

学校编码: 10384

分类号\_\_\_\_\_密级\_\_\_\_\_

学号: B200442020

UDC \_\_\_\_\_

厦 门 大 学

博 士 学 位 论 文

金融资产动态相关性方法及应用研究

Methodical and Applicable Research about Dynamic  
Correlation of Financial Assets

张 蕾

指导教师姓名: 郑振龙 教授

专业名称: 金融工程

论文提交日期: 2007年 4 月

论文答辩时间: 2007年 月

学位授予日期: 2007年 月

答辩委员会主席: \_\_\_\_\_

评 阅 人: \_\_\_\_\_

2007年 4 月

金融资产动态相关性方法及应用研究

张蕾

指导教师: 郑振龙教授

厦门大学



# 厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日



## 论文摘要

在许多金融计量问题中，例如衍生产品定价、风险管理、套期保值和最优投资组合选择等，模拟和预测二阶矩的相关性和各个市场收益之间波动性的动态相关关系有很重要的意义。目前研究动态相关关系的角度有三类，第一类是利用多维时间序列的角度进行研究，第二类是利用多维随机过程进行研究，第三类是利用 Copula 理论研究随机变量之间的相关程度。本文主要从多维时间序列的角度对相关分析进行理论方法的研究和实证分析。

论文第一章对本文的选题意义、方法的创新和缺陷进行了总结。论文的第二章回顾了多维时间序列相关性模型的发展历程以及估计方法和诊断检验，并对各种模型的优缺点和适用性进行了比较评述。论文的第三章和第四章运用各种多维时间序列进行预测分析和风险分析，实证结果表明，同其他模型相比，用 ADCC 模型拟合中国股指收益率的方差协方差矩阵效果较好，这个结果可以为资产配置决策、风险管理提供理论性的指导。

论文的第五章提出了一种处理高维金融时间序列的新方法—基于 Cholesky 分解方法的 SCC（序列条件相关）方法，并进行理论研究和实证分析。近年来出现了大量多维 GARCH 模型来模拟资产组合的波动性及相关性，但是在估计多维 GARCH 模型中仍存在着不尽如人意的地方，资产数量过多困扰着估计时的最优化问题，即使随着模型的逐渐改善，参数估计仍然十分困难，因此许多文献都停留在二维或三维 GARCH 模型的估计，很难推广到高维 GARCH 模型。本文提出了一种处理高维相关矩阵的估计方法—SCC 方法，此方法非常灵活，可以把高维相关矩阵的估计问题转化为二维相关矩阵的估计，并且允许二维相关矩阵的估计采用各种灵活的二维 GARCH 模型或其他新的方法进行分别估计，然后再组合为一个高维相关矩阵。这种新方法未来的发展不仅仅局限于和二维 GARCH 模型相结合，它也可以与任何描述二维动态相关性的最新理论，譬如动态的 Copula 理论相结合，从而更好地解决高维动态相关性的问题，是一种理论的突破。

论文的最后一章对全文进行了总结，并提出了今后所做的工作，和未来关于动态相关性发展的方向。

总之，论文研究了动态相关性的预测以及在风险上的应用，同时针对金融资产高维相关性目前存在的估计上的困难，提出了一种基于 Cholesky 分解方法的 SCC 方法，并给出了理论证明和完善。

**关键词：**多维 GARCH；SCC 方法；预测和风险

## ABSTRACT

Modeling the temporal dependence in the second order moments and forecasting future volatility have key relevance in many financial econometric issues such as risk evaluation, derivatives pricing, and optimal portfolio choice. We review different specifications of dynamic correlation. They differ in various aspects. We distinguish three approaches for constructing correlation models: (1) multivariate time series; (2) multivariate stochastic process; (3) copula theory.

The dissertation includes five sections. After the brief introduction (chapter one), in chapter two, the dissertation reviews the overseas development of multivariate correlation, and then comments the advantage, disadvantage and applicability of different models. Then from chapter three to chapter four, the dissertation uses the different multivariate GARCH model, including ADCC, CCC, and Riskmetrics method, to study the forecast and risk management of Chinese Indices Portfolio. The empirical results show that: ADCC multivariate GARCH model is best among them. This result offers theoretic support to portfolio and risk management.

Chapter five presents a new approach to the modeling of the conditional correlation matrix in the large cross-sectional dimension based on Cholesky method. Many multivariate GARCH models have been developed in the recent years to model the conditional second moments. However, all of them must make the trade-off between parameters' parsimony and richness in the description of the second order moments dynamics. In fact, the number of parameters of a fairly rich multivariate volatility model soon becomes large enough to render estimation infeasible. The key feature of the SCC based on Cholesky is the decomposition of the conditional correlation matrix into the product of a sequence of matrices with desirable characteristics. Then a highly dimensional and intractable optimization problem is converted into a series of simple and feasible estimations. Moreover, the latest methods dealing with two dimensional correlation matrix can combine with SCC to solve complex problems.

In a word, the dissertation studies forecast of dynamic correlation and application in financial risk. And then the dissertation introduces SCC method based on Cholesky decomposition to handle a highly dimensional correlation matrix and gives theory result.

**Key Words:** multivariate GARCH; SCC method; forecast and risk



# 目录

<b>1. 导 言</b> .....	<b>1</b>
1. 1 选题意义 .....	1
1. 2 篇章结构 .....	2
1. 3 创新和缺陷 .....	3
<b>2. 多维相关性模型的回顾</b> .....	<b>4</b>
<b>2. 1 多维 GARCH 模型的发展</b> .....	<b>5</b>
2. 1. 1 一维 GARCH 模型的一般化.....	6
2. 1. 2 一维 GARCH 模型的线性混合.....	10
2. 1. 3 一维 GARCH 模型的非线性混合.....	11
2. 1. 4 MGARCH 模型的杠杆效应.....	16
2. 1. 5 其它的多维波动和相关性模型.....	17
<b>2. 2 估计方法</b> .....	<b>18</b>
2. 2. 1 两步估计.....	19
2. 2. 2 基于遗传算法的似然函数的估计方法.....	20
2. 2. 3 多维 GARCH 模型的半参数有效估计.....	20
<b>2. 3 诊断性检验</b> .....	<b>21</b>
2. 3. 1 Portmanteau 统计量 .....	21
2. 3. 2 基于残差的诊断.....	22
2. 3. 3 Lagrange 乘子检验 .....	22
<b>3. 多维相关性模型预测分析</b> .....	<b>24</b>
<b>3. 1 波动模型预测的基本符号</b> .....	<b>24</b>
<b>3. 2 波动性预测的意义</b> .....	<b>27</b>
3. 2. 1 普通的预测应用.....	27
3. 2. 2 金融应用.....	29
<b>3. 3 单维波动性预测方法</b> .....	<b>33</b>
3. 3. 1 利用历史数据进行预测.....	33

3. 3. 2 随机波动性的预测.....	34
3. 3. 3 已实现波动性的预测.....	36
3. 3. 4 GARCH 类模型的预测.....	37
<b>3. 4 多维相关性预测 .....</b>	<b>41</b>
3. 4. 1 指数平滑和 Riskmetrics 方法 .....	42
3. 4. 2 BEKK 模型的预测.....	43
3. 4. 3 DCC 模型的预测 .....	43
3. 4. 4 多变量随机波动模型和因子模型.....	43
<b>3. 5 预测效果检验方法 .....</b>	<b>44</b>
<b>3. 6 实证分析 .....</b>	<b>45</b>
3. 6. 1 数据处理.....	46
3. 6. 2 ADCC 模型估计 .....	46
3. 6. 3 CCC 多维 GARCH 模型.....	50
3. 6. 4 预测结果.....	50
<b>4. 基于多维 GARCH 模型动态投资组合的风险分析 .....</b>	<b>53</b>
4. 1 在险价值基本原理 .....	53
4. 2 风险的评价方法 .....	54
4. 2. 1 部分评价法.....	55
4. 2. 2 全额评价法.....	57
4. 3 风险模型 .....	58
4. 3. 1 VaR.....	59
4. 3. 2 CVaR.....	60
4. 3. 3 ER(Expected Regret) .....	60
4. 4 在险价值的准确性检验方法 .....	60
4. 4. 1 动态系数检验.....	61
4. 4. 2 LR 似然比检验 .....	61
4. 5 实证分析 .....	61
4. 5. 1 正态分布的检验.....	62
4. 5. 2 t 分布的检验 .....	63

<b>5. SCC 模型 .....</b>	<b>66</b>
<b>5. 1 SCC 方法的起源 .....</b>	<b>66</b>
5. 1. 1 条件相关模型.....	66
5. 1. 2 SCC 的介绍 .....	67
5. 1. 3 数值例子.....	70
<b>5. 2 Cholesky 分解方法.....</b>	<b>72</b>
<b>5. 3 渐近理论 .....</b>	<b>75</b>
5. 3. 1 传统渐近定理.....	75
5. 3. 2 两阶段估计.....	77
<b>5. 4 SCC 的对数似然函数 .....</b>	<b>78</b>
<b>5. 5 模型的估计 .....</b>	<b>81</b>
<b>5. 6 SCC 模型估计的渐近性定理 .....</b>	<b>84</b>
<b>5. 7 实证分析 .....</b>	<b>86</b>
<b>6. 结论及今后研究的方向 .....</b>	<b>93</b>
<b>6. 1 总结 .....</b>	<b>93</b>
<b>6. 2 未来的方向 .....</b>	<b>93</b>
6. 2. 1 理论方面.....	93
6. 2. 2 方法的创新.....	94
<b>参考文献: .....</b>	<b>96</b>
<b>致 谢.....</b>	<b>103</b>

# CONTENTS

<b>1.Introduction.....</b>	<b>1</b>
<b>1. 1 Significance of the Research .....</b>	<b>1</b>
<b>1. 2 Framework of the Research.....</b>	<b>2</b>
<b>1. 3 Innovations and Shortcomings .....</b>	<b>3</b>
<b>2.Multivariate Correlation Models .....</b>	<b>4</b>
<b>2. 1 Development of Multivariate GARCH Models.....</b>	<b>5</b>
2. 1. 1 Generalizations of the Univariate Standard GARCH Model.....	6
2. 1. 2 Linear Combinations of Univariate GARCH Models .....	10
2. 1. 3 Nonlinear Combinations of Univariate GARCH Models.....	11
2. 1. 4 Leverage Effects in MGARCH Models.....	16
2. 1. 5 Alternative Multivariate Correlation Models.....	17
<b>2. 2 Estimate Method .....</b>	<b>18</b>
2. 2. 1 Two-Step Estimation.....	19
2. 2. 2 Estimation based on Genetic algorithm .....	20
2. 2. 3 Semiparameter Estimation .....	20
<b>2. 3Diagnostic Checking .....</b>	<b>21</b>
2. 3. 1 Portmanteau Statistics.....	21
2. 3. 2 Residual-Based Diagnostics.....	22
2. 3. 3 Lagrange Multiplier Tests .....	22
<b>3. Correlation Forecast Analysis .....</b>	<b>24</b>
<b>3. 1 Basic Notation for Forecast.....</b>	<b>24</b>
<b>3. 2 Uses of Volatility Forecast.....</b>	<b>27</b>
3. 2. 1 Generic Forecasting Applications .....	27
3. 2. 2 Financial Applications .....	29
<b>3. 3 Forecasting Methods for Univariate Volatility .....</b>	<b>33</b>
3. 3. 1 Standard Historical Volatility Forecast .....	33

3. 3. 2 Forecast based on Stochastic Volatility.....	34
3. 3. 3 Forecast based on Realized Volatility .....	36
3. 3. 4 Forecast based on GARCH models .....	37
<b>3. 4 Multivariate Correlation Forecast .....</b>	<b>41</b>
3. 4. 1 Exponential Smoothing and Riskmetrics.....	42
3. 4. 2 Forecast based on BEKK.....	43
3. 4. 3 Forecast based on DCC.....	43
3. 4. 4 Forecast based on Multivariate Stochastic Volatility and Factor Models.....	43
<b>3. 5 Evaluation for Forecast Result .....</b>	<b>44</b>
<b>3. 6 Empirical Study .....</b>	<b>45</b>
3. 6. 1 Data .....	46
3. 6. 2 Estimation for ADCC.....	46
3. 6. 3 CCC.....	50
3. 6. 4 Forecast Result.....	50
<b>4. Risk Analysis of Dynamic Portfolio based on Multivariate GARCH Model.....</b>	<b>53</b>
<b>4. 1 Principle of Risk at Value.....</b>	<b>53</b>
<b>4. 2 Valuation of Risk.....</b>	<b>54</b>
4. 2. 1 Part Valuation.....	55
4. 2. 2 Full Valuation.....	57
<b>4. 3 Risk Models .....</b>	<b>58</b>
4. 3. 1 VaR .....	59
4. 3. 2 CVaR.....	60
4. 3. 3 ER(Expected Regret) .....	60
<b>4. 4 Veracity Test for VaR.....</b>	<b>60</b>
4. 4. 1 Dynamic Quantile Test.....	61
4. 4. 2 Failing Rate Test .....	61

<b>4. 5 Empirical Study .....</b>	<b>61</b>
4. 5. 1 Test based on Normal Distribution .....	62
4. 5. 2 Test based on t Distribution .....	63
<b>5. SCC.....</b>	<b>66</b>
<b>5. 1 Origin of SCC .....</b>	<b>66</b>
5. 1. 1 Conditional Correlation .....	66
5. 1. 2 Introduction of SCC .....	67
5. 1. 3 Numerical Example .....	70
<b>5. 2 Cholesky Decomposition .....</b>	<b>72</b>
<b>5. 3 Asymptotic Theory.....</b>	<b>75</b>
5. 3. 1 Traditional Asymptotic Theory .....	75
5. 3. 2 Two Stage Estimation .....	77
<b>5. 4 Log-likelihood of SCC .....</b>	<b>78</b>
<b>5. 5 Estimator of SCC .....</b>	<b>81</b>
<b>5. 6 Asymptotic Theory of SCC Estimator .....</b>	<b>84</b>
<b>5. 7 Empirical Analysis .....</b>	<b>86</b>
<b>6. Conclusion .....</b>	<b>93</b>
<b>6. 1 Conclusion .....</b>	<b>93</b>
<b>6. 2 Research Direction in Future.....</b>	<b>93</b>
6. 2. 1 Theory .....	93
6. 2. 2 New Approach .....	94
<b>Reference: .....</b>	<b>96</b>
<b>Thanks.....</b>	<b>103</b>

## 1. 导 言

在许多金融计量问题中，例如衍生产品定价、风险管理、套期保值和最优投资组合选择等，模拟二阶矩的相关性和各个市场之间波动性的相关关系有很重要的意义。动态的相关性研究已经成为近来金融学家的重要研究课题，同时也为个人或金融机构投资者做出更准确的决策提供了理论支持。

### 1. 1 选题意义

对于金融市场波动特性的研究，是分析资本资产定价、金融风险防范等问题的基础，对于金融市场进行定量研究，前提是对于金融市场的波动特性进行准确的刻画。波动性代表了资产收益的风险，那么，一个市场的波动性对另外一个市场的波动性有无影响？一个市场的冲击的到来是否增加另外一个市场的波动性，增加多少？同样大小的正的冲击和负的冲击的影响是相同的吗？资产收益的相关性是随时间可变的吗？从长期来看，资产收益的相关性是逐渐增加的吗？诸如这样的问题必须用多变量模型来研究，同时也必须引入动态方差协方差矩阵和相关性设定的问题。此外动态相关性还在金融其他领域中有着重要的作用，这些作用体现在：

一、多变量 GARCH 模型可以用来计算时变的套期保值比率。传统的研究利用 OLS 来计算常数套期保值比率，即现货的收益对于期货收益的回归的斜率，因为这个比率等价于估计现货和期货收益的协方差与期货收益的方差的比率。Lien 和 Tse(2002)<sup>[1]</sup>对此作了一个总结，用双变量 GARCH 模型直接设定了现货和期货收益的条件方差—协方差矩阵。

二、可以用来计算动态的  $\beta$ 。某一个资产的超额收益可以表示为市场收益的线性函数，假定它是常数， $\beta$  可以用 OLS 进行估计，而多维 GARCH 模型可以用来计算随时间而变的  $\beta$  系数。Bollerslev(1988)<sup>[2]</sup>，Hafner 和 Herwartz(1998)<sup>[3]</sup>给出了理论和实证分析。

三、可以用来计算投资组合的收益和投资组合的风险值。当考虑到多资产

的投资组合的收益和风险时，资产之间相关性的估计是一个很重要的因素。因此可以利用多维 GARCH 模型来计算投资组合的收益，从而进而计算投资组合的风险。Giot 和 Laurent(2003)<sup>[4]</sup>用三维 GARCH 模型计算时变的相关性和风险。

四、对于涉及多资产的衍生产品的定价具有重要的作用。价值依赖于多个基本标的资产的结构化产品譬如彩虹期权的价格对于基本资产的收益间的相关性也很敏感，因此对衍生产品定价时可以更合理地利用多维 GARCH 模型引入动态相关性的估计。

多维 GARCH 模型最初发展于 20 世纪 80 年代后期，随后由于估计和经济意义解释上所出现的种种困难，经历了一段时间的平静。随后随着计算机技术的发展，从 90 年代中期至今多维 GARCH 模型处于理论和应用的繁荣时期，在风险管理、衍生产品的定价和套期保值等方面都有很重要的理论和实际价值。Franses 和 Van Dijk(2000)<sup>[5]</sup>, Gouriou(1997)<sup>[6]</sup>, Bauwens (2006)<sup>[7]</sup>对 ARCH 族模型从不同方面进行了综述。

## 1. 2 篇章结构

论文共分六章。在对国内和国外多维 GARCH 模型进行回顾评价的基础上，作者利用已有的各种多维 GARCH 模型以及其他业界部门常用的工具对中国主要股指市场的预测、投资组合和风险管理进行了研究。并且在此基础上，提出了基于 Cholesky 分解方法的处理高维相关性的研究方法—SCC 方法，并给予了理论和实证的研究。本文的具体结构安排如下：

第一章，导言。本章主要针对多维波动性研究提出了选题依据和研究意义。简要介绍本文的研究思路和逻辑框架，指出了论文在多维波动性方面一些理论和应用的研究成果以及未来需要进一步研究的问题。

第二章，对多维波动性的发展进行了回顾和评述。在这一部分，作者分类回顾了多维波动性模型相关的理论进展，并给出了综合对比及评述。由于在估计上的一些困难，国内的研究还停留在一些方法简单但缺乏经济含义的一些模型的实证方面，缺乏综合的比较和理论的突破。

第三章，给出多维相关性模型预测分析。作者主要利用了各种多维 GARCH 模型进行了预测的实证分析，同时利用了现有的 Riskmetrics 方法进行对比研究，



并对预测结果进行了比较和评价，结果表明非对称的多维 GARCH 模型由于抓住了金融资产波动的非对称性和群聚性，对未来的预测效果最佳。

第四章，利用多维波动性模型对资产的投资组合的风险进行了研究。在研究中，作者发现了对于中国主要股指市场利用非对称的多维 GARCH 模型进行风险管理的效果比其他的方法要好。

第五章，对于由 Palandri (2005)<sup>[8]</sup>提出的处理金融资产高维相关性的 SCC 模型，论文不仅给出了基于 Cholesky 分解的理论创新，而且给出了实证分析和结果的比较。

第六章，给出了资产动态相关性未来的研究方向及创新之处。

### 1.3 创新和缺陷

本篇论文对于研究资产动态相关性的创新之处主要有：

(1) 本文对于多维时间序列的动态相关性研究进行了系统的综合和比较；  
(2) 从未来预测和风险管理两方面对资产动态相关性的应用进行了研究，目前国内的文献由于估计方法的困扰，较少涉及到超过二维资产的动态相关性方面，论文在此进行了一些尝试；

(3) 提出了基于 Cholesky 分解的处理高维相关性的一种新方法—SCC 方法，并进行了理论证明和实证研究，表明了此方法的确是处理高维动态条件相关性的行之有效的方法，并且它具有很强的灵活性，可以和其他有效的处理二维相关性的方法结合起来，对于金融资产高维相关性的发展具有重要的意义。

同时，由于多维时间序列研究方法的局限性和研究水平的限制，论文存在如下一些不足之处：

(1) 对于多维 GARCH 模型以及本文提出的 SCC 模型，理论上还未证明两步估计的有效性，这是本文进一步努力的方向；

(2) SCC 模型的最大优势在于高维金融时间序列相关矩阵的计算，而本文所用到的动态相关性的计算还是停留在四维相关性的计算上，未明显体现出 SCC 模型的优势；

(3) 对于把 SCC 方法和 Copula 理论结合起来更好解决高维资产动态条件相关性，本文还未研究，作者认为这对于解决相关性估计的难题有重要的意义。

## 2. 多维相关性模型的回顾

Campbell,Lo 和 MacKinlay (1997)<sup>[9]</sup>认为金融理论和实践中最具有特点的是不确定性。金融经济学与微观经济学最显著的不同在于波动性是内在不可观测的，不确定的，并且是随机的。根据数据的取值时间，我们可以把波动性模型分为离散时间和连续时间模型。众所周知，随着计算机和电子技术的发展，金融市场资产交易和定价的流动性加强，几乎接近于连续交易的方式。但是，构造一个离散时间模型仍是非常有用的，它克服了连续时间模型在估计和预测上的困难。因此，至今所出现的离散时间模型虽与实际中的连续交易有不相符之处，但是从估计参数和推断性的角度讲，更容易处理。并且，离散模型对未来的预测上提供了更多的方法。因此，本文首先对拟合资产收益波动性较好的 GARCH 模型作简单的回顾，然后重点对多维时间序列模型—多维 GARCH 模型进行回顾和评述。

在单维离散的情况下，本文引入这样一个过程：

$$y_t = y(t) \quad t = 1, 2, \dots$$

我们假定资产收益率  $y_t$  的二阶矩存在，同时相应的条件期望和方差为：

$$\mu_{t|t-1} = E[y_t | \psi_{t-1}]$$

$$\sigma_{t|t-1}^2 = \text{Var}[y_t | \psi_{t-1}] = E[(y_t - \mu_{t|t-1})^2 | \psi_{t-1}]$$

$\psi_{t-1}$  指信息集，反映了直到  $t-1$  时刻的所有相关信息。

正是因为条件期望并入了  $t-1$  时刻前的所有信息，因此它和无条件期望有所不同，即  $\mu_{t|t-1} \neq E(y_t)$ ，同时  $\sigma_{t|t-1}^2 \neq \text{Var}(y_t)$ 。这种条件和无条件矩的差别为利用时间序列进行预测奠定了基础。

在 ARCH 模型提出之前，描述股票价格等金融现象最常用的模型是随机游走模型。但这一模型在应用中有明显的局限性，因为金融市场中一个常见现象是股价或其他类似金融产品的变化率不仅随时间  $t$  变化，而且常常在某一时段出现偏高或偏低的情况，而这一特征很难由随机游走模型解释。

1982 年 Engle<sup>[10]</sup>首先提出了用 ARCH 模型来刻画英国通货膨胀率中存在的

条件异方差。从此，描述均值的传统的时间序列，例如自回归移动平均模型（ARMA）已经扩展到了本质上类似的描述方差的时间序列模型。GARCH 模型描述和预测金融时间序列的波动性已被广泛接受。自 ARCH 模型始创以来，经历了两次突破，一是 Bollerslev(1986)<sup>[11]</sup>提出广义 ARCH(Generalized ARCH)，即 GARCH 模型，从此以后，几乎所有的 ARCH 模型新成果都是在 GARCH 模型基础上得到的；第二次则是由于长记忆在经济学上的研究取得突破，分整研究被证明更有效地刻画了某些长记忆性经济现象，与 ARCH 模型相结合所诞生的一系列长记忆 ARCH 模型的研究从 1996 年至今方兴未艾。对 GARCH 模型的一些综述性文章，请参见 Bollerslev,Chou 和 Kroner (1992)<sup>[12]</sup>，Bera 和 Higgins(1993)<sup>[13]</sup>，Bollerslev,Engle 和 Nelson (1994)<sup>[14]</sup>，Pagan(1996)<sup>[15]</sup>，Palm(1996)<sup>[16]</sup> 等。

多维 GARCH 模型自 Bollerslev,Engle 和 Wooldridge(1988)<sup>[2]</sup>提出了多维 GARCH 模型以来，经过了 20 世纪 80 年代到 90 年代的发展，随后有了一段沉寂，随着计算机技术的发展，现在又处于较快的发展时期。Franses 和 van Dijk (2000)<sup>[5]</sup>，Gourieroux (1997)<sup>[6]</sup> 总结了一部分的多维 GARCH 模型。而 Bauwens, L., S. Laurent 和 J.V.K. Rombouts<sup>[7]</sup>对多维 GARCH 模型做了较完整的综述。

## 2. 1 多维 GARCH 模型的发展

$\{y_t\}$  是  $N \times 1$  维随机过程， $\psi_{t-1}$  是  $\sigma$ -域， $\mu_t(\theta)$  是条件期望向量， $\theta$  是有限的参数向量，可以把  $y_t$  满足的等式写为：

$$\begin{aligned} y_t &= \mu_t(\theta) + \varepsilon_t \\ \varepsilon_t &= H_t^{1/2}(\theta)\eta_t \end{aligned}$$

其中， $H_t^{1/2}(\theta)$  是  $N \times N$  正定矩阵， $H_t^{1/2}$  是  $H_t$  的 Cholesky 分解， $I_N$  是  $N$  维单位矩阵， $\eta_t$  满足：

$$E(\eta_t) = 0 \quad \text{Var}(\eta_t) = I_N$$

为了对  $y_t$  有更清楚的了解，可以得到：

$$\begin{aligned} \text{Var}(y_t | \psi_{t-1}) &= \text{Var}_{t-1}(y_t) = \text{Var}_{t-1}(\varepsilon_t) \\ &= H_t^{1/2} \text{Var}_{t-1}(\eta_t) (H_t^{1/2})' \\ &= H_t \end{aligned}$$

在下面的章节，我们着重对  $H_t$  的设定作简单的回顾。我们根据构建多维 GARCH 模型不同的方法把  $H_t$  的设定分为三类：第一类模型是 Bollerslev(1986)<sup>[11]</sup> 提出的单维 GARCH 模型的多维化扩展，包括 VEC, BEKK, 因子模型, Riskmetrics 等；第二类模型是单维 GARCH 模型的线性组合，包括正交化模型等；第三类模型是单维 GARCH 模型的非线性组合，包括动态条件相关模型和 Copula-GARCH 模型等。以下我们将分别加以回顾和讨论。

### 2. 1. 1 一维 GARCH 模型的一般化

这一分类的模型是一维 GARCH 模型向多变量 GARCH 模型直观上的扩展。当我们考虑几个时间序列的的向量 ARMA 模型时，参数的数量急速增加，因此，当我们从一维 GARCH 模型扩展到多维 GARCH 模型时，我们面对的主要问题也是参数数量增多所带来的估计上的困难。特别需要强调的是  $H_t$  作为方差协方差矩阵，它的对称性和正定性必须保证。这一类的模型很容易得到平稳性和遍历性的理论证明，请参见 He 和 Terasvirta(2002)<sup>[17]</sup>。

#### 2. 1. 1. 1 VEC 模型

在 VEC 模型中， $H_t$  的每一个元素是滞后的误差项、误差项的交叉乘积和  $H_t$  的滞后项的线性函数。

**定义 2-1** VEC(1,1)模型被定义为

$$h_t = c + Az_t + Gh_{t-1} \quad (2.1)$$

其中  $h_t = \text{vech}(H_t)$      $z_t = \text{vech}(\varepsilon_t \varepsilon_t')$

$\text{vech}(\cdot)$  表示一个  $N \times N$  矩阵的下三角部分依次排列的算子，是一个  $N(N+1)/2 \times 1$  的向量。 $A$  和  $G$  是秩为  $(N+1)N/2$  的方阵， $c$  是  $N(N+1)/2 \times 1$  参数向量。

模型的参数数量是  $N(N+1)(N(N+1)+1)/2$  (例如, 当  $N=3$ , 参数数量为 78), 这就意味着在实际应用中, 只能用双变量 VEC 模型。为了克服这个缺点, Bollerslev(1988)<sup>[2]</sup>提出了对角 VEC(DVEC)模型,  $A$  和  $G$  被假定为对角矩阵。这个限制将参数的数量减少为  $N(N+5)/2$  (例如, 当  $N=3$ , 参数数量为 12)。可以看出, 即使在对角的假设下, 高维的情况在实际中也很难估计。

为了保证对角 VEC 模型的条件方差协方差矩阵是正定的, 需要引入 Hadamard 乘积<sup>①</sup> (用  $\square$  表示)。定义  $N \times N$  矩阵  $A^\circ, G^\circ$ , 和  $C^\circ$ , 满足  $A = \text{diag}[\text{vech}(A^\circ)]$ <sup>②</sup>,  $G = \text{diag}[\text{vech}(G^\circ)]$ , 和  $c = \text{vech}(C^\circ)$ 。对角模型可以被写作  $H_t = C^\circ + A^\circ \square (\varepsilon_{t-1} \varepsilon'_{t-1}) + G^\circ \square H_{t-1}$ 。

Attanasio(1991)<sup>[18]</sup>证明了只要  $C^\circ, A^\circ, G^\circ$  和初始的方差矩阵  $H_0$  是正定的, 就可以保证  $H_t$  是正定的。Ding 和 Engle(2001)<sup>[19]</sup>提出的对角 VEC 模型包含的  $C^\circ, A^\circ, G^\circ$  可以简化为正的标量, 叫作标量 VEC 模型。

Riskmetrics(1996)<sup>[20]</sup>用指数加权移动平均模型 (EWMA) 来预测方差协方差矩阵。研究波动过程的业界工作者观测到波动模型经常接近于单位根过程。鉴于此, Riskmetrics 把方差协方差定义为 IGARCH 模型:

$$h_{ij,t} = (1-\lambda)\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \lambda h_{ij,t-1} \quad (2.2)$$

(2.2)可以转化为 VEC 模型:  $h_t = (1-\lambda)\eta_{t-1} + \lambda h_{t-1}$ 。

这其实就是一个标量 VEC 模型。Riskmetrics(1996)<sup>[20]</sup>提出的衰减因子  $\lambda$  并不是估计得到的, 而是根据经验事先设定好的, 当用于拟合每天的数据时等于 0.94, 当用于拟合每个月的数据时等于 0.97。显而易见, 这样的模型很容易估计。但是不事先考虑所用到的数据的特征, 一概采用相同的动态性来描述各类时间序列, 也不是一种合理的做法。

由于 VEC 的设定很难保证  $H_t$  的正定性, Engle 和 Kroner (1995)<sup>[21]</sup>提出了

<sup>①</sup>  $A = (a_{ij})$  和  $B = (b_{ij})$  是  $m \times n$  矩阵, 那么  $A \square B$  是每个元素是  $(a_{ij}b_{ij})$  的  $m \times n$  矩阵。

<sup>②</sup> 当  $v$  是  $m$  维向量, 那么  $\text{diag}(v)$  是  $m \times m$  对角矩阵, 而  $v$  在主对角线上。

一种新的模型，也叫作 BEKK 模型（Baba,Engle,Kraft 和 Kroner 四人名字首写字母的简称）这种模型的好处是很容易保证  $H_t$  的正定性。

### 2. 1. 1. 2 BEKK 模型

**定义 2-2** BEKK(1,1,K)模型被定义为

$$H_t = C^* C^* + \sum_{k=1}^K A_k^* \varepsilon_{t-1} \varepsilon_{t-1}' A_k^* + \sum_{k=1}^K G_k^* H_{t-1} G_k^* \quad (2.3)$$

其中  $C^*$ ,  $A_k^*$  和  $G_k^*$  是  $N \times N$  矩阵， $C^*$  是上三角矩阵。

从定义可以看出，BEKK 模型的参数当中，并没有直接给出不同的滞后项对于  $H_t$  的元素的影响大小。BEKK(1,1,1)中的参数的数量为  $N(5N+1)/2$ ，有时为了减少参数的数量，我们可以使  $A_k^*$  和  $G_k^*$  为对角矩阵。通过比较可以看出，对于二维模型，DVEC 模型包含 9 个参数，BEKK 模型包含 7 个参数。这是因为在 BEKK 模型中控制协方差等式的动态性的参数和控制方差等式的动态性的参数是相同的。

估计 VEC 甚至 BEKK 模型的困难在于未知参数的数量。毋庸置疑，当维数超过 3 时，估计变得相当困难。因此就出现了下述的因子模型。

### 2. 1. 1. 3 Factor 模型

Engle(1990b)<sup>[22]</sup>提出了股票收益的协变性（co-movements）被几个共同的隐藏变量所驱动的思想，这几个变量叫做因子（factor）。Bollerslev 和 Engle(1993)<sup>[23]</sup>用这些参数来模拟条件方差的普通持续性。因子模型可以被看作特殊的 BEKK 模型。我们给出 Lin(1992)<sup>[24]</sup>的定义：

**定义 2-3** 若对于  $k=1, \dots, K$ ， $A_k^*$  和  $G_k^*$  的秩为 1，且都有相同的左和右特征向量<sup>①</sup>  $\lambda_k$  和  $\omega_k$ ，也就是说  $A_k^*$  和  $G_k^*$  满足：

<sup>①</sup> 对于  $m \times m$  矩阵  $A$ ， $1 \times m$  向量  $v \neq 0$  满足  $vA = \lambda v$ ，那么  $v$  叫做  $A$  的相对于特征值  $\lambda$  的左特征向量。

$$A_k^* = \alpha_k \omega_k \lambda_k' \text{ 和 } G_k^* = \beta_k \omega_k \lambda_k' \quad (2.4)$$

其中,  $\alpha_k$  和  $\beta_k$  是标量,  $\lambda_k$  和  $\omega_k$  是  $N \times 1$  向量, 满足

$$\omega_k' \lambda_i = \begin{cases} 0 & k \neq i \\ 1 & k = i \end{cases} \quad (2.5)$$

$$\sum_{n=1}^N \omega_{kn} = 1 \quad (2.6)$$

那么, 在定义 2-2 中, **BEKK(1,1,K)** 是一个因子 - GARCH 模型, 叫作 F-GARCH(1,1,k)。

如果我们把 (2.4) 和 (2.5) 代入 (2.3) 中, 定义  $\Omega = C^* C^*$ , 得到:

$$H_t = \Omega + \sum_{k=1}^K \lambda_k \lambda_k' (\alpha_k^2 \omega_k' \varepsilon_{t-1} \varepsilon_{t-1}' \omega_k + \beta_k^2 \omega_k' H_{t-1} \omega_k) \quad (2.7)$$

$K$  因子 GARCH 模型意味着  $H_t$  的时变性部分已经缩减为  $K$  秩, 而  $H_t$  仍保持满秩, 因为  $\Omega$  被假定为正定的。向量  $\lambda_k$  叫做  $k$  阶因子荷载, 标量  $\omega_k' \varepsilon_t$  叫做  $k$  阶因子 (也记为  $f_{kt}$ )。F-GARCH(1,1,1) 的参数的数量为  $N(N+5)/2$ 。式(2.7)中括弧里面的表达式可以用其他的单维 GARCH 模型代替。

例如, F-GARCH(1,1,2) 模型的条件方差矩阵可以写作:

$$\begin{aligned} H_t = \Omega &+ \lambda_1 \lambda_1' [\alpha_1^2 \omega_1' \varepsilon_{t-1} \varepsilon_{t-1}' \omega_1 + \beta_1^2 \omega_1' H_{t-1} \omega_1] \\ &+ \lambda_2 \lambda_2' [\alpha_2^2 \omega_2' \varepsilon_{t-1} \varepsilon_{t-1}' \omega_2 + \beta_2^2 \omega_2' H_{t-1} \omega_2] \end{aligned} \quad (2.8)$$

其中参数向量  $\lambda_k = (\lambda_{k1}, \lambda_{k2}, \dots, \lambda_{kN})'$ ,  $\omega_k$  是  $N \times 1$  维,  $\alpha_k^2$  和  $\beta_k^2$  是标量参数。

令  $\sigma_{k,t}^2 = \omega_k' H_t \omega_k$ , 我们可以把式 (2.8) 写作更熟悉的公式:

$$h_{ijt} = \tau_{ij} + \lambda_{1i} \lambda_{1j} \sigma_{1t} + \lambda_{2i} \lambda_{2j} \sigma_{2t}, \quad \forall i, j = 1, \dots, N$$

$$\sigma_{k,t}^2 = \omega_k + \alpha_k^2 f_{k,t-1}^2 + \beta_k^2 \sigma_{k,t-1}^2 \quad k = 1, 2$$

其中,  $\tau_{ij} = \omega_{ij} - \lambda_{1i} \lambda_{1j} \omega_1 - \lambda_{2i} \lambda_{2j} \omega_2$ ,  $\omega_k = \omega_k' \Omega \omega_k$ ,  $\sigma_{kt}$  定义为一维的 GARCH

(1, 1) 模型。式 (2.8) 中的条件方差的持续性用  $\sum_{k=1}^2 (\alpha_k^2 + \beta_k^2)$  表示。换句话说讲,

$H_t$  的元素的动态性是相同的。我们可以把  $H_t$  写作:

$$H_t = \Omega^* + \lambda_1 \lambda_1' \sigma_{1,t}^2 + \lambda_2 \lambda_2' \sigma_{2,t}^2 \quad (2.9)$$

其中,  $\Omega^* = \Omega - \lambda_1 \lambda_1' \omega_1 - \lambda_2 \lambda_2' \omega_2$ ,  $E_{t-1}(f_{1t} f_{2t}') = \omega_1 \Omega \omega_2'$  (因为  $w_k' \lambda_l = 0, k \neq l$ )。

式 (2.9) 中的两因子模型可以从  $\varepsilon_t = \lambda_1 f_{1t} + \lambda_2 f_{2t} + e_t$  得到,  $e_t$  代表了个别风险, 具有常数方差矩阵, 并且与两因子不相关。每一个因子  $f_{kt}$  的条件期望是 0, 条件方差是 GARCH(1,1) 过程,  $K$  因子模型可以写作:

$$\varepsilon_t = \Lambda f_t + e_t$$

其中,  $\Lambda$  是一个  $N \times K$  矩阵,  $f_t$  是  $K \times 1$  向量。

### 2. 1. 2 一维 GARCH 模型的线性混合

在这个分类中, 我们考虑正交的模型和隐藏因子模型, 它们都是单维 GARCH 模型的线性混合。这些单维 GARCH 模型可以不是标准的 GARCH 模型, 它们可以是各种 GARCH 模型的扩展模型, 譬如, Nelson(1991)<sup>[25]</sup> 的 EGARCH 模型和 Ding(1993)<sup>[26]</sup> 的 AGARCH 模型。

在正交 GARCH 模型中, 被观测到的数据假定是  $N$  个单维 GARCH 模型的正交转换。线性转换矩阵是标准化之后的收益向量的无条件方差协方差矩阵的特征向量组成的正交矩阵, 而且这个矩阵是可逆的。这个正交模型被称为因子模型, 其中的因子是单维 GARCH 过程。

Kariya(1988)<sup>[27]</sup> 和 Alexander 和 Chibumba(1997)<sup>[28]</sup> 提出的正交 GARCH 模型中,  $N \times N$  时变方差协方差矩阵  $H_t$  是由  $m \leq N$  个单维 GARCH 模型构成的。

**定义 2-4** 正交-GARCH 模型被定义为:

$$V^{-1/2} \varepsilon_t = u_t = \Lambda_m f_t \quad (2.10)$$

其中  $V = \text{diag}(v_1, v_2, \dots, v_N)$ ,  $v_i$  是  $\varepsilon_{it}$  的无条件方差。  $\Lambda_m$  是  $N \times m$  维矩阵, 并且  $\Lambda_m = P_m \text{diag}(l_1^{1/2} \dots l_m^{1/2})$ ,  $l_1 \geq \dots \geq l_m > 0$  是  $u_t$  的无条件相关矩阵的  $m$  个最大的特征值,  $P_m$  是与之相对应的  $N \times m$  的特征向量组成的矩阵。向量  $f_t = (f_{1t} \dots f_{mt})'$  满足



$$E_{t-1}(f_t) = 0, \text{Var}_{t-1}(f_t) = \sum_i \text{diag}(\sigma_{f_{1t}}^2, \dots, \sigma_{f_{mt}}^2)$$

$$\sigma_{f_{it}}^2 = (1 - \alpha_i - \beta_i) + \alpha_i f_{i,t-1}^2 + \beta_i \sigma_{f_{i,t-1}}^2, i = 1, \dots, m$$

因此,  $H_t = \text{Var}_{t-1}(\varepsilon_t) = V^{1/2} V_t V^{1/2}$ , 其中,  $V_t = \text{Var}_{t-1}(u_t) = \Lambda_m \sum_i \Lambda_m'$

此模型中包含的参数为  $V, \Lambda_m$  和 GARCH 因子中的参数  $(\alpha_i, \beta_i)$ , 总的参数数量为  $N(N+5)/2(m=N)$ 。在实际中,  $V$  和  $\Lambda_m$  用样本的相应部分来代替,  $m$  通过用对标准残差  $\hat{u}_t$  进行主成分分析的方法来进行选择。

### 2. 1. 3 一维 GARCH 模型的非线性混合

这一节收集了由单维 GARCH 模型经过非线性混合所组成的多维 GARCH 模型。这些模型允许把这类多维 GARCH 模型分拆为单个的条件方差和单个序列之间的动态相关。关于这类模型的平稳性、遍历性和矩的理论性质或许不能像前面的模型那样直接获得, 但是, 这类模型参数相对较少, 并且比较容易估计。这类模型从总体上分为两大类——条件相关模型和 Copula-MGARCH 模型。条件相关模型分为常数条件相关 (CCC) 多维 GARCH 模型, 动态条件相关 (DCC) 多维 GARCH 模型以及灵活的多维 GARCH 模型 (F-多维 GARCH)。

条件相关模型对于条件方差协方差矩阵的设定是分开进行的。首先, 我们可以针对每一个序列选择一个适合的 GARCH 类的条件方差, 并且各个序列允许选择不同的条件方差模型, 其次, 在此基础上, 我们再来估计条件相关矩阵。

Bollerslev(1990)<sup>[29]</sup>提出了条件相关矩阵是常数的 MGARCH 的模型。

#### 2. 1. 3. 1 CCC 模型

**定义 2-5** 常数条件相关模型 (CCC) 可以被定义为

$$H_t = D_t R D_t = (\rho_{ij} \sqrt{h_{ii} h_{jj}}) \quad (2.11)$$

其中,  $D_t = \text{diag}(h_{1t}^{1/2} \dots h_{Nt}^{1/2})$ , 即  $D_t$  是对角矩阵, 每个对角线上的元素是单维 GARCH 模型的方差的平方根。  $R = (\rho_{ij})$  表示常数相关矩阵, 它是一个对称正定

矩阵,  $\rho_{ii} = 1$ 。

最初 CCC 模型中的对角矩阵  $D_t$  所包含的条件方差都是具有 GARCH(1, 1) 形式的:

$$h_{iit} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{iit,t-1}, i = 1, \dots, N$$

因此, 最初的 CCC 模型中的参数有  $N(N+5)/2$  个。  $H_t$  是对称正定的当且仅当所有的  $N$  个条件方差都是正的, 同时相关矩阵  $R$  是正定的。

但是, 条件相关系数是常数似乎不现实。Engle(2002)<sup>[30]</sup>, Tse 和 Tsui(2002)<sup>[31]</sup> 提出了 CCC 模型的一般化, 即把条件相关矩阵转化为时变的, 这个模型叫做动态条件相关模型 (DCC)。

### 2. 1. 3. 2 DCC 模型

Engle(2002)<sup>[30]</sup>提出了  $DCC_E$  模型的定义如下:

定义 2-6  $DCC_E$  模型定义为:

$$H_t = D_t R_t D_t$$

$$R_t = \text{diag}(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2}) Q_t \text{diag}(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2}) \quad (2.12)$$

其中  $Q_t = (q_{ij,t})$  为  $N \times N$  维对称正定矩阵,  $Q_t = (1 - \alpha - \beta) \bar{Q} + \alpha u_{t-1} u_{t-1}' + \beta Q_{t-1}$ 。

$u_{it} = \varepsilon_{it} / \sqrt{h_{iit}}$ ,  $\bar{Q}$  是  $u_t$  的  $N \times N$  无条件方差方差协方差矩阵,  $\alpha, \beta$  是非负参数, 满足  $\alpha + \beta < 1$ 。

Tse 和 Tsui(2002)<sup>[31]</sup>提出了不同的  $DCC_T$  模型:

定义 2-7  $DCC_T$  模型定义为:

$$H_t = D_t R_t D_t$$

$$D_t = \text{diag}(h_{11t}^{1/2} \dots h_{NNt}^{1/2})$$

$$R_t = (1 - \theta_1 - \theta_2) R + \theta_1 \Psi_{t-1} + \theta_2 R_{t-1}, \theta_1 + \theta_2 < 1 \quad (2.13)$$

$R$  是一个对称的  $N \times N$  正定参数矩阵，且  $\rho_{ii} = 1$ 。 $\Psi_{t-1}$  是  $\varepsilon_t$  的相关矩阵，其中，

$$\tau = t - M, t - M + 1, \dots, t - 1, \text{ 它的第 } i, j \text{ 个元素是 } \psi_{ij,t-1} = \frac{\sum_{m=1}^M u_{i,t-m} u_{j,t-m}}{\sqrt{(\sum_{m=1}^M u_{i,t-m}^2)(\sum_{m=1}^M u_{j,t-m}^2)}}, \text{ 其}$$

中， $u_{it} = \varepsilon_{it} / \sqrt{h_{it}}$ ， $\Psi_{t-1} = B_{t-1}^{-1} L_{t-1} L_{t-1}' B_{t-1}^{-1}$ ， $B_{t-1}$  是  $N \times N$  对角矩阵，对角元素是

$$\left(\sum_{h=1}^M u_{i,t-h}^2\right)^{1/2}, \quad L_{t-1} = (u_{t-1}, \dots, u_{t-M}) \text{ 是 } N \times M \text{ 矩阵， } u_t = (u_{1t} \dots u_{Nt})'$$

为了更清楚地比较  $DCC_E$  和  $DCC_T$  模型两者之间的差别，我们把双变量模型的相关系数写出来：

对于  $DCC_T$ ：

$$\rho_{12t} = (1 - \theta_1 - \theta_2) \rho_{12} + \theta_2 \rho_{12,t-1} + \theta_1 \frac{\sum_{m=1}^M u_{1,t-m} u_{2,t-m}}{\sqrt{(\sum_{m=1}^M u_{1,t-m}^2)(\sum_{m=1}^M u_{2,t-m}^2)}}$$

对于  $DCC_E$ ：

$$\rho_{12t} = \frac{(1 - \alpha - \beta) \bar{q}_{12} + \alpha u_{1,t-1} u_{2,t-1} + \beta q_{12,t-1}}{\sqrt{((1 - \alpha - \beta) \bar{q}_{11} + \alpha u_{1,t-1}^2 + \beta q_{11,t-1})(1 - \alpha - \beta) \bar{q}_{22} + \alpha u_{2,t-1}^2 + \beta q_{22,t-1}}}$$

比较上述两种模型， $DCC_T$  中的条件相关系数是过去条件相关系数的加权和。实际上  $DCC_E$  中的  $Q_t$  更类似于 GARCH 模型，然后再把  $Q_t$  转化为相关矩阵。

对于  $DCC_E$  和  $DCC_T$  模型，我们可以通过检验  $\theta_1 = \theta_2 = 0$  或者  $\alpha = \beta = 0$  来检验条件相关系数是否为常数。

$DCC$  模型的缺点是， $DCC_T$  中的  $\theta_1$  和  $\theta_2$  和  $DCC_E$  中的  $\alpha$ ， $\beta$  都是标量，即所有的条件相关都服从相同的动态性。

如果条件方差为 GARCH (1, 1) 模型，那么  $DCC_T$  和  $DCC_E$  中的参数数量为  $(N+1)(N+4)/2$  个。

### 2. 1. 3. 3 F-GARCH

Ledoit (2003) [32]提出了一种灵活的多维 GARCH 模型, 适用于高维的 GARCH 模型, 而且并没有对参数施加太多限制。

首先分别计算每个收益变量的方差, 即求出  $H_t$  的对角元素:

$$\max_{c_{ii}, a_{ii}, b_{ii} \leq 1} \prod_{t=1}^T \frac{1}{\sqrt{2\pi h_{ii,t}}} e^{-x_{ii,t}^2 / (2h_{ii,t})}$$

其中,  $h_{ii,t} = c_{ii} + a_{ii}x_{i,t-1}^2 + b_{ii}h_{ii,t-1}$

然后, 利用参数估计  $\hat{c}_{ii}, \hat{a}_{ii}$  和  $\hat{b}_{ii}$ , 求出  $H_t$  离对角元素为:

$$\max_{c_{ij}, a_{ij}, b_{ij}} \prod_{t=1}^T \frac{1}{2\pi \sqrt{\det(H_{ij,t})}} e^{-x'_{ij,t} H_{ij,t}^{-1} x_{ij,t} / 2}$$

其中,  $X_{ij,t} = \begin{bmatrix} x_{i,t} \\ x_{j,t} \end{bmatrix}$ ,  $H_{ij,t} = \begin{bmatrix} \hat{h}_{ii,t} & h_{ij,t} \\ h_{ij,t} & \hat{h}_{jj,t} \end{bmatrix}$

$$h_{ij,t} = c_{ij} + a_{ij}x_{i,t-1}x_{j,t-1} + b_{ij}h_{ij,t-1}$$

但是  $H_t$  并不能保证正定性, 因此, 根据定理 2-1, 通过一定的计算, 得到真正的方差协方差矩阵。

**定理 2-1** 给定对称矩阵  $A$ , 其中  $\text{diag}(A) > 0$ , 可以找到一个对称, 半正定矩阵  $M$ , 满足  $\text{diag}(M) = \text{diag}(A)$ , 而且可以使  $\|A - M\|_F$  最小化。其中  $\|\cdot\|_F$  是

Frobenius 范数。  $\|U\|_F = \sqrt{\sum_{i=1}^N \sum_{j=1}^N u_{ij}^2}$

此外, Sklar(1959)[33]提出  $N$  维联合分布函数可以分解为  $N$  个边际分布和 Copula 函数, 其中 Copula 函数可以完整地叙述  $N$  个变量的依赖关系。Nelsen(1999)[34]对 Copula 作了完整的总结。Patton(2000)[35], Jondeau 和 Rockinger(2001)[36]提出了 Copula-GARCH 模型, 这些模型可以分解为用 GARCH 模型描述的条件方差、每个序列的边际分布和条件 Copula 函数。

例如, 单个资产收益率为 t-GARCH 的 Copula-GARCH 模型为:

考虑如下的 GARCH(1,1)模型:

$$\begin{aligned}
 y_t &= \mu_t + \varepsilon_t \\
 \varepsilon_t &= \sigma_t \eta_t, \eta_t \sim t_\nu \\
 \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2
 \end{aligned} \tag{2.14}$$

其中,  $y_t$  为收益率序列,  $\mu_t$  为  $y_t$  的均值项, 通过选取合适的 ARMA 或者 ARIMA 模型来确定  $\mu_t$  的形式,  $\eta_t$  为 i.i.d. 自由度为  $\nu$  的  $t$  分布,  $\nu, \alpha_0, \alpha_1$  和  $\beta$  为待估计的参数。

在估计出参数  $\nu, \alpha_0, \alpha_1$  和  $\beta$  之后, 就得到了  $y_{T+1}$  的条件分布:

$$\begin{aligned}
 P(y_{T+1} \leq r | \psi_T) &= P(\varepsilon_{T+1} \leq (r - \mu) | \psi_T) = P(\sigma_{T+1} \eta_{T+1} \leq (r - \mu) | \psi_T) \\
 &= P(\eta_{T+1} \leq \frac{r - \mu}{\sqrt{\alpha_0 + \alpha_1 \varepsilon_T^2 + \beta \sigma_T^2}}) = t_\nu\left(\frac{r - \mu}{\sqrt{\alpha_0 + \alpha_1 \varepsilon_T^2 + \beta \sigma_T^2}}\right)
 \end{aligned}$$

其中,  $t_\nu(\cdot)$  为自由度为  $\nu$  的  $t$  分布函数,  $\psi_t$  是到时刻  $t$  为止的信息集。

$N$  个资产中, 设资产  $i$  的等间隔收益率观测样本为  $\{y_{1i}, y_{2i}, \dots, y_{Ti}\}, i = 1, \dots, N$ 。

根据前面的 t-GARCH 模型, 可以得到  $y_i = y_{T+1,i}$  的条件分布  $F_i(\cdot)$ 。

$N$  维正态 Copula:

$$C_\Sigma^{Normal}(u_1, \dots, u_N) = \Phi_\Sigma(\varphi^{-1}(u_1), \dots, \varphi^{-1}(u_N)), \tag{2.15}$$

这里  $\Sigma$  表示多维正态分布中的相关系数矩阵, 也是正态 Copula 的参数。  $\Phi_\Sigma$  为相关系数矩阵为  $\Sigma$  的  $N$  维正态分布函数,  $\varphi^{-1}$  为标准正态分布函数的反函数。

由于  $y_i$  的边缘分布为  $F_i(y)$ , 则由样本  $\{y_{t1}, y_{t2}, \dots, y_{tN}\}, t = 1, \dots, T$ , 令

$$\begin{aligned}
 u_t &= (u_{1,t}, u_{2,t}, \dots, u_{N,t}) = (F_1(y_{t1}), F_2(y_{t2}), \dots, F_N(y_{tN})), \\
 \xi_t &= (\varphi^{-1}(u_{1,t}), \varphi^{-1}(u_{2,t}), \dots, \varphi^{-1}(u_{N,t})), t = 1, 2, \dots, T
 \end{aligned} \tag{2.16}$$

把 (2.16) 式中的  $u_i$  代入 (2.15) 中, 得到  $C_\Sigma^{Normal}(u_t) = \Phi_\Sigma(\xi_t)$ ,  $\xi_t$  就是相关系数矩阵为  $\Sigma$  的  $N$  维正态分布样本, 由此得到  $\Sigma$  的最小二乘估计为:

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \xi_t' \xi_t$$

### 2. 1. 4 MGARCH 模型的杠杆效应

Black(1976)<sup>[37]</sup>认为对于股票收益来讲，负的冲击比正的的冲击对波动性的影响较大。描述这种非对称性的单维模型有 Nelson(1991)<sup>[25]</sup>的 EGARCH 模型，Glosten(1993)<sup>[38]</sup>的 GIR 模型和 Zakoian(1994)<sup>[39]</sup>的门类 ARCH 模型。对于多维的序列，同样的观点也存在：方差和协方差对于正的冲击和负的冲击的反应不同。在多维的框架之下，外界的冲击可以被定义为  $\varepsilon_t$  或者  $\eta_t$ 。

考虑到误差的符号的模型是由 Kroner 和 Ng(1998)<sup>[40]</sup>提出的非对称动态协方差(ADC)模型。

$$\theta_{ijt} = \bar{\omega}_{ij} + a_i' \varepsilon_{t-1} \varepsilon_{t-1}' a_j + g_i' H_{t-1} g_j + b_i' v_{t-1} v_{t-1}' b_j$$

其中，  $v_t = \max[0, -\varepsilon_t]$

Kroner 和 Ng(1998)<sup>[40]</sup>用 ADC 模型研究了大公司和小公司的收益，他们发现关于大公司的坏消息既可以引起大公司的波动性的变化也可以引起小公司波动性的变化，而且坏消息的到来可以增加条件协方差。但是小公司的消息却只能产生较小的效果。

以上内容从多维时间序列角度研究动态相关性，我们用表格对主要的模型进行比较：

表 2.1 多维 GARCH 模型优缺点比较

模型	优点	缺点
VECH	全面、准确刻画波动的相关特性，可以研究波动溢出效应，经济意义明确	参数多，估计困难
BEKK	可以保证 $H_t$ 的正定性，具有较少的参数个数	参数经济意义不够明确
CCC	模型形式简化，参数较少，有一定的经济意义	常数相关约束可能并不成立
DCC	具有动态性，形式简化，允许分开估计	相同的持续性

## 2. 1. 5 其它的多维波动和相关性模型

除了多维 GARCH 模型以外，还有随机波动性 (SV) 模型和已实现的波动性来描述多维相关性。Harvey(1994)<sup>[41]</sup>关于多维随机波动模型的研究认为条件方差协方差矩阵依赖于不可观测的或者潜在的过程，而不是多维 GARCH 模型通常认为的过去的观测值。一个多维随机波动模型由 N 个单维随机波动模型构成：

$$\begin{aligned} \varepsilon_{it} &= c_i z_{it} \exp(0.5h_{it}) \quad i=1, \dots, N, \quad t=1, \dots, T \\ h_t &= \beta h_{t-1} + \zeta_t, \quad |\beta| < 1 \end{aligned}$$

其中  $c_i$  是比例参数，创新向量  $z_t = (z_{1t}, \dots, z_{Nt})'$ ， $E(z_t) = 0$ ， $Var(z_t) = \Sigma_z$ 。

写作向量的形式为：

$$h_t = \phi h_{t-1} + \eta_t, \quad h_t = (h_{1t}, \dots, h_{Nt})'$$

其中， $\eta_t \square i.i.d.N(0, \Sigma_\eta)$ 。

SV 模型和 ARCH 类模型是广泛应用于经济、金融时间序列波动性分析和金融实证研究的有力工具。从数据生成过程的角度来看，ARCH 类模型描述的是离散的可观测时间序列的波动特征，其波动过程由过去的观测值和过去误差的平方项线性表示。而 SV 模型则是一类随机微分方程的离散化表示形式。

由于 SV 模型的波动是由一个不可观测的随机过程决定的，它被认为更加适合于金融领域的实际研究，但是由于这一类模型的参数估计比较困难，因此在一定程度上影响了它的实际应用。如果能够找到这两类模型之间的联系，无疑为充分利用这两类模型的优点并为深入了解其数据生成过程的内在机制提供便利条件。

另一个描述多维波动和相关性模型的方向就是 Andersen et al(2003)<sup>[42]</sup>提出的已实现的波动性，每天的方差和协方差用日间的累积收益来计算，这种方法比较适用于预测的目的。但是这种模型的主要缺点是提取日间收益是比较耗费成本的，并且不一定所有的资产的日间收益都可以获得。而且，日间收益数据的处理和计算也是比较麻烦的。

## 2. 2 估计方法

构建多维 GARCH 模型的估计方法所作的假设是：标准残差项（即创新项）是独立同分布（i.i.d.），因此，接下来必须确定标准残差项的密度函数，假设为  $g(z_t(\theta))$ 。估计问题转化为求样本的  $\log$  似然函数  $L_T(\theta)$

（ $L_T = \sum_{t=1}^T \log f(y_t | \theta, I_{t-1})$ ）的最大值。其中  $f(y_t | \theta, I_{t-1}) = |H_t|^{-1/2} g(H_t^{-1/2}(y_t - \mu_t))$ ,

而  $|H_t|^{-1/2}$  是 Jacobian 项，把创新项转化为观测值。

在大部分的文献中一般采用多变量正态分布，那么样本  $\log$  似然函数为：

$$L_T(\theta) = -\frac{1}{2} \sum_{t=1}^T \log |H_t| - \frac{1}{2} \sum_{t=1}^T (y_t - \mu_t)' H_t^{-1} (y_t - \mu_t) \quad (2.17)$$

但是，在实际处理每天或每周的数据时，标准残差项为正态分布的假设一般被拒绝。特别是，大多数金融资产收益的峰度大于 3，这就意味着它们比正态分布有更多的极端值。而且，它们的无条件分布的尾部也更厚。

所幸的是，Bollerslev 和 Wooldridge(1992)<sup>[43]</sup>证明了即使数据生成过程（DGP）不是条件 Gaussian 分布，那么最大化  $\log$  似然函数也可以获得  $\theta$  的一致性估计。这个估计叫做伪最大似然估计（QMLE）。Jeantheau(1998)<sup>[44]</sup>证明了多维 GARCH 模型的 Gaussian QMLE 估计的强相合性。

Harvey et al. (1992)<sup>[45]</sup> 和 Fiorentini et al. (2003)<sup>[46]</sup>认为多维 Gaussian 密度函数一个很自然的替换是 Student 密度函数，同时也增加了一个参数—自由度  $\nu$ 。当  $\nu$  趋向于无穷时，Student 密度函数趋于正态密度函数。当  $\nu$  趋于 0 时，密度函数的尾部变得越来越厚。Student 密度函数定义为：

$$g(\eta_t | \theta, \nu) = \frac{\Gamma(\frac{\nu+N}{2})}{\Gamma(\frac{\nu}{2})[\pi(\nu-2)]^{N/2}} \left[1 + \frac{\eta_t' \eta_t}{\nu-2}\right]^{-\frac{N+\nu}{2}} \quad (2.18)$$

但是多维 Student 密度函数却不能抓住创新项的偏度。为了正确拟合收益的偏度和峰度，多变量混合正态密度函数（Vlaar 和 Palm, 1993<sup>[47]</sup>）、一般的双曲线分布（Barndorff-Nielsen 和 Shephard, 2001<sup>[48]</sup>）和多变量偏—Student 密度函数（Bauwens 和 Laurent, 2002<sup>[49]</sup>）纷纷与 MGARCH 结合起来，更加准确地拟合实际



中的多维收益率的波动性和相关性。

另外一种思路是 Hafner 和 Rombouts(2004)<sup>[50]</sup>提出的半参数估计方法。估计分为两步，首先通过 QML 得到创新项的一致估计，然后利用第一步的结果对  $g(\cdot)$  进行非参数估计，最后利用  $\hat{g}(\cdot)$  定义的似然函数估计 GARCH 模型的参数。

### 2. 2. 1 两步估计

多变量 GARCH 模型的 ML 和 QML 估计的渐近性质在较严格的假设条件下成立，但在放松假设条件下进行渐近性质的证明还是存在着一定的困难。Jeantheau(1998)<sup>[44]</sup>证明了 QMLE 的一致性。Gourieroux(1997)<sup>[6]</sup>证明了在较强的假设条件下的 QMLE 的渐近正态性。Comte 和 Lieberman(2003)<sup>[51]</sup>证明了 BEKK 的渐近性质。DCC 模型的一个有用的特点是可以两步估计方法进行一致的估计。Engle 和 Sheppard (2001)<sup>[52]</sup>的研究表明 DCC<sub>E</sub> 的对数似然函数可以写作均值、波动性部分和相关部分的和，如下所示：

$$\begin{aligned} L &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |H_t| + \varepsilon_t' H_t^{-1} \varepsilon_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |D_t R_t D_t| + \varepsilon_t' D_t^{-1} R_t^{-1} D_t^{-1} \varepsilon_t) \\ &= [-\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |D_t|^2 + \varepsilon_t' D_t^{-2} \varepsilon_t)] + [-\frac{1}{2} \sum_{t=1}^T (\log |R_t| + \eta_t' R_t^{-1} \eta_t - \eta_t' \eta_t)] \end{aligned}$$

参数估计方法：

第一步：估计上式的第一部分

$$\begin{aligned} QL1_T(\theta_1^*) &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |D_t|^2 + \varepsilon_t' D_t^{-2} \varepsilon_t) \\ &= -\frac{1}{2} \sum_{t=1}^T \sum_{i=1}^2 (\log(2\pi) + \log(h_{ii,t}) + \frac{\varepsilon_{i,t}^2}{h_{ii,t}}) \end{aligned} \quad (2.19)$$

其实就是单个 GARCH 似然函数的和。

第二步：对第二部分进行估计

$$QL2_T(\theta_2^* | \theta_1^*) = -\frac{1}{2} \sum_{t=1}^T (\log |R_t| + \eta_t' R_t^{-1} \eta_t - \eta_t' \eta_t) = -\frac{1}{2} \sum_{t=1}^T (\log |R_t| + \eta_t' R_t^{-1} \eta_t) \quad (2.20)$$

Engle 和 Sheppard(2001)<sup>[52]</sup>证明了两步估计得到的  $\hat{\theta}_1^*$  和  $\hat{\theta}_2^*$  虽不是有效的，但是一致的估计。Brooks et al. (2003)<sup>[53]</sup>对估计不同的多维 GARCH 模型的各种软件包进行了回顾和总结，随着对软件技术研究的深入，给应用性研究者提供新的技术和方法。

### 2. 2. 2 基于遗传算法的似然函数的估计方法

传统的估计多维 GARCH 模型的最优化算法，采用 BHHH 的方法，利用目标函数的梯度信息进行迭代和优化，但是多元 GARCH 模型的参数较多，在优化的搜索过程中容易产生震荡，导致局部最优解，且结果对初始值的选取比较敏感。为解决传统算法的问题，樊智(2003)<sup>[54]</sup>认为，引入人工智能方法是一种较好的途径，他采用的遗传算法具有以下优点：

(1) 遗传算法的并行性，不易陷入局部最优解。

(2) 搜索过程中，基本不用目标函数的梯度信息，仅用适应值函数来评估个体，避免了梯度算法中由于搜索的震荡或梯度的不存在而使似然函数陷入局部最优解。

(3) 可以选择不同的编码和操作方式，设置不同的交叉和变异概率，算法具有更好的适应性。

### 2. 2. 3 多维 GARCH 模型的半参数有效估计

多维 GARCH 模型的残差序列的分布选择不一定符合实际情况，值得庆幸的是 Bollerslev, Wooldridge(1992)<sup>[43]</sup>和 Jeantheau(1998)<sup>[44]</sup>已经证明了即使基本的分布不是正态的，伪极大似然法得到的参数估计仍然是相合的。然而，错误假定下残差序列的分布所得到的参数的极大似然函数法估计相对于真正的分布下的极大似然函数估计的效率损失还没有人研究。王传美(2005)<sup>[55]</sup>假定方差协方差矩阵有一种特定的参数形式，采用一种非参数方法，让数据来决定信息列的分布。对于单维 GARCH 模型，Engle, Gonzalez Rivera(1991)<sup>[56]</sup>，Linton(1993)<sup>[57]</sup>，Drost 及 Klaassen(1997)<sup>[58]</sup>得到了单维 GARCH 模型半参数有效估计。

## 2. 3 诊断性检验

因为估计一个 MGARCH 模型是较费时间的。因此，在估计前检验数据是否符合多变量 GARCH 模型的性质显得非常重要。同样，在估计后，对 MGARCH 的设定性检验（模型拟合数据是否适当）也是十分重要的。对单维 GARCH 模型的诊断性检验已经比较完整，而多变量模型的设定性检验体系还显得不很成熟。

在现有的 MGARCH 模型中，我们可以把设定性检验分为两类，第一类是分别针对一维 GARCH 的检验，第二类是把多变量 GARCH 模型作为一个整体向量进行检验。但是，必须注意的是，虽然，一维 GARCH 模型的检验可以为多变量 GARCH 模型的检验提供一个指导，但是由于同时期扰动项的相关性使单个等式的统计量不是独立的，因此多维 GARCH 需要一个联合检验（Dufour 2003<sup>[59]</sup>）。Ding 和 Engle(2001)<sup>[19]</sup>证明，在多维 GARCH 模型中，标准残差  $\eta_t = H_t^{-1/2} \varepsilon_t$  必须满足：

- (1)  $E(\eta_t \eta_t') = I_N$
- (2)  $Cov(\eta_{it}^2, \eta_{jt}^2) = 0$ , 对于  $i \neq j$
- (3)  $Cov(\eta_{it}^2, \eta_{j,t-k}^2) = 0$ , 对于  $k > 0$

条件(1)用来发现条件期望的错误设定；条件(2)用来检验条件分布是否 Gaussian；条件(3)用来检验  $H_t$  的动态设定的充分性。

正如 Tse(2002) 提出的，条件异方差的诊断检验可以分为三类：Bos-Pierce-Ljung 的 portmanteau 检验，基于残差的检验和 Lagrange 乘子检验。

### 2. 3. 1 Portmanteau 统计量

广泛应用于检验 ARCH 效应的诊断性方法是 Box-Pierce/Ljung-Box portmanteau 检验。根据 Hosking(1980)<sup>[60]</sup>，多变量的情况下 Ljung-Box 统计量为

$$HM(M) = T^2 \sum_{j=1}^M (T-j)^{-1} tr\{C_{Y_t}^{-1}(0)C_{Y_t}(j)C_{Y_t}^{-1}(0)C_{Y_t}'(j)\} \quad (2.21)$$

其中， $Y_t = vech(y_t, y_t')$ ， $C_{Y_t}(j)$  是  $j$  阶的样本自协方差矩阵。这个检验在原假设：

不存在 ARCH 效应之下， $HM(M)$  渐近服从  $\chi^2(K^2M)$ 。

### 2. 3. 2 基于残差的诊断

这个检验是用标准残差  $\hat{\eta}_i$  的交叉乘积对一些解释变量作回归，然后对回归系数的显著性作检验。这个检验的缺陷在于，因为是先估计模型，然后再做回归，因此回归系数依赖于被估计的参数，因此通常的 OLS 理论不再适用。Tse(2002)<sup>[31]</sup> 在此框架内建立了 OLS 估计的渐近分布。Tse 所建立的回归：

$$\begin{aligned}\hat{\eta}_{it}^2 - 1 &= \hat{d}_{it}' \delta_i + \zeta_{it}, \quad i = 1, \dots, N \\ \hat{\eta}_{it} \hat{\eta}_{jt} - \hat{\rho}_{ijt} &= \hat{d}_{ijt}' \delta_{ij} + \zeta_{ijt}, \quad 1 \leq i < j \leq N\end{aligned}$$

其中， $\hat{\eta}_{it} = \hat{\varepsilon}_{it} / \sqrt{\hat{h}_{it}}$ ， $i = 1, \dots, N$  是标准残差， $\hat{\rho}_{ijt} = \hat{h}_{ijt} / \sqrt{\hat{h}_{it} \hat{h}_{jt}}$  是  $y_{it}$  和  $y_{jt}$  之间估计得到的条件相关系数。 $\delta_i$  和  $\delta_{ij}$  是回归系数。 $\hat{d}_{it}$  和  $\hat{d}_{ijt}$  分别是  $d_{it} = (u_{i,t-1}^2, \dots, u_{i,t-M}^2)'$  和  $d_{ijt} = (u_{i,t-1} u_{j,t-1}, \dots, u_{i,t-M} u_{j,t-M})'$  的估计值。

基于残差的诊断分析的好处在于它们集中了引起 ARCH 效应的可能原因的不同方面。

### 2. 3. 3 Lagrange 乘子检验

在多维 GARCH 模型中，为了减少参数的数量，通常的做法是引入约束。例如 Bollerslev(1990)<sup>[29]</sup> 提出的常数条件相关 (CCC) 假定条件相关矩阵为常数。Tse(2002)<sup>[31]</sup> 提出了检验常数相关的检验方法。原假设是  $h_{ijt} = \rho_{ij} \sqrt{h_{it} h_{jt}}$ ，其中条件方差是 GARCH(1,1)，备选假设是  $h_{ijt} = \rho_{ijt} \sqrt{h_{it} h_{jt}}$ 。在原假设的条件下，LM 统计量渐近服从  $\chi^2(N(N-1)/2)$ 。

Engle 和 Sheppard(2001)<sup>[52]</sup> 提出了检验常数相关假设的另外一种方法。原假设是： $H_0 : R_t = \bar{R}$ ，备选假设是：

$$H_1 : \text{vech}(R_t) = \text{vech}(\bar{R}) + \beta_1^* \text{vech}(R_{t-1}) + \dots + \beta_p^* \text{vech}(R_{t-p})。$$

这个检验很容易实现，因为  $H_0$  意味着  $X_t = \beta_0^* + \beta_1^* X_{t-1} + \dots + \beta_p^* X_{t-p} + u_t^*$  的回

归系数为 0，其中， $X_t = \text{vech}^u(\hat{\eta}_t \hat{\eta}_t' - I_N)$ ， $\hat{\eta}_t = \hat{R}^{-1/2} \hat{D}_t^{-1} \hat{\varepsilon}_t$  是  $N \times 1$  维标准残差向量， $D_t = \text{diag}(h_{11t}^{1/2} \cdots h_{NNt}^{1/2})$ ， $\text{vech}^u$  是一个向量算子，但是它只选择主对角线上的元素。

### 3. 多维相关性模型预测分析

对多种资产收益的未来相关关系进行准确的预测是研究投资和风险管理的基础。同时，评价各种多维 GARCH 模型表现好坏的方法之一就是看能否对未来做出很好的预测。

针对单维 GARCH 模型族的预测能力，国内的研究有惠晓峰（2003）<sup>[61]</sup> 基于 GARCH 模型进行的人民币汇率预测和刘国旗（2000）<sup>[62]</sup> 对中国股市的预测。对于多维 GARCH 模型的预测能力，国外的研究有 Engle(2001)<sup>[52]</sup> 利用 DCC 多维 GARCH 模型对 S&P 500 和道琼斯工业股价平均数指数的预测。本文利用非对称的 DCC 模型和 CCC 多维 GARCH 模型对中国主要股指之间的相关性进行预测，以期为中国股市的投资管理和风险管理提供理论支持。

多维相关性模型的预测不仅涉及到单个市场收益的波动性预测，同时还必须对各个市场收益之间的相关性进行预测。因此，本文先介绍一些基本符号。

#### 3.1 波动模型预测的基本符号

波动性对参与金融市场的任何人都显的相当重要。在一般的情况下，简单来讲，波动性和风险是等价的。证券价格的高波动性意味着证券的价格是不可依赖的，是具有风险的。但是，对于投资者来讲，特别是对于衍生产品的投资者来讲，获得一个准确的波动性信息，能够对未来的波动性进行准确的预测，从而合理地投资组合进行管理显得至关重要。例如 Black& Scholes<sup>[63]</sup> 提出的期权定价理论认为波动性是决定期权价值的最重要的角色之一。虽然标的资产收益的波动性仅仅是 BS 期权定价公式五个参数（股票价格，执行价格，期权到期的时间，波动性，利率）之一，但是因为它并不象其他的四个参数可以直接在市场上观测到，因此对波动性进行准确的预测显得尤其重要。

在单维离散的情况下，本文引入这样一个过程：

$$y_t = y(t) \quad t = 1, 2, \dots$$

我们假定  $y_t$  的二阶矩存在，同时相应的条件期望和方差为：

$$\mu_{t|t-1} = E[y_t | \psi_{t-1}]$$

$$\sigma_{t|t-1}^2 = \text{Var}[y_t | \psi_{t-1}] = E[(y_t - \mu_{t|t-1})^2 | \psi_{t-1}] \quad (3.1)$$

其中， $\psi_t$  指信息集，反映了直到  $t-1$  时刻的所有相关信息。

正是因为条件期望并入了一步向前预测的信息，因此它和无条件期望有所不同，即  $\mu_{t|t-1} \neq E(y_t)$ ，同时  $\sigma_{t|t-1}^2 \neq \text{Var}(y_t)$ 。这种条件和无条件矩的差别为时间序列的预测奠定了基础。

正如前面所述，大部分波动性估计和预测的应用和发展都集中在金融经济学中，因此，本文集中在对收益波动性的预测上。令  $p(t)$  指资产的对数价格，离散区间  $[t-h, t]$ ， $h > 0$  的收益为：

$$r(t, h) = p(t) - p(t-h)$$

当  $h=1$  时， $r(t) = p_t - p_{t-1}$  为日收益。通常在离散情况下，也可写为：

$$r_t = p_t - p_{t-1} \quad (3.2)$$

考虑离散样本收益过程  $r_t$  被分解为条件期望收益和一个残差项，其中残差项被表示为标准白噪声过程乘以条件波动性，表示为

$$r_t = \mu_{t|t-1} + \varepsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} \eta_t \quad (3.3)$$

$\eta_t$  是期望为 0，方差为 1 的序列不相关过程。这种方法的方便性体现在如果所有相关的信息都是已知的并且模型是正确设定的，波动性就可以计算出来了。

(3.1) 式提供了无偏的一步预测，通过与收益残差项的平方实现值进行对比，从而对相应的模型的设定和预测评价，做这种评价的依据为：

$$\varepsilon_t^2 = \sigma_{t|t-1}^2 \eta_t^2 = \sigma_{t|t-1}^2 + \sigma_{t|t-1}^2 (\eta_t^2 - 1) \quad (3.4)$$

上式第二项的期望为 0，因此可以证明条件方差  $\sigma_{t|t-1}^2$  的无偏估计是  $\varepsilon_t^2$ 。但是收益的残差项平方  $\varepsilon_t^2$  相对于  $\sigma_{t|t-1}^2$  有很大的噪声，即个别误差  $\text{Var}(\eta_t^2)$ 。实际上，对于日收益和周收益，这个个别误差有不同的大小。因此，即使和收益残差项的平方项相比，条件方差  $\sigma_{t|t-1}^2$  被看作波动性正确的预测，收益残差项的平

方项也不是一个判断预测好坏的可靠标准。

在实际中，另外一个描述资产收益的方法是连续时间过程。当对数价格服从连续时间过程，收益过程可写作标准随机微分等式：

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) \quad t \geq 0 \quad (3.5)$$

其中， $\mu(t)$ 指的是漂移项， $\sigma(t)$ 指点波动率， $W(t)$ 指标准布朗运动，在无穷小的时间间隔之内：

$$r(t, \Delta) \equiv p(t) - p(t - \Delta) \approx \mu(t - \Delta) \cdot \Delta + \sigma(t - \Delta) \Delta W(t)$$

其中， $\Delta W(t) \equiv W(t) - W(t - \Delta) \sim N(0, \Delta)$ ，当 $\Delta = 1$ ，漂移项 $\mu(\tau) \equiv \mu_{t|t-1}$ ， $\sigma(\tau) \equiv \sigma_{t|t-1}$ ，其中， $t-1 < \tau \leq t$ ，当假设 $z_t \sim i.i.d.N(0,1)$ ，可以简化为离散收益的等式(3.3)。

需要说明的一点是，在时间间隔 $[t-1, t]$ 内， $\mu(\tau)$ 和 $\sigma(\tau)$ 不是常数，间隔期为1的收益一般表达式为：

$$r(t) = p(t) - p(t-1) = \int_{t-1}^t \mu(s)ds + \int_{t-1}^t \sigma(s)dW(s) \quad (3.6)$$

假定收益的创新项是布朗运动过程，并且是独立同分布(i.i.d.)，那么收益总的变化量应该等于过去累积的即时方差，因此引出了整合的波动性：

$$IV(t) \equiv \int_{t-1}^t \sigma^2(s)ds \quad (3.7)$$

等式(3.7)对于日或周区间间隔也成立，对于小的时间间隔 $\Delta$

$$Var(r(t) | \psi_{t-1}) = E\left[\sum_{j=1}^{1/\Delta} \sigma^2(t - j/\Delta) \cdot \Delta | \psi_{t-1}\right] = E[IV(t) | \psi_{t-1}] \quad (3.8)$$

可以看出，整合方差的衡量更接近于离散样本收益的条件方差 $\sigma_{t|t-1}^2$ ，同时它也可以作为在一步预测期间的已实现的波动性，它简单地反映了在收益区间即时波动性的累积效应。换句话讲，整合的方差是事后的已实现值，并且直接和事前的波动性预测相比。从另一方面讲，整合的波动性中并没有出现误差项，因此，可以作为评价事前波动性预测效果的事后标准。

本文在用各种模型进行预测之前，先强调几个理论特点和实际操作出现的



问题:

- (1) 即使潜在的模型是连续时间模型,波动性模型的估计和预测一般局限于离散时间间隔;
- (2) ARCH 模型和随机波动模型的差别在于信息集的假设上。SV 模型的波动是由一个不可观测的随机过程决定的,ARCH 类模型描述的是离散的可观测时间序列的波动特性;
- (3) 要认识到事前预测和事后已实现的波动性的区别。

### 3. 2 波动性预测的意义

波动性预测在学术文献中已有广泛应用,本章节重点讨论它在实际中的应用,特别在金融领域中的应用。因此,我们没有假设特定的波动性预测模型,而是先讨论一般的统计预测应用,然后再进一步讨论在金融应用中的细节问题。

#### 3. 2. 1 普通的预测应用

假定收益变量可以写为:

$$y_{t+1} = \mu_{t+1|t} + \sigma_{t+1|t} z_{t+1}, \quad z_{t+1} \square i.i.d.F, \quad (3.9)$$

其中,  $\{y_{t+1}\}$  表示离散的单变量随机过程,  $F$  表示  $z_{t+1}$  的分布,  $z_{t+1}$  的期望为 0, 方差为 1。

##### 3. 2. 1. 1 点预测

首先定义预测损失函数,此函数将事前预测  $\hat{y}_{t+1|t}$  和事后已实现的值  $y_{t+1}$  映射到一个损失函数  $L(y_{t+1}, \hat{y}_{t+1|t})$ , 用以反映已实现的和预测值的差别。损失函数确切的形式依赖于具体的预测的使用。但是,在大部分情况下,损失函数可以被合理地写作:

$$e_{t+1} = y_{t+1} - \hat{y}_{t+1} \quad (3.10)$$

从而  $L(y_{t+1}, \hat{y}_{t+1|t}) = L(e_{t+1})$  叫做预测误差损失函数。

在实际应用中大多数文献采用对称二次预测误差损失函数,最优的点预测

是此过程的条件期望形式:

$$\hat{y}_{t+1|t} = \text{Arg min}_y E[(y_{t+1} - \hat{y})^2 | \psi_t] = \mu_{t+1|t} \quad (3.11)$$

从上式可以看出在条件期望直接依赖于条件波动性的情况下, 波动性预测与计算最优点预测才是相关的, 否则不相关。这个规则在金融市场上非常适合, 因为, 预期收益一般都与市场风险因子的波动性相关。

对于非对称的损失函数, 波动性是最优点预测的关键因素。考虑非线性非对称损失函数:

$$L(e_{t+1}) = a|e_{t+1}|I(e_{t+1} > 0) + b|e_{t+1}|I(e_{t+1} \leq 0) \quad (3.12)$$

其中,  $a, b > 0$ ,  $I(\cdot)$  是示性函数, 当预测误差为正或为负时具有不同的损失函数。那么最优预测可以写作:

$$\hat{y}_{t+1|t} = \mu_{t+1|t} + \sigma_{t+1|t} F^{-1}(a/(a+b)) \quad (3.13)$$

具体的讨论请参见 Granger(1969)<sup>[64]</sup>

### 3. 2. 1. 2 区间预测

Engle(1982)<sup>[10]</sup>提出根据条件期望预测构造准确的区间预测, 区间预测包括上下界。区间预测被写为:

$$\hat{y}_{t+1|t} = \{\mu_{t+1|t} + \sigma_{t+1|t} F^{-1}(p/2), \mu_{t+1|t} + \sigma_{t+1|t} F^{-1}(1-p/2)\} \quad (3.14)$$

从上式可以看出, 波动性预测在区间预测中起着重要的作用。

### 3. 2. 1. 3 包括符号的概率预测

一个资产组合管理者最关心的预测问题是: 股票指数的收益是否大于已知的无风险债券收益。一个评级机构对某个公司进行预测, 判断未来公司的资产是否大于还是小于公司的债务, 是否会引发破产。这类问题关心的是变量是否超出一个事先固定的值:

$$L(y_{t+1}, \hat{y}_{t+1|t}) = (I(y_{t+1} > c) - \hat{y}_{t+1|t})^2 \quad (3.15)$$

让预期损失函数的一阶倒数为 0, 从而使预期损失函数最小, 得到:

$$\hat{y}_{t+1|t} = E[I(y_{t+1} > c) | \psi_t] = P(y_{t+1} > c | \psi_t) = 1 - F((1 - \mu_{t+1|t}) / \sigma_{t+1|t}) \quad (3.16)$$

从上式可以看出，波动预测在区间预测中也有着不可忽略的作用。

### 3. 2. 1. 4 密度预测

在许多情况下，对变量的整个条件密度的预测是我们所关注的问题，这个预测采用概率分布函数的形式，为

$$\hat{y}_{t+1|t} = f_{t+1|t}(y) \equiv f(y_{t+1} = y | \mu_{t+1|t}, \sigma_{t+1|t}) = f(y_{t+1} = y | \psi_t) \quad (3.17)$$

### 3. 2. 2 金融应用

风险和预期收益之间的相互权衡构成了现代金融理论的重要组成部分。波动性的预测是资产定价和风险管理中重要的内容。

#### 3. 2. 2. 1 VaR 和 ES

考虑  $N$  个风险资产构成的向量  $R_{t+1}$ ，它们在资产组合中相应的权重为  $W_t$ ，投资组合的收益为：

$$r_{w,t+1} = \sum_{i=1}^N w_{i,t} r_{i,t+1} \equiv W_t' R_{t+1} \quad (3.18)$$

如果我们用单维过程模拟投资组合的收益，得到：

$$r_{w,t+1} = \mu_{w,t+1|t} + \sigma_{w,t+1|t} z_{w,t+1} \quad z_{w,t+1} \square i.i.d. F_w \quad (3.19)$$

那么，VaR 可以写作：

$$VaR_{t+1|t}^p = \mu_{w,t+1|t} + \sigma_{w,t+1|t} F_w^{-1}(p) \quad (3.20)$$

另外的一种风险测度 ES 提供了关于分布函数尾部的额外信息，定义为当损失超过 VaR 时损失的预期值。表示为：

$$ES_{t+1|t}^p \equiv E[r_{w,t+1} | r_{w,t+1} < VaR_{t+1|t}^p] = \mu_{w,t+1|t} + \sigma_{w,t+1|t} EF_w^p \quad (3.21)$$

其中， $z_{w,t}$  是独立同分布的，因子  $EF_w^p$  是常数，并且依赖于分布函数  $F_w$  的形状。动态的波动性在风险测度 VaR 和 ES 中起到了类似的效果。

上述分析假定投资组合收益过程是单维的，并且是给定权重的函数。但是在实际操作中，计算最优投资组合权重就不再适用。当计算在风险 VaR 的约束

之下使预期收益最大化时，这时就需要一个多维的相关性模型。这时，我们用另外一种方法，假定单独模拟每个资产的收益，那么收益向量可以写作：

$$R_{t+1} = M_{t+1|t} + \Omega_{t+1|t}^{1/2} Z_{t+1} \quad Z_{t+1} \square i.i.d.F \quad (3.22)$$

其中， $M_{t+1|t}$  和  $\Omega_{t+1|t}$  分别代表条件期望收益向量和方差协方差矩阵， $Z_{t+1}$  服从独立同分布的分布  $F$ 。随机过程向量  $Z_t$  的各个元素是独立的，并且每个元素的期望为 0，方差为 1。那么投资组合收益  $W_t' R_{t+1}$  的期望和方差分别为：

$$\begin{aligned} \mu_{w,t+1|t} &= W_t' M_{t+1|t} \\ \sigma_{w,t+1|t}^2 &= W_t' \Omega_{t+1|t} W_t \end{aligned} \quad (3.23)$$

当  $Z_{t+1} \square N(0, I)$  时，多变量正态分布的线性混合也是正态分布，因此  $r_{w,t+1} \equiv W_t' R_{t+1} \square N(\mu_{w,t+1|t}, \sigma_{w,t+1|t}^2)$ ，但是这个性质在其他多变量分布的情况下并不成立，因此在不符合正态分布假设的情况下，并不能给出 VaR 和 ES 的解析形式，必须用 Monte Carlo 模拟方法。

### 3. 2. 2. 2 协方差风险：时变的 Betas 和条件 Sharpe 比率

上述讨论主要从统计角度衡量投资组合风险，接下来将从更加基础的角度研究给定风险下资产的预期收益问题。假定没有套利机会，那么存在一个随机贴现因子  $SDF_{t+1}$  可以用来对任何资产进行定价，表示为：

$$E[SDF_{t+1}(1+r_{i,t+1}) | \psi_t] = 1 \quad (3.24)$$

无风险资产的收益必须满足：

$$1+r_{f,t} = E[SDF_{t+1} | \psi_t]^{-1} \quad (3.25)$$

从 (3.25) 式可以推出：任何风险资产的预期超额收益与风险资产和随机贴现因子的协方差成比例，即

$$E[r_{i,t+1} - r_{f,t} | \psi_t] = -(1+r_{f,t})Cov(SDF_{t+1}, r_{i,t+1} | \psi_t) \quad (3.26)$$

假定随机贴现因子与市场收益是线性相关的，即  $SDF_{t+1} = a_t - b_t(1+r_{M,t+1})$ 。

根据  $E[SDF_{t+1}(1+r_{M,t+1}) | \psi_t] = 1$  和  $1+r_{f,t} = E[SDF_{t+1} | \psi_t]^{-1}$ ，得到：

$$\begin{aligned} a_t &= (1+r_{f,t})^{-1} + b_t \mu_{M,t+1|t} \\ b_t &= (1+r_{f,t})^{-1} (\mu_{M,t+1|t} - r_{f,t}) / \sigma_{M,t+1|t}^2 \end{aligned} \quad (3.27)$$

其中,  $\mu_{M,t+1|t} \equiv E[1+r_{M,t+1} | \psi_t]$ ,  $\sigma_{M,t+1|t}^2 = \text{Var}[r_{M,t+1} | \psi_t]$ , 由于市场收益的矩的动态性使得 SDF 也是时变的。与 Markowitz(1952)<sup>[65]</sup>和 Sharpe(1964)<sup>[66]</sup>提出的经典的 CAPM 模型类似, 条件预期超额收益必须满足:

$$E[r_{i,t+1} - r_{f,t} | \psi_t] = \beta_{i,t} (\mu_{M,t+1|t} - r_{f,t}) \quad (3.28)$$

其中, 条件  $\beta$  被定义为  $\beta_{i,t} \equiv \text{Cov}(r_{M,t+1}, r_{i,t+1} | \psi_t) / \sigma_{M,t+1|t}^2$ , 而且预期的风险调整收益, 也叫作条件 Sharpe 比率, 可写作:

$$SR_t = E[r_{i,t+1} - r_{f,t} | \psi_t] / \text{Var}(r_{i,t+1} | \psi_t)^{1/2} = \text{Corr}(r_{M,t+1}, r_{i,t+1} | \psi_t) / \sigma_{M,t+1|t}$$

这个简单的资产定价模型描述了各种资产的预期收益受到整个市场收益的期望和动态的波动性影响, 同时也受到市场与各种资产的动态协方差的影响。

### 3. 2. 2. 3 时变协方差的资产配置

假定一个投资者在一定的投资组合收益的约束之下, 通过最优化问题使投资组合的方差最小, 这时需要假定一个动态协方差矩阵。这个问题可以转化为投资者通过选择一个投资组合权重  $W_t$ , 解决二次规划的问题:

$$\min W_t' \Omega_{t+1|t} W_t \text{ s.t. } W_t' M_{t+1|t} = \mu_p \quad (3.29)$$

从相应的一阶条件可知, 风险资产的投资组合权重为:

$$W_t^* = \frac{\Omega_{t+1|t}^{-1} M_{t+1|t}}{M_{t+1|t}' \Omega_{t+1|t}^{-1} M_{t+1|t}} \mu_p$$

无风险资产的权重为:

$$w_{f,t}^* = 1 - \sum_{i=1}^N w_{i,t}^*$$

从 (3.29) 式中得出投资组合的 Sharpe 比率为:

$$SR_t = \mu_p / \sqrt{W_t^{*'} \Omega_{t+1|t} W_t^*}$$

正如 CAPM 定价模型中所讨论的一样, 动态的波动性和协方差对于动态的资产配置都显得十分重要。

### 3. 2. 2. 4 动态波动性的期权定价

上述波动性的应用仅适用于具有线性回报的基础证券，譬如股票，债券，外汇和期货协议。而在这一节中我们考虑欧式看涨期权的定价，其中，执行价格为  $K$ ， $T$  时刻到期。由于执行期权在未来将获得一个非线性的回报，因此需要特殊的工具对期权进行定价和风险管理。

在 Black-Scholes-Merton(BSM)定价公式中，收益被假定为具有常数波动的正态分布，同时假设无风险收益是常数并且没有交易成本。在这种情况下，期权的价格是：

$$c_t = BSM(s_t, \sigma^2, K, r_f, T) = s_t \Phi(d) - K \exp(-r_f T) \Phi(d - \sigma \sqrt{T})$$

$s_t$  表示现在资产的价格， $d = (\ln(s_t / K) + T(r_f + \sigma^2 / 2)) / (\sigma \sqrt{T})$ ， $\Phi(\cdot)$  表示累积正态分布函数。

同时，当把理论价格与真实的市场价格作对比时，BSM 定价公式中的常数波动性的假设会引起系统的定价误差，它体现在波动性微笑上，表明了对于实值和虚值的期权，系统性的定价过低。因此，学术界引入了随机波动性，增加了比正态分布更厚的尾部，从而得到了比常数波动性 BSM 模型得出的实值和虚值期权更高的价格，改善了 BSM 定价的误差。

Hull 和 White(1987)<sup>[67]</sup>提出了基础资产收益服从独立的随机波动因子，假定这个额外的风险因子在均衡中不能被定价，那么 Hull 和 White 的看涨期权价格等于预期的 BSM 价格，所用到的波动性采用了未来的整合波动性。定义期权的波动性等于在期权的剩余时间内的点波动性的积分，写作：

$$IV(T, t) = \int_t^T \sigma^2(\mu) du \quad (3.30)$$

其中， $IV(T, t) = IV(T) + IV(T-1) + \dots + IV(t+1)$  给出了多时期区间的整合的方差的概念，Hull 和 White 定价公式可以简化为：

$$C_t = E[BSM(IV(T, t)) | \psi_t] \quad (3.31)$$

在离散时间的情形下，整合的波动性可以表示为相应的每一时期的条件方差的和，记为：

$$IV(T, t) \approx \sum_{\tau=t}^{T-1} \sigma_{\tau+1|\tau}^2$$

虽然波动性的模拟和预测在金融领域中已有很广泛的应用，但是最初 Engle(1982)提出的 ARCH 模型是用来模拟通货膨胀的动态性的。可见，模拟动态的波动性除了在金融领域外，在经济学和社会科学的其他领域都有重要的应用价值，本文将简单地对波动性估计和预测的其他范围做以总结，以便于互相交流和参考。Rich 和 Tracy(2004)<sup>[68]</sup>研究了劳动市场中劳动力契约的久期和通货膨胀的波动性具有反向关系。Lastrapes(1989)<sup>[69]</sup>研究了汇率波动性和美国货币政策之间的关系。Ruge-Murcia(2004)<sup>[70]</sup>建立模型研究了失业的条件方差与通货膨胀率的负相关关系。

### 3. 3 单维波动性预测方法

在本章节中，我们给出各种预测市场收益的波动性的方法。

#### 3. 3. 1 利用历史数据进行预测

当资产的价格遵从常数波动对数正态漂移过程时，波动率可以很容易地从历史数据中得到。考虑到资产价格服从  $\frac{dS}{S} = \mu dt + \sigma dz$  的历史价格  $\{S_0, S_1, \dots, S_T\}$ ，对数收益率为  $R_t = \ln(S_t / S_{t-1}), t = 1, \dots, T$ 。

那么  $R_t$  的均值的估计可以写为：

$$\bar{R} = \frac{\sum R_t}{T}, \quad (3.32)$$

$R_t$  的方差可以表示为  $v^2 = \frac{\sum (R_t - \bar{R})^2}{T-1}$ ，按年计算的方差为：

$$\sigma = \sqrt{Nv^2} \quad (3.33)$$

其中  $N$  为在一年中观测到的价格数据的个数。

从历史数据对波动性进行预测最常用的方法是正确地选择样本区间和过去价格的样本数量，然后应用 (3.32) 和 (3.33)。但是值得说明的是需要根据不

同的预测区间调整历史样本的长度，这个问题需要在实际中进行研究。

考虑从过去  $k$  个价格样本中估计波动性，从而对未来  $T$  个时期进行波动性预测。这种方法叫做  $(k, T)$  模型。波动性估计如下式：

$$\sigma_t = \left( \frac{\sum_{s=1}^k (R_{t+1-s} - \bar{R})^2}{k-1} \right)^{1/2} \quad (3.34)$$

我们进行如下假定，详细说明利用历史数据进行预测的方法。

假设： $t_{beg}$  和  $t_{end}$  代表了给定数据的开始和结束的日期，

$t_{1st}$  和  $t_{last}$  代表了波动性预测的开始和结束的日期。

具体的做法如下：用  $t_{1st}$  时刻前  $k$  个月的收益计算历史波动性，假设收益均值为 0，这个历史波动性就作为未来区间上  $t_{1st}$  的预测波动性。未来的  $T$  时期的已实现的波动性我们可以计算出来，从而得到预测误差。从开始的时刻  $t_{1st}$  往前推，重复上述过程，弃除样本开始的一个数据点，加入后一个样本点，这种方法一直持续，对于所有的数据（从  $t_{1st}$  到  $t_{last}$ ）都计算出来预测的波动性和预测的误差，整个过程重复  $k$  次。

这种方法导致预测误差的自相关，究其原因因为每个月的预测和已实现的波动性的计算都是从前面用过的数据样本中来，并且只与预测上个月的数据差前后两个不同的数值。

### 3. 3. 2 随机波动性的预测

从更广泛的意义来看，随机波动性是指条件方差的时间序列中加入了随机因素。例如 GARCH 模型可以看作是随机波动模型。而真正的随机波动模型是把不可观测到的对收益的方差的冲击包含在动态的波动性的特征里面，从这个角度来讲，方差是内在的不可观测到的。即使我们掌握了过去所有的信息和知识，我们也不可能确切知道现在波动性的大小，因此  $t$  时刻的随机波动性的大小随着同时期或未来信息（ $t+j, j>0$ ）的变化而变化。而 ARCH 模型在过去信



息已知的条件下仅仅把条件方差作为可观测到的。

虽然两种模型有种种不同之处,但是我们可以把它们看作两个互补的模型,而不是作为相互对立的两种模型。从实际预测的角度进行比较,很难区分标准 GARCH 和 SV 模型的预测效果的好坏。有些经济学家认为 SV 具有模型上的优势,但是 ARCH 模型易于估计的特点也是非常具有竞争性的。随着计算机和模拟技术的发展,例如 MCMC 等方法也使得 SV 的估计和预测得到了很大的改善。

连续时间的资产收益模型起源于 Merton(1969)<sup>[71]</sup>的投资组合选择和 Black Scholes(1973)<sup>[63]</sup>的期权定价理论。一般地,无套利资产价格过程为:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + j(t)dq(t), \quad t \in [0, T] \quad (3.35)$$

其中,  $\mu(t)$  是连续的,局部有界的过程,波动过程  $\sigma(t)$  是严格正的,  $W(t)$  表示标准布朗运动,  $q(t)$  是一个跳跃示性函数,当无跳跃时取 0,有跳跃时取 1。  $j(t)$  指跳跃的大小。

单一时期的收益可以表示为:

$$r(t) = p(t) - p(t-1) = \int_{t-1}^t \mu(\tau)d\tau + \int_{t-1}^t \sigma(\tau)dW(\tau) + \sum_{t-1 \leq \tau < t} \kappa(\tau) \quad (3.36)$$

在此框架之内,收益变化的方差为:

$$QV(t) = \int_{t-1}^t \sigma^2(s)ds + \sum_{t-1 < s \leq t} \kappa^2(s) \quad (3.37)$$

对于单一的 SV 扩散过程,并且当  $\sigma(\tau)$  是独立的 Wiener 过程条件下,上式可以简化为:

$$IV(t) = \int_{t-1}^t \sigma^2(s)ds$$

$$r(t) | \{\mu(\tau), \sigma(\tau); t-1 \leq \tau \leq t\} \sim N\left(\int_{t-1}^t \mu(\tau)d\tau, \int_{t-1}^t \sigma^2(\tau)d\tau\right)$$

如果条件期望和方差都是随机的,那么我们不能简单地对未来的波动性进行预测,而应该根据已掌握的信息形成期望。对于短时期收益的条件方差可以忽略,对于长时期,可得到:

$$\text{Var}(r(t+h)|\psi_t) \approx E\left[\int_{t+h-1}^{t+h} \sigma^2(\tau)d\tau|\psi_t\right] \equiv E[IV(t+h)|\psi_t] \quad (3.38)$$

由于波动性是隐藏的，在 SV 的估计和预测上就出现了问题：估计时不可采用 MLE。而用许多其它的方法会造成估计结果有效性的丧失，同时预测又依赖于参数的估计，因此将会造成预测结果更不准确。

SV 的似然函数为：

$$f(\underline{r}; \theta) = \int f(\underline{r}, \underline{s}; \theta) d\underline{s} = \prod_{t=1}^T f(r_t | \psi_t; \theta) = \prod_{t=1}^T \int f(r_t | s_t; \theta) f(s_t | \psi_{t-1}; \theta) ds_t \quad (3.39)$$

其中， $[0, T]$  时期的收益和波动向量分别为  $\underline{r} = (r_1, \dots, r_T)$  和  $\underline{s} = (s_1, \dots, s_T)$

对 SV 最初的处理是 Andersen(1994)<sup>[72]</sup>运用标准的 GMM 方法估计离散时间的 SV 模型，但是作预测时必须知道状态变量的确切估计，但是 GMM 方法却不能提供这样的估计，因此，用 GMM 方法不能提供 SV 模型的预测。另外一种方法是 Kalman 滤波技术，这种技术虽然可以对参数和状态变量进行估计，但是这种技术最大的缺陷在于不适用非线性和非正态的 SV 模型。另外处理 SV 的数值方法为 Markov 链 Monte Carlo 方法。MCMC 方法代表了一种 Bayesian 方法，Jacquier (1994)<sup>[73]</sup>用 MCMC 方法来处理离散的 SV 模型。Bayesian 方法把模型的参数作为随机变量而不是固定的系数，除此之外 MCMC 方法把隐藏的状态向量也作为额外的参数。因此，它的主要的目标就是求出建立在数据集上的参数和状态变量向量  $\phi = (\theta, \underline{s})$  的联合分布  $f(\phi | \underline{r})$ ，这个密度函数看似不可处理。而 MCMC 方法利用联合分布可以分解为若干个相关联条件分布的思想，把联合分布进行拆分。

### 3. 3. 3 已实现波动性的预测

已实现波动性的估计和预测与其他方法有两个关键不同之处，第一是它不需要一个固定的模型就可以来描述波动性，因此，它提供了一个比较预测效果的标准；第二是它用每天的时间序列数据直接构造波动性，有效地利用日间高频数据的信息。

已实现波动性起源于连续时间扩散过程：

$$dp(t) = \mu(t)dt + \sigma(t)dW(t), t \in [0, T]$$

进行离散化后可得,  $r(t, \Delta) = p(t) - p(t - \Delta) = \mu(t - \Delta)\Delta + \sigma(t - \Delta)\Delta W(t)$ , 其中,  
 $\Delta W(t) = W(t) - W(t - \Delta) \sim N(0, \Delta)$ 。

$$\begin{aligned} \text{由于 } \sum_{j=1}^{1/\Delta} \text{Var}[r^2(t-1+j \cdot \Delta, \Delta) | \psi_{t-1+j \cdot \Delta}] &= 2 \sum_{j=1}^{1/\Delta} \sigma^4(t-1+j \cdot \Delta) \cdot \Delta^2 = 2\Delta \int_{t-1}^t \sigma^4(s) ds \\ \frac{\sum_{j=1}^{1/\Delta} \text{Var}[r^2(t-1+j \cdot \Delta) | \psi_{t-1+j \cdot \Delta}]}{\sum_{j=1}^{1/\Delta} E[r^2(t-1+j \cdot \Delta) | \psi_{t-1+j \cdot \Delta}]} &= 2\Delta \frac{\int_{t-1}^t \sigma^4(s) ds}{\int_{t-1}^t \sigma^2(s) ds} \end{aligned} \quad (3.40)$$

文献中一般把 (3.40) 称为信号噪声, 显然, 当  $\Delta \rightarrow 0$ , 即数据为高频数据时, 信号噪声为 0, 因此, 构造已实现波动性

$$RV(t, \Delta) = \sum_{j=1}^{1/\Delta} r^2(t-1+j \cdot \Delta, \Delta) \quad (3.41)$$

Andersen, Bollerslev, Diebold 和 Ebens (2001)<sup>[74]</sup>, Andersen, Bollerslev, Diebold 和 Labys (2003)<sup>[75]</sup> 认为一旦使用已实现的波动性, 预测方法就非常直接了, 因为在预测过程中已经没有隐藏的变量。但目前利用已实现的波动性进行预测还存在着很多缺陷, 例如, 日间收益样本频率的选取等问题, 使预测问题还处于起步阶段。

### 3. 3. 4 GARCH 类模型的预测

GARCH 模型一经提出, 就由于它突破了传统方差模型方式并更好地与实践相结合, 而显示了强大的生命力, 并成为金融计量学研究条件方差的重要手段。同时, GARCH 类模型在估计上的优势也可以为波动性的预测提供一个意义直观明确 ( $t$  时刻预期的未来的方差与长期走向的方差大小、过去时刻方差的预期值和过去的冲击的大小相关)、计算方便的预测方法。

但是用 GARCH 作预测的缺点有三个, 第一个是用 GARCH 模型作预测必须有一个大量的数据进行稳健性估计; 第二个是这些 GARCH 模型存在一个一般性的问题, 那就是任何模型越复杂, 模型所包括的参数越多, 则模型对给定样本拟合的效果越好, 得到的样本外预测的效果也越好; 第三个问题就是 GARCH 模型原则上只能应用于一步向前预测, 它们并不能在较长的区间内进

行预测。例如，考虑一个 GARCH (1, 1) 模型：

$$\begin{aligned}\sigma_t^2 &= C + a_1\sigma_{t-1}^2 + b_1\varepsilon_{t-1}^2 \\ E_t[\sigma_{t+1}^2] &= C + a_1\sigma_t^2 + b_1E_t[\varepsilon_t^2] \\ &= C + (a_1 + b_1)\sigma_t^2 \\ &\vdots \\ E_t[\sigma_{t+k}^2] &= C \sum_{s=0}^{k-1} (a_1 + b_1)^s + (a_1 + b_1)^k \sigma_t^2\end{aligned}$$

因为  $t+1$  时刻方差的预测包含了  $t$  时刻残差的平方，而这个值是未知的。显而易见，当进行多步预测时，预测没有融合任何关于未来残差的新信息，从而方差的预测从长期来看收敛于  $\sigma_{Long Run}^2 = \frac{C}{1 - a_1 - b_1}$ ，而收敛速度依赖于  $(a_1 + b_1)$ 。

### 3. 3. 4. 1 GARCH(1, 1) 模型的预测

GARCH(1,1)模型如下：

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\sigma_{t-1}^2, 0 \leq \alpha_1, \beta_1 \leq 1, (\alpha_1 + \beta_1 < 1)$$

假定预测开始为  $K$  时刻。那么对于 1 步向前预测，有  $\sigma_{K+1}^2 = \alpha_0 + \alpha_1\varepsilon_K^2 + \beta_1\sigma_K^2$

其中， $K$  时刻的  $\varepsilon_K$  和  $\sigma_K^2$  是已知的。因此，1 步向前预测是：

$$\sigma_K^2(1) = \alpha_0 + \alpha_1\varepsilon_K^2 + \beta_1\sigma_K^2 \quad (3.42)$$

对于多步预测，利用  $\varepsilon_t^2 = \sigma_t^2\eta_t^2$ ，得出：

$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2 + \alpha_1\sigma_t^2(\eta_t^2 - 1)$$

则当  $t = K+1$  时，等式变为  $\sigma_{K+2}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_{K+1}^2 + \alpha_1\sigma_{K+1}^2(\eta_{K+1}^2 - 1)$ 。

因为  $E(\eta_{K+1}^2 - 1 | \Psi_K) = 0$ ，建立在预测开始时刻  $K$  基础上的 2 步向前的波动

性预测为  $\sigma_K^2(2) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_K^2(1)$ 。依此类推，得到：

$$\sigma_K^2(l) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_K^2(l-1), l > 1$$

重复这个步骤，我们可以得到  $l$  步的预测公式为：

$$\sigma_K^2(l) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{l-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{l-1}\sigma_K^2(1) \quad (3.43)$$

因此,  $\sigma_K^2(l) \rightarrow \frac{\alpha_0}{1-\alpha_1-\beta_1}, l \rightarrow \infty$

### 3.3.4.2 EGARCH 模型的预测

我们使用最简单的 EGARCH(1,0)模型来描述 EGARCH 的多步预测方法。假设一个市场收益的波动服从 EGARCH(1,0), 我们有:

$$\begin{aligned}\ln(\sigma_t^2) &= (1-\alpha_1)\alpha_0 + \alpha_1 \ln(\sigma_{t-1}^2) + g(\varepsilon_{t-1}) \\ g(\varepsilon_{t-1}) &= \theta\varepsilon_{t-1} + \gamma(|\varepsilon_{t-1}| - \sqrt{2/\pi})\end{aligned}$$

两边取指数, 得到:

$$\begin{aligned}\sigma_t^2 &= \sigma_{t-1}^{2\alpha_1} \exp[(1-\alpha_1)\alpha_0] \exp[g(\varepsilon_{t-1})] \\ g(\varepsilon_{t-1}) &= \theta\varepsilon_{t-1} + \gamma(|\varepsilon_{t-1}| - \sqrt{2/\pi})\end{aligned}$$

令  $K$  是预测的初始值, 那么 1 步预测为:

$$\sigma_{K+1}^2 = \sigma_K^{2\alpha_1} \exp[(1-\alpha_1)\alpha_0] \exp[g(\varepsilon_K)], \quad (3.44)$$

其中, 右边的系数都是已知的。对于 2 步预测:

$$\sigma_{K+2}^2 = \sigma_{K+1}^{2\alpha_1} \exp[(1-\alpha_1)\alpha_0] \exp[g(\varepsilon_{K+1})] \quad (3.45)$$

取条件期望, 得到  $\hat{\sigma}_K^2(2) = \hat{\sigma}_K^{2\alpha_1}(1) \exp[(1-\alpha_1)\alpha_0] E_K \{ \exp[g(\varepsilon_{K+1})] \}$ 。

其中,  $E_K$  表示了建立在开始时刻  $K$  的条件期望。

$$\begin{aligned}E\{\exp[g(\varepsilon)]\} &= \int_{-\infty}^{\infty} \exp[\theta\varepsilon + \gamma(|\varepsilon| - \sqrt{2/\pi})] f(\varepsilon) d\varepsilon \\ &= \exp(-\gamma\sqrt{2/\pi}) \left[ \int_0^{\infty} e^{(\theta+\gamma)\varepsilon} \frac{1}{\sqrt{2\pi}} e^{-\varepsilon^2/2} d\varepsilon \right] \\ &\quad + \int_{-\infty}^0 e^{(\theta-\gamma)\varepsilon} \frac{1}{\sqrt{2\pi}} e^{-\varepsilon^2/2} d\varepsilon \\ &= \exp(-\gamma\sqrt{2/\pi}) [e^{(\theta+\gamma)^2} \Phi(\theta+\gamma) + e^{(\theta-\gamma)^2/2} \Phi(\gamma-\theta)]\end{aligned}$$

其中,  $f(\varepsilon)$  和  $\Phi(x)$  是概率密度函数和标准正态分布函数。因此, 可得 2 步波动性预测为:

$$\begin{aligned}\hat{\sigma}_K^2(2) &= \hat{\sigma}_K^{2\alpha_1}(1) \exp[(1-\alpha_1)\alpha_0 - \gamma\sqrt{2/\pi}] \\ &\quad \times [\exp\{(\theta+\gamma)^2/2\} \Phi(\theta+\gamma) + \exp\{(\theta-\gamma)^2/2\} \Phi(\gamma-\theta)]\end{aligned}$$

重复这样的步骤, 我们可以得到  $l$  步的预测公式为:

$$\hat{\sigma}_k^2(l) = \hat{\sigma}_k^{2\alpha_1} (l-1) \exp(\omega) \times [\exp\{(\theta + \gamma)^2 / 2\} \Phi(\theta + \gamma) + \exp\{(\theta - \gamma)^2 / 2\} \Phi(\gamma - \theta)] \quad (3.46)$$

其中,  $\omega = (1 - \alpha_1)\alpha_0 - \gamma\sqrt{2/\pi}$ 。

### 3. 3. 4. 3 非对称性 GARCH

在金融领域中, Glosten, Jagannathan 和 Runkle (1993)<sup>[38]</sup>与 Zakoian (1994)<sup>[39]</sup>发现正的收益带来的冲击和负的收益带来的冲击对未来的条件方差有不同的影响。这种非对称性叫做杠杆效应, 也可以解释为: 增加的波动性需要一个未来预期的收益的增加来补偿风险, 因此造成现在价格的下跌。因此 Engle 和 Ng (1993)<sup>[76]</sup>提出的非对称的 GARCH 模型, GJR 或门类 GARCH 模型等都是描述金融领域中的非对称性的。

GJR(1,1)和 TGARCH(1,1)模型的条件方差等式为:

$$\sigma_{t|t-1}^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I(\varepsilon_{t-1} < 0) + \beta \sigma_{t-1|t-2}^2 \quad (3.47)$$

$I(\cdot)$  是示性函数,  $\gamma > 0$  说明过去负的收益带来的冲击对未来的条件方差有一个较大的影响, 在预测时, 有一个很重要的假设:  $P(z_t = \sigma_{t|t-1}^{-1} \varepsilon_t < 0) = 0.5$ , 因此, 式 (3.47) 变为:

$$\sigma_{t+h|t}^2 = \sigma^2 + (\alpha + 0.5\gamma + \beta)^{h-1} (\sigma_{t+1|t}^2 - \sigma^2) \quad (3.48)$$

从长期来看, 无条件方差为:

$$\sigma^2 = \omega(1 - \alpha - 0.5\gamma - \beta)^{-1}$$

### 3. 3. 4. 4 长记忆模型

GARCH, EGARCH, TARCH 和 AGARCH 模型的共同特点是对波动性的冲击以指数率衰减。以 GARCH 模型为例, 用  $t$  时刻的冲击对于  $h$  步向前方差预测的效应可以表示为:

$$\frac{\partial \sigma_{t+h}^2}{\partial \varepsilon_t^2} = \alpha(\alpha + \beta)^{h-1} \quad (3.49)$$

$$\text{或者 } \frac{\partial \sigma_{t+h}^2}{\partial \varepsilon_t^2} = \kappa \delta^h \quad \text{其中, } 0 < \delta < 1$$

指数衰减率对于短时期的预测非常有效, 但是大量的研究, 包括 Ding, Granger 和 Engle(1993)<sup>[77]</sup>与 Andersen, Bollerslev (1997)<sup>[78]</sup>的研究表明对于较长的滞后项, 平方收益和绝对值收益的自相关系数以较慢的双曲线速率衰减, 在波动性预测方面, 长期的波动性预测应该较好, 这时它的响应效果为:

$$\frac{\partial \sigma_{t+ht}^2}{\partial \varepsilon_t^2} \approx \kappa h^\delta \quad (3.50)$$

其中,  $h$  为较大的值,  $0 < \delta < 1$ 。

具有代表性的模型为 Baillie, Bollerslev 和 Mikkelsen (1996)<sup>[79]</sup>提出的分整 FIGARCH (1,d,1) 模型, 这里,  $0 < d < 1$ , 条件方差定义为:

$$\sigma_{t|t-1}^2 = \omega + \beta \sigma_{t-1|t-2}^2 + [1 - \beta L - (1 - \alpha L - \beta L)(1 - L)^d] \varepsilon_t^2$$

当  $d = 0$  时, 可以转化为标准的 GARCH 模型; 当  $0 < d < 1$  时, 从此模型得出的对波动性预测的冲击将以较慢的双曲线速率衰减。实际的预测可以很容易地用递归的方法得出:

$$\sigma_{t+h|t+h-1}^2 = \omega(1 - \beta)^{-1} + \lambda(L) \sigma_{t+h-1|t+h-2}^2, \quad (3.51)$$

其中,  $\sigma_{t+h|t+h-1}^2 = \varepsilon_t^2, h < 0$ ,  $\lambda(L) = 1 - (1 - \beta L)^{-1}(1 - \alpha L - \beta L)(1 - L)^d$  中的系数可以通过下列递归方法得到:

$$\lambda_1 = \alpha + d \quad \lambda_j = \beta \lambda_{j-1} + [(j-1-d)j^{-1} - (\alpha + \beta)] \delta_{j-1} \quad j = 2, 3, \dots \quad (3.52)$$

其中,  $\delta_j = \delta_{j-1}(j-1-d)j^{-1}$ 。

高阶的 FIGARCH 模型的预测可以用类似的算法得到。同理, 非对称性也可以引入, 得到相应的分整 EGARCH (FIEGARCH)。

### 3. 4 多维相关性预测

多维相关性预测的思路大多数可以采取单维波动性预测的方法, 但是有两个问题需要注意: 第一是在整个预测区间, 方差协方差矩阵的正定性必须满足; 第二是根据对数据的分析和诊断检验的结果, 选择一个参数数量相对较少的但

合适的多维相关性模型进行预测。

在对多维相关性预测时，分为两种情况，第一种情况是直接对方差协方差矩阵进行预测，第二种情况是先对单个市场的方差进行预测，再对市场收益的相关性进行预测，然后再组合成对方差协方差矩阵的预测。

考虑由  $N$  个资产组成的投资组合：

$$r_{w,t-1} = \sum_{i=1}^N w_{i,t} r_{i,t-1}$$

投资组合的一步向前预测方差满足：

$$\sigma_{w,t+1|t}^2 = \sum_{i=1}^N \sum_{j=1}^N w_{i,t} w_{j,t} \{\Omega_{t+1|t}\}_{ij} = w_t' \Omega_{t+1|t} w_t$$

其中  $\Omega_{w,t+1|t}^2$  代表  $N \times N$  维的收益的方差协方差矩阵。因此一个投资组合的收益的方差的预测需要对投资组合中各个资产之间的方差协方差矩阵中的  $\frac{1}{2}N(N+1)$  个元素进行预测。

### 3. 4. 1 指数平滑和 Riskmetrics 方法

指数加权移动平均技术，又被称作 Riskmetrics 方法，在实际部门得到了广泛应用。令  $Y_t = R_t$  为  $N \times 1$  维的资产收益，那么方差协方差矩阵的估计为：

$$\hat{\Omega}_t = \gamma Y_t Y_t' + (1 - \gamma) \hat{\Omega}_{t-1} = \gamma \sum_{i=1}^{\infty} (1 - \gamma)^{i-1} Y_t Y_t' \quad (3.53)$$

其中， $Y_t$  中的每个元素的期望等于 0。调整参数  $\gamma$  在应用于每日数据时取 0.06，应用于每月数据时取 0.04。同时，所估计出来的方差协方差矩阵必须保证是对称正定的。有一点需要说明的是，我们可以允许  $\hat{\Omega}_t$  中的元素有不同的衰减率，特别是可以用较小的  $\gamma$  来描述对角线以外的元素，因此时变的相关系数为：

$$\hat{\rho}_{ij,t} = \frac{\{\hat{\Omega}_t\}_{ij}}{\{\hat{\Omega}_t\}_{ii}^{1/2} \{\hat{\Omega}_t\}_{jj}^{1/2}} \quad (3.54)$$

同理， $h$  步向前预测等于建立在现在对未来的条件方差协方差矩阵：

$$\text{Var}(Y_{t+h} | \psi_t) = \Omega_{t+h|t} \approx \hat{\Omega}_t \quad (3.55)$$



### 3. 4. 2 BEKK 模型的预测

由于 BEKKE 模型是单维 GARCH 模型直观并且一般化的扩展，因此它的预测方法比较简单。Pojarliev(2001,2003)<sup>[80,81]</sup>提出了 BEKK 模型的 1 步预测方法：

$$\hat{H}_t^{(1)} = \hat{C}^* \hat{C}^* + \sum_{k=1}^K \hat{A}_k^* \varepsilon_t \varepsilon_t' \hat{A}_k^* + \sum_{k=1}^K \hat{G}_k^* H_{t1} \hat{G}_k^* \quad (3.56)$$

### 3. 4. 3 DCC 模型的预测

DCC 模型是一个非线性过程：

$$Q_{t+r} = (1 - \alpha - \beta) \bar{Q} + \alpha (\varepsilon_{t+r-1} \varepsilon_{t+r-1}') + \beta Q_{t+r-1}$$

其中  $E_t[\varepsilon_{t+r-1} \varepsilon_{t+r-1}'] = E_t[R_{t+r-1}]$ ,  $R_{t+r} = Q_{t+r}^*{}^{-1} Q_{t+r} Q_{t+r}^*{}^{-1}$ ,  $\bar{Q}$  是标准残差的无条件方差协方差矩阵，因此， $r$  步的相关性预测不能通过一个简便的方法求出，Engle (2001) <sup>[52]</sup> 发展了两种技术对此进行简化。第一种技术是利用逼近  $E_t[\varepsilon_{t+1} \varepsilon_{t+1}'] \approx Q_{t+1}$ ，得到  $Q$  的  $r$  步预测是：

$$E_t[Q_{t+r}] = \sum_{i=0}^{r-2} (1 - \alpha - \beta) \bar{Q} (\alpha + \beta)^i + (\alpha + \beta)^{r-1} Q_{t+1} \quad (3.57)$$

$$R_{t+r} = Q_{t+r}^*{}^{-1} Q_{t+r} Q_{t+r}^*{}^{-1} \quad (3.58)$$

另外一种方法是，利用  $\bar{Q} \approx \bar{R}$  和  $E_t[Q_{t+1}] \approx E_t[R_{t+1}]$ ，我们可以直接得到：

$$E_t[R_{t+r}] = \sum_{i=0}^{r-2} (1 - \alpha - \beta) \bar{R} (\alpha + \beta)^i + (\alpha + \beta)^{r-1} R_{t+1} \quad (3.59)$$

### 3. 4. 4 多变量随机波动模型和因子模型

一个更加简约且可行的多变量预测模型是因子结构的使用。因子结构在金融领域有着重要的作用，譬如著名的套利定价理论 (APT)。多变量随机波动模型和多变量因子 GARCH 模型最早由 Diebold 和 Nerlove (1989) <sup>[82]</sup> 提出。考虑一个简单的单因子模型， $N \times 1$  资产收益向量  $R_t$  由因子  $f_t$  决定：

$$R_t = a + bf_t + e_t \quad (3.60)$$

其中,  $a$  和  $b$  是  $N \times 1$  维参数向量, 假定  $e_t$  为 *i.i.d.* 序列, 方差协方差矩阵为  $\Lambda$ ; 假定因子  $f_t$  是条件异方差, 条件方差为  $\sigma_{f_t}^2 = \text{Var}(f_t | \psi_t)$ ,  $R_t$  的条件方差协方差矩阵为:

$$\Omega_{t|t-1} = \text{Var}(R_t | \psi_{t-1}) = bb' \sigma_{f_t}^2 + \Lambda \quad (3.61)$$

对照前面的多维 GARCH 模型, 因子 GARCH 模型大大减少了参数的数量。为了进一步认识到因子的意义, 令  $b_i$  和  $\lambda_{ij}$  分别指  $b$  和  $\Lambda$  的第  $i$  个和第  $ij$  个元素, 从式 (3.61) 可得第  $i$  个和第  $j$  个观测值之间的相关系数为:

$$\rho_{ij,t} = \frac{b_i b_j \sigma_{f_t}^2 + \lambda_{ij}}{(b_i^2 \sigma_{f_t}^2 + \lambda_{ii})^{1/2} (b_j^2 \sigma_{f_t}^2 + \lambda_{jj})^{1/2}} \quad (3.62)$$

如果相应的因子荷载有相同的符号, 即  $b_i b_j > 0$ , 那么条件相关性将会随着因子波动性的增加而逐渐增强, 接近于 1。

### 3. 5 预测效果检验方法

预测方差协方差矩阵, 除了上述提到的方法以外, 目前还有许多其他的方法。因此对预测结果的评价和比较就显得相当重要和不可缺少。Diebold(1998)<sup>[83]</sup> 认为设定一个普遍接受的评价准则是不可能的。特别是如何对非线性波动性预测模型进行评价, 使经济学家创造了很多特殊的经济计量准则来解决这类问题。

本文采用两种方法对模型的预测进行评价。第一种方法是从统计学的角度, 利用各种多维 GARCH 模型和其他方法的预测与已实现值通过一定的方法进行比较。

根据 Ledoit, Santa-Clara 和 Wolf(2003)<sup>[32]</sup>, 一个评价模型优劣的方法是利用已实现的值构造的方差协方差矩阵与预测的方差协方差矩阵进行对比。

首先构造一个残差在  $K$  时期已实现的方差协方差矩阵  $\hat{\Omega}$ :

$$\hat{\Omega}^{(K)} = \frac{1}{K} \sum_{k=1}^K y_{t-1+k} \cdot y_{t-1+k}' \quad \forall t = 1, \dots, T - K + 1$$

其中， $y_t$  是残差。

在同一  $K$  时期模型关于方差协方差矩阵的预测  $\hat{H}$  是：

$$\hat{H}_t^{(K)} = \frac{1}{K} \sum_{k=1}^K (\hat{H}_{t-1+k} | I_{t-1}) \quad \forall t = 1, \dots, T - K + 1$$

预测和已实现方差—协方差的距离可以通过平均绝对值偏差 (MAD) 和平均平方误差 (MSE) 来衡量：

$$MAD = \frac{1}{T} \left[ \sum_{r=1}^M \sum_{c=r}^M |\hat{\Omega}_{r,c,t}^{(K)} - \hat{H}_{r,c,t}^{(K)}| \right] \quad (3.63)$$

$$MSE = \frac{1}{T} \left[ \sum_{r=1}^M \sum_{c=r}^M (\hat{\Omega}_{r,c,t}^{(K)} - \hat{H}_{r,c,t}^{(K)})^2 \right] \quad (3.64)$$

另外一种方法是从金融应用的角度，利用相关矩阵的预测值构造投资组合来比较投资组合的收益和标准差的大小。

### 3. 6 实证分析

以往的国内外大多文献集中在研究单个金融市场收益的波动性，那么各个金融市场之间有无相关性呢？这些相关性随着时间的推移和各种政策制度的实施又是怎样变化呢？以中国股市的发展为例，目前深沪两市共有 B 股上市公司 100 多家，其中大部分上市公司同时又有 A 股上市，A、B 股市场并存与分割可以说是我国股市有别于境外成熟规范市场的一个主要特点。由于 A、B 股所代表的股权及其背后的经济内涵的同一性，A、B 股市场分割所产生的制度缺陷是不可避免的：其一，A、B 股市场间的价格扭曲现象十分严重；其二，非国内居民投资 A 股、而国内居民投资 B 股的非法套利交易难以避免，这也在很大程度上降低了市场的透明度；其三，A、B 股市场分割有损于提高市场的流动性和效率；其四，存在 B 股投资者，特别是境外投资者侵占 A 股投资者利益的市场空间，这与证券市场三公原则相矛盾。因此，为了考察中国十年间四个主要股指市场收益的相关性的动态变化，本文采用多维 GARCH 模型利用四个主要股指市场的数据进行分析。

### 3. 6. 1 数据处理

总样本为 1992 年 12 月 28 日至 2006 年 7 月 12 日上证 A 股、上证 B 股、深综 A 指和深综 B 指的每日收盘价，共 3049 个数据。本文采用样本外预测，从 1992 年 12 月 28 日至 2006 年 3 月 31 日的数据为估计样本，从 2006 年 3 月 31 日至 7 月 12 日为预测样本。对于估计样本，本文对其取对数收益：

$$x_t = 100 \times (\ln P_t - \ln(P_{t-1})) \quad (3.65)$$

其数据的描述性统计量如表 3.1 所示：

表 3.1 对数收益数据统计量

	均值	标准差	偏度	峰度	Jarque-Bera
上证A股	0.003592	2.380739	1.488071	27.29316	68811.84
上证B股	0.015034	2.332258	1.225764	24.11885	51925.33
深综A指	-0.009563	2.234408	0.346519	7.862732	2771.528
深综B指	0.016808	2.311948	0.365589	10.00474	5697.915

四个市场的收益率全部通过了单位根检验，符合 GARCH 模型对变量序列的平稳性要求，同时关于异方差的 ARCH-LM 检验，四个市场均在 1% 置信水平拒绝同方差的原假设，因此，采用 ADCC 多维 GARCH 模型建模方法是较为合理的。

### 3. 6. 2 ADCC 模型估计

总结各类研究多维 GARCH 模型的文献，我们可以比较出各种模型的优缺点，Bollerslev (1988)<sup>[2]</sup>提出的 VECM 模型最大的缺点是不能保证  $H_t$  的正定性；由 Bollerslev(1990)<sup>[29]</sup>提出的常数相关的多维 GARCH 模型又违反了相关性的时变性；而 Engle 和 Kroner(1995)<sup>[21]</sup>提出的多维 GARCH(1,1)-BEKK 模型虽然保持了  $H_t$  的正定性，但是模型中参数的意义不能用经济意义解释。本文采用了 Cappiello, Engle 和 Shepard (2003)<sup>[84]</sup>以及 Lin(2004)<sup>[85]</sup>年提出的非对称动态条件相关 (ADCC) 的多维 GARCH 模型，模型如下：

$$\varepsilon_t \sim N(0, H_t)$$

$$H_t = D_t R_t D_t$$

其中  $D_t$  是对角矩阵，对角线上的元素是单维 GARCH (p,q) 的方差， $R_t$  是非对称动态条件相关矩阵，具体的 ADCC 模型如下：

$$Q_t = (\bar{Q} - A' \bar{Q} A - B' \bar{Q} B - C' \bar{N} C) + A' u_{t-1} u'_{t-1} A + B' Q_{t-1} B + C' n_{t-1} n'_{t-1} C$$

$$\bar{Q} = T^{-1} \sum_{t=1}^T u_t u'_t, \quad \bar{N} = T^{-1} \sum_{t=1}^T n_t n'_t, \quad \text{其中 } n_{i,t} = 1_{[u_{i,t} < 0]} u_{i,t}$$

$$A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & b_n \end{bmatrix} \quad C = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}$$

为使其简化， $A = a_1$ ， $B = b_1$ ， $C = c_1$

$$u_{i,t} = \frac{\varepsilon_{i,t}}{\sqrt{h_{i,t}}}$$

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}$$

$$Q_t^{*-1} = \begin{bmatrix} \sqrt{q_{1,1,t}} & 0 & \cdots & 0 \\ 0 & \sqrt{q_{2,2,t}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sqrt{q_{n,n,t}} \end{bmatrix}^{-1}, \quad n=4。$$

其中  $H_t$  是方差协方差矩阵， $D_t = \text{diag}[\sqrt{h_{i,t}}]_{(4,4)}$ ， $h_{i,t}$  是第  $i$  个市场的方差， $Q_t$  是正定矩阵， $R_t$  是相关系数矩阵， $\varepsilon_t$  是残差向量， $u_{i,t}$  是标准残差， $n_t$  是  $n_{i,t}$  组成的向量， $A$ 、 $B$ 、 $C$  为参数值。 $n_{t-1} n'_{t-1}$  表示了动态条件相关系数的非对称性。由于中国股指市场在数据样本的取值区间经历了国际形势和国内政策变化的影响，因此本文在估计单个市场的波动率时设置了两个虚拟变量， $x_1$  代表了 1997 年亚洲金融危机对于中国主要股指市场条件方差的影响的虚拟变量， $x_2$  代表了 2001 年 2 月 16 日公布 B 股市场向国内购买者开放对于中国主要股指市场条件

方差的影响的虚拟变量，当然我们也选取了代表其他重要政策变动的虚拟变量，但是结果并不显著，没有在此列出。单个市场的波动率估计等式为：

$$h_{i,t}^2 = \alpha_0 + \alpha_1 \varepsilon_{i,t-1}^2 + \phi \varepsilon_{i,t-1}^2 d_{t-1} + \theta h_{i,t-1}^2 + c_1 x_1 + c_2 x_2$$

其中，反映杠杆效应的变量  $d_t = \begin{cases} 1 & \varepsilon_t < 0 \\ 0 & \text{其他} \end{cases}$ ， $x_1 = \begin{cases} 1 & \text{1997年7月1日之后} \\ 0 & \text{其他} \end{cases}$ ，和

$x_2 = \begin{cases} 1 & \text{2001年2月16日之后} \\ 0 & \text{其他} \end{cases}$ 。

$h_{i,t}$  是第  $i$  个市场的条件方差， $\varepsilon_{i,t}$  是第  $i$  个市场的残差。 $\alpha_0$ 、 $\alpha_1$ 、 $\phi$ 、 $\theta$ 、 $c_1$ 、 $c_2$  为参数值。

四个市场波动率和相关矩阵的估计结果如表 3.2、表 3.3 所示：

表 3.2 四个市场波动率的参数估计

	$\alpha_0$	$\alpha_1$	$\phi$	$\theta$	$c_1$	$c_2$
sha	0.97878 (2.3989)	0.083878 (2.9681)	0.12802 (3.3977)	0.78960 (16.468)	-0.84974 (-2.2562)	
sza	0.91096 (2.5093)	0.10645 (3.1543)	0.12015 (3.0241)	0.77248 (15.702)	-0.76621 (-2.2883)	
shb	0.17370 (3.1435)	0.20824 (6.0618)		0.74033 (19.389)	0.40225 (3.2884)	-0.30760 (-2.7207)
szb	0.18459 (2.6453)	0.23937 (4.3490)		0.69988 (10.772)	0.43410 (2.5139)	-0.23617 (-1.6684)

注：无特别说明，表中数字代表参数估计值，下面括弧中的数字代表 t 统计量。

表 3.3 相关矩阵的参数估计

	$a_1$	$b_1$	$c_1$
ADCC	0.96594 (790.10)	0.030409 (24.970)	-0.0034888 (-2.0912)

四个市场之间的动态相关系数图如图 3.1 所示：

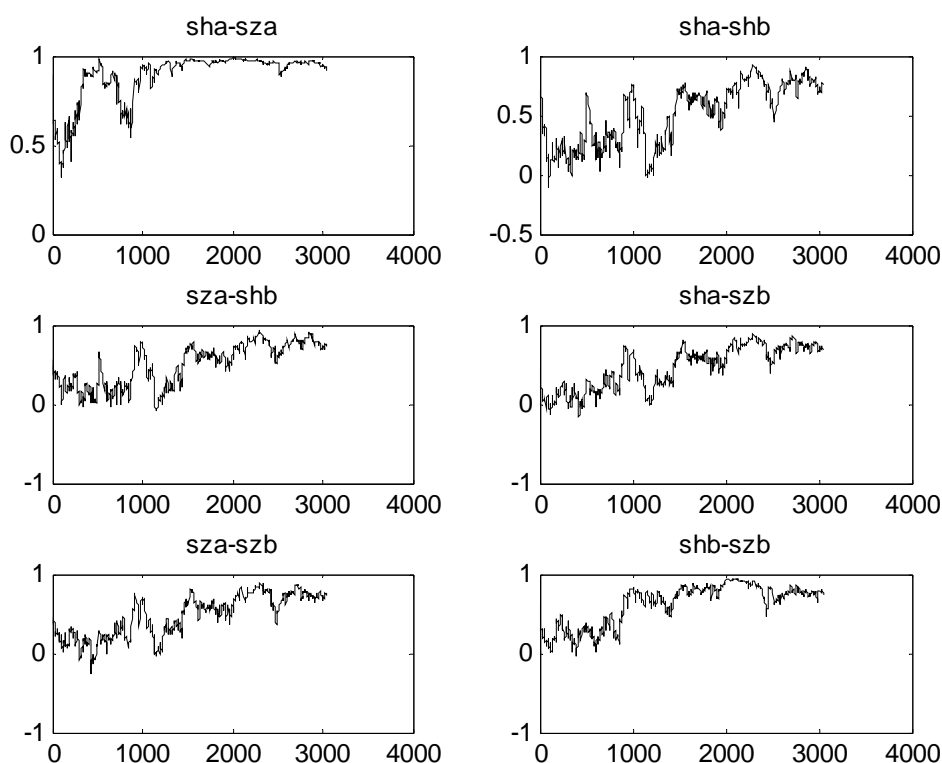


图 3.1 动态相关系数图

从图 3.1 可以得出以下结论，首先，从 1997 年亚洲金融危机以来中国主要股指市场之间的条件相关性总体上呈现逐渐增强的趋势，其次，从图中可以明显的看到，自从 2001 年 2 月 16 日公布 B 股市场向国内购买者开放，以及 2003 年 5 月 26 日 A 股市场向国外购买者开放以后，中国主要股指市场的相关性也明显增强。从上海 A 股和上海 B 股的相关关系和深圳 A 股和深圳 B 股的相关关系来看，A、B 股市场长期分割的现象逐渐得到改善。表明中国主要股指市场的分离程度已经得到缓和。第三，可以看到当国家放宽对 A 股和 B 股市场购买者的政策出台时，四个市场间的动态相关系数就有突然的升高，然后又逐渐呈现下降平稳的趋势。说明中国主要股指市场在得到利好消息后，各个主要股指市场反应相当，相关性增强。

### 3. 6. 3 CCC 多维 GARCH 模型

与 ADCC 模型不同，同时也是为了进行对比，我们用 CCC 多维 GARCH 模型进行研究，CCC 多维 GARCH 模型可以表示为：

$$\varepsilon_t \sim N(0, H_t)$$

$$H_t = D_t R D_t$$

其中  $D_t$  是对角矩阵，对角线上的元素是单维 GARCH (p,q) 的方差， $R$  是常数相关矩阵。

CCC 多维 GARCH 的相关矩阵估计结果如表 3.4 所示：

表 3.4 CCC 模型相关矩阵估计系数

	上证 A 股	上证 B 股	深综 A 指	深综 B 指
上证 A 股	1			
上证 B 股	0.5082 (42.790)	1		
深综 A 指	0.8953 (334.45)	0.4996 (41.622)	1	
深综 B 指	0.4582 (35.926)	0.6277 (67.728)	0.4730 (37.693)	1

### 3. 6. 4 预测结果

本文采用两种方法对模型的预测进行评价。第一种方法利用 ADCC 模型和 CCC 模型的预测的动态条件相关矩阵与已实现值通过一定的方法进行比较，第二种方法利用预测得到的相关矩阵构造最小方差投资组合，比较所得到的投资组合的标准差的大小。

根据文中式 (3.63)，(3.64)，对预测结果进行评价和比较的结果如表 3.5 所示：



表 3.5 评价和比较

	ADCC 多维 GARCH	CCC 多维 GARCH
MAD (25 天)	0.00074648930	0.00097194314
MSE (25 天)	0.0014952724	0.0017069670
MAD (50 天)	0.0018064028	0.0022133214
MSE (50 天)	0.0021882232	0.0024020902

从预测结果分析可以看出，不管是用 MAD 方法还是 MSE 方法，用 ADCC 模型作预测的结果都优于 CCC 多维 GARCH 模型所得到的结果。因此，用 ADCC 模型拟合中国股指收益率的方差协方差矩阵比常数相关模型拟合的效果要好。众所周知，方差协方差矩阵在许多领域，例如资产配置决策和风险管理都有很重要的应用。因此这个预测结果可以为以后许多工作进行铺垫，政府部门可以用该模型提高股市监管能力，投资者可用其进行规避市场风险。

另外一种评价模型好坏的方法是利用相关矩阵的预测值构造投资组合来比较投资组合的收益和标准差的大小。为了避免设定条件预期收益的值，本文采取最小方差投资组合得到的标准差来比较各种模型预测结果。

CAMPBELL(1997)<sup>[9]</sup>提出没有无风险资产并且允许卖空的最小方差投资组合的最优化问题如下：

$$\min w' H_t w, \text{ 满足条件 } w' \mathbf{1} = 1$$

$$\text{其解为 } w_{GMV} = \frac{1}{C} H^{-1} \mathbf{1}$$

其中  $\mathbf{1}$  是单位 1 向量， $C = \mathbf{1}' H^{-1} \mathbf{1}$ 。

这个最优化问题的解  $w_{GMV}' H_t w_{GMV}$  称为最小方差投资组合。根据 ADCC 多维 GARCH 模型和 CCC 多维 GARCH 模型预测所得的方差协方差矩阵  $H_t$ ，进而构造最小方差投资组合得到的标准差和平均收益如表 3.6 所列：

表 3.6 平均收益和标准差

	ADCC 多维 GARCH	CCC 多维 GARCH
标准差 (50 天平均)	1.3225	1.3275
平均收益 (50 天平均)	0.1607	0.1557

从表 3.6 中可以看出,ADCC 多维 GARCH 模型所得到的标准差略小于 CCC 多维 GARCH 模型的标准差,但是在预测区间的平均收益方面却是明显好于 CCC 多维 GARCH 模型。从而再次验证了用 ADCC 多维 GARCH 模型来预测中国主要股指之间的动态方差协方差矩阵效果比较好。

## 4. 基于多维 GARCH 模型动态投资组合的风险分析

近年来在财务金融界最重要的革命性议题是在险价值方法的诞生，它大幅修正过去仅以获利数字评价绩效的观念，明确提供量化的数字以作为适当资本准备金额及获利调整的标准，通过在险价值使投资人及管理者更能掌握头寸的风险，也因此刺激金融商品市场的更加活跃。

过去对于风险的概念，常仅限于概率概念，通常仅有参考的作用，直到近年来在险价值概念的蓬勃发展，风险的概念才逐渐落实到管理的层面，主要在于在险价值明确量化了投资组合的风险。在其他的文献中，大部分都计算一个市场收益的风险，而本文主要研究投资组合收益的 VaR 估计及检验。

在险价值是指商品价格改变的特定概率分布下，投资组合在特定持有期间，和既定置信水平下的最大可能损失。在险价值的由来为美国摩根银行（以下简称 J.P.Morgan）前任主席丹尼斯·威瑟斯要求业务部门于每日营业结束后交一页报告，摘要公司暴露于市场行情变动的风险，以提供未来 24 小时公司潜在损失的估计值。

过去在金融理论与实务上，对于风险的表达大多为描述性质，即使部分有量化指标或数字，但是距离实务上风险评估的“价值”仍有一段距离。一般衡量风险的相关指标有三种类型，但对于信息使用者或风险管理者而言，仍显得不够合理。传统风险衡量的类型略述如下：第一种是“敏感度”，例如系统风险( $\beta$ )、存续期间(duration)等，前者是个别证券对市场价值的敏感度，后者是债券价值对市场利率的敏感度，两者共同缺点是，无法知道整体的绝对影响；第二种是“波动”，最常用以衡量波动的指标是标准差；第三种是“概率”，风险与结果固然有密切关系，概率却扮演重要角色，比如搭乘飞机远比搭乘汽车安全，背后考虑的问题是针对于概率着眼，然而未必合乎许多人衡量风险时所使用的标准。因此使用者对风险衡量的水准日益增加。

### 4.1 在险价值基本原理

如前所述，在险价值是指在商品价格变动的特定概率分布下，投资组合在

此特定持有期间，在既定置信水平下的最大可能损失。首先，假设“预期每日损益”服从正态分布，以平均数为中心，若置信水平为 95%，即预期每日损益变动额出现在平均数上下两个标准差的概率为 95%。

理论上衡量在险价值的方式非常多，但一般实务上发展较为成熟的评估软件则属摩根银行发展的风险矩阵。J.P.Morgan 基本上以指数加权平均法来计算在险价值，主要的思想在于计算观察值的权数会随着时间不同而异，也就是说距观察时点愈近其权数愈大，距观察时点愈远其权数则愈小。而计算在险价值 VaR 步骤大致可归为下列五个步骤：

- (1) 确认风险资产类型；净现值 (NPV) 可能产生变动的资产，例如，远期外汇代表未来可依特定汇率交易的合约，当国内出口商 30 天后有 100 万美元的销售收入，则可以预售 30 天期远期外汇来进行汇率避险；
- (2) 确认影响该风险性资产价值的因子；影响该资产价值变动的风险因子，可能包含一个或一个以上的因子。譬如，资产本身因子：风险性资产本身价格变化；传统风险因子：权益、利率、汇率、商品；计量风险因子：套利定价模型、会计  $\beta$ ；
- (3) 选定风险预测模型及风险因子在模型中的风险暴露额；确认模型计算风险因子类别后，该风险性资产受到风险因子影响的金额；
- (4) 执行风险的模型运算；
- (5) 产生总风险估计值。

## 4. 2 风险的评价方法

对风险的评价方法分为部分评价法和全额评价法 (full valuation)，一般地，部分评价法事先假定单个资产或资产组合的收益服从给定的分布，从而建立一定的模型，依据一定的规则来评价风险；而全额评价法往往采用模拟的方法，构造资产收益服从的分布，避免了资产收益的概率分布假设错误所带来的风险。

#### 4. 2. 1 部分评价法

##### (1) 历史移动平均法

对单个资产而言，本方法将单个资产的收益率的分布假定为正态分布，先计算出过去一段时间风险因子价格变动的标准差，再求出风险值。历史移动平均法的 VaR 如下：

$$VaR = Z_{\alpha} \sigma_p \sqrt{T} \quad (4.1)$$

其中， $Z_{\alpha}$  表示置信度为  $\alpha$  时，标准正态分布中单侧的临界值， $\sigma_p$  为单个资产收益的波动率，而  $T$  表示持有该资产的时间。

根据正态分布的特性，若单个资产收益服从正态分布，那么资产组合收益也服从正态分布。假设投资组合中各个资产的收益向量价值变动为  $\Delta S$ ，其中  $\mu$  为收益向量的期望向量， $H$  表示收益向量变动的方差协方差矩阵， $\sigma_{ij}$  表示第  $i$  个资产收益与第  $j$  个资产收益之间的协方差。 $\Delta S$ 、 $\mu$  和  $H$  可以表示为：

$$\Delta S \sim N(\mu, H)$$

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}_{N \times 1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H = \begin{pmatrix} \sigma_1^2 & \sigma_{21} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{pmatrix}_{N \times N} \quad (4.2)$$

由于投资组合的价值变化是单个资产价值变化的线性组合，因此投资组合价值的变动  $\Delta P \sim N(\mu_p, \sigma_p^2)$ ，其中， $\mu_p = w' \mu$ ， $\sigma_p = \sqrt{w' H w}$ ， $w$  为分配到各资产的权重。有了投资组合价格变动率的方差之后，就可以计算风险值，VaR 计算公式为：

$$VaR = Z_{\alpha} \sigma_p \sqrt{T} \quad (4.3)$$

其中， $Z_{\alpha}$  表示置信度为  $\alpha$  时，标准正态分布中单侧的临界值， $T$  表示持有该资产的时间。

##### (2) 指数移动平均法

指数移动平均法又称为 Riskmetrics 方法，其计算公式仍是以正态分布为前

提假设，本方法基本上是以指数加权移动平均法来计算风险值，主要思路体现在权数会随着时间不同而异，距观察时点越近的其权数越大，距观察时点越远则权数越小。指数移动平均法的变动率计算公式如下：

$$\sigma_{t+1} = \sqrt{(1-\lambda) \sum_{s=t-k}^t \lambda^{t-s} (X_s - \mu)^2} \quad (4.4)$$

其中， $\sigma_t$  = 从第  $t$  天起开始的投资组合的标准差；

$k$  = 移动平均所包含的天数；

$X_s$  = 投资组合价值在第  $s$  天的收益率；

$\mu$  = 投资组合价值平均变动；

$\lambda$  = 衰退因子，估计值越久，权数越小。

对于多维的情况，指数加权移动平均的方差协方差计算公式为：

$$\sigma_{xy,t}^2 = \sum_{i=0}^{\infty} w_i X_{t-i} Y_{t-i} \quad (4.5)$$

衰退因子的选取，隐含着对于距目前时间越近的信息，给予越重的权数，在 Riskmetrics 技术手册中，衰退因子为 0.94。

### (3) GARCH 模型

请参照第二章的模型。

### (4) 多因子风险模型

Rosenberg(1973)<sup>[86]</sup>发展多因子风险预测模型以预测证券风险，将过去财务研究所使用的历史指标以及会计信息预测系统风险的方式加以整合。

过去发展的证券投资风险的市场模型，将风险分为系统及非系统风险，如：

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i \quad (4.6)$$

$R_i$  = 第  $i$  种证券的收益率；

$\alpha$  = 常数项；

$\beta_i$  = 系统风险系数；

$R_m$  = 市场收益；

$\varepsilon_i$  = 证券的非系统风险报酬；

传统的方法皆假设公司的系统风险在各期都是固定的，但 Rosenberg 认为证券的市场行为及其公司特性会产生改变，因此证券的系统风险及非系统风险也会随之改变。假设  $\beta$  和  $\alpha$  都受到多种因素影响， $W$  为影响因素的组合，表示如下：

$$\beta_i = \sum_{j=1}^j b_{ij} W_j + \eta_i + \tilde{b}_i \tilde{W}_j + \varepsilon_i \quad (4.7)$$

其中， $W_j$  = 共同因素组合（ $j$  为共同因素数量）。

将（4.7）式代入到（4.6）式，则市场模型转变为受因子影响的多因子风险模型，如下所示：

$$R_i = \alpha + [\sum_j b_{ij} W_j + \eta_i] R_m + \varepsilon_i = \alpha + \sum_j b_{ij} (W_j R_m) + u_i$$

多因子风险模型认为任何种类的资产报酬，在不同时间下其本身特性（如产业属性不同），或是外在特征（如经济情况变化）的改变，会导致本身波动特性有不同变化。虽然多变量估计方式可能考虑的因子很多，但是通过多变量计算技术，可以大量浓缩因子特性到几个简单的因素，如此一来也能解决因子数量过多运算负荷过重，或线性相关的情形，虽然多因子风险模型符合直觉及运算的优点，但是该模型对于因子的选取、处理及运算都相当复杂，加上主观判断是不可避免的缺点，因此该方法的模型风险也相对较高。

#### 4. 2. 2 全额评价法

##### （1）历史模拟法

历史模拟法顾名思义，即是完全由实际的历史资料中，计算资产组合风险值的一种方法。在方法的操作上，历史模拟法利用所持有的资产组合，过去一段期间的历史价格时间序列，搭配目前持有资产的头寸，重新构建资产组合未来报酬值的分配之后，再经过由小到大顺序排序后，依百分位数计算特定置信水平下的风险值。

##### （2）拔靴法

拔靴法在操作过程中，并不需要知道样本的分布如何，此法是从一组历史

资料中进行多次重复抽样方式，来赋予每个样本相同的概率，利用重复抽样的方式估计资产组合未来报酬值的分布，以此计算风险值。拔靴法与历史模拟法不同的地方在于，历史模拟法是直接利用这些样本来构造资产组合未来报酬值的分布，因此只有一种价格路径；而拔靴法则是从这些历史资料中进行多次抽样，来模拟真实的分布，以改善历史模拟法的缺点，也就是拔靴法的价格变动值是经由重复抽样而来，所以除了产生价格变动值的方法不同之外，其余各步骤流程均可以参照历史模拟法的计算步骤。缺点是，使用不同长短的历史资料期间，将获得不同的风险值，如若使用较长的区间，可能会掺杂太过老旧的资料，而无法反映现在，若使用较短的区间，则可能降低估计的正确性。

### (3) 蒙特卡罗模拟法

蒙特卡罗模拟法是假设资产价格的变动服从某种随机过程，利用电脑模拟，在目标时间范围内，产生随机价格的路径，并依次构造资产报酬分布，进而推算风险值。这样的做法不但涵盖变量的所有可能状况，也可以处理非正态模型。蒙特卡罗模拟法和历史模拟法的基本概念类似，但不同之处在于，蒙特卡罗模拟法的价格变动率是从所指定的随机模型中任意抽取出来的。蒙特卡罗模拟包括三个步骤：先选择适合描述资产价格路径的随机过程；然后依随机过程模拟虚拟的资产价格路径；最后综合模拟结果，构造资产报酬值分布，并以此计算投资组合的风险值。优点是可涵盖非线性资产的价格风险、波动性风险、甚至可计算信用风险，并且可处理肥尾、不对称等非正态分布和极端状况等特殊情况。缺点是需要繁杂的电脑技术和大量的样本数，较费时，而且，若选择的代表价格变动的随机模型不当，会导致模型风险的产生。

## 4.3 风险模型

在过去，对风险的表达大多是描述性质，但是，随着使用者对风险衡量水准日益增加，因此，可量化的风险值越来越受重视。其实，寻求量化的风险值可以看作是寻找一个对应关系  $\rho$ ，从随机变量空间  $X$  映射到一个非负的实数  $R$ 。即  $\rho: X \rightarrow R$ 。作为标量的风险值允许排序，从而用来对比不同投资各自的风险值的大小。当然，映射  $\rho$  若是一致的风险测度必须符合一定的条件：



- (1) 正的齐次性：对于所有随机变量  $x$ ，正的实数  $\lambda$ ，有  $\rho(\lambda x) = \lambda\rho(x)$ ；
- (2) 次可加性：对于所有的随机变量  $x$  和  $y$ ， $\rho(x + y) \leq \rho(x) + \rho(y)$ ；
- (3) 单调性：对于所有的随机变量  $x$  和  $y$ ， $x \leq y \Rightarrow \rho(x) \leq \rho(y)$ ；
- (4) 传递不变性：对于所有的随机变量  $x$ ，实数  $\alpha$ ，和无风险率  $r_0$ ， $\rho(x + \alpha r_0) = \rho(x) - \alpha$ 。

以上条件具有一定的经济解释，若  $\rho$  不满足次可加性，即  $\rho(x) + \rho(y) < \rho(x + y)$ ，那就意味着为了降低风险，需要将一个公司分为两个不同的部分。但是，从协方差满足次可加性，可以得出次可加性在 Markowitz 投资组合理论中也是一个基本的特性，从而说明了新的投资是不会增加风险的。

传递不变性隐含着在随机收益  $x$  的基础上增加一个确定性的收益  $\alpha r_0$ ，那么风险  $\rho(x)$  将减少  $\alpha$ 。

#### 4. 3. 1 VaR

为了衡量交易资产投资组合的市场风险，许多银行都相应采取以在险价值为基础的内部模型。在策略性风险控管方面，企业可事先模拟各种状况下的风险值，提供部门就每一项投资组合，估计投入的资金或其他资源、可能产生损失的概率、部门所能承受损失的程度、部门预计承受的损失、如何调度资金以抛补既有风险等有效的、量化的可行方案。同时，风险值的概念可以应用于每一个交易对手，估计交易对手信用风险暴露的程度以及本身所能承受的最大损失。

综合而言，在险价值从各种角度评价企业所面临的潜在风险，不但提供企业风险管理信息，作为企业设立头寸限额的参考，更可以便利企业规划资源分配和衡量操作绩效。

在险价值 (VaR) 是在给定的置信度内，某种金融资产或资产组合在未来一段有效期内的最坏预期损失值，其数学表达式为：

$$\Pr(\Delta V_T < -VaR) = X\% \quad (4.8)$$

也就是说, VaR 是金融资产或者组合在  $1-X\%$  的置信度下持有期为  $T$  的最大价值损失。

对于单一资产的 VaR 的计算, 国内外已有很多研究。但是对于投资组合的 VaR 的计算, 却很少涉及。这是因为在计算投资组合的 VaR 时, 不仅牵涉到单个资产的波动性, 而且涉及到了各种资产之间的相关性的计算。

#### 4. 3. 2 CVaR

CvaR 的经济意义清楚, 数学性质比较好, 并且 CvaR 除了使相容风险测度外, 还是凸风险测度, 容易进行优化处理。

$$CVaR_k = \phi_k(x) = (1-k^{-1}) \int_{f(x,y) \geq \alpha_k(x)} f(x,y) p(y) dy \quad (4.9)$$

其中,  $k$  是置信水平,  $\alpha_k(x)$  为阈值,  $y$  是超过  $\alpha_k(x)$  的投资组合损失的变量,  $x$  为投资组合的权重,  $p(y)$  是概率密度函数。

#### 4. 3. 3 ER(Expected Regret)

ER 和 CvaR 相关度很高, 用来描述超过阈值的损失的预期值, 定义为:

$$G_\alpha(x) = \int_{y \in R^m} [f(x,y) - \alpha]^+ p(y) dy \quad (4.10)$$

其中,  $[u]^+ = \max\{0, u\}$ 。

人们在 VaR 的基础上根据风险必须满足的一致性条件出发, 构造了上述许多新的测量风险的方法, 除此之外还有 ES(expected shortfall), TCE(tail conditional expectation), TM(tail mean), WCE(worst conditional expectation), spectral risk measures。但是对于这些方法的检验等统计理论还没有完善, 因此本文对于投资组合的风险还停留在 VaR 上。

### 4. 4 在险价值的准确性检验方法

VaR 模型的准确性检验是指 VaR 模型的测量结果对实际损失的覆盖程度, 其检验方法包括 Kupiec 失败率检验法 (1995) [87] 和 Engle(2002) [88] 提出的动态系数检验。

#### 4. 4. 1 动态系数检验

Engle 和 Manganelli(2002)<sup>[88]</sup>提出了一种评价 VaR 的检验方法,此检验可以验证根据各种模型所计算的 VaR 对于实际发生风险的覆盖程度。

定义 hit 为:

$$hit_t = I\{w'x_t < -VaR(q)_t\} \quad (4.11)$$

$$hit_{t+1} = \delta_0 + \sum_{k=1}^8 \delta_k hit_{t-k+1} + \delta_9 VaR_t + \varepsilon_t \quad (4.12)$$

其中  $I(\cdot)$  是示性函数,  $q$  是 VaR 的系数,  $hit = 1$  表示投资组合的收益小于 VaR。

动态系数检验的原假设  $H_0$ : 模型正确设定的情况下, hit 序列应该具有期望  $q$ , 并且与它的滞后值不相关。

#### 4. 4. 2 LR 似然比检验

LR 似然比检验也叫作 Kupiec 失败率检验,假定实际考察天数为  $T$ , 失败天数为  $N$ , VaR 置信度为  $\alpha$ , 失败率为  $p$ 。假定 VaR 估计具有时间独立性,则失败天数  $N$  服从参数为  $T$  和  $p$  的二项分布,即  $N \sim B(T, p)$ 。

原假设  $H_0$ :  $p = p^*$  下, 似然比

$$LR = -2 \ln[(1 - p^*)^{T-N} p^{*N}] + 2 \ln[(1 - N/T)^{T-N} (N/T)^N] \sim \chi^2(1) \quad (4.13)$$

因此,通过计算,在 5% 的显著性水平下,如果  $LR > 3.8415$ , 我们拒绝原假设。

#### 4. 5 实证分析

在以往的投资决策中,我们一般使用预期收益率的估计值、收益率的均方差的无偏估计和协方差的无偏估计分别作为两种证券的均值、均方差和协方差,从而计算两种证券的相关系数。但是,相关经验研究表明,预期收益率、均方差、协方差以及相关系数会随着时间的变化而变化。Goetzman、Li & Rouwenhorst (2002)<sup>[89]</sup>提供了世界主要股票市场在过去 150 多年间的收益相关性的历史描述,他们的研究发现在过去 150 多年中世界股票市场收益之间的相关性呈现出

巨大的时变性。本章节采用数据总样本为 1992 年 12 月 28 日至 2006 年 3 月 31 日上证 A 股、上证 B 股、深综 A 指和深综 B 指的每日收盘价，共 3049 个数据。在做 VaR 模型的准确性检验时，我们用 2006 年 3 月 31 日之前 2500 个数据。数据的处理和模型的估计请参照章节 3.6。

#### 4. 5. 1 正态分布的检验

投资组合的方差为  $\hat{h}_{w,t} = w' \hat{H}_t w$ ，若假设投资组合的收益服从正态分布，则在 5% 的水平下， $VaR = 1.65 \sqrt{\hat{h}_t}$ ；在 1% 的水平下， $VaR = 2.33 \sqrt{\hat{h}_t}$ 。投资组合的方差中用到的权重  $w$  可以取为等权重  $w_t = k^{-1} \mathbf{1}$  和最小方差投资组合权重  $w_t = \frac{H_t^{-1} \mathbf{1}}{W_t}$ ，其中  $\mathbf{1}$  是 1 向量， $W_t = \mathbf{1}' H_t^{-1} \mathbf{1}$ 。在本文中，用 EW 代表等权重的投资组合，用 MVP 代表最小方差投资组合权重的投资组合，用 RiskM 代表 Riskmetrics 方法。

##### 1. 动态系数检验

具体方法是用 ADCC 模型估计出来的方差协方差矩阵和不同的权重来计算投资组合未来一天的 5%，1% 的置信水平下的 VaR 值。

表 4.1 动态系数检验结果

正态分布	Hits 百分比 (5%)	Q 统计量 (5%)	P 值 (5%)	Hits 百分比 (1%)	Q 统计量 (1%)	P 值 (1%)
ADCC(EW)	4.9	11.662	0.167	1.72	6.2805	0.616
RiskM(EW)	5.76	43.671	0	1.96	12.357	0.136
ADCC(MVP)	5.4	11.164	0.193	2.56	12.089	0.147
RiskM(MVP)	7.32	36.283	0	3.32	67.718	0

根据观测值的数量，在表 4.1 中列出了 8 阶滞后的 Q 统计量和 P 值。在 8 阶滞后的情况下，我们可以看到，利用 ADCC 模型得到的 VaR，不管是等权重还是最优投资组合权重计算出来的 VaR 的 Hits 都是不错的；同时可以很明显的

看出，用 Riskmetrics 的指数移动平均所得到的 Hits 偏大，这就说明了，对于中国主要指数组合，用 Riskmetrics 的方法来估计投资组合的方差  $\hat{h}_{w,t}$  是不合适的，犯了低估方差的错误，使得 Hits 偏大。

并且从 Q 统计量和 P 值来看，虽然在 5%显著水平下，Q 统计量都小于  $Q(8)=15.507$ ，但是，Riskmetrics (EW)和 Riskmetrics (MVP)的 P 值都是小于 0.01，也就是说残差序列相互独立的概率并不大。

根据 Engle 和 Manganeli(2002)动态系数检验表明用 Riskmetrics 来估计方差进而计算风险对于中国股指市场是不合适的。

## 2. LR 似然比检验 (Kupiec 失败率检验)

表 4.2 给出了 5%的置信水平的检验结果：

表 4.2 LR 似然比检验结果

相关矩阵估计方法	5%			1%		
	期望天数	失败天数	LR	期望天数	失败天数	LR
ADCC(EW)	125.00	122.5	0.0530	25	43	10.771
ADCC(MVP)	125.00	135	0.8217	25	64	42.939
RiskM(EW)	125.00	144	2.9043	25	49	18.182
RiskM(MVP)	125.00	183	24.937	25	83	84.564

从以上两种检验方法的结果来看，在 5%的置信水平下，利用 ADCC 模型，等权重和最小方差投资组合权重的投资组合的风险 VaR 值对实际损失的覆盖程度相当好。

其实投资组合收益的风险的这个结果和单个市场得到的结果有类似之处。譬如，在罗付岩（2005）<sup>[90]</sup>的文章中，利用上海综指的数据计算风险 VaR 值，正态分布在 1%的置信水平被拒绝。在 5%的置信水平下，所计算的 VaR 值对实际损失的测量结果相对比较好。

## 4. 5. 2 t 分布的检验

研究表明，很多金融资产的收益是符合厚尾分布的，从本文的数据统计量也可以看出各个股票的指数具有尖峰厚尾，t 分布比正态分布具有更厚的尾部，

可以描述收益率的厚尾特征，因此根据厚尾分布的性质构造投资组合  $t$  分布下的 VaR 具有现实意义。

具体算法如下：

1. 由于  $t$  分布牵涉到自由度的选择问题，因此把 3049 个数据平均分为两部分，用前 1524 天的观测值求出  $t$  分布的自由度  $\nu$ ；

2. 根据 ADCC 模型和 Riskmetrics 方法估计从 1525 天至 3049 天收益的方差协方差矩阵，计算等权重和最优投资组合权重两种方法的投资组合的方差  $\hat{h}_{w,t}$ ，根据  $\nu$  和  $\hat{h}_{w,t}$  得到后 1525 天  $t$  分布下 VaR 值；

3. 列出第 1525 天—3049 天的收益向量，从而求出等权重和最优投资组合权重实际的收益，从而得到 hit 序列。

首先根据  $T=1524$  天<sup>①</sup>之前的数据来计算  $t$  分布的自由度  $\nu$ ，那么 student's  $t$  分布的对数似然函数为：

$$L(r_1, r_2, \dots, r_T; \Theta) = T \ln \left[ \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi} \Gamma(\nu/2)} (\nu-2)^{-1/2} \right] - \frac{\nu+1}{2} \sum_{i=1}^T \left[ 1 + \frac{\hat{r}_i^2}{\hat{h}_i (\nu-2)} \right] - (1/2) \sum_{i=1}^T \ln \hat{h}_i$$

其中  $T$  为样本数， $\nu$  为自由度， $r_i$  为第  $i$  天的残差项， $\Theta$  为参数空间。

假设投资组合收益服从  $t$  分布，那么经过调整自由度后的 VaR 计算如下：

1%的  $t$  分布的 VaR 是  $t_{\nu,0.01} \sqrt{\hat{h}_t (\hat{\nu}-2)/\hat{\nu}}$ ；5%的  $t$  分布的 VaR 是  $t_{\nu,0.05} \sqrt{\hat{h}_t (\hat{\nu}-2)/\hat{\nu}}$ 。

表 4.3  $t$  分布检验结果

t 分布	Hits 百分比 (5%)	Q 统计量 (5%)	P 值 (5%)	Hits 百分比 (1%)	Q 统计量 (1%)	P 值 (1%)
ADCC(EW)	6.7906	7.377	0.497	0.57184	0.3743	0.99
RiskM(EW)	7.5054	16.806	0.032	0.78628	0.5878	0.99
ADCC(MVP)	7.3624	2.5119	0.961	1.4296	3.9235	0.864
RiskM(MVP)	9.6497	12.815	0.118	2.2159	16.794	0.032

<sup>①</sup>本文用了其他的数据分法，把 3049 个数据分为  $k$  个和  $(3049-k)$  个数据，其中， $k=800,1000,1200$ ，用前  $k$  天的数据计算  $t$  分布的自由度，得到的结果是大体相同，也就是说算法与计算  $t$  分布的自由度  $\nu$  的数据的分法没有关系，因此算法具有稳健性。

在假设投资组合收益服从  $t$  分布的情况下，从 hit 序列的独立性来看，用 Riskmetrics 方法除了 1% 置信水平下的 hit 序列的 P 值较大以外，其余的 P 值都是小于 0.2 的，拒绝了 hit 序列不相关。因此，用 ADCC 模型比用 Riskmetrics 方法要好。同时在 1% 的置信水平下，用  $t$  分布得到的等权重 hit 的百分数较小，说明假设投资组合为  $t$  分布并且等权重的情况下 VaR 值较高，从而设置的门槛也高，使得 hits 的百分数集体偏小。

对比表 4.1 和表 4.3，ADCC 模型的最优投资组合权重在 1% 的置信水平下效果较好。但是 5% 的置信水平下，用等权重的投资组合结合 ADCC 模型得到的 hits 的百分数较好。不管投资组合收益在正态分布假设和  $t$  分布假设的情况下，用 ADCC 模型所得到的风险值都好于 Riskmetrics 方法。

其实投资组合收益的风险的这个结果和单个市场得到的结果有类似之处。譬如，在罗付岩（2005）的文章中，利用上海综指的数据计算风险 VaR 值，正态分布在 1% 的置信水平被拒绝。1% 的置信水平下， $t$  分布表现过于保守，高估了市场风险。在 5% 的置信水平下，两者的差别不大。

## 5. SCC 模型

在许多金融计量问题中(如风险管理、衍生产品定价和最优投资组合选择),模拟二阶矩的相关性和预测未来的波动性有很重要的意义。近年来出现了大量多维 GARCH 模型来模拟资产组合的波动性及相关性,但是在估计多维 GARCH 模型中仍存在着不尽如人意的地方,资产的数量过多困扰着估计时的最优化问题,即使随着模型的逐渐改善,参数估计仍然十分困难,因此许多文献都停留在二维或三维 GARCH 模型的估计,很难推广到高维 GARCH 模型。本文的这一章引入一种处理高维相关矩阵的估计方法--SCC(序列条件相关)方法,此方法非常灵活,可以把高维相关矩阵的估计问题转化为二维相关矩阵的估计,并且允许二维相关矩阵的估计采用各种灵活的二维 GARCH 模型或其他新的方法进行分别估计,然后再组合为一个高维相关矩阵。

### 5.1 SCC 方法的起源

SCC 方法起源于条件相关的方法,其中条件相关包括常数条件相关(CCC)和动态条件相关(DCC)。

#### 5.1.1 条件相关模型

Bollerslev(1990)<sup>[29]</sup>提出了多维 GARCH 模型的条件相关(CC)模型,CC 模型分别对条件方差和条件相关性进行估计,并且条件方差  $h_{i,t}$  是滞后项  $h_{i,t-j}$  和残差项  $\varepsilon_{i,t-j}$  的函数。这就意味着多维 CC 模型可以被表示成 M 个单维波动性模型,因此可以允许相当灵活的一维波动性模型的估计,譬如 GARCH、随机波动性、非参数方法和已实现的波动性等。其中条件相关(CC)多维 GARCH 模型中的常数条件相关(CCC)模型给出的条件方差-协方差的形式为:

$$H_t = D_t R D_t$$

其中,  $D_t$  是对角矩阵,每个对角线上的元素是一维 GARCH 的方差的平方根。 $R$  是常数相关矩阵,  $R_{ii} = 1$ 。但是常数条件相关的主要缺点是常数相关的



假设在许多实际问题中是不现实的。

Engle(2001)<sup>[52]</sup>提出的动态条件相关模型把 CCC 进一步一般化, DCC 模型的好处是估计可以分成两步走: 第一步, 可以用一维的 GARCH 模型进行估计, 得到矩阵  $D_t$ , 同时对数据进行标准化; 第二步对于标准化后的数据, 估计条件相关模型的参数, 构造  $R_t$ 。和 CCC 所不同的是, 它的相关矩阵  $R_t$  的构造打破了常数的限制,  $R_t$  的结构如下:

$$Q_t = (1 - \sum_{m=1}^M \alpha_m - \sum_{n=1}^N \beta_n) \bar{Q} + \sum_{m=1}^M \alpha_m (\varepsilon_{t-m} \varepsilon'_{t-m}) + \sum_{n=1}^N \beta_n Q_{t-n}$$

$$R_t = Q_t^{-1} Q_t Q_t^{-1}$$

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{q_{kk}} \end{bmatrix}$$

其中,  $\bar{Q}$  是标准残差的无条件方差协方差矩阵,  $Q_t$  是为了得到正确的相关矩阵所用到的一个正定矩阵,  $R_t$  是相关矩阵,  $Q_t^*$  是一个对角矩阵, 对角线上的元素等于  $Q_t$  的对角线元素的平方根,  $\alpha_m$ 、 $\beta_n$  是待估参数。

### 5. 1. 2 SCC 的介绍

Alessandro (2006)<sup>[91]</sup>提出了 SCC 方法, 这种方法利用了条件相关 (CC) 模型的思想, 把条件方差—协方差矩阵进行分解, 只是分解方法与 CCC 和 DCC 方法的分解方法有所不同。CCC 和 DCC 方法的分解只是把把方差和相关矩阵分开, 从而允许了两步估计, SCC 方法进一步分离了相关性和偏相关, 从而允许了一个真正的多步估计程序。这个思路使条件方差—协方差矩阵写成了以下形式:

$$H_t = D_t R D_t$$

$$\Rightarrow H_t = D_t K_{1,2,t} K_{1,3,t} \cdots K_{M-1,M,t} K'_{M-1,M,t} \cdots K'_{1,3,t} K'_{1,2,t} D_t \quad (5.1)$$

$$\text{其中, } K_{p,q,t}[i, j] = \begin{cases} \rho & i = q, j = p \\ (1 - \rho^2)^{1/2} & i = q, j = q, \\ I[i, j] & \text{otherwise} \end{cases}$$

$$K_{p,q}^{-1}[i, j] = \begin{cases} -\rho(1 - \rho^2)^{-1/2} & i = q, j = p \\ (1 - \rho^2)^{-1/2} & i = q, j = q \\ I[i, j] & \text{otherwise} \end{cases}, \rho \text{ 是在被 } K_{p,q}^{-1} \text{ 相乘之前的相}$$

关矩阵的第  $(p, q)$  个元素。

这种特殊的分解保证了方差—协方差矩阵的正定性和对称性，并且这种分解允许多步的模拟和方差的估计。我们暂且把  $K_{i,j,t}$  叫作  $K$  矩阵，那么多维 GARCH 模型的相关矩阵的复杂估计可以转化为一系列简单的估计。

为了弄清楚  $K$ —矩阵的分解，令  $y_t$  是  $t$  时刻  $N$  维观察向量， $H_t$  是它的方差协方差矩阵。那么条件方差协方差矩阵分解的步骤是：

(1) 模拟  $N$  维的  $y_t$  序列的条件方差，并且用各自的条件方差对数据  $y_t$  进行标准化：

$$\hat{\varepsilon}_{i,t} = \frac{y_{i,t}}{\hat{h}_{i,t}^{1/2}} \quad \forall i = 1, \dots, N, \forall t = 1, \dots, T$$

则  $\hat{\varepsilon}_t$  表现出同方差，并且仍旧表现出随时间可变的相关性。

(2) 对于序列  $\hat{\varepsilon}_{i,t}$ ，模拟此序列的第一项同余下的  $N-1$  项的相关性，得出下列偏相关估计  $\{\rho_{1,2,t}\}_{t=1}^T, \{\rho_{1,3,t}\}_{t=1}^T, \dots, \{\rho_{1,N,t}\}_{t=1}^T$ ，然后从第 2 项至第  $N$  项抽出对  $\hat{\varepsilon}_{1,t}$  的影响，换句话说这些相关估计可以进一步使数据标准化

$$\tilde{\varepsilon}_{i,t} = \frac{\hat{\varepsilon}_{i,t} - \rho_{1,i,t} \hat{\varepsilon}_{1,t}}{(1 - \rho_{1,i,t}^2)^{1/2}} \quad \forall i > 1, \text{ 使得转化后的序列 } \tilde{\varepsilon}_{i,t} \text{ 的第 } 2, 3, \dots, N \text{ 项与它的第一项条件不相关。}$$

(3) 抛弃  $\tilde{\varepsilon}_{i,t}$  的第一项，重复第二个步骤，求出  $\tilde{\varepsilon}_{i,t}$  的第二项分别与余下的  $N-2$  项的偏相关估计，并对序列  $\tilde{\varepsilon}_{i,t}$  的第二项至第  $N$  项进行标准化；然后是第三个序列和余下的  $N-3$  个序列相关性的估计和后面序列的标准化。

(4) 重复以上步骤, 完全标准化后的序列, 它的方差—协方差矩阵是单位阵。而且被估计的相关性可以插入到相应的  $K$ —矩阵中, 这些  $K$ —矩阵的乘积可以得到相关矩阵  $R_t$ , 再结合  $D_t$  就可以得到  $H_t$ 。

对于以上的步骤, 本文给出了严格的数学证明。

**定理 5-1** 经过标准化后的数据  $\tilde{\varepsilon}_{i,t} (i > 1)$  的期望是 0, 方差是 1, 并且与  $\hat{\varepsilon}_{1,t}$  不相关。

**证明:**  $E(\hat{\varepsilon}_{i,t} - \rho_{1,i,t}\hat{\varepsilon}_{1,t}) = 0,$

$$\begin{aligned} & V(\hat{\varepsilon}_{i,t} - \rho_{1,i,t}\hat{\varepsilon}_{1,t}) \\ &= E[(\hat{\varepsilon}_{i,t} - \rho_{1,i,t}\hat{\varepsilon}_{1,t})^2] - [E(\hat{\varepsilon}_{i,t} - \rho_{1,i,t}\hat{\varepsilon}_{1,t})]^2 \\ &= E[\hat{\varepsilon}_{i,t}^2 - 2\rho_{1,i,t}\hat{\varepsilon}_{1,t}\hat{\varepsilon}_{i,t} + \rho_{1,i,t}^2\hat{\varepsilon}_{1,t}^2] \\ &= 1 - \rho_{1,i,t}^2, \text{ 因此, } V\left(\frac{\hat{\varepsilon}_{i,t} - \rho_{1,i,t}\hat{\varepsilon}_{1,t}}{(1 - \rho_{1,i,t}^2)^{1/2}}\right) = 1. \end{aligned}$$

$$E(\tilde{\varepsilon}_{i,t}\hat{\varepsilon}_{1,t}) = E\left(\frac{\hat{\varepsilon}_{i,t} - \rho_{1,i,t}\hat{\varepsilon}_{1,t}}{(1 - \rho_{1,i,t}^2)^{1/2}}\hat{\varepsilon}_{1,t}\right) = 0.$$

**定理 5-2** 经过以上步骤 (1) — (4) 的分解, 条件相关矩阵  $R_t$  可以被分解为:

$$R_t = K_{1,2,t}K_{1,3,t} \cdots K_{N-1,N,t}K'_{N-1,N,t} \cdots K'_{1,3,t}K'_{1,2,t} \quad (5.2)$$

**证明:** 经过变换,  $K_{N-1,N,t}^{-1} \cdots K_{1,N,t}^{-1} \cdot K_{1,N-1,t}^{-1} \cdots K_{1,2,t}^{-1} \cdot \hat{\varepsilon}_t$  的序列是一个方差协方差为单位阵的序列。即  $V(K_{N-1,N,t}^{-1} \cdots K_{1,N,t}^{-1} \cdot K_{1,N-1,t}^{-1} \cdots K_{1,2,t}^{-1} \cdot \hat{\varepsilon}_t) = I$ 。

设  $A = K_{N-1,N,t}^{-1} \cdots K_{1,N,t}^{-1} \cdot K_{1,N-1,t}^{-1} \cdots K_{1,2,t}^{-1}$ , 则  $V(A\hat{\varepsilon}) = I$ , 又因为  $E(A\hat{\varepsilon}) = 0$ , 因此  $E[(A\hat{\varepsilon})(A\hat{\varepsilon})'] = I$ , 即  $A \cdot E(\hat{\varepsilon}\hat{\varepsilon}') \cdot A' = I$ , 也即  $A \cdot R \cdot A' = I$ 。

**定理 5-3**<sup>①</sup> 对于  $M \times M$  相关矩阵  $R$ , 存在唯一的序列矩阵  $K_{1,2}^{-1}K_{1,3}^{-1} \cdots K_{N-1,N}^{-1}$  满足:

<sup>①</sup>请参照 Alessandro (2006) [91]。

$$I = K_{N-1,N}^{-1} \cdots K_{1,3}^{-1} \cdot K_{1,2}^{-1} \cdot R \cdot (K_{1,2}^{-1})' \cdot (K_{1,3}^{-1})' \cdots (K_{N-1,N}^{-1})' \quad (5.3)$$

定理 5-4<sup>①</sup>: 按照一定的顺序, 任何  $K$  序列的乘积都是相关矩阵。

### 5. 1. 3 数值例子

为了更好地理解 SCC 的分解, 下面给出一个数值例子:

$$R = \begin{pmatrix} 1 & 0.3 & 0.2 & -0.1 \\ 0.3 & 1 & 0.4 & -0.5 \\ 0.2 & 0.4 & 1 & -0.6 \\ -0.1 & -0.5 & -0.6 & 1 \end{pmatrix} \rightarrow K_{1,2}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -0.31448545 & 1.0482848 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_{1,3} = K_{1,2}^{-1} R (K_{1,2}^{-1})'$$

$$R_{1,3} = \begin{pmatrix} 1 & 0 & 0.2 & -0.1 \\ 0 & 1 & 0.35641684 & -0.49269387 \\ 0.2 & 0.35641684 & 1 & -0.6 \\ -0.1 & -0.49269387 & -0.6 & 1 \end{pmatrix}$$

$$\rightarrow K_{1,3}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0.20412415 & 0 & 1.0206207 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_{1,4} = K_{1,3}^{-1} R_{1,3} (K_{1,3}^{-1})'$$

$$R_{1,4} = \begin{pmatrix} 1 & 0 & 0 & -0.1 \\ 0 & 1 & 0.36376642 & -0.49269387 \\ 0 & 0.36376642 & 1 & -0.59196002 \\ -0.1 & -0.49269387 & -0.59196002 & 1 \end{pmatrix}$$

$$\rightarrow K_{1,4}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.10050378 & 0 & 0 & 1.0050378 \end{pmatrix}$$

$$R_{2,3} = K_{1,4}^{-1} R_{1,4} (K_{1,4}^{-1})'$$

<sup>①</sup>请参照 Alessandro (2006) [91]。

$$R_{2,3} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0.36376642 & -0.49517597 \\ 0 & 0.36376642 & 1 & -0.59494221 \\ 0 & -0.49517597 & -0.59494221 & 1 \end{pmatrix}$$

$$\rightarrow K_{2,3}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -0.39052093 & 1.0735486 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_{2,4} = K_{2,3}^{-1} R_{2,3} (K_{2,3}^{-1})'$$

$$R_{2,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -0.49517597 \\ 0 & 0 & 1 & -0.44532279 \\ 0 & -0.49517597 & -0.44532279 & 1 \end{pmatrix}$$

$$\rightarrow K_{2,4}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0.56995872 & 0 & 1.1510226 \end{pmatrix}$$

$$R_{3,4} = K_{2,4}^{-1} R_{2,4} (K_{2,4}^{-1})'$$

$$R_{3,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -0.51257658 \\ 0 & 0 & -0.51257658 & 1 \end{pmatrix}$$

$$\rightarrow K_{3,4}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0.59696226 & 1.1646304 \end{pmatrix}$$

$$K_{3,4}^{-1} R_{3,4} (K_{3,4}^{-1})' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

为了更清楚 SCC 方法的分解步骤，本文以三维收益向量为例用图示的方法给出进一步的说明：

$$\begin{aligned}
 & \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{pmatrix} \xrightarrow{\frac{y_{i,t}}{\hat{h}_{i,t}^{1/2}}} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \hat{\epsilon}_{2,t} \\ \hat{\epsilon}_{3,t} \end{pmatrix} \xrightarrow{\tilde{\epsilon}_{2,t} = \frac{\hat{\epsilon}_{2,t} - \rho_{1,2,t} \hat{\epsilon}_{1,t}}{(1 - \rho_{1,2,t}^2)^{1/2}}} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \tilde{\epsilon}_{2,t} \\ \hat{\epsilon}_{3,t} \end{pmatrix} \\
 & \xrightarrow{\tilde{\epsilon}_{3,t} = \frac{\hat{\epsilon}_{3,t} - \rho_{1,3,t} \hat{\epsilon}_{1,t}}{(1 - \rho_{1,3,t}^2)^{1/2}}} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \tilde{\epsilon}_{2,t} \\ \tilde{\epsilon}_{3,t} \end{pmatrix} \xrightarrow{\bar{\epsilon}_{3,t} = \frac{\tilde{\epsilon}_{3,t} - \tilde{\rho}_{2,3,t} \tilde{\epsilon}_{2,t}}{(1 - \tilde{\rho}_{2,3,t}^2)^{1/2}}} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \tilde{\epsilon}_{2,t} \\ \bar{\epsilon}_{3,t} \end{pmatrix}
 \end{aligned}$$

写成 K 矩阵 ( $K_{i,j,t}$ ) 的形式如下:

$$\begin{aligned}
 & \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{pmatrix} \xrightarrow{\frac{y_{i,t}}{\hat{h}_{i,t}^{1/2}}} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \hat{\epsilon}_{2,t} \\ \hat{\epsilon}_{3,t} \end{pmatrix} \xrightarrow{K_{1,2,t}^{-1} \hat{\epsilon}_t} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \tilde{\epsilon}_{2,t} \\ \hat{\epsilon}_{3,t} \end{pmatrix} \\
 & \xrightarrow{K_{1,3,t}^{-1} K_{1,2,t}^{-1} \hat{\epsilon}_t} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \tilde{\epsilon}_{2,t} \\ \tilde{\epsilon}_{3,t} \end{pmatrix} \xrightarrow{K_{2,3,t}^{-1} K_{1,3,t}^{-1} K_{1,2,t}^{-1} \hat{\epsilon}_t} \begin{pmatrix} \hat{\epsilon}_{1,t} \\ \tilde{\epsilon}_{2,t} \\ \bar{\epsilon}_{3,t} \end{pmatrix}
 \end{aligned}$$

根据定理 5.1 可得:

$$E(\hat{\epsilon}_{1,t} \tilde{\epsilon}_{2,t}) = 0, \quad E(\hat{\epsilon}_{1,t} \tilde{\epsilon}_{3,t}) = 0;$$

$$E(\tilde{\epsilon}_{2,t} \tilde{\epsilon}_{3,t}) = 0;$$

$$E(\hat{\epsilon}_{1,t} \bar{\epsilon}_{3,t}) = 0。即变换后的序列彼此不相关。$$

## 5. 2 Cholesky 分解方法

令  $A$  是一个  $N \times N$  的对称正定矩阵,  $E_i$  是一个初等矩阵, 把矩阵  $A$  的第一行乘以  $-a_{i1}/a_{11}$  加到  $A$  的第  $i$  行, 使  $A$  的第  $(i,1)$  元素变为 0。令  $F_1 = E_N \cdots E_2$ , 那

么  $F_1 A F_1'$  是分块对角矩阵, 分块矩阵的维数分别是 1 和  $N-1$ 。同理, 令  $F_2$  是初等矩阵的乘积, 这些初等矩阵的目的是使  $F_1 A F_1'$  的第  $(i, 2)$  元素变为 0, 其中  $i > 0$ 。那么,  $F_2 F_1 A F_1' F_2'$  将是分块对角矩阵, 分块矩阵的维数分别是 1, 1,  $N-2$ 。重复以上步骤, 存在  $F_1, F_2, \dots, F_{N-1}$ , 满足  $F_{N-1} \dots F_1 A F_1' \dots F_{N-1}' = \Lambda$ , 其中  $\Lambda$  是一个对角矩阵。令  $T = (F_{N-1} \dots F_2 F_1)^{-1} \Lambda^{1/2}$ , 那么  $A = T T'$ , 其中  $T$  是下三角矩阵。

以上是 Cholesky 分解最一般的方法, 除此之外 Cholesky 分解方法有很多, SCC 方法其实就是一种特殊的 Cholesky 分解, 经过这种分解, 那么 SCC 方法的渐近性定理就迎刃而解。

考虑一个典型的金融计量模型:  $y_t = \mu_t + H_t^{1/2} u_t$

其中,  $y_t$  是  $n$  维的资产收益率向量,  $\mu_t$  是条件期望向量,  $H_t$  是条件方差协方差矩阵,  $\varepsilon_t = H_t^{1/2} u_t$  是残差向量。条件方差协方差矩阵  $H_t$  可以进一步分解为:  $H_t = D_t R_t D_t$ , 其中  $D_t$  是对角矩阵, 对角线上的元素是  $H_t$  对角线元素的平方根,  $R_t$  是相关矩阵, 若采用 CCC (Bollerslev, 1990)<sup>[29]</sup>, 那么  $R_t$  是常数相关矩阵, 若采用 DCC 模型(Engle, 2002)<sup>[30]</sup>, 那么  $R_t$  是动态条件相关矩阵。基于 Cholesky 分解相关矩阵  $R_t$  的过程, 可以将其写成:

$$R_t = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} & \rho_{2n} & \cdots & 1 \end{bmatrix} = K_t K_t' \quad (5.4)$$

$K_t$  是下三角矩阵。为简化符号方便说明起见, 在不混淆一般的情况下, 可省略时点  $t$  下标的使用, 首先写出相关系数矩阵  $R$  的转换过程如下:

$$K^{-1} R (K^{-1})' = K_{n-1}^{-1} \dots K_2^{-1} \cdot K_1^{-1} \cdot R \cdot (K_1^{-1})' \cdot (K_2^{-1})' \dots (K_{n-1}^{-1})' = I \quad (5.5)$$

$$K_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{-\rho_{12}}{\sqrt{1-\rho_{12}^2}} & \frac{1}{\sqrt{1-\rho_{12}^2}} & 0 & \cdots & 0 \\ \frac{-\rho_{13}}{\sqrt{1-\rho_{13}^2}} & 0 & \frac{1}{\sqrt{1-\rho_{13}^2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{-\rho_{1n}}{\sqrt{1-\rho_{1n}^2}} & 0 & 0 & \cdots & \frac{1}{\sqrt{1-\rho_{1n}^2}} \end{bmatrix}$$

$$K_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & \frac{-\rho_{23|1}}{\sqrt{1-\rho_{23|1}^2}} & \frac{1}{\sqrt{1-\rho_{23|1}^2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \frac{-\rho_{2n|1}}{\sqrt{1-\rho_{2n|1}^2}} & 0 & \cdots & \frac{1}{\sqrt{1-\rho_{2n|1}^2}} \end{bmatrix}$$

...

$$K_{n-1}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & \frac{-\rho_{n-1,n|n-2}}{\sqrt{1-\rho_{n-1,n|n-2}^2}} & \frac{1}{\sqrt{1-\rho_{n-1,n|n-2}^2}} \end{bmatrix}$$

以上转换矩阵之中  $\rho_{ijk}$  是剔除第 1 个至第  $k$  个变量影响之后,  $i$  和  $j$  变量的  $k$  阶偏相关系数。按变量序列排列  $i > k, j > i$  ( $i=1,2,\dots,n-1; j=2,\dots,n; k=0,1,\dots,n-2$ ),  $\rho_{ij} = \rho_{ij0}$  则是  $i$  和  $j$  变量的相关系数。Palandri (2006)<sup>[91]</sup> 所提出的 SCC (Sequential Conditional Correlations)与上述 Cholesky 分解相关系数矩阵思路相同, 具有对称正定及唯一性, 以下称此为 SCC(序列条件相关)方法, 主要的用途是将相关矩阵  $R$  进一步分离出相关性和偏相关, 从而允许一个简单的二维多步估计程序。先计算两两变量偏相关, 最后再还原得到整个相关矩阵  $R$ 。



### 5.3 渐近理论

证明多维 GARCH 模型的渐近理论有两种方法，一种思路是利用传统的渐近定理，另外一种是利用前人证明的已知的结论，譬如两阶段估计。

#### 5.3.1 传统渐近定理

Basawa, Feigin 和 Heyde(1976)<sup>[92]</sup>给出了对于一般随机过程 MLE 的渐近正态性的充分条件，这些条件是：

$$(1) -\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'} \xrightarrow{P} C_1, T \rightarrow \infty, \text{ 其中 } C_1 \text{ 是非随机正定矩阵;}$$

$$(2) \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial l_t(\theta_0)}{\partial \theta} \xrightarrow{d} N(0, C_0), T \rightarrow \infty; \text{ 其中 } C_0 \text{ 是非随机的;}$$

$$(3) \text{ 存在常数 } M < \infty, \text{ 使得对所有 } i, j, k, \text{ 对所有的 } \theta \in \Theta, E \left| \frac{\partial^3 l_t(\theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| < M。$$

$$\text{其中, } f_n(\theta) = -L_n(\theta) = \frac{1}{2n} \sum_{t=1}^n (\ln |H_t| + X_t' H_t^{-1} X_t) = \frac{1}{2n} \sum_{t=1}^n l_t(\theta)$$

与本文所研究的 ARCH 模型体系的渐近定理相关知识的研究概况如下：  
Ling(1995)<sup>[93]</sup>研究了一维的 GARCH 模型的严平稳性及高阶矩的存在性，给出了一些简单的充分条件，奠定了该模型的理论基础。在理论研究方面，Sang-Won Lee 和 Bruce E.Hansen(1994)<sup>[94]</sup>研究了 GARCH(1,1)模型伪极大似然估计的渐近理论，包括相合性和渐近正态性。在证明过程中把  $h_t$  的初值设为是常数，文中得到的结论如下：

对于模型 GARCH(1,1):

$$y_t = \mu + \varepsilon_t, t = 1, 2, \dots, n$$

$$\varepsilon_t = \eta_t h_t^{1/2}$$

$$E(\eta_t | \psi_{t-1}) = 0 \quad a.s. \quad E(\eta_t^2 | \psi_{t-1}) = a.s.$$

$$h_t = \omega_0(1 - \beta_0) + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 h_{t-1} \quad \psi_t = \sigma(\varepsilon_t, \varepsilon_{t-1}, \dots)$$

定理 5-5<sup>[94]</sup> 假设：

- (1)  $\eta_t$  是严平稳和遍历的;
- (2)  $\eta_t^2$  是非退化的;
- (3) 对某一个  $\delta > 0$ , 存在  $S_\delta < \infty$  满足  $E(\eta_t^{2+\delta} | \psi_{t-1}) \leq S_\delta < \infty \quad a.s.$ ,
- (4)  $\sup_t E(\ln(\beta_0 + \alpha_0 z_t^2) | \psi_{t-1}) < 0 \quad a.s.$

则  $\hat{\theta}_n \xrightarrow{p} \theta_0$

引理 5-6<sup>[94]</sup> 假设:

- (1)  $\eta_t$  是严平稳和遍历的;
- (2)  $\eta_t^2$  是非退化的;
- (3) 对某一个  $\delta > 0$ , 存在  $S_\delta < \infty$  满足  $E(\eta_t^{2+\delta} | \psi_{t-1}) \leq S_\delta < \infty \quad a.s.$ ;
- (4)  $\sup_t E(\ln(\beta_0 + \alpha_0 \eta_t^2) | \psi_{t-1}) < 0 \quad a.s.$ ;
- (5)  $E(\eta_t^4 | \psi_{t-1}) \leq \kappa < \infty \quad a.s.$ ;
- (6)  $\theta_0$  在  $\Theta$  的内部;

则  $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, V_0)$ 。

其中  $V_0 = B_0^{-1} A_0 B_0^{-1}$ ,  $B_0 = B(\theta_0) = -E\nabla^2 l_t(\theta_0)$

$$A_t = E(\nabla l_t(\theta_0) \nabla l_t(\theta_0)'), l_t(\theta) = -\left( \ln h_t(\theta) + \frac{e_t^2}{h_t(\theta)} \right)$$

Daniel Straumann 和 Thomas Mikosch (2003)<sup>[95]</sup>用随机迭代等式证明了 GARCH(1,1)模型 MLE 的相合性和渐近正态性。

对于多维的情形, Jeantheau 研究了它们的 MLE 的强相合性。对于 Bollerslev, Engle 和 Woodridge(1988)<sup>[2]</sup>提出的多维 GARCH 模型, 刘继春(2001)<sup>[96]</sup>证明此模型的 MLE 的渐近正态性和相合性。对于 Baba, Engle, Kraft 和 Kroner(1995)<sup>[21]</sup>提出的多维 GARCH-BEKK 模型, F.Comte 和 O.Lieberman(2003)<sup>[97]</sup>给出了此模型的 MLE 的渐近正态性和相合性的证明。

### 5.3.2 两阶段估计

对于 Engle 和 Sheppard (2001)<sup>[52]</sup>提出的 DCC 模型, 对数似然函数可以分为两部分进行估计:

$$\begin{aligned} QL1_T(\phi | r_t) &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log |D_t|^2 + r_t' D_t^{-2} r_t) \\ &= -\frac{1}{2} \sum_{t=1}^T \sum_{i=1}^2 (\log(2\pi) + \log(h_{ii,t}) + \frac{r_{i,t}^2}{h_{ii,t}}) \end{aligned}$$

$$QL2_T(\psi | \hat{\phi}, r_t) = -\frac{1}{2} \sum_{t=1}^T (\log |R_t| + \eta_t' R_t^{-1} \eta_t - \eta_t' \eta_t) = -\frac{1}{2} \sum_{t=1}^T (\log |R_t| + \eta_t' R_t^{-1} \eta_t)$$

其中, 参数  $\theta = (\phi, \psi) = (\phi_1, \phi_2, \dots, \phi_k, \psi)$ ,  $\phi_i = (\omega, \alpha_{1i}, \dots, \alpha_{pi}, \beta_{1i}, \dots, \beta_{qi})$ 。

White(1996)<sup>[98]</sup>提出的两阶段 QMLE 估计成立的条件是:

A1 (1) 对于参数集合  $\Phi$  的  $\phi$ ,  $E(\log f_1(r_t, \phi))$  存在并且有限,  $t=1, 2, \dots$ ;

(2)  $\{\log f_1(r_t, \phi)\}$  服从强大数定律;

(3) 对于参数集  $\theta = (\phi, \psi)$ ,  $E(\log f_2(r_t, \theta))$  存在且有限;

(4)  $\{\log f_2(r_t, \theta)\}$  服从强大数定律;

A2 (1)  $\theta_0 = (\phi_0, \psi_0)$  是唯一确定的,  $\Theta$  是紧集,  $\theta_0$  满足 DCC 正定性的条件;

(2)  $\{\bar{L}_{1T}(\phi) = E(T^{-1} \sum_{t=1}^T \log f_1(r_t, \phi))\}$  是  $O(1)$  的;

(3)  $\{\bar{L}_{2T}(\theta) = E(T^{-1} \sum_{t=1}^T \log f_2(r_t, \theta))\}$  是  $O(1)$  的;

A3 (1) 对于  $\phi$ ,  $\nabla \bar{L}_{1T}(\phi) = E(\nabla L_{1T}(r^T, \phi)) < \infty$ , 其中  $r^T = (r_1, r_2, \dots, r_T)$ , 是  $T$  维观测值向量;

(2) 对于  $\theta$ ,  $\nabla \bar{L}_{2T}(\theta) = E(\nabla L_{2T}(r^T, \theta)) < \infty$ ;

A4 (1) 对于  $\phi$ ,  $\nabla^2 \bar{L}_{1T}(\phi) = E(\nabla^2 L_{1T}(r^T, \phi)) < \infty$ ;

(2)  $E(\nabla^2 L_{1T}(r^T, \cdot))$  在  $\Phi$  上对于  $T=1, 2, \dots$  是一致连续的;

(3)  $\{\nabla^2 \log f_1(r_t, \phi)\}$  服从强大数定律;

(4) 对于  $\theta$ ,  $\nabla^2 \bar{L}_{2T}(\theta) = E(\nabla^2 L_{2T}(r^T, \theta)) < \infty$ ;

(5)  $E(\nabla^2 L_{2T}(r^T, \cdot))$  在  $\Theta$  上对于  $T=1, 2, \dots$  是一致连续的;

(6)  $\{\nabla^2 \log f_2(r_t, \phi)\}$  服从强大数定律;

A5 (1)  $\{A_{11,T} = \nabla_{\phi\phi} \bar{L}_{1T}(\phi_0)\}$  是  $O(1)$ , 并且是负定的;

(2)  $\{A_{22,T} = \nabla_{\psi\psi} \bar{L}_{2T}(\theta_0)\}$  是  $O(1)$ , 并且是负定的;

A6 (1)  $\{T^{-1/2} \nabla_{\phi}' \ln f_1(r_t, \phi_0), T^{-1/2} \nabla_{\psi}' \ln f_2(r_t, \phi_0, \psi_0)\}$  服从中心极限条件。

定理 5-8<sup>[98]</sup> (渐近正态性) 在 White(1996)<sup>[98]</sup> 提出的 A1-A6 的假设条件下,

对于  $f_1$  和  $f_2$ , 那么

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, C_1^{-1} C_0 C_1^{-1}), \quad n \rightarrow \infty$$

其中,  $C_0 = \text{Var}[\sum_{t=1}^T \{T^{-1/2} \nabla_{\phi}' \ln f_1(r_t, \phi_0), T^{-1/2} \nabla_{\psi}' \ln f_2(r_t, \phi_0, \psi_0)\}]$ ,

$$C_1 = \begin{bmatrix} \nabla_{\phi\phi} \ln f_1(\phi_0) & 0 \\ \nabla_{\phi\psi} \ln f_2(\theta_0) & \nabla_{\psi\psi} \ln f_2(\theta_0) \end{bmatrix}.$$

## 5. 4 SCC 的对数似然函数

假设模型误差为多维正态分布的随机变量, 写成对数似然函数 (不含常数项) 为

$$L = -\frac{1}{2} \sum_{t=1}^T (\log |H_t| + \varepsilon_t' H_t^{-1} \varepsilon_t) \quad (5.6)$$

式中  $\varepsilon_t = H_t^{1/2} u_t = y_t - \mu_t$ 。根据协方差矩阵  $H_t = D_t R_t D_t'$  的分解, 可以展开如下:

$$\begin{aligned} L &= -\frac{1}{2} \sum_{t=1}^T (\log |D_t R_t D_t'| + \varepsilon_t' D_t^{-1} R_t^{-1} D_t^{-1} \varepsilon_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (\log |D_t|^2 + u_t' u_t) - \frac{1}{2} \sum_{t=1}^T (\log |R_t| + u_t' R_t^{-1} u_t - u_t' u_t) \end{aligned}$$

其中  $u_t = D_t^{-1} \varepsilon_t = \varepsilon_t / \sigma_t$  为标准化误差,  $\sigma_t$  是  $D_t$  的对角向量。有关方差部分的对数似然函数为:

$$L_v = -\frac{1}{2} \sum_{t=1}^T (\log |D_t|^2 + u_t' u_t) = -\frac{1}{2} \sum_{t=1}^T \sum_{i=1}^n [(\log \sigma_{it}^2 + u_{it}^2)] \quad (5.7)$$

有关相关性部分的对数似然函数 (不含常数项) 为:

$$L_c = -\frac{1}{2} \sum_{t=1}^T (\log |R_t| + u_t' R_t^{-1} u_t) \quad (5.8)$$

对以上似然函数, 为避免高维相关系数矩阵  $R_t$  或逆矩阵  $R_t^{-1}$  的直接计算, 可将其按上节介绍的 Cholesky 分解方法将其逐步分离, 每个步骤只须对二维矩阵运算。为使清晰起见, 以下将专注每个  $t$  时点相关性部分的似然函数的分解, 因而省略使用  $t$  下标的描述。

因为  $R = KK'$ , 所以  $R^{-1} = (K^{-1})' K^{-1} = (K_1^{-1})' (K_2^{-1})' \cdots (K_{n-1}^{-1})' K_{n-1}^{-1} \cdots K_2^{-1} K_1^{-1}$ 。

$R^{-1}$  可以依次分解 (或对角单位化):

设  $R_1^{-1} = R^{-1}$

$$R_1^{-1} = (K_1^{-1})' R_2^{-1} K_1^{-1}$$

$$R_2^{-1} = (K_2^{-1})' R_3^{-1} K_2^{-1}$$

$$R_3^{-1} = (K_3^{-1})' R_4^{-1} K_3^{-1}$$

...

$$R_{n-1}^{-1} = (K_{n-1}^{-1})' R_n^{-1} K_{n-1}^{-1}$$

$$R_n^{-1} = I$$

令每个时点  $t$  有关相关性部份的对数似然函数为:

$$\begin{aligned}
 L_1 &= -\frac{1}{2}(\log|R| + u'R^{-1}u) \\
 &= -\frac{1}{2}(\log|K_1R_2K_1'| + u'(K_1^{-1})'R_2^{-1}K_1^{-1}u) \\
 &= -\frac{1}{2}\sum_{j>1}^n(\log|\Omega_{1j}| + u'_{[1,j]}\Omega_{1j}^{-1}u_{[1,j]}) \\
 &\quad -\frac{1}{2}[\log|R_2| + (u^{(1)})'R_2^{-1}u^{(1)} - \sum_{j>1}^n u'_{[1,j]}\Omega_{1j}^{-1}u_{[1,j]}]
 \end{aligned} \tag{5.9}$$

其中  $u^{(1)} = K_1^{-1}u$  为剔除第一个变量影响（一阶偏相关）之后标准化误差；对每个  $j$  大于 1（ $j > 1$ ）的变量， $\Omega_{1j} = \begin{bmatrix} 1 & \rho_{1j} \\ \rho_{1j} & 1 \end{bmatrix}$  是为  $u_{[1,j]}$  的二维方差协方差矩阵。

定义  $L_{c,1} = -\frac{1}{2}\sum_{j>1}^n(\log|\Omega_{1j}| + u'_{[1,j]}\Omega_{1j}^{-1}u_{[1,j]})$ 。  $L_{c,1}$  为第一个及第  $j$  个变量（ $j > 1$ ）的总和对数似然函数（不含常数项）。继续剔除第二个变量影响（二阶偏相关）之后的对数似然函数为：

$$\begin{aligned}
 L_2 &= -\frac{1}{2}(\log|R_2| + (u^{(1)})'R_2^{-1}u^{(1)}) \\
 &= -\frac{1}{2}(\log|K_2R_3K_2'| + (u^{(1)})'(K_2^{-1})'R_3^{-1}K_2^{-1}u^{(1)}) \\
 &= -\frac{1}{2}\sum_{j>2}^n(\log|\Omega_{2j}| + u^{(1)'}_{[2,j]}\Omega_{2j}^{-1}u^{(1)}_{[2,j]}) \\
 &\quad -\frac{1}{2}[\log|R_3| + (u^{(2)})'R_3^{-1}u^{(2)} - \sum_{j>2}^n u^{(1)'}_{[2,j]}\Omega_{2j}^{-1}u^{(1)}_{[2,j]}]
 \end{aligned} \tag{5.10}$$

其中  $u^{(2)} = K_2^{-1}u^{(1)}$  为剔除第一及第二个变量影响（二阶偏相关）之后的标准化误差；对每个  $j$  大于 2（ $j > 2$ ）的变量， $\Omega_{2j} = \begin{bmatrix} 1 & \rho_{2,j|1} \\ \rho_{2,j|1} & 1 \end{bmatrix}$  为  $u^{(1)}_{[2,j]}$  的二维方差协方差矩阵。

定义  $L_{c,2} = -\frac{1}{2}\sum_{j>2}^n(\log|\Omega_{2j}| + u^{(1)'}_{[2,j]}\Omega_{2j}^{-1}u^{(1)}_{[2,j]})$  为第二个及第  $j$  个变量（ $j > 2$ ）的总和对数似然函数（不含常数项）。

...

最后

$$\begin{aligned}
 L_{n-1} &= -\frac{1}{2}(\log|R_{n-1}| + (u^{(n-2)})' R_{n-1}^{-1} u^{(n-2)}) \\
 &= -\frac{1}{2}(\log|K_{n-1} K_{n-1}'| + (u^{(n-2)})' (K_{n-1}^{-1})' K_{n-1}^{-1} u^{(n-2)}) \\
 &= -\frac{1}{2}(\log|\Omega_{n-1,n}| + u_{[n-1,n]}^{(n-2)'} \Omega_{n-1,n}^{-1} u_{[n-1,n]}^{(n-2)}) \\
 &\quad - \frac{1}{2}[\log|I| + (u^{(n-1)})' u^{(n-1)} - u_{[n-1,n]}^{(n-2)'} \Omega_{n-1,n}^{-1} u_{[n-1,n]}^{(n-2)}]
 \end{aligned} \tag{5.11}$$

其中  $u^{(n-1)} = K_{n-1}^{-1} u^{(n-2)}$ ,

$$\Omega_{n-1,n} = \begin{bmatrix} 1 & \rho_{n-1,n|n-2} \\ \rho_{n-1,n|n-2} & 1 \end{bmatrix}^{\circ}$$

定义  $L_{c,n-1} = -\frac{1}{2}(\log|\Omega_{n-1,n}| + u_{[n-1,n]}^{(n-2)'} \Omega_{n-1,n}^{-1} u_{[n-1,n]}^{(n-2)})$  为第  $n-1$  个及第  $n$  个变量的对数似然函数 (不含常数项)。

最后清除从第一个到及第  $n-1$  个变量影响 ( $n-1$  阶偏相关) 的似然函数, 形成一个常数项:

$$L_n = -\frac{1}{2}(\log|I| + (u^{(n-1)})' u^{(n-1)}) = -\frac{1}{2}(u^{(n-1)})' u^{(n-1)} \tag{5.12}$$

因此模型参数的估计是按序列作成  $n(n-1)/2$  的偏相关系数, 再经由  $K = K_1 \cdots K_{n-1}$  及  $R = KK'$  矩阵分解关系, 最后完成相关系数矩阵的计算, 这个过程是对每一时点  $t$  完成, 因而达到时变的结果。

总结以上所述, 有关相关系数矩阵部分的对数似然函数可以写成:

$$L_c = \sum_t \sum_{i=1}^{n-1} L_{c,i,t} = -\frac{1}{2} \sum_t \sum_{i=1}^{n-1} \sum_{j>i}^n (\log|\Omega_{ij,t}| + u_{[i,j],t}^{(i-1)'} \Omega_{ij,t}^{-1} u_{[i,j],t}^{(i-1)}) + \text{常数项} \tag{5.13}$$

其中  $u_t^{(i)} = K_{ii}^{-1} u_t^{(i-1)}$ ; 对每个  $j$  大于  $i$  ( $j > i$ ) 的变量, 有

$$\Omega_{ij,t} = \begin{bmatrix} 1 & \rho_{ij|i-1,t} \\ \rho_{ij|i-1,t} & 1 \end{bmatrix}^{\circ}$$

## 5.5 模型的估计

对以上似然函数的估计, 引入有关参数向量  $\theta = (\theta_{\mu}', \theta_{\sigma}', \theta_{\rho}')'$ , 其中  $\theta_{\mu}$  是有关

均值部分的参数向量,  $\theta_\sigma$  是有关方差部分的参数向量,  $\theta_\rho$  是有关相关部分的参数向量。写成:

$$L(\theta) = L_v(\theta_\mu, \theta_\sigma) + L_c(\theta_\rho | \theta_\mu, \theta_\sigma)$$

由 (5.7) 式,  $L_v(\theta_\mu, \theta_\sigma) = -\frac{1}{2} \sum_t \sum_i [\log \sigma_{it}^2(\theta_\sigma) + u_{it}^2(\theta_\mu, \theta_\sigma)]$ ,

其中  $u_{it}(\theta_\mu, \theta_\sigma) = \varepsilon_{it} / \sigma_{it} = (y_{it} - \mu_{it}(\theta_\mu)) / \sigma_{it}(\theta_\sigma)$ 。

对每一个变量  $i$ , 似然函数的第一部分可以写成:

$$L_{v,i}(\theta_\mu, \theta_\sigma) = -\frac{1}{2} \sum_t \left[ \log \sigma_{it}^2(\theta_\sigma) + \left( \frac{y_{it} - \mu_{it}(\theta_\mu)}{\sigma_{it}(\theta_\sigma)} \right)^2 \right] \quad (5.14)$$

为描述时变方差函数形式, 取相对简单而普及的 GARCH(1,1)模型为例,

$$\sigma_{it}^2 = \alpha_0 + \alpha_1 \varepsilon_{it-1}^2 + \alpha_2 \sigma_{it-1}^2$$

可知  $\theta_\sigma = (\alpha_0, \alpha_1, \alpha_2)'$ ,  $\alpha_0 > 0, \alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_1 + \alpha_2 < 1$ 。

其他许多拓展 GARCH 模型如考虑不对称性、结构变化及加入外生影响变量等, 都是可行的。

至于似然函数的第二部分 (5.13), 是在完成第一部分参数向量  $(\theta'_\mu, \theta'_\sigma)'$  估计之后进行的, 即在条件  $(\theta_\mu, \theta_\sigma)$  之下的似然函数 (不含常数项),

$$L_c(\theta_\rho | \theta_\mu, \theta_\sigma) = -\frac{1}{2} \sum_{t=1}^T \sum_{i=1}^{n-1} \sum_{j>i}^n (\log |\Omega_{ij,t}| + u_{[i,j],t}^{(i-1)} \Omega_{ij,t}^{-1} u_{[i,j],t}^{(i-1)}) \quad (5.15)$$

其中偏相关系数向量  $\theta_\rho = (\rho'_1, \rho'_{2|1}, \dots, \rho'_{n-1|n-2})'$ ,

$$\begin{aligned} \rho_1 &= (\rho_{12}, \rho_{13}, \dots, \rho_{1n})', \\ \rho_{2|1} &= (\rho_{23|1}, \rho_{24|1}, \dots, \rho_{2n|1})', \\ &\dots \\ \rho_{n-1|n-2} &= \rho_{n-1,n|n-2}. \end{aligned}$$

因为其可加特性, 可以按序列取两个变量, 即是对每一个变量  $i$  及其对应变量  $j$  ( $j > i$ ), 求解以下各别  $(i, j)$  的条件似然函数,



$$L_{c,ij}(\rho_{ij|i-1} | \rho_1, \rho_{2|1}, \dots, \rho_{i-1|i-2}, \theta_\mu, \theta_\sigma) = -\frac{1}{2} \sum_t (\log |\Omega_{ij,t}| + u_{[i,j],t}^{(i-1)'} \Omega_{ij,t}^{-1} u_{[i,j],t}^{(i-1)})$$

其中  $u_{[i,j],t}^{(i-1)}$  及  $\Omega_{ij,t}^{-1} = \text{Var}(u_{[i,j],t}^{(i-1)})^{-1}$  均为条件于  $(\rho_1, \rho_{2|1}, \dots, \rho_{i-1|i-2}, \theta_\mu, \theta_\sigma)$  的  $\rho_{ij|i-1}$  的函数。

两个变量  $i$  及  $j$  的时变相关系数函数，可以根据 Engle(2002) 的 DCC 模型形式写成：

$$Q_{ij,t} = (1 - \delta_1 - \delta_2) \bar{Q}_{ij} + \delta_1 u_{[i,j],t}^{(i-1)} u_{[i,j],t}^{(i-1)'} + \delta_2 Q_{ij,t-1}$$

使用  $Q_{ij,t}$  是为了得到正确的相关矩阵所用到的一个正定矩阵， $\bar{Q}_{ij}$  是标准残差变量  $u_{[i,j],t}^{(i-1)}$  的无条件方差协方差矩阵。 $\delta_1 > 0, \delta_2 > 0$  是待估参数,  $\delta_1 + \delta_2 < 1$ 。理论上  $Q_{ij,t} = \Omega_{ij,t}$  是两个变量  $i$  及  $j$  的二维条件协方差矩阵。实证应用上因为抽样及计算误差则是稍有出入。因此：

$$\begin{bmatrix} 1 & \rho_{ij|i-1,t} \\ \rho_{ij|i-1,t} & 1 \end{bmatrix} = Q_{ij,t}^{*-1} Q_{ij,t} Q_{ij,t}^{*-1}$$

$Q_{ij,t}^*$  是一个对角线上的元素等于  $Q_{ij,t}$  的对角线元素平方根的对角矩阵。

因此相关系数的估计是在序列中取两个变量作成个别对数似然函数，求解相对应的偏相关系数。每个步骤只是一个简单二维的条件优化过程，最后将所估计得的各阶偏相关系数代入各个  $K_{1t}, \dots, K_{n-1,t}$  矩阵，进而得到相关系数矩阵：

$$R_t = K_t K_t' = K_{1t} \dots K_{n-1,t} K_{n-1,t}' \dots K_{1t}'$$

值得一提的是，SCC 的方法并不限于使用 DCC 形式来描述动态相关。只要能保证相关系数值落在 -1 及 +1 区间的其他形式，如 Fisher 转换，或更复杂的向量 DCC 及 BEKK 等二维模型均为可行的选择，这体现了 SCC 方法的灵活性。

以上使用的似然函数是来源于变量为正态分布的假设，这在许多金融计量模型中是有争议的。文献之中（如 Bollerslev, 1986<sup>[11]</sup>）提到使用 QML 的估计方法，是正态分布的假设不成立时，最大似然函数值的求解方法。有关 QML 一致性及渐近性等极限理论可参考 White(1996)<sup>[98]</sup> 的证明，在多维 GARCH 模型的证明可见 Engle(2001)<sup>[52]</sup> 及 Comte 和 Lieberman (2003)<sup>[97]</sup> 的文献。由

Engle(2002)<sup>[30]</sup>提出的对 DCC 模型两阶段的 QML 的估计方法, 本文延伸及拓展为多步骤的 SCC 方法, 只要 SCC 的两两序列的似然函数最大, 所得到的整个 SCC 模型的似然函数 (5.13) 式最大, 是证明 SCC 最大似然估计的渐近性等极限定理的关键所在。

## 5. 6 SCC 模型估计的渐近性定理

SCC 模型的对数似然函数可以写为:

$$L = L_v + L_c \quad (1^*)$$

$$L_1 = L_c = L_{c1} + L_2 \quad (2^*)$$

.....

$$L_{n-2} = L_{c,n-2} + L_{n-1} \quad (n-1^*)$$

$$L_{n-1} = L_{c,n-1} + L_n \quad (n^*)$$

得到:

$$L = L_v + L_{c1} + L_{c2} + \dots + L_{c,n-2} + L_{c,n-1} + L_n \quad (5.16)$$

**定理 5-9 (渐近正态性)** 在 White(1996)<sup>[98]</sup>提出的定理 5-8 中的 A1-A6 的假设条件下, 有

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, C_1^{-1} C_0 C_1^{-1}), \quad n \rightarrow \infty$$

其中,  $C_0 = \text{Var}[\sum_{t=1}^T \{T^{-1/2} \nabla_{\phi}' \ln f_1(r_t, \phi_0), T^{-1/2} \nabla_{\psi}' \ln f_2(r_t, \phi_0, \psi_0)\}]$ ,

$$C_1 = \begin{bmatrix} \nabla_{\phi\phi} \ln f_1(\phi_0) & 0 \\ \nabla_{\phi\psi} \ln f_2(\theta_0) & \nabla_{\psi\psi} \ln f_2(\theta_0) \end{bmatrix}^{\circ}$$

**证明:** 对于等式 (1<sup>\*</sup>), 似然函数  $L$  可以允许两阶段估计, 第一个阶段对各个市场的收益序列估计单维 GARCH 模型, 第二阶段利用第一阶段估计结果求出的标准残差向量对相关矩阵进行估计。令模型的参数向量为  $\theta$ , 分为两部分  $\theta = (\phi, \psi)$ , 其中  $\phi$  代表所有各个市场收益的 GARCH 模型的参数, 那么第一阶段的似然函数为  $L_v(\phi | y_t)$ , 其实就是对于每个资产收益的单个 GARCH 的似然

函数的和。一旦第一阶段的参数估计出来，我们用第一阶段估计出来的结果，对第二阶段的似然函数进行估计，第二阶段的似然函数为  $L_c(\phi|\psi, y_t)$ ，此时要求最大化  $L_c(\phi|\hat{\psi}, y_t)$ ，即  $\max L_c(\phi|\psi, y_t)$ 。我们利用 QMLE 的两阶段渐近定理，可以得到此对数似然函数 QMLE 的渐近正态性。

同理，对于等式 (2\*)，似然函数  $L_1$  可以允许两阶段估计，第一个阶段是对标准残差向量  $u_t$  中第一项分别与其他各项的相关性估计似然函数的和，第二阶段是，剔除第一个变量影响（一阶偏相关）之后标准化误差  $u^{(1)}$  的相关矩阵的估计。令模型的参数向量为  $\phi$ ，分为两部分  $\phi = (\phi_1, \phi_{-1})$ ，其中  $\phi_1$  是对标准残差向量  $u_t$  中第一项分别与其他各项的相关性估计中的参数， $\phi_{-1}$  是  $\phi$  中除去  $\phi_1$  中的参数后剩余的参数。那么第一阶段的似然函数为  $L_{c1}(\phi_1|\hat{\psi}, y_t)$ ，我们利用第一阶段估计出来的结果，对第二阶段的似然函数进行估计，第二阶段的似然函数为  $L_2(\phi_{-1}|\phi_1, \hat{\psi}, y_t)$ 。那么利用 White(1996)<sup>[98]</sup> 的两阶段 QMLE 的估计的渐近分布的证明，参数估计的渐近正态性成立。同理对于等式 (3\*) 我们利用 QMLE 的两阶段渐近定理，可以得到此对数似然函数 QMLE 的渐近正态性。

.....

依此类推，对于等式 (n\*)，似然函数  $L_{n-1}$  可以允许两阶段估计，第一个阶段是为第  $n-1$  个及第  $n$  个变量的对数似然函数（不含常数项）；第二阶段是，清除从第一个到及第  $n-1$  个变量影响（ $n-1$  阶偏相关）的似然函数，形成一个常数项。令模型的参数向量为  $\phi_{-(n-2)}$ ，分为两部分  $\phi_{-(n-2)} = (\phi_{n-1}, \phi_{-(n-1)})$ ，其中  $\phi_{n-1}$  是为第  $n-1$  个及第  $n$  个变量的对数似然函数（不含常数项）的参数， $\phi_{-(n-1)}$  是  $\phi_{-(n-2)}$  中除去  $\phi_{n-1}$  中的参数后剩余的参数。那么第一阶段的似然函数为  $L_{c,n-1}(\phi_{n-1}|\hat{\phi}_{-(n-2)}, \hat{\phi}_{(n-2)}, \dots, \hat{\phi}_{-1}, \hat{\phi}_1, \hat{\psi}, y_t)$ ，接着对第二阶段似然函数进行估计，第二阶段的似然函数为  $L_n(\phi_{-(n-1)}|\hat{\phi}_{n-1}, \hat{\phi}_{-(n-2)}, \hat{\phi}_{(n-2)}, \dots, \hat{\phi}_{-1}, \hat{\phi}_1, \hat{\psi}, y_t)$ ，那么利用 White(1994) 的两阶段 QMLE 的估计的渐近分布的证明，此似然函数的 QMLE

参数估计的渐近正态性成立。

## 5. 7 实证分析

本文估计 SCC 的样本数据从从 1992 年 12 月 28 日至 2004 年 12 月 31 日上证 A 股(sha), 上证 B 股(shb), 深综 A 指(sza), 深综 B 指(szb)每日的收盘价, 取对数收益:

$$y_t = 100 \times (\ln P_t - \ln(P_{t-1}))$$

其数据的描述统计表于表 5.1:

表 5.1 数据统计量

$y_t$	均值	标准差	偏度	峰度	Jarque-Bera
sha	0.003592	2.380739	1.488071	27.29316	68811.84
sza	0.015034	2.332258	1.225764	24.11885	51925.33
shb	-0.009563	2.234408	0.346519	7.862732	2771.528
szb	0.016808	2.311948	0.365589	10.00474	5697.915

由于中国股指市场在数据样本的取值区间经历了国际形势和国内政策变化的影响, 因此本文设置了两个虚拟变量。  $x_1$  代表了 1997 年亚洲金融危机对于中国主要股指市场条件方差的影响得虚拟变量,  $x_2$  代表了 2001 年 2 月 16 日公布 B 股市场向国内购买者开放对于中国主要股指市场条件方差的影响的虚拟变量, 当然我们也选取了代表其他重要政策变动的虚拟变量, 但是结果并不显著, 没有在此列出。单个市场的波动率估计等式为:

$$\sigma_{i,t}^2 = \alpha_0 + (\alpha_1 + \phi d_{t-1}) \varepsilon_{i,t-1}^2 + \alpha_2 \sigma_{i,t-1}^2 + c_1 x_{1,t} + c_2 x_{2,t}$$

其中  $d_t = \begin{cases} 1 & \varepsilon_t < 0 \\ 0 & \text{其他} \end{cases}$ ,  $x_1 = \begin{cases} 1 & \text{1997年7月1日之后} \\ 0 & \text{其他} \end{cases}$ , 和  $x_2 = \begin{cases} 1 & \text{2001年2月16日之后} \\ 0 & \text{其他} \end{cases}$ 。

$\sigma_{i,t}^2$  是第  $i$  个市场的条件方差,  $\varepsilon_{i,t}$  是第  $i$  个市场的残差,  $d_t$  是反映市场的非对称因素。

估计结果如表 5.2 所示：

表 5.2 四个市场波动率参数估计

	$\alpha_0$	$\alpha_1$	$\phi$	$\alpha_2$	$c_1$	$c_2$
sha	1.0789 (2.2094)	0.095568 (2.7815)	0.12714 (3.0272)	0.77211 (13.211)	-0.93762 (-2.0972)	
sza	1.0053 (2.4085)	0.11977 (3.0924)	0.12363 (2.7791)	0.75251 (13.436)	-0.84913 (-2.2138)	
shb	0.18785 (2.9400)	0.22413 (5.4543)		0.72279 (15.644)	0.43331 (3.1364)	-0.33707 (-2.6087)
szb	0.18460 (2.6301)	0.25954 (4.2074)		0.68871 (10.096)	0.44639 (2.4667)	-0.23138 (-1.5325)

注：如无特别说明，表中数字代表参数估计值，下面括弧中的数字代表 t 统计量。

四个股指市场条件波动率如图 5.1 所示：

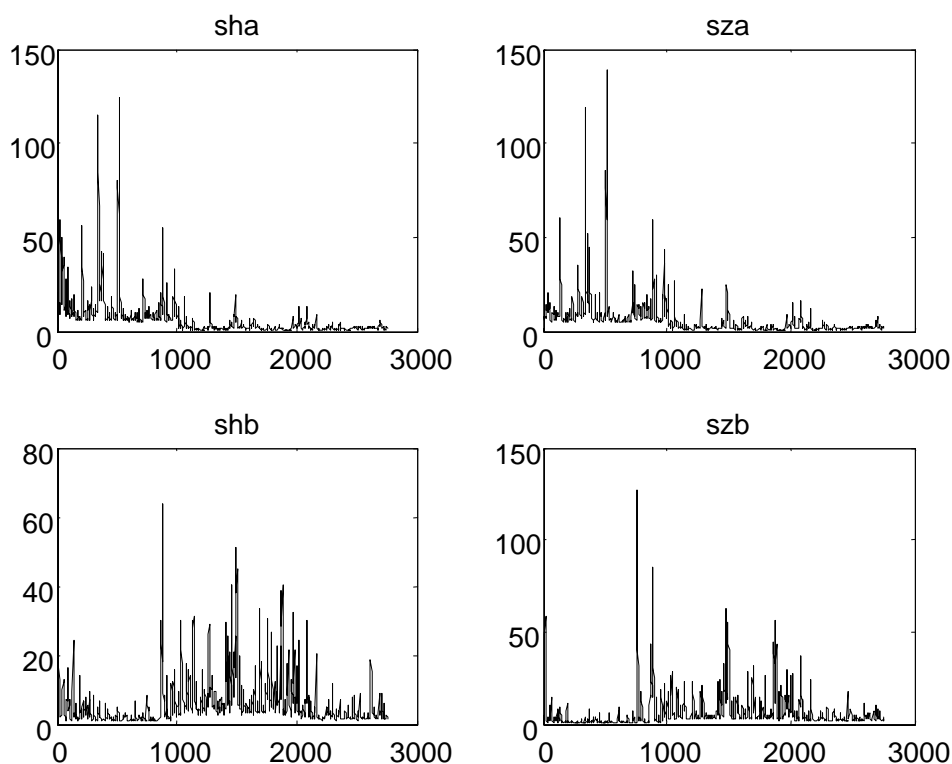


图 5.1 四个市场波动率

针对四个市场收益相关性的计算，具体如下：首先对  $\varepsilon_t$  进行标准化，得到  $u_{it} = \frac{\varepsilon_{i,t}}{\hat{\sigma}_{it}}$ ,  $i=1,2,3,4$ ，用我们所熟知的二维 GARCH 模型分别估计  $u_t$  中的第 1 项与第 2, 3 和 4 项的动态条件相关系数，构成  $K_{1t}$ ，从第 2, 3 和 4 项中分别消除第 1 项的影响，即数据再次标准化，得到  $K_{1t}^{-1}u_t$ ；然后用二维 GARCH 模型分别估计  $K_{1t}^{-1}u_t$  中的第 2 项与第 3 和第 4 项的动态条件相关系数，构成  $K_{2t}$ ，同理从第 3 和第 4 项中分别消除第 2 项的影响，即得到  $K_{2t}^{-1}K_{1t}^{-1}u_t$ ；最后，用二维 GARCH 模型估计  $K_{2t}^{-1}K_{1t}^{-1}u_t$  中的第 3 项与第 4 项的动态条件相关系数，构成  $K_{3t}$ ，同样地从第 4 项中消除第 3 项的影响，得到  $K_{3t}^{-1}K_{2t}^{-1}K_{1t}^{-1}u_t$ 。

用上述方法构造的  $K$  矩阵，即可得到四个市场之间的动态相关矩阵。至于对于两两市场间的动态条件相关系数的估计，我们可以采取二维 GARCH 模型动态条件相关矩阵的各种方法，然后用本文介绍的 SCC 方法得到相应的高维市场间的动态条件相关系数矩阵：

$$R_t = K_{1t} \cdots K_{n-1,t} K'_{n-1,t} \cdots K'_{1t}$$

本文用来估计两两动态条件相关的模型分别采用了 DCC 和 ADCC 模型，他们分别为： $Q_{ij,t} = (1 - \delta_1 - \delta_2)Q_{ij}^0 + \delta_1 e_{i,t-1} e_{j,t-1} + \delta_2 Q_{ij,t-1}$

$$Q_{ij,t} = (1 - \delta_1 - \delta_2 - \delta_3)Q_{ij}^0 + (\delta_1 + \delta_3 d_{t-1}) e_{i,t-1} e_{j,t-1} + \delta_2 Q_{ij,t-1}, \quad d_t = \begin{cases} 1 & e_{i,t} < 0, e_{j,t} < 0 \\ 0 & \text{其他} \end{cases}$$

以下列出了分别用 CCC、DCC、ADCC 以及 DCC、ADCC 和 SCC 方法相结合得到的中国四个主要股指市场之间的动态条件相关系数参数估计。估计结果请参照表 5.3，表 5.4 和表 5.5，同时列出了 ADCC 和 SCC 相结合得到的动态条件相关系数图，见图 5.2：

表 5.3 DCC 和 DCC+SCC 模型参数估计

	$\delta_1$	$\delta_2$
DCC	0.18204 (7.6643)	0.98209 (197.24)

DCC+SCC		
$\rho_{12}$	0.081558 (3.9930)	0.91126 (40.390)
$\rho_{13}$	0.045182 (3.5581)	0.95008 (63.523)
$\rho_{14}$	0.032434 (1.8006)	0.96511 (46.578)
$\rho_{23}(1)$	0.026845 (1.5400)	-0.43048 (-1.2790)
$\rho_{24}(1)$	0.023465 (1.8427)	-0.14907 (-0.20960)
$\rho_{34}(2)$	0.018280 (1.9481)	0.98012 (90.024)

表 5.4 ADCC 和 ADCC+SCC 模型参数估计

	$\delta_1$	$\delta_2$	非对称参数 $\delta_3$
ADCC	0.032501 (24.770)	0.96382 (738.58)	0.0036689 (2.0704)
ADCC+SCC			
$\rho_{12}$	0.28634 (8.3570)	0.95554 (86.249)	-2.2184e-006 (-0.95145)
$\rho_{13}$	0.20261 (6.6489)	0.97747 (134.59)	1.7350e-006 (0.21772)
$\rho_{14}$	0.18939 (4.5522)	0.98049 (105.58)	-4.2890e-009 (-0.00026089)
$\rho_{23}(1)$	0.15106 (2.2510)	3.9850e-005 (0.30894)	0.17111 (1.7825)

$\rho_{24}(1)$	0.13534 (1.1308)	0.66517 (0.13445)	0.071260 (0.053454)
$\rho_{34}(2)$	0.15397 (4.3380)	0.98714 (149.76)	3.2702e-007 (0.21913)

表 5.5 CCC 的参数估计

CCC	sha	sza	shb	szb
sha	1			
sza	0.88825 (295.72)	1		
shb	0.47596 (36.068)	0.46710 (35.055)	1	
szb	0.42699 (30.387)	0.44329 (32.062)	0.61132 (60.224)	1

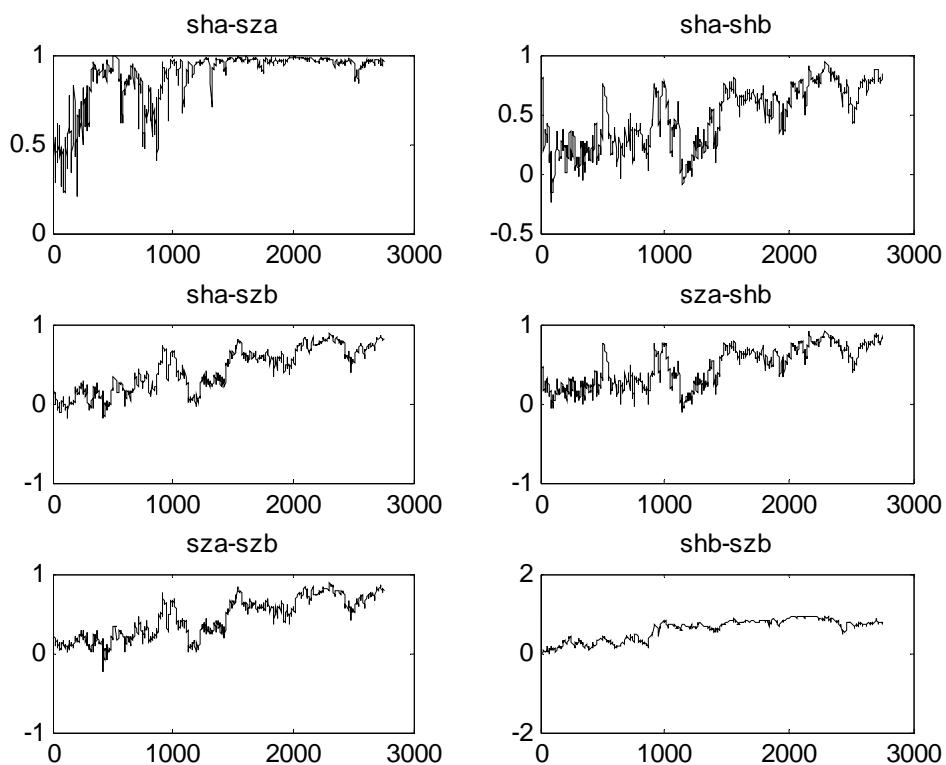


图 5.2 动态条件相关图



在对数据建模<sup>①</sup>的过程中，一个重要的步骤是用所采取的模型进行预测，对预测结果的评价和比较就显得相当重要和不可缺少。本文采用样本外预测，总样本从 1992 年 12 月 28 日至 2005 年 3 月 9 日每天上证 A 股，上证 B 股，深综 A 指，深综 B 指的收盘价。从 1992 年 12 月 28 日至 2004 年 12 月 31 日的样本为估计样本，从 2005 年 1 月 4 日至 3 月 9 日为预测样本。

评价多维 GARCH 的设定，由于相关矩阵是不可观测的，一个最自然的方法就是把模型的估计值与相对应的已实现值对比。首先利用真实的收益率向量构造在  $K$  时期已实现的相关性矩阵  $\hat{Q}_t^{(K)}$ ：

$$\hat{Q}_t^{(K)} = \frac{1}{K} \sum_{k=1}^K (y_{t-1+k} - \bar{y})(y_{t-1+k} - \bar{y})' \quad \forall t = 1, \dots, T - K + 1,$$

其中， $y_t$  是  $t$  时刻四个市场收益率向量， $\bar{y}$  是  $t$  时刻四个市场收益率的平均值。

$K$  时期模型平均相关性矩阵的预测  $\hat{R}_t$  为：

$$\hat{R}_t^{(K)} = \frac{1}{K} \sum_{k=1}^K (\hat{R}_{t-1+k} | I_{t-1}) \quad \forall t = 1, \dots, T - K + 1,$$

其中， $\hat{R}_t$  是  $t$  时刻的相关性矩阵的预测。

对于已实现的和预测的相关性矩阵的差别可以通过 MAD 和 MSE 来计算。

$$MAD = \sum_{r=1}^n \sum_{c=r}^n |\hat{Q}_{r,c,t}^{(K)} - \hat{R}_{r,c,t}^{(K)}|, \quad n = 4$$

$$MSE = \sum_{r=1}^n \sum_{c=r}^n (\hat{Q}_{r,c,t}^{(K)} - \hat{R}_{r,c,t}^{(K)})^2, \quad n = 4$$

本文进行的预测分为两步，首先分别对单个市场收益的波动进行预测，采取了一步向前预测，即滚动预测和迭代预测相结合的方法，具体做法是在估计样本的基础上，依次加入一个预测样本进行估计，根据估计出来的参数值，迭代得出下一个预测值。第二步根据预测得到的方差，进行相关性预测，采取的方法也是一步向前预测方法。最后，根据条件方差的不同分解，把方差矩阵和相关性矩阵进行组合，得到  $H_t$ 。

<sup>①</sup> 本文所用到的程序都采用 GAUSS 6.0 编写。

所得到的预测结果的比较如表 5.6，表 5.7 所示：

表 5.6 预测评价

模型对比	K=5		K=10	
	MAD	MSE	MAD	MSE
CCC	0.20454	0.96567	0.13594	0.77373
Riskmetrics <sup>①</sup>	0.18648256	0.86356634	0.16780130	0.83506120
DCC	0.10532303	0.57187963	0.15222087	0.75537384
ADCC	0.096554482	0.53319963	0.14015389	0.71725884
DCC+SCC	0.087454937	0.50785246	0.11141036	0.61813584
ADCC+SCC	0.087256121	0.51004046	0.11122442	0.61641484

表 5.7 似然函数值

	CCC	DCC	ADCC	DCC+SCC	ADCC+SCC
对数似然值	-2284.2	4736.0	4736.1	4798.5	4799.6

表 5.6 列出了用不同的方法对于中国主要股指市场进行预测得到的 MAD 和 MSE 值，表 5.7 列出了利用各个模型估计所得到的对数似然函数值，通过对比，我们可以看出，对于预测时期 K=5,10 时，SCC 方法与 DCC，ADCC 相结合得到的 MAD,MSE 最小，因此预测结果最好；对于模型估计的对数似然函数值，SCC 方法分别与 DCC，ADCC 结合得到的对数似然函数值比其他的模型要大，说明了 SCC 方法拟合中国主要股指市场的相关性是最为准确和有效的。

<sup>①</sup>模型 Riskmetrics 方法方差协方差公式单步向前预测如下：

$$\sigma_{ii,t}^2 = \lambda\sigma_{ii,t-1}^2 + (1-\lambda)x_{i,t}^2, \quad \sigma_{ij,t}^2 = \lambda\sigma_{ij,t-1}^2 + (1-\lambda)x_{i,t}x_{j,t}$$

$$\text{相关系数为 } Q_{ij,t} = \frac{\sigma_{ij,t}}{\sqrt{\sigma_{ii,t}}\sqrt{\sigma_{jj,t}}}, \quad \lambda = 0.94。$$

## 6. 结论及今后研究的方向

### 6.1 总结

金融风险的规避与防范一直是金融理论与实践的重要研究课题。而衡量投资组合或某项资产的风险离不开对资产收益的波动性和资产收益之间相关性的估计与对未来的预测。同时，金融风险及动态投资组合方法对长期投资分析和实践有着重要的意义。但只有解决了一般多维 GARCH 模型和多维 SV 模型的建模及有关波动率的估计问题，对于动态投资组合的研究与实践才能取得实质性的进展。

本文对单维波动性和多维的方差协方差矩阵的预测方法以及在风险管理上的应用进行了一个比较全面、系统的研究，并且针对原有的多维 GARCH 模型或多维 SV 模型参数过多导致的估计困难，作者根据 Cholesky 分解方法提出了一种处理高维相关性模型的 SCC(Sequential Conditional Correlation)方法，并给出了模型 MLE 估计渐近性质理论上的证明。值得一提的是，SCC 的方法具有很强的灵活性，可以与各种估计二维相关性的模型结合起来，解决高维相关性的问题，是一种新方法的探索。

### 6.2 未来的方向

前面对金融资产动态相关性方法和应用进行了研究，那么接下来将对近年来被广泛关注的一些新课题进行总结。

#### 6.2.1 理论方面

由于目前的研究能力以及文献限制，在研究过程中还有许多问题需要进一步的完善。这些工作包括：

1. 多维 GARCH 模型的统计推断和诊断性检验，即在实际中，对于具有不同统计特征的数据，如何通过一定的金融计量方法选择某一种多维 GARCH 模型来模拟收益向量的动态相关性；
2. 提高对多维 GARCH 模型统计推断的软件设计；

3. DCC 多维 GARCH 模型两步估计有效性的条件;
4. 不仅通过实证的方法比较各种模型的优劣,而且要通过模拟的方法给出模型的适用环境和对比;
5. SCC 方法的提出就是着眼于解决高维相关性的问题,而文中只用 SCC 方法估计了四个市场的收益的动态相关性,未来的研究应该把 SCC 方法应用于高维相关性的估计和预测上,进一步体现 SCC 方法的优势。
6. SCC 方法具有灵活性,未来应该把 SCC 方法与研究相关性的最新方法结合起来更好地解决实际问题。

### 6. 2. 2 方法的创新

本文主要以多维时间序列模型为主线讨论了相应的模型估计、预测方法以及在风险中的应用等问题。那么接下来将以金融时间序列分析为中心讨论近年来几个被广泛关注的课题。

第一,高频时间序列分析与建模的研究。高频时间序列通常指以小时、分钟甚至秒为频率所采集的金融类数据以及记录每笔交易的时间序列,有时为了区别,后者也叫做超高频时间序列。在金融市场中,信息连续地影响着证券市场价格的变化。高频数据比低频数据包含着更多的信息,如市场微观结构的信息及重要的长期日间现象等信息。因此,高频时间序列分析与建模是一个全新的研究领域,而且具有重要的理论和实践意义,对于金融理论工作者和实际从业人员都非常重要。连续时间金融模型是金融计量分析中的另一重要领域,它为金融产品价格及其波动建立相应的模型,并应用于期权以及其他衍生金融产品的定价等。高频数据为我们更准确地理解金融市场提供了丰富的信息,而高频数据尤其是超高频数据的性质比低频数据更接近于连续时间的情况,所以把高频数据和连续时间模型结合起来研究是一个值得探讨的问题。关于金融高频时间序列的研究,国际上还仅处于起步阶段,有许多问题尚未解决。目前,随着信息技术的发展,国内证券市场的高频数据可以很方便地获得。对于高频数据在金融分析方面的研究,特别对中国金融市场高频特性的研究,相信会得到广泛的开展,对金融市场的微观结构特征也会得到进一步的揭示。

第二, Copula 理论是近年来金融分析中的新方法,这种方法的特点在于它不仅可以有效地描述随机变量之间的相关程度,并且能够反映它们之间的相关模式,同时对于他们的联合分布函数有一个描述。基于 Copula 理论的特点,可以运用条件 Copula 理论<sup>[99]</sup>建立具有时变、偏斜、尖峰、厚尾等分布特性的多变量金融时间序列模型,以处理非对称相关性、投资组合、资产定价等问题。由于维数的限制,其建模方法受到较大的局限性,仅仅限于二维模型,那么是否可以把本文所研究的 SCC 方法把高维的相关性估计和预测分解为二维相关性的估计问题,然后用 Copula 来研究二维相关性的估计,这是下一步研究的重点。

## 参考文献:

- [1] Lien D, Tse YK. Some recent developments in futures hedging [J]. *Journal of Economic surveys*, 2002, 16: 357-396.
- [2] Bollerslev T., Engle RF, Wooldridge JM. A capital asset pricing model with time varying covariances [J]. *Journal of Political Economy*, 1988, 96: 116-131.
- [3] Hafner C, Herwartz H. Time-varying market price of risk in the CAPM: Approaches, empirical evidence and implications [J]. *Finance*, 1998, 19: 93-112.
- [4] Giot P, Laurent S. Value at Risk for long and short positions [J]. *Journal of Applied Econometrics*, 2003, 18: 641-664.
- [5] Fanses PH, van Dijk D. *Nonlinear Series Models in Empirical Finance* [M]. Cambridge University Press: Cambridge, 2000.
- [6] Gouriéroux C. *ARCH Models and Financial Applications* [M]. Springer-Verlag: New York, 1997.
- [7] Bauwens, L., S. Laurent, and J.V.K. Rombouts. Multivariate GARCH Models: A Survey [J]. *Journal of Applied Econometrics* 2006, 21: 79-110.
- [8] Palandri Alessandro. Sequential Conditional Correlations: Inference and Evaluation [Z]. 2005, 10, working paper.
- [9] Campbell.J.Y., A.W.Lo, A.C.MacKinlay. *The Econometrics of Financial Markets* [M]. Princeton.NJ: Princeton University Press, 1997.
- [10] Engle R F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K..Inflation [J]. *Econometrica*, 1982, 50: 987-1008.
- [11] Bollerslev T. Generalized autoregressive conditional heteroskedasticity [J]. *Journal of Economics*, 1986(31): 307-327.
- [12] Bollerslev, T., R. Chou, and K. Kroner. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence [J]. *Journal of Econometrics*, 1992, 52: 5-59.
- [13] Bera, A., and M. Higgins. ARCH Models: Properties, Estimation and Testing [J]. *Journal of Economic Surveys*, 1993, 7: 305-362.
- [14] Bollerslev, T., R. Engle, and D. Nelson. ARCH Models. in *Handbook of Econometrics* [M]

- ed. by R. Engle, and D. McFadden, 1994, chap. 4, pp. 2959–3038. North Holland Press, Amsterdam.
- [15] Pagan, A. The Econometrics of Financial Markets [J]. *Journal of Empirical Finance*, 1996, 3: 15–102.
- [16] Palm, F. GARCH Models of Volatility. in *Handbook of Statistics [M]*, ed. by G. Maddala, and C. Rao, 1996, pp. 209–240. Elsevier Science, Amsterdam.
- [17] He C, Terasvirta T. An application of the analogy between vector ARCH and vector random coefficient autoregressive models [Z]. SSE/EFI working paper Series in Economics and Finance No.516, Stockholm School of Economics, 2002.
- [18] Attanasio O. Risk, time-varying second moments and market efficiency [J]. *Review of Economic Studies*, 1991, 58: 479-494.
- [19] Ding Z, Engle RF. Large scale conditional covariance matrix modeling, estimation and testing [Z]. Working paper FIN-01-029, NYU Stern School of Business, 2001.
- [20] J.P.Morgan. Riskmetrics Technical Document [M], 4th edn. New York, 1996.
- [21] Engle R, Kroner FK. Multivariate simultaneous generalized ARCH [J]. *Econometric Theory*, 1995, 11: 122-150.
- [22] Engle RF, Ng VK, Rothschild M. Asset pricing with a factor-ARCH covariance structure: empirical estimates for treasury bills [J]. *Journal of Econometrica*, 1990, 45: 213-238.
- [23] Bollerslev T, Engle RF. Common persistence in conditional variances [J]. *Econometrica*, 1993, 61: 167-186.
- [24] Lin WL. Alternative estimators for factor GARCH models-a Monte Carlo comparison [J]. *Journal of Applied Econometrics*, 1992, 7: 259-279.
- [25] Nelson DB. Conditional heteroskedasticity in asset returns: a new approach [J]. *Econometrica*, 1991, 59: 349-370.
- [26] Ding Z, Granger C W J, Engle R F. A long memory property of stock market returns and a new model [J]. *Journal of Empirical Finance*, 1993, 1: 83-106.
- [27] Kariya T. MTV model and its application to the prediction of stock prices [M]. In *Proceedings of the Second International Tampere Conference in Statistics*, Pullila T, Puntanen S (eds). University of Tampere, Finland, 1988.
- [28] Alexander CO, Chibumba AM. Multivariate orthogonal factor GARCH [M]. Mimeo,

University of Sussex, 1997.

- [29] Bollerslev T. Modeling the coherence in short-run nominal exchange rates:a multivariate generalized ARCH model [J]. *Review of Economics and Statistics*, 1990, 72: 498-505.
- [30] Engle RF. Dynamic conditional correlation-a simple class of multivariate GARCH models [J]. *Journal of Business and Economic Statistics*, 2002, 20: 339-350.
- [31] Tse YK, Tsui AKC. A multivariate GARCH model with time-varying correlations [J]. *Journal of Business and Economic Statistics*, 2002, 20: 351-362.
- [32] Olivier Ledoit, Pedro Santa-Clara, and Michael Wolf. Flexible Multivariate GARCH modeling with an application to international stock markets [J]. *The review of Economics and Statistics*, 2003, 85, 3: 735-747.
- [33] Sklar A. Fonctions de repartition à n dimension et leurs marges [M]. *Publications de l'Institut Statistique de l'Université de Paris*, 1959, 8: 229-231.
- [34] Nelsen RB. *An Introduction to Copulas* [M]. Springer-Verlag: New York, 1999.
- [35] Patton A. Modelling time-varying exchange rate dependence using the conditional copula [Z]. University of California, San Diego, Discussion Paper 01-09, 2000.
- [36] Jondeau E, Rockinger M. The copula-GARCH model of conditional dependencies: an international stock-market application [J]. *Journal of International Money and Finance*, 2001.
- [37] Black F. Studies of stock market volatility change [J]. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 1976: 177-181.
- [38] Glosten LR, Jagannathan R, Runkle DE. On the relation between expected value and the volatility of the nominal excess return on stocks [J]. *Journal of Finance*, 1993, 48: 1779-1801.
- [39] Zakoian J-M. Threshold heteroskedasticity models [J]. *Journal of Economic Dynamics and Control*, 1994, 15: 931-955.
- [40] Kroner FK, Ng VK. Modelling asymmetric comovements of asset returns [J]. *The Review of Financial Studies*, 1998, 11: 817-844.
- [41] Harvey AC, Ruiz E, Shephard N. Multivariate stochastic variance models [J]. *Review of Economic Studies*, 1994, 61: 247-264.
- [42] Andersen TG, Bollerslev T, Diebold FX, Labys P. Modeling and forecasting realized volatility [J]. *Econometrica*, 2003, 71: 579-625.
- [43] Bollerslev T, Wooldridge JM. Quasi-maximum likelihood estimation and inference in



- dynamic models with time-varying covariances [J]. *Econometric Reviews*, 1992,11:143-172.
- [44] Jeantheau T. Strong consistency of estimation for multivariate ARCH models [J]. *Econometric Theory*, 1998, 14: 70-86.
- [45] Harvey AC, Ruiz E, Shephard N. Unobservable component time series models with ARCH disturbances [J]. *Journal of Econometrics*, 1992, 52: 129-158.
- [46] Fiorentini G, Sentana E, Calzolari G. Maximum likelihood estimation and inference in multivariate conditionally heteroskedastic dynamic regression models with Student t innovation [J]. *Journal of Business and Economic Statistics*, 2003, 21: 532-546.
- [47] Vlaar PJG, Palm FC. The message in weekly exchange rates in the European Monetary System: mean reversion, conditional heteroskedasticity and jumps [J]. *Journal of Business and Economic Statistics*, 1993, 11: 351-360.
- [48] Barndorff-Nielsen OE, Shephard N. Normal modified stable processes [J]. *Theory of Probability and Mathematics Statistics*, 2001, 65: 1-19.
- [49] Bauwens L, Laurent S. A new class of multivariate skew densities, with application to GARCH models [J]. *Journal of Business and Economic Statistics*, 2002.
- [50] Hafner CM, Rombouts JVK. Semiparametric multivariate volatility models [R]. *Econometric Institute Report 21*. Erasmus University Rotterdam, 2004.
- [51] Comte F, Lieberman O. Asymptotic theory for multivariate GARCH processes [J]. *Journal of Multivariate Analysis*, 2003, 84: 61-84.
- [52] Engle RF, Sheppard K. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH [M]. Mimeo, UCSD, 2001.
- [53] Brooks C, Burke SP, Persaud G. Multivariate GARCH models: software choice and estimation issues [J]. *Journal of Applied Econometrics*, 2003, 18: 725-734.
- [54] 樊智, 张世英. 多元 GARCH 建模及其在中国股市分析中的应用 [J]. *管理科学学报*, 2003, 6 (2): 68-73.
- [55] 王传美, 童恒庆. 多维 GARCH 模型的半参数有效估计 [J]. *应用数学*, 18 (2): 260-264.
- [56] Engle R, Gonzalez Rivera G. Semiparametric ARCH models [J]. *Journal of Business and Economic Statistics*, 1991,9(4): 345-359.
- [57] Linton, Oliver. Adaptive estimation in ARCH models [J]. *Econometric Theory*, 1993, 9(4):

539--569

- [58] Feike C. Drost, Chris A. J. Klaassen and Bas J. M. Werker. Adaptive estimation in time-series models [J]. *The Annals of Statistics*, 1997, 25, (2): 786–817
- [59] Dufour J-M, Khalaf L, Beaulieu M-C. Finite-sample diagnostics for multivariate regression with applications to linear asset pricing models [Z]. Working Paper No. 06-2003, CIREQ, 2003
- [60] Hosking JRM. The multivariate portmanteau statistic [J]. *Journal of American Statistical Association*, 1980, 75: 602–608.
- [61] 惠晓峰. 基于时间序列GARCH模型的人民币汇率预测 [J]. *金融研究*, 2003, 5: 99–105.
- [62] 刘国旗. 非线性 GARCH 模型在中国股市波动预测中的应用研究 [J]. *统计研究*, 2000, 1: 49–52.
- [63] Black, F. and M. Scholes. The Pricing of Options and Corporate Liabilities [J]. *Journal of Political Economy*, 1973, 81: 637-654.
- [64] Granger, C.W.J. Prediction with a Generalized Cost of Error Function [J]. *Operational Research Quarterly*, 1969, 20: 199-207.
- [65] Markowitz, H. Portfolio Selection [J]. *Journal of Finance*, 1952, 7: 77-91.
- [66] Sharpe, W. Capital Asset Prices - A Theory of Market Equilibrium Under Conditions of Risk [J]. *Journal of Finance*, 1964, 19: 425-442.
- [67] Hull, J. and A. White. The Pricing of Options on Assets with Stochastic Volatilities [J]. *Journal of Finance*, 1987, 42: 281-300.
- [68] Rich, R. and J. Tracy. Uncertainty and Labor Contract Durations [J]. *Review of Economics and Statistics*, 2004, 86: 270-287.
- [69] Lastrapes, W.D. Exchange Rate Volatility and US Monetary Policy: An ARCH Application [J]. *Journal of Money, Credit and Banking*, 1989, 21: 66-77.
- [70] Ruge Murcia, F.J. The Inflation Bias When the Central Bank Targets the Natural Rate of Unemployment [J]. *European Economic Review*, 2004, 48: 91-107.
- [71] Merton, R.C. Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case [J]. *Review of Economics and Statistics*, 1969, 51: 247-257.
- [72] Andersen, T.G. Stochastic Autoregressive Volatility: A Framework for Volatility Modeling [J]. *Mathematical Finance*, 1994, 4: 75-102.

- [73] Jacquier, E., N.G. Polson and P.E. Rossi. Bayesian Analysis of Stochastic Volatility Models [J]. *Journal of Business & Economic Statistics*, 1994, 12: 371-389.
- [74] Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens. The Distribution of Stock Return Volatility [J]. *Journal of Financial Economics*, 2001, 61: 43-76.
- [75] Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys. Modeling and Forecasting Realized Volatility [J]. *Econometrica*, 2003, 71: 579-625.
- [76] Engle, R.F. and V.K. Ng. Measuring and Testing the Impact of News on Volatility [J]. *Journal of Finance*, 1993, 48: 1749-1778.
- [77] Ding, Z., C.W.J. Granger and R.F. Engle. A Long Memory Property of Stock Market Returns and a New Model [J]. *Journal of Empirical Finance*, 1993, 1: 83-106.
- [78] Andersen, T.G. and T. Bollerslev. Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns [J]. *Journal of Finance*, 1997, 52: 975-1005.
- [79] Baillie, R.T., T. Bollerslev and H.O. Mikkelsen. Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity [J]. *Journal of Econometrics*, 1996,74: 3-30.
- [80] Pojarliev, M.and W.Polasek. Applying multivariate Time Series Forecasts For Active Portfolio Management [J]. *Financial Markets and Portfolio Management*,2001,15: 201—211.
- [81] Pojarliev, M.and W.Polasek. Portfolio Construction By Volatility Forecasts: Does the Covariance Structure Matter? [J].*Financial Markets And Portfolio Management*, 2003,17: 103-116.
- [82] Diebold, F.X. and M. Nerlove. The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model [J]. *Journal of Applied Econometrics*, 1989: 1-21.
- [83] Diebold, F.X., Gunther, T.A., Tay, A.S. Evaluating density forecasts [J]. *International Economic Review*, 1998, forthcoming.
- [84] Cappiello, L., R. F. Engle and K. Shepard. Asymmetric Dynamics in the correlations of Global Equity and Bond Returns [J]. *Technical Report 2004*, European Central Bank.
- [85] Lin, K.-P., A. J. Menkveld, and Z. Yang. China and the World Equity Markets: A Review of the First Decade [Z]. 2004, working paper.
- [86] B Rosenberg. Extra-Market components of Covariance in Security Returns [J]. *The Journal of Financial and Quantitative Analysis*, 1974: 263-274.

- [87] Kupiec P H. Techniques for verifying the accuracy of measurement models [J]. Journal of Derivatives, 1995, 3: 73-80.
- [88] Engle, Robert and Simone Manganelli. CAViaR: Conditional Value At Risk By Regression Quantiles [Z]. NBER Working Paper, 2002.
- [89] Goetzman, W., L. Li and K. Rouwenhorst. Long-Term Global Market Correlations [Z], Working Paper, Yale ICF, 2002.
- [90] 罗付岩、唐邵玲. 基于 fattailed-garch 的 VaR 模型 [J]. 系统工程, 2005, 11, 29-33.
- [91] Palandri Alessandro. Sequential Conditional Correlations: Inference and Evaluation [Z]. 2006, 10, working paper.
- [92] I.V. Basawa, P.D. Feigin, C.C. Heyde, Asymptotic properties of maximum likelihood estimators for stochastic processes [J]. Sankhya Ser. A 38 1976: 259-270.
- [93] Ling.S. On the probabilistic properties of a double-threshold ARCH conditional heteroskedasticity model [J]. Technical report. University of Hong Kong Dept.of Statistics,1995.
- [94] Sang-Won Lee and Bruce E.Hansen. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator [J]. Econometric Theory, 1994,10: 29-52.
- [95] Daniel Straumann and Thomas Mikosch. Quasi-maximum-likelihood estimation in heteroscedastic time series: a stochastic recurrence equations approach [M].to be appear 2003.
- [96] 刘继春著, 多维广义自回归条件异方差时间序列模型 [D], 吉林大学博士论文, 2001.
- [97] F.Comte and O.Lieberman. Asymptotic theory for multivariate GARCH processes [J]. J.Multivariate analysis, 2003, 84: 61-84.
- [98] White, H. Estimation,Inference,and Specification Analysis [M].Cambrige University Press,Cambridge,1996.econometric Society Monographs.
- [99] Patton, A.J. Modelling time-varying exchange rate dependence using the conditional copula [Z]. Working paper of London School of Economics& Political Science, 2001.

## 致 谢

值此毕业论文完成之际，谨向我的导师郑振龙教授表示由衷的感谢和敬意！郑老师在金融学领域渊博的专业知识，对待金融思维的把握程度，在金融新事物方面敏锐的洞察能力，严谨的治学作风和敏锐的学术思想使我终生受益。尤其是郑老师对待学术的热情和执著对我产生了深深的影响，郑老师授课时的那种热情，一开始讲课往往就忘记了时间的存在；郑老师在学术上总是站在最前沿，把我们直接带到最尖端的方向。郑老师批改我们论文时的一丝不苟以及在seminar 讨论时闪现出的思想火花，都让我感到由衷的钦佩和景仰，也是我一生学习的榜样和楷模。郑老师在学术上对我们严格要求，在生活上却对我们爱护有加，每年不论中秋元旦，还是平时节假，郑老师总会让师母为我们准备和安排好丰盛的菜肴和节目，与我们同乐，亦师亦友。每每这个时候，师母也总是想尽一切办法使我们放松，缓解我们平时学习的压力，总给我们以“严父慈母”两相宜的感觉。

同时我也要特别感谢美国波特兰州立大学经济学系林光平教授，林教授在我们一年级下学期为我们经济学院博士班讲授《高级计量经济学 II》课程，从那时起，我就和林老师建立了联系。林老师在计量经济学的高深造诣使我感到十分的钦佩。在这两年多来，不论林老师来国内授课，还是林老师身在美国讲学，我们一直保持着密切的联系，不论我在学术上遇到什么问题，向他请教，林老师总是在第一时间给我答复。同时我和林老师也建立了固定的学术交流，共同合作学术论文。

我从理科背景转到厦门大学金融系学习感到十分的荣幸和自豪。金融系的每个老师学有所专，业有所长，是一个非常优秀的科研教学组合。能有机会倾听他们的授课，让我受益匪浅。在这里，请允许我罗列出他们的名字：张亦春教授、林宝清教授、江曙霞教授、朱孟楠教授、邱崇明教授、魏巍贤教授、郑鸣教授、陈国进教授、李晓峰教授、何孝星教授、郑荣鸣教授、陈善昂副教授、黄华副教授等等，我衷心的感谢他们。同时也感谢金融系里为我们学生做了很多后勤工作让我们安心学习的蔡淑昭老师、张忠正老师以及许多我叫不上名字的老师。

当然还有我一辈子也忘不的那些兄弟姐妹们，他们来自于不同的专业领域，他们个个各有所长，技艺精湛，在某个方面都能独树一帜，研究颇深。我在他们那里学到了很多知识，和他们的友谊也是我一生最重要的宝贵财富。有师兄师姐陈蓉、林海、康朝峰、陈惠玲、邱文华、黄兴李、王保合、郑泽星、唐革榕、冯玲；有一起入门的兄弟姐妹陈淼鑫、包苏昱、马喜德、杨睥阳；还有师弟师妹：刘晓曙、胡韡、何凯浩、王珏、黄蕙舟、黄文彬、杨伟、任婕茹、邹春昊、俞琳、贺涛、张睿、刘晓颖、林舒等等。我感谢他们。

最后要感谢我的奶奶、我的父母、爱人以及我的亲朋，他们博大无私的爱是我一直坚持下来的勇气和动力，我深爱着他们。

张蕾

2007年5月18日