 1) r_t 是不相关的；
- 2) $|r_t|, r_t^2$ 的自相关函数缓慢衰减，衰减的速度较稳定 ARMA 自相关函数要慢的多；
- 3) 收益绝对值的自相关在幂等于 1 的时候最大，这就是泰勒效应，即 $\text{corr}(|r_t|, |r_{t-k}|) > \text{corr}(|r_t|^p, |r_{t-k}|^p)$ ， $p \neq 1$ 。
- 4) 收益常常表现出厚尾的边缘分布特征。

5.5.2 正态逆高斯扩散过程的离散化

对 NIG 模型采用 Milstein 方法近似：

$$X_{t+\Delta t} = X_t + \mu(X_t, \theta) \Delta t + \sigma(X_t, \theta) \Delta W_t + \frac{1}{2} \sigma(X_t, \theta) \frac{\partial \sigma(X_t, \theta)}{\partial X_t} \left[(\Delta W_t)^2 - \Delta t \right] \quad (5-25)$$

其中：

$$\frac{\partial \sigma(X_t, \theta)}{\partial x_t} = \frac{1}{2} \sigma \exp \left(-\frac{1}{2} \beta (x_t - \mu) \right) \frac{1}{\sqrt{g_\delta(x_t - \mu) K_1}} \left(-\beta g_\delta(x_t - \mu) + \frac{(x_t - \mu)}{g_\delta(x_t - \mu)} + \alpha (x_t - \mu) \left(\frac{1}{\alpha g_\delta(x_t - \mu)} + \frac{K_0}{K_1} \right) \right) \quad (5-26)$$

其中 $K_i = K_i(\alpha g_\delta(x - \mu))$ ， $i = 0, 1$ 。

(5-26) 式可以转换为：

$$X_{t+\Delta} - X_t - \mu(X_t, \theta)\Delta + g(X_t, \theta)\Delta = \sigma(X_t, \theta)\sqrt{\Delta}\varepsilon + g(X_t, \theta)\Delta\varepsilon^2 \quad (5-27)$$

其中 $g(X_t, \theta) = \frac{1}{2}\sigma(X_t, \theta)(\partial\sigma(X_t, \theta)/\partial X_t)$ ，令

$$a = \sigma(X_t, \theta)\sqrt{\Delta}, \quad b = g(X_t, \theta)\Delta \quad (5-28)$$

这样 (5-25) 表示为：

$$Y = a\varepsilon + b\varepsilon^2 = b\left[\left(\varepsilon + \frac{a}{2b}\right)^2 - \frac{a^2}{4b^2}\right] \quad (5-29)$$

依据 3.2.2 中的公式 (3-15) - (3-18) 可得对数似然函数为：

$$\log p_M(\theta|X) = \sum_{i=1}^n \left[\log \left\{ f\left(\frac{y_i}{b} + \frac{a^2}{4b^2}\right) \right\} - \log(b) \right] \quad (5-30)$$

这里， $y_i = x_t - x_{t-1} - \mu(x_{t-1}, \theta)\Delta + g(x_{t-1}, \theta)\Delta$ 。另外上面的似然对数可以直接做为 ML 估计中的近似似然函数，也可以做为 MCMC 方法中的改进近似对数后验。

5.5.3 实证研究

5.5.3.1 参数的估计结果

本节中，使用 MCMC 方法来估计 NIG 扩散模型，数据分别采用上海和深圳股市的周综合指数，即从 1992 年 5 月 3 日到 2004 年 4 月 30 日共 12 年的上海股市周综合指数和深圳股市周综合指数。

参数先验分布利用文献 [117] 中介绍的先验分布，各个参数的先验分别假定为： $\kappa \sim N(0, 10)$ ， $\alpha \sim \Gamma(1, 20)$ ， $\delta^2 \sim \Gamma(0.05, 20)$ ， $\mu \sim N(5, 10)$ ， $\beta \sim U(-\alpha, \alpha)$ ， $\sigma^2 \sim IG(5, 0.05)$ 。

按照随机游走 MH 算法对 NIG 扩散模型进行参数估计，每个参数共产生 60000 个的样本，剔除前面的 10000 个模拟样本，记录的模拟样本为后面的 50000 个。其中上海综指参数的样本曲线见图 5.5，上海和深圳综合指数的参数估计结果为表 5.4 所示。

表 5.4 总结了个参数的遍历平均 (Mean)、标准离差 (SD)、95% 的贝叶斯置信区间 (CI)、蒙特卡罗标准误差 (MCSE) 和 SIF；贝叶斯置信区间说明了各参数估计的有效性。

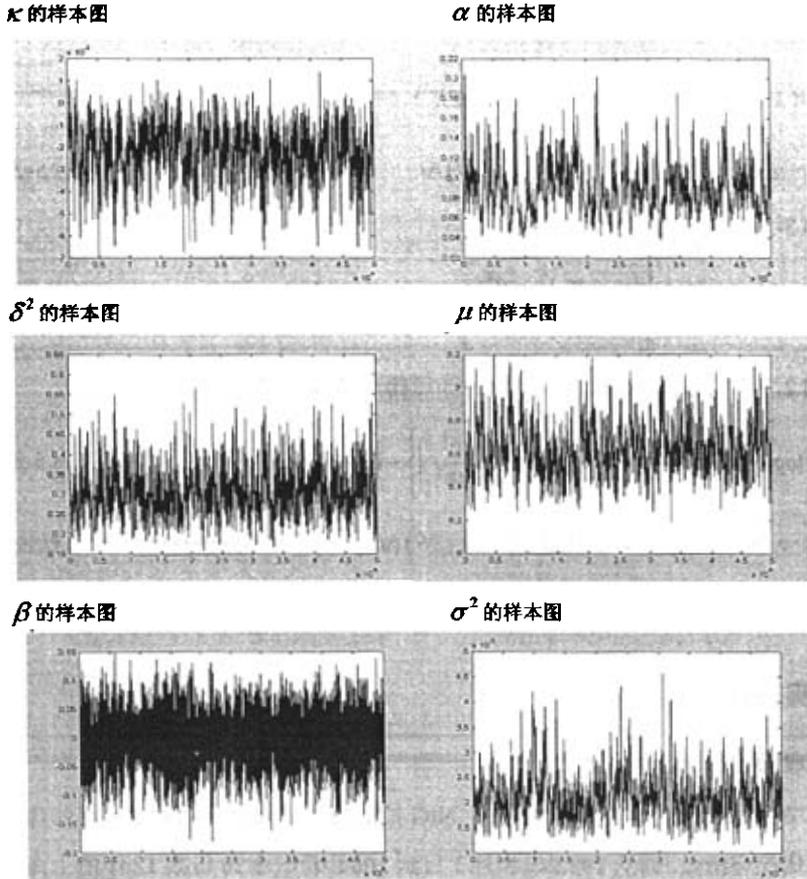


图 5.5：上海股市综指参数的样本曲线

表 5.4：Milstein 方法下 NIG 扩散模型的 MCMC 结果

数据	参数	均值	置信区间	SD	MCSE	SIF	AC
上海	κ	-0.003278	-0.003281,-0.00327	1.087e-006	5.214e-005	94.995	0.20
	α	0.206444	0.202463, 0.210426	0.00203148	0.00323852	98.136	0.25
	δ^2	0.255833	0.251503,0.260163	0.00220917	0.00232845	92.709	0.23
	μ	7.85132	7.82939, 7.87326	0.0111924	0.00579377	89.958	0.21
	β	0.107626	0.096752, 0.1185	0.00554802	0.00260517	61.164	0.29
	σ^2	0.001951	0.001950, 0.001951	1.426e-007	2.225e-005	93.591	0.22
深圳	κ	-0.002239	-0.002242, -0.00224	1.242e-006	6.944e-005	104.09	0.24
	α	0.088416	0.087283, 0.089550	0.00057847	0.00196124	102.46	0.24
	δ^2	0.304953	0.299149, 0.310756	0.00296096	0.0029884	100.80	0.21
	μ	8.63030	8.58676, 8.67385	0.0222164	0.0100352	106.64	0.24
	β	-0.006615	-0.011951, -0.00128	0.00272245	0.00119069	26.037	0.27
	σ^2	0.002117	0.002118,0.002118	1.865e-007	3.186e-005	102.09	0.22

其中，SD 是参数抽样样本的标准离差；MCSE 是蒙特卡罗模拟标准误差，通过公式 (3-33) — (3-37) 来计算；SIF 是模拟无效因子，SIF 是用来检验模拟样本的收敛效果的，通过公式 (3-38) — (3-39) 来计算；在计算 MCSE、SIF 时，令 $f(\theta^{(i)}) = \theta^{(i)}$ 、 $m = 50$ 、 $n = 1000$ 。AC 为在 MCMC 算法中的样本接受概率。

从表 5.4 可以看出，各个参数的估计值都是显著不为 0。注意到对称参数 β ，上海股市综指的 β 为正，即收益分布向右倾斜，而深圳股市综指的 β 为负，从而说明了资产收益分布的有偏性这一经验特征。另外，对于倾斜度参数 α 而言，上海股市综指为 0.206444 较深圳股市综指 0.088 的高，这说明上海股市综指和深圳股市综指的收益分布都不同程度的存在尖峰厚尾的特征，但是上海股市的尖峰厚尾特征更为明显。所以，NIG 模型很好的揭示了收益分布的尖峰厚尾、弱对称的经验特征。

5.5.3.2 先验分布的鲁棒性检查

为了检查先验分布对参数估计结果的鲁棒性，从两个方面来改变参数先验分布。首先，假定参数的先验分布就是上一小节中的形式，但是改变先验分布的参数，结果与原来非常相似。另外一方面，我们用均匀分布代替上一小节中的参数先验分布，采用上面使用的 MCMC 估计方法，得到的结果见表 5.5。

表 5.5：参数先验分布为均匀分布的 NIG 扩散模型的 MCMC 结果

数据	参数	均值	置信区间	SD	MCSE	SIF	AC
上海 股市 综合 指数	κ	-0.003478	-0.003480,-0.00347	1.067e-006	5.414e-005	124.99	0.21
	α	0.216444	0.212463, 0.220426	0.00233148	0.00353852	158.13	0.24
	δ^2	0.235833	0.231503,0.240163	0.00230917	0.00242845	122.70	0.22
	μ	7.95132	7.92939, 7.97326	0.0101924	0.00589377	149.95	0.22
	β	0.097626	0.086752, 0.1085	0.00534802	0.00360517	61.165	0.27
	σ^2	0.0020506	0.002050, 0.002050	1.226e-007	2.425e-005	143.59	0.23
深圳 股市 综合 指数	κ	-0.002439	-0.002441,-0.00243	1.252e-006	7.044e-005	164.09	0.23
	α	0.090416	0.089282, 0.090550	0.00056847	0.00206124	132.46	0.25
	δ^2	0.314953	0.309149, 0.320756	0.00306096	0.0030884	145.80	0.21
	μ	8.7303	8.68676, 8.77385	0.0242164	0.0110352	156.64	0.23
	β	-0.00861	-0.013950, -0.00427	0.00262245	0.00149069	46.037	0.26
	σ^2	0.003117	0.003117,0.003118	1.765e-007	3.258e-005	172.09	0.23

通过与表 5.4 的结果相比较，我们可以发现参数的遍历均值与置信区间没有明显的差异，也就是说参数的估计值不会随参数先验分布的改变而发生明显的变化。那么这也就说明参数的后验分布不会随着参数的先验分布发生变化。

5.6 本章小结

本章中，首先将跳跃因子引入到抛物线扩散模型中。接着，用基于 Milstein 的方法获得参数后验分布的离散密度，使用 MCMC 方法来估计抛物线跳跃扩散模型。最后用所得的估计值模拟了资产收益序列。通过上面的结果，不但发现 MCMC 方法较其他方法，如 ML 方法，更适合还有隐含变量的模型估计；还说明了抛物线跳跃扩散模型 (HJD) 能够很好的拟合资产收益的经验特征，如有偏、尖峰厚尾以及泰勒效应，即 HJD 模型更能反映资产收益的序列特征。最后还用 MCMC 方法估计了广义抛物线扩散模型的一种特例：正态逆高斯扩散模型。

第六章 总结和展望

计量经济学和金融工程学作为金融市场研究的主要理论工具，在近几十年得到了蓬勃的发展，其理论与方法广泛应用到经济与金融的各个领域。回顾计量经济学的发展历程，可以看到有以下一些特点：第一、实证金融与时间序列计量经济学的结合，使得金融计量学（Financial Econometrics）成为计量经济学的一个分支独立出来。学者们在这方面的研究主要围绕着资产定价、投资组合、风险管理等方面来进行，这些也仍将是未来的研究热点，尤其是其中的市场宏观结构与流动性、资产收益波动与相关性、投资组合与估价等方面有着更为广阔的前景。当然在这些研究中，传统的计量经济方法与贝叶斯计量经济方法都起着极为重要的作用；第二、预测的理论与方法研究对于宏观经济、微观经济、金融计量等领域来说，仍有着巨大的发展前景。其实，所有计量经济学方面的文章都与预测有着直接或间接的联系。预测可以分为结构型预测与非结构型预测，结构预测在微观经济学文献中较为多见；宏观经济学方面的一些研究则将非结构预测与结构预测相结合，金融计量学用的更多的仍是结构建模。由于有了大量的计量经济模型，学术界对预测的研究就已经取得很多的成果；第三、信息技术，特别是计算机技术的发展为计量经济理论与实践的发展提供了可靠的技术保障。对金融市场波动模型的研究及其参数估计而言，这一点显得尤为重要。在计算机尚未普及的时候，人们对于有关问题的研究与相关模型的估计，完全是借助于封闭的解析方法，限制了许多估计方法的使用，所以有很大的局限性；随着计算机技术的发展，学者们开始转向数值方法与模拟技术，使得模型的估计方法更加多样，模型的估计变得更加容易。第四，由于引入新的计算方法和工具，如：神经网络，遗传算法，数据挖掘、小波分析，信息技术等，大量金融数据的分析和计算工作能够进行，尤其对多维波动模型的建立和参数估计更为重要。

6.1 论文工作总结

6.1.1 连续时间模型及估计方法的总结

本文首先讨论了连续时间线性金融模型的进展情况,连续时间线性模型通过对BS模型添加随机波动因子或者跳跃因子从而体现不同的资产收益分布经验特征。接着总结了连续时间非线性模型的进展,相对于线性模型来说,由于非线性模型的复杂性导致这一类模型的进展相对缓慢,主要有两大类线性模型CEV模型和广义抛物线扩散模型。在连续时间模型不断发展的同时,模型的各种估计方法也在不断改进,本章的最后讨论了连续时间模型的估计方法,并总结了各种估计方法的优缺点和适用范围。

6.1.2 MCMC 方法的介绍及应用

连续时间模型在过去几十年中取得了巨大的进步,在其中关于连续时间模型的估计方法的不断更新是促进连续时间模型发展的巨大推动力,而其中基于贝叶斯原理的MCMC方法是最近十几年来金融计量经济学领域的热门研究领域,基于贝叶斯原理的MCMC估计方法是从正比于参数后验密度的密度函数中取样,从而确定参数的估计值。而其中的关键在于获得参数后验分布的正比形式,而这就需要贝叶斯统计分析。

本文首先介绍了有关贝叶斯统计方法,依据贝叶斯原理,参数的后验分布正比于模型的似然函数与参数的先验分布的乘积。其次讨论了如何采用Euler或Milstein方法对连续时间模型进行离散近似从而获得模型的近似似然函数。接着给出了MCMC算法和所取样本的收敛性检验,本文提出的基于MH算法的MCMC方法较使用Winbugs软件的Gibbs算法具有明显的优点,适用的范围更广;并使用C++语言开发出了能够解决含有隐含变量的连续时间模型的估计的MH算法。最后通过对双指数跳跃扩散模型的估计来验证基于MH算法的MCMC估计方法的有效性,并对沪深股市进行了实证分析。

6.1.3 连续时间变结构模型的研究

模型结构变化是社会经济系统中模型的基本特征,尤其在一个新兴金融市场

上, 结构变化是一个普遍存在的现象; 这种结构变化有可能是模型中参数的改变, 也可能是模型形式的变化, 基本原因在于外在因素的影响。为了更好的进行分析和预测, 因此需要充分考虑模型结构的变化。

本文提出了应用 MCMC 方法的连续时间变结构模型的单一变结构点的定位方法, 并用仿真试验验证了该方法的有效性。接着提出了连续时间多变结构点模型的变结构点定位方法; 该方法在确定变结构点位置的同时, 又能估计相应的模型参数。用该方法对上海股市综合指数的收益序列进行了变结构分析, 理论与实证结果表明该方法是有有效且可行的。

同时将该方法从 BS 模型扩展到了连续时间随机波动模型, 并给出了各参数和隐含波动过程的后验分布, 并给出了相应的 MCMC 算法。讨论了变结构点的检验, 同时应用该方法对连续时间随机波动模型进行了实证分析, 说明所提方法的可行性。

6.1.4 抛物线跳跃扩散模型的研究

Bibby 与 Sørensen (1997) 提出了用抛物线扩散模型来描述资产的收益分布, 之后 Tse、Zhang 和 Yu (2003) 用基于贝叶斯原理的 MCMC 方法对该模型的参数进行了估计。但是越来越多的实证研究显示汇率和股票市场经常收到外在因素的冲击, 从而导致金融产品市场价格的跳跃 (jump), 所以有必要把跳跃成分引入到抛物线扩散模型。

本文将跳跃因子引入到抛物线扩散模型中, 提出了抛物线跳跃扩散模型。接着, 用基于 Milstein 的方法获得参数后验分布的离散密度, 使用 MCMC 方法来估计抛物线跳跃扩散模型。并用所得的估计值模拟了资产收益序列。通过上面的结果, 不但发现 MCMC 方法较其他方法, 如 ML 方法, 更适合还有隐含变量的模型估计; 还说明了抛物线跳跃扩散模型 (HJD) 能够很好的拟合资产收益的经验特征, 如有偏、尖峰厚尾以及泰勒效应, 即 HJD 模型更能反映资产收益的序列特征; 最后用 MCMC 方法估计了正态逆高斯扩散模型。

6.2 研究展望

前面我们已经总结了本文的主要工作, 重点介绍了使用基于贝叶斯原理的

MCMC 方法来研究金融工程中资产收益的连续时间模型。虽然连续时间资产收益模型在过去 20 多年的研究中已经取得了巨大进步，但是还有众多的问题需要我们去研究和开拓。总的来说可以从这几个方面来考虑：

第一，模型的估计问题。连续时间模型是建立在市场存在“连续交易机会”的假设基础上，但是市场并非连续开放。因此，用什么样的数据来进行模型估计将是非常重要的，是采用收盘价还是实际交易价格将是模型估计的问题之一。另外，我们所获得的数据都是离散时间的，如何用离散时间数据来估计连续时间模型也将是一项挑战性很强的课题，即模型估计方法的选择问题。

第二，需要提出一类能包容现有的线性和非线性模型并扩展的综合模型，以及这一类模型的建模方法。对于连续时间模型的检验和选择问题，在理论上应用一些贝叶斯模型的选择方法进行比较，如似然比检验、贝叶斯因子以及 DIC 准则；从实证分析角度，利用不同的模型进行预测，并用预测的准确性来进行模型比较。

另外在寻找连续时间模型建模方法上，如何利用高频数据或超高频数据来进行模型的选择也将是一项挑战性的课题。

第三，连续时间模型持续性的探讨。波动持续性是金融市场一个重要现象，它反映了历史信息对波动影响的持续性和非线性特征，同时也体现了金融风险的持续性。在低频数据中通过收益的绝对值或平方来研究波动的持续性，在高频数据中，通常“已实现波动”（realized volatility）来探讨波动的持续性特征；那么在连续时间模型中如何来体现金融波动风险的持续性是需要研究的内容之一。

另外，Markowitz 于 1952 年建立的资产组合理论，首次对金融风险进行了定量刻画，即通过不同的资产组合来降低金融风险。利用协整的思想，讨论多元连续时间模型以及波动的协同持续性，并应用到动态资产组合的研究中。

第四，连续时间模型的变结构分析，主要研究连续时间随机波动模型的变结构问题。模型结构变化是社会经济系统中模型的基本特征，尤其在一个新兴金融市场上，结构变化是一个普遍存在的现象；这种结构变化有可能是模型中参数的改变，也可能是模型形式的变化，基本原因在于外在因素的影响。为了更好的进行分析和预测，因此要充分考虑模型结构的变化。本文仅仅考虑了最简单的 BS 模型和随机波动模型，考虑更能描述资产收益分布经验特征的复杂模型的变结构

问题尚需要进一步研究。

第五，金融市场作为一个整体、一个系统，其各部分或各个子系统之间必然存在着相互联系或相互影响，其波动会从一个市场传导到另一市场，即存在着波动传导性(transmission of volatility or volatility spillover)。Engle, Ito and Lin (1990)^[165]，Karolyi (1995)^[166]，Jeong (1999)^[167]，Brooks and Henry (2000)^[168]等用向量GARCH模型、向量SV模型作为反映多变量时间序列波动特征的模型应用于金融市场间的波动传导分析。而关于利用向量连续时间模型进行金融市场间波动传导分析的文献尚未见报道，这将一个具有挑战性的一项课题。

6.3 结束语

近些年来，随着金融市场的不断发展变化，金融计量学已经成为最活跃的领域之一。本文使用基于贝叶斯原理的MCMC方法对金融工程中资产收益的连续时间模型做了初步研究，深感连续时间模型的复杂，但我们深信在错综复杂的表面之下，却蕴含着事物发展的内在规律和内在机理。寻找金融市场的本质特征是我们追求的目标，探索金融市场的未知领域是我们不断学习的动力，随着金融计量学理论的不断深入和完善，金融计量学必将有着更为广阔的应用前景。

参考文献

-
- [1] Markowitz H. Portfolio Selection[J]. *Journal of Finance*, 1952, 7(1): 77-91
- [2] Sharp W F. Capital asset prices: A Theory of Market Equilibrium Under Conditions of Risk [J]. *Journal of Finance*. 1964, 19: 425-442
- [3] Litner J. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets [J]. *Review of Economics and Statistics*. 1965, 47: 13-37
- [4] Mossin J. Equilibrium in a Capital Asset Market [J]. *Econometrica*. 1966, 34: 768-83.
- [5] Black, F. and Scholes, M. The pricing of options and corporate liabilities[J], *Journal of Political Economy*. 1973, 81: 637-659
- [6] Ross, Stephen A. Arbitrage Theory of Capital Asset Pricing [J]. *Journal of Economic Theory*. 1976, 13: 341-60
- [7] Merton, Robert C. Theory of Rational Option Pricing[J]. *Bell Journal of Economics*. 1973, 4: 141-183
- [8] Bhattacharya, Sudipto, and George M. Constantinides, eds., Theory of valuation: Frontiers of modern financial theory, Volume 1. Rowman and Littlefield, Totowa, N.J., 1989.
- [9] Constantinides, George M., Theory of valuation: Overview and recent developments, in Sudipto Bhattacharya and George M. Constantinides, eds.: Theory of Valuation: Frontiers of Modern Financial Theory, Volume 1. Rowman and Littlefield, Totowa, N.J., 1989.
- [10] Harrison, J. M., Brownian Motion and Stochastic Flow Systems[M]. Wiley, New York, 1985.
- [11] Malliaris, A., Stochastic Methods in Economics and Finance[M]. North-Holland, Amsterdam, 1982.
- [12] Ingersoll, Jonathan, Theory of Financial Decision Making[M]. Rowman and Littlefield, Totowa, N.J., 1987.
- [13] Dothan, Michael, Prices in Financial Markets[M]. Oxford University Press, New York, 1990.

- [14] Duffie, Darrell, *Securities Markets: Stochastic Models*[M]. Academic Press, Boston, 1988.
- [15] Duffie, Darrell, *Dynamic Asset Pricing Theory*[M]. Princeton University Press, Princeton, N.J., 1996.
- [16] Karatzas, I., J. P. Lehoczky, and S. E. Shreve. Optimal portfolio and consumption decisions for a small investor on a finite horizon[J]. *SIAM Journal of Control and Optimization*, 1987, 25: 1157–1186.
- [17] Karatzas, I., and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, 113 ~Springer-Verlag, New York, 1988.
- [18] Cox, John C., and Chi-fu Huang, Optimal consumption and portfolio policies when asset prices follow a diffusion process[J]. *Journal of Economic Theory*, 1989a, 49: 33–83.
- [19] Cox, John C., and Chi-fu Huang, Options pricing theory and its applications, in Sudipto Bhattacharya and George M. Constantinides, eds.: *Theory of Valuation: Frontiers of Modern Financial Theory*, Volume 1 ~Rowman and Littlefield, Totowa, N.J., 1989b.
- [20] Duffie, Darrell, and Chi-fu Huang, Implementing Arrow–Debreu equilibria by continuous trading of few long-lived securities[J]. *Econometrica*, 1985, 53: 1337–1356.
- [21] Back, Kerry, Insider trading in continuous time[J], *Review of Financial Studies*. 1992, 5: 387–409.
- [22] Back, Kerry, Asymmetric information and options[J], *Review of Financial Studies*, 1993, 6: 435–472.
- [23] Xia, Yihong, Learning and predictability: The effect of parameter uncertainty on optimal dynamic asset allocation, Working paper, Anderson School of Management, UCLA, 1999.
- [24] Detemple, Jerome, Asset pricing in a production economy with incomplete information[J], *Journal of Finance*, 1986, 41: 383–391.
- [25] He, Hua, and N. D. Pearson, Consumption and portfolio with incomplete markets and shortsale constraints: The finite-dimension case[J], *Journal of Economic Theory*, 1991, 54: 259–304.
- [26] Veronesi, Pietro, How does information quality affect stock returns? [J]. *Journal of Finance*, forthcoming, 1999.

- [27] Ho, Thomas, and S. Lee, Term structure movements and the pricing of interest rate contingent claims[J], *Journal of Finance*, 1986, 41: 1011-1029.
- [28] Brennan, Michael J., and Eduardo Schwartz, Evaluating natural resource investments[J], *Journal of Business*, 1985, 58: 35-57
- [29] McDonald, Robert, and Dan Siegel, The value of waiting to invest[J], *Quarterly Journal of Economics*, 1986, 101: 707-728.
- [30] Duffie, Darrell, and Kenneth Singleton, Simulated moments estimation of Markov models of asset prices[J], *Econometrica*, 1993, 61: 929-952.
- [31] Ait-Sahalia Y. Maximum likelihood estimation of discretely sampled diffusions: a closed form approximation approach[J]. *Econometrica*. 2002,70(1):223-262
- [32] Singleton, Kenneth, Estimation of affine diffusion models based on the empirical characteristic function, *Working paper*, Graduate School of Business, Stanford University, 1999.
- [33] Lo, Andrew. Maximum likelihood estimation of generalized Ito processes with discretely sampled data [J]. *Econometric Theory*. 1988, 4:231-47
- [34] Hansen L P and Scheinkman J A. Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes[J]. *Econometrica*. 1995,63 (4) :767-804
- [35] Ait-Sahalia, Y. Nonparametric pricing of interest rate derivative securities. *Econometrica*.1996,64,527-560.
- [36] Gallant, AR and Tauchen, GE. Which Moments to Match? [J]. *Econometric Theory* . 1996, 12(4): 657--681
- [37] Stanton, R. A nonparametric model of term structure dynamics and the market price of interest rate risk[J]. *Journal of Finance*.1997, 52:1973-2002
- [38] Bandi F and Phillips P C B. Econometric estimation of diffusion models. Unpublished working paper, 1998, Yale University.
- [39] Bandi F and Phillips P C B. Accelerated Asymptotics for Diffusion Model Estimation. Unpublished working paper, 2000, Yale University.
- [40] Chacko, George, and Luis Viceira, Spectral GMM estimation of continuous time processes, Working paper, Graduate School of Business, 1999, Harvard University
- [41] Tim Bollerslev, Hao Zhou. Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility.
- [42] 王春峰, 万海辉, 李刚. 基于MCMC的金融市场风险VaR的估计[J]. 管理科学学报. 2000, 3(2): 54-62

- [43] 朱崇军. MCMC 样本确定的后验密度的收敛性[J]. 数学杂志. 2002, 22(3): 345-348
- [44] 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, 6(4): 63-72
- [45] 林静, 韩玉启, 朱慧明. 基于 MCMC 稳态模拟的指数回归模型及其应用[J]. 运筹与管理. 2005, 14(4): 95-100
- [46] 孟利锋, 张世英, 何信. 厚尾 SV 模型的贝叶斯分析及其应用研究[J]. 西北农林科技大学学报. 2003 年 6 月
- [47] 孟利锋. 随机波动模型及其建模方法研究: [博士学位论文], 天津: 天津大学, 2004
- [48] 许启发, 张世英. Box-Cox-SV 模型及其对金融时间序列刻画能力研究[J]. 系统工程学报, 2005, 20(4): 352-360
- [49] 黄大海, 郑丕谔. 基于 MCMC 方法的两类波动模型的应用比较[J]. 系统工程学报. 2004, 19(4): 413-417
- [50] 吴振翔, 缪柏其. 扩散方程参数的 MCMC 估计[J]. 中国科学院研究生院学报. 2004, 21(3): 310-314
- [51] 胡素华, 张世英, 张彤. 资产价格的抛物线跳跃扩散模型[J]. 系统工程理论与实践. 2006, 26(3): 1-10
- [52] 胡素华, 张彤, 张世英. NIG 扩散模型的 MCMC 估计[J]. 系统工程理论方法应用. 2006, 15(2): 133-138
- [53] 胡素华, 张世英, 张彤. 金融工程中资产收益连续时间模型评述[J]. 中国管理科学. 2006, 14(2): 24-32
- [54] 胡素华, 张世英, 张彤. 双指数跳跃扩散模型 MCMC 估计[J]. 系统工程学报. 2006, 21(2): 113-118
- [55] 周彦, 张世英. 基于 MCMC 方法的连续时间 SV 模型建模研究. 工业工程. 待刊
- [56] 周彦. 连续时间随机波动建模及其在中国股市的研究: [硕士学位论文], 天津: 天津大学, 2006
- [57] Merton, Robert C. Lifetime portfolio selection under uncertainty: The continuous time case[J]. *Review of Economics and Statistics*. 1969, 51: 247-257
- [58] Scott, Louis. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of Fourier inversion methods [J]. *Mathematical Finance*. 1997, 7: 413-426

- [59] 杨智元,陈浪南.基于跳跃过程的指数期权模型.经济研究[J].2001,2:61-66
- [60] Hull, John, and Alan White. The impact of default risk on the prices of options and other derivative securities [J]. *Journal of Banking and Finance*. 1995, 19:299-322
- [61] 郭磊,吴冲锋.基于价值分解的 CAPM “异常现象” 理论解释.中国管理科学[J].2005,1:14-18
- [62] Hull, John C. and Alan White. The pricing interest rate caps and floors using the Hull-White model [J].*Journal of Financial Engineering*. 1994, 2: 287-296.
- [63] Campbell, John Y., and Luis Viceira. Consumption and portfolio decisions when expected returns are time varying [J]. *Quarterly Journal of Economics*. 1999,114: 433-495.
- [64] 郭文旌.跳跃扩散股价的最优投资组合选择.控制理论与应用[J]. 2005,2: 171-176
- [65] Kou S G. A jump-diffusion model for option pricing[J]. *Management Science*. 2002, 48(8): 1086-1101
- [66] 沈根祥. 股票收益随机波动模型研究. 中国管理科学[J].2003,2:16-19
- [67] Merton, Robert C. Option Pricing when Underlying Stock Returns Are Discontinuous[J]. *Journal of Financial Economics*. 1976,3: 125-144
- [68] Hull, J.C. and White, A. The pricing of options on assets with stochastic volatilities[J]. *Journal of Finance*. 1987, 42: 281-300.
- [69] Heston, Steven L. A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options [J]. *The Review of Financial Studies*. 1993,6: 327-343
- [70] Bates, D. S. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options[J]. *The Review of Financial Studies*.1996, 9: 69-107
- [71] Andersen, T. G, L. Benzoni, and J. Lund. An Empirical Investigation of Continuous Time Models for Equity Returns[J]. *Journal of Finance*.2002, 57(3):1239-1284.
- [72] Duffie, D., Jun Pan, and K. J. Singleton. Transform Analysis and Asset Pricing for Affine Jump-Diffusions[J]. *Econometrica*.2000, 68: 1343-1376
- [73] Eraker, B., M. J. Johannes, and N. G. Polson. The Impact of Jumps in Returns and Volatility [J]. *Journal of Finance*. 2003,53(3): 1269-1300.
- [74] Eraker, B. Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices [J]. *Journal of Finance*. 2004,59(3): 1367-1404

- [75] Melino Angelo, and Stuart M. Turnbull. Pricing Foreign Currency Options with Stochastic Volatility[J]. *Journal of Econometrics*. 1990,45: 239-265
- [76] Benzoni, L. Pricing Options under Stochastic Volatility: an Econometric Analysis [J]. *Working Paper*. University of Minnesota.1998
- [77] Eberlein, E. and Keller, U. Hyperbolic distributions in finance[J]. *Bernoulli*, 1995, 1:281- 299
- [78] Barndorff-Nielsen, O. E. and Shephard, N. Modeling by Levy processes for financial econometrics, in O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (eds), *Levy Processes -Theory and Applications*, Birkh user, Basel, 2001: 283-318
- [79] Schoutens, W. *Levy Processes in Finance - Pricing Financial Derivatives* [J]. *Wiley Series in Probability and Statistics*. John Wiley and Sons, Chichester. 2003
- [80] Barndorff-Nielsen, O. E. and Shephard, N. *Continuous Time Approach to Financial Volatility*[M]. Cambridge University Press. 2005, To appear
- [81] Rydberg, T. H. The normal inverse Gaussian Levy process: Simulation and approximation [J]. *Communication in Statistics: Stochastic Models*. 1997,13: 887-910
- [82] Bibby B M and Sørensen M. A hyperbolic diffusion model for stock prices [J]. *Finance and Stochastic*. 1997, 1: 25-41
- [83] Rydberg T H. Generalized hyperbolic diffusion processes with applications in finance [J]. *Mathematical Finance*. 1999, 9(2): 183—201
- [84] Cont R and Tankov P. *Financial Modeling with Jump Processes*. Chapman & Hall/CRC, London. 2004
- [85] Emmer S and Kl uppelberg C. Optimal portfolios when stock prices follow an exponential Levy process[J]. *Finance and Stochastics*. 2004,8:17-44
- [86] Cox, J. C. and S. Ross. The Valuation of Options for Alternative Stochastic Processes[J], *Journal of Financial Economics*.1976,3:145-166.
- [87] Cox J C. The Constant Elasticity of Variance Option Pricing Model [J]. *The Journal of Portfolio Management*. 1996,2:15-17
- [88] Vasicek Oldrich. An equilibrium characterization of the term structure[J], *Journal of Financial Economics*. 1977, 5, 177-18
- [89] Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross. A Theory of the Term Structure of Interest Rates[J]. *Econometrica*. 1985,53: 385-407
- [90] Chan K C, G Andrew Karolyi, Francis A Longstaff, Anthony B. Sanders An empirical comparison of alternative models of the short term interest rate[J]. *Journal of Finance*. 1992, 1209-1227

- [91] Nelson, Daniel B. Modeling stock market volatility changes[J], *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, 1989, 93-98.
- [92] Nelson Daniel B. ARCH models as diffusion approximations[J]. *Journal of Econometrics*. 1990, 45, 7-39.
- [93] Nelson Daniel B. Conditional heteroskedasticity in asset returns: A new approach[J], *Econometrica*. 1991, 59, 347-370.
- [94] Liu, J. Generalized Method of Moments Estimation of Affine Diffusion Processes. Working Paper, Graduate School of Business, Stanford University, 1997
- [95] Pearson Niel and Tongsheng Sun. Exploiting the conditional density in estimating the term structure: An application to the CIR model[J], *Journal of Finance*. 1994, 49:1279-1304.
- [96] Ait-Sahalia, Yacine. Transition densities for interest rates and other nonlinear diffusions[J]. *Journal of Finance*. 1999b, 54, 499-547.
- [97] Gallant, Tauchen, Efficient method of moments, Working Paper, University of North Carolina, 2001
- [98] Broze, Laurence, Olivier Scaillet, and Jean Michel Zakoian. Quasi-indirect inference for diffusion processes[J]. *Econometric Theory*. 1998, 14:161-186
- [99] Brandt, Michael, and Pedro Santa-Clara. Simulated likelihood estimation of multivariate diffusions with an application to interest rates and exchange rates with stochastic volatility, *Working paper*, 1999, UCLA
- [100] Gallant, A. Ronald, and George Tauchen. Estimation of continuous-time models for stock returns and interest rates[J]. *Macroeconomic Dynamics*. 1997a, 1, 135-168.
- [101] Gallant, A. Ronald, and George Tauchen. Reprojecting partially observed systems with application to interest rate diffusions, *Journal of American Statistical Association*. 1998, 93, 1-24
- [102] Andersen, Torben G., and Jesper, Lund. Estimating continuous-time stochastic volatility models of the short-term interest rate[J]. *Journal of Econometrics*. 1997, 77:343-377
- [103] Benzoni, Luca. Pricing options under stochastic volatility: An econometric analysis. *Working paper*, 1999, University of Minnesota
- [104] Parzen, E. On estimation of a probability density function and mode[J], *The Annals of Mathematical Statistics*, 1962, 33: 1065-1076

- [105] Singleton, K. J., Estimation of affine asset pricing models using the empirical characteristic function[J], *Journal of Econometrics*, 2001, 102: 111-141.
- [106] Chacko, G; Viceira, L. M., Spectral GMM estimation of continuous-time processes[J], *Journal of Econometrics*, 2003, 116: 259-292
- [107] Bandi, F. and T.H. Nguyen (2000a), Fully nonparametric estimators for diffusions: a small sample analysis, *Working Paper*, Graduate School of Business, The University of Chicago
- [108] Jiang, G.J. and J.L. Knight, A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model[J], *Econometric Theory*.1997,13: 615-645
- [109] Chapman D and Pearson N. Is the short rate drift actually nonlinear? [J] *Journal of Finance*. 2000, 55: 355-388.
- [110] Ait-Sahalia Y. Testing continuous-time models of the spot interest rate[J]. *Review of Financial Studies*. 1996, 9:385-426.
- [111] Jiang, G.J. and J.L. Knight. Estimation of continuous-time processes via the empirical characteristic function[J]. *Journal of Business and Economic Statistics*. 2002, 20:198-212
- [112] Pritsker, Matt. Nonparametric density estimation and tests of continuous time interest rate models[J]. *Review of Financial Studies*. 1998,11, 449-487.
- [113] Pedersen Asger. A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations [J]. *Scandinavian Journal of Statistics* . 1995. 22(1):55-71
- [114] Eraker, Bjorn. MCMC Analysis of Diffusion Models with Application to Finance [J]. *Journal of Business and Economic Statistics*. 2001, 19: 177-191.
- [115] Jones, Christopher S. Bayesian Estimation of Continuous-Time Finance Models [J]. *Working Paper*. Rochester University. 1998
- [116] Elerian, Ola, Siddhartha Chib, and Neil Shephard. Likelihood Inference for Discretely Observed Non-linear Diffusions [J]. *Econometrica*. 2001,69: 959-993
- [117] Y K Tse、Xibin Zhang and Jun Yu. Estimation of hyperbolic diffusion using the Markov chain Monte Carlo method [J]. *Quantitative Finance* .2003.3(1):1-12
- [118] 孟利锋, 张世英, 何信.具有杠杆效应 SV 模型的贝叶斯分析及其应用[J]. *系统工程*.2004 年 3 月。
- [119] Gelfand, A. E., and A. F. M. Smith, Sampling-based approaches to calculating marginal densities[J], *Journal of American Statistic Association*, 1990, 85: 398-409

-
- [120] Metropolis, N., and S. Ulam, The Monte Carlo method[J], *Journal of American Statistic Association*, 1949, 44: 335-341
- [121] Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. Teller, and H. Teller, Equations of state calculations by fast computing machines[J], *Journal of Chemical Physics*, 1953, 21:1087-1091
- [122] Koopman, S. J., N. Shephard, and J. A. Doornik, Statistical algorithms for models in state space using `ssfPack2.2`[J], *Econometrics Journal*, 1999, 2:107-160
- [123] Elerian O. A note on the existence of a closed-form conditional transition density for the Milstein scheme [J]. *Economics Discussion Paper*. 1998-W18 Nuffield College, Oxford
- [124] Tierney L. Markov chains for exploring posterior distributions[J]. *Annals of Statistics*. 1994, 22: 1701-1762
- [125] Hasting, W. K. Monte Carlo sampling methods using Markov Chains and their applications [J]. *Biometrika*. 1970, 57: 97-109
- [126] Geyer, C. J., Practical Markov chain Monte Carlo[J], *Statistical Science*, 1992, 7: 473-511
- [127] Gelman, A., and D. B. Rubin, Inferences from iterative simulation using multiple sequences[J], *Statistical Science* 1992, 7:457-511
- [128] Raftery, A. E., and S. Lewis, Comment: One long run with diagnostics: Implementation strategies for Markov Chain Monte Carlo[J], *Statistical Science*, 1992b, 7: 493-497
- [129] Chib S and Greenberg E. Understanding the Metropolis–Hastings algorithm [J]. *American Statistician*. 1995, 49:327–35
- [130] Chib S and Greenberg E. Bayes inference for regression models with ARMA (p,q) errors[J]. *Journal of Econometrics*. 1994,64:183:206
- [131] Roberts G O, Gelman A and Gilks W R. Weak convergence and optimal scaling of random walking metropolis algorithms. Technical Report. 1994, University of Cambridge
- [132] Muller A. A general approach to posterior integration and Gibbs sampling. 1993, manuscript
- [133] Roberts G O. Markov chain concepts related to sampling algorithms. *Markov Chain Monte Carlo in Practice*, ed W R Gilks, S Richardson and D J Spiegelhalter (London: Chapman and Hall) .1996:45–57

- [134] Kim S, Shephard N and Chibs. Stochastic volatility: likelihood inference and comparison with ARCH models [J]. *Review of Economic Studies*. 1998, 65(2): 361-93
- [135] Meyer R and J Yu. BUGS for a Bayesian analysis of stochastic volatility models [J]. *Econometrics Journal* . 2000, 3(2):198-215
- [136] Gilks W R, Richardson S and Spiegelhatler D J. Introducing Markov chain Monte Carlo *Markov Chain Monte Carlo in Practice* ed W R Gilks, S Richardson and D J Spiegelhalter (London: Chapman and Hall) .1996, 45-57
- [137] Fama E F. The behavior of stock market prices[J]. *Journal of Business*. 1965,38:34-105
- [138] Mandelbrot B. The variation of certain speculative prices[J]. *Journal of Business*. 1963,36: 394-419.
- [139] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation[J]. *Econometrica*.1982,50:987-1007.
- [140] Bollerslev T. Generalized autoregressive conditional heteroskedasticity[J]. *Journal of Econometrics*. 1986,31:307-327.
- [141] Kim D and Kon S J. Sequential parameter nonstationarity in stock market returns[J]. *Review of Quantitative Finance and Accounting*. 1996, 6:103-131.
- [142] Lamoureux C G and Lastrapes W D. Persistence in variance, structural change, and the GARCH model[J]. *Journal of Business and Economic Statistics*. 1990,8:225-234
- [143] Andrews D W K and Fair R. Inference in econometric models with structure change[J]. *Review of Economic Studies*. 1988, 55: 615-640
- [144] Ghysels E and Hall A. A test for structural stability of Euler conditions parameters estimated via the generalized methods of moments estimator[J]. *International Economic Review*. 1990, 31:355-364
- [145] Bai Jushan. Likelihood ratio tests for multiple structural changes[J]. *Journal of Econometrics*. 1999, 91: 299-323
- [146] Kim D and Kon S J. Structural change and time dependence in models of stock Returns [J]. *Journal of Empirical Finance*. 1999, 6:283-308.
- [147] Gagliardini, P., Trojani, F., Urga, G. Robust GMM tests for structural breaks. *Journal of Econometrics*, 10.1016/j.jeconom.2004.09.006.
- [148] 柯珂, 张世英. ARCH模型的诊断分析[J]. *管理科学学报*, 2001, (4)2: 12-18.

- [149] 黄违洪, 张世英. 模型结构变化点检测法[J]. 应用数学学报, 1987, 10(3):267-275
- [150] 苏卫东, 张世英. 变截距 SV 模型及其在上海股市的实证[J]. 系统工程理论方法应用. 2003,1:57-63
- [151] 徐梅, 张世英. 基于小波变换的 LMSV 模型变结构研究[J]. 系统工程学报. 2005,20 (3) :232-238
- [152] Fox A J, Outliers in time series[J]. J. Roy. Statist. Soc. B, 1972(48):39-47
- [153] 韦博成, 鲁国斌, 史建清, 统计诊断引论[M]. 南京: 东南大学出版社, 1991
- [154] Kim D. A Bayesian significance test on the stationarity of regression parameters[J]. *Biometrika*. 1991, 78: 667—675.
- [155] Box G E P and Tiao G C. Bayesian inference in statistical analysis. Reading Mass : Addison Wesley, 1973
- [156] 白昆, 张世英. SV模型的变结构研究及应用[J]. 系统工程. 2003,2:7-11
- [157] Black, F. and Scholes, M. The pricing of options and corporate liabilities[J], *Journal of Political Economy*. 1973,81:637-659
- [158] Perreauh L, Bemier J, Parent E. Bayesian change point analysis in hydrometeorological time series[J]. *Journal of Hydrology*, 2000, 235(3): 221—241.
- [159] Rafterty A E. Approximate bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika*. 1986, 83: 251-266
- [160] 余素红, 张世英. SV 和 GARCH 模型拟合优度的似然比检验[J]. 系统工程学报. 2004, 19(6): 625-629
- [161] Kuchler, U., Neumann, K., Sørensen M., A.Streller. Stock returns and hyperbolic distributions. *Discussion Paper No. 23*, Sonderforschungsbereich 373. Humboldt Universit at zu Berlin. 1994
- [162] E.F. Fama. The behaviour of stock market prices[J]. *Journal of Business*. 1965. 38(1):34—105
- [163] S.Taylor. Modelling financial time series[M]. Chichester: John Wiley & Sons. 1986
- [164] Ryden T, Terasverta T and Asbrink S., Stylized facts of daily return series and the Hidden Markov model[J], *J. Appl.Econometrics*,1998, 13,217-44
- [165] Engle Robert F., Takatoshi Ito, and Wen-Ling Lin. Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market[J]. *Econometrica*. 1990,58, 525-542

- [166] Karolyi G.A. A multivariate GARCH model of international transmissions of stock returns and volatility: the case of the United States and Canada[J]. *Journal of Business and Economic Statistics*. 1995, 13, 11-25.
- [167] Jeong J.G. Cross-border transmission of stock price volatility: evidence from the overlapping trading hours[J]. *Global Finance Journal*. 1999,10, 53-70
- [168] Brooks C., O.T.Henry. Linear and non-linear transmission of equity return volatility : evidence from the US, Japan and Australia[J]. *Economic Modeling*. 2000, 17, 497-513

发表论文和科研情况说明

发表的论文:

- [1] 胡素华、张世英、张彤, “资产价格的抛物线跳跃扩散模型”, 《系统工程理论与实践》, 2006年3月 (EI: 06189857426)
- [2] 胡素华、张世英、张彤, “金融工程中资产收益连续时间模型评述”, 《中国管理科学》, 2006年4月
- [3] 胡素华、张世英、张彤, “双指数跳跃扩散模型 MCMC 估计”, 《系统工程学报》, 2006年4月
- [4] 胡素华、张彤、张世英, “NIG 扩散模型的 MCMC 估计”, 《系统工程理论方法应用》, 2006年4月

参与的科研项目:

- 1、参加张彤副教授主持的国家自然科学基金资助项目《基于连续时间模型的贝叶斯分析》(No:70301006)
- 2、参加张世英教授主持的国家自然科学基金资助项目《多变量矩序列长期均衡关系及动态金融风险规避策略研究》(No:70471050)

致 谢

在论文完成之际，首先向我的恩师张世英教授表示由衷的感谢！三年来，先生对我的论文研究工作倾注了大量的心血！让学生深深地懂得：什么是学问和如何做学问，这是学生以后从事学术研究最珍贵的财富。三年中，先生言传身教、诲人不倦、严谨治学、为人谦虚，让我深切地体会到名师风范和高风亮节，这是学生一辈子努力的目标。先生对待科学的态度、从事科学研究的精神始终鞭策着我，让学生不敢有丝毫的懈怠。在先生的悉心指导下，我得以站在学科最前沿，顺利完成博士学业。谨此，向恩师致以深深的敬意和感激！

张彤副教授悉心指导我们完成了实验室的科研工作，在学习上和生活中都给予了我很大的关心和帮助，在此向张彤老师表示衷心的感谢。

张喜彬副教授对我论文的选题都提出了许多的宝贵意见，在此表示衷心的感谢。

感谢许启发、郭彬同学，一起度过了三年美好的求学时光！

感谢徐梅、孟利锋、何信、韦艳华、刘惠萍、张庶萍、郭名媛、唐勇、张瑞锋、战雪丽、耿克红、李松臣博士、王晓天博士后以及所有参加学术讨论的同学！是你们让我的学术研究得以完善。感谢管理学院03春博班的同学们，三年的时间里与你们在一起学习、生活、交流，这将是一段美好的回忆！

感谢我的父母、岳父母及家人给予我无私的爱和对我继续求学给予的所有帮助！尤其感谢我的爱人赵绮黎女士给我学业上的帮助和三年中默默承担了大量的家务和照顾年幼的儿子！