

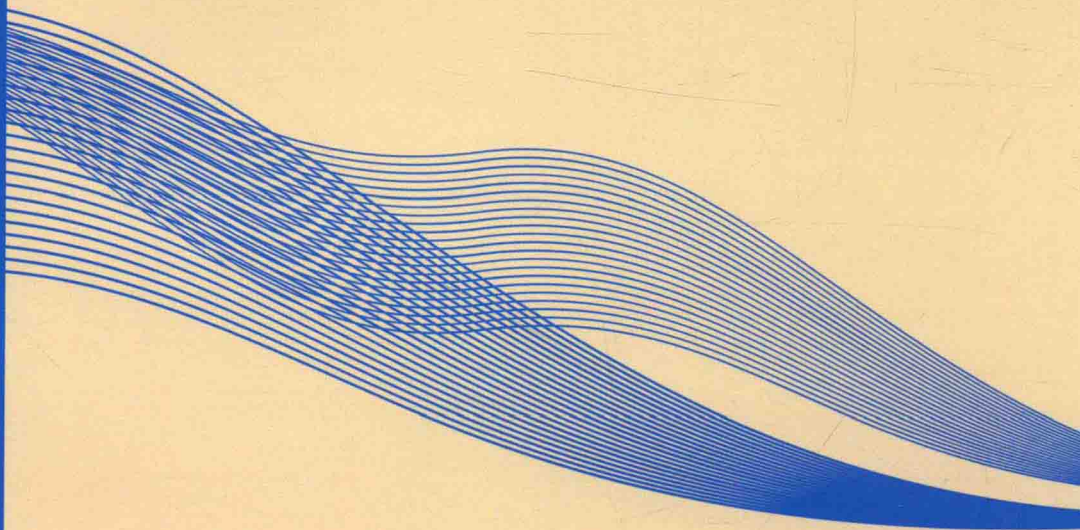
国家自然科学基金面上项目71471119

函数型数据分析的 金融应用与实证分析

Functional Data Analysis
in Financial Application and Empirical Study



隋钰冰◎著



中国书籍出版社
China Book Press

国家自然科学基金面上项目71471119

函数型数据分析的 金融应用与实证分析

Functional Data Analysis
in Financial Application and Empirical Study

隋钰冰◎著



中国书籍出版社
China Book Press

图书在版编目 (CIP) 数据

函数型数据分析的金融应用与实证分析 / 隋钰冰著.

—北京: 中国书籍出版社, 2017.3

ISBN 978-7-5068-6076-5

I. ①函… II. ①隋… III. ①数理统计—应用—金融市场—研究 IV.

①F830.9

中国版本图书馆 CIP 数据核字 (2017) 第 032910 号

函数型数据分析的金融应用与实证分析

隋钰冰 著

责任编辑 刘 娜

责任印刷 孙马飞 马 艺

封面设计 田新培

出版发行 中国书籍出版社

地 址 北京市丰台区三路居路 97 号 (邮编: 100073)

电 话 (010) 52257143 (总编室) (010) 52257153 (发行部)

电子邮箱 chinabp@vip.sina.com

经 销 全国新华书店

印 刷 北京市媛明印刷厂

开 本 170 毫米 × 240 毫米 1/16

字 数 124 千字

印 张 11.75

版 次 2017 年 3 月第 1 版 2017 年 3 月第 1 次印刷

书 号 ISBN 978-7-5068-6076-5

定 价 38.00 元

版权所有 翻印必究

目 录

第 1 章 导论	1
1.1 研究背景和意义	1
1.2 国内外研究现状	3
1.3 研究思路与研究方法	9
1.4 特色与创新之处	10
1.5 内容与结构安排	12
第 2 章 函数型数据分析	15
2.1 引言	15
2.2 函数型数据	16
2.2.1 函数型数据	16
2.2.2 函数型数据预处理	18
2.3 函数型数据分析的一般方法	26
2.4 金融数据的函数型表达	29
2.5 本章小结	37
第 3 章 函数型主成分分析与聚类分析	39
3.1 引言	39
3.2 理论背景	40
3.2.1 函数型主成分分析的原理	40
3.2.2 Mercer 引理与 Karhunen-Loeve 展开式	50
3.3 实证研究	52

3.3.1	模型设定	52
3.3.2	数据与变量	56
3.3.3	实证结果分析	59
3.4	聚类分析	68
3.4.1	聚类分析的原理	68
3.4.2	K-means 聚类分析的算法	71
3.4.3	聚类分析的实证结果	73
3.5	本章小结	75
第 4 章	基于函数型数据的股票市场波动率研究	77
4.1	引言	77
4.2	文献回顾	79
4.2.1	股票日间波动率	79
4.2.2	股票日内波动率	82
4.3	理论背景	87
4.4	实证分析	90
4.4.1	数据与变量	91
4.4.2	模型设定	92
4.4.3	算法	97
4.4.4	模型估计	100
4.4.5	实证结果分析	103
4.5	模型预测	104
4.5.1	样本内与样本外预测	104
4.5.2	预测效果的评价	106
4.6	本章小结	107
第 5 章	函数型数据分析与金融风险管理	109
5.1	引言	109

5.2 文献回顾	113
5.2.1 风险价值的相关文献	113
5.2.2 基于分位数的风险价值	118
5.3 理论背景	120
5.3.1 相对风险测度方法	121
5.3.2 绝对风险测度方法	122
5.4 基于函数型数据的多元风险管理	124
5.4.1 模型设定	125
5.4.2 数据与变量	128
5.4.3 算法	129
5.4.4 实证结果分析	133
5.5 本章小结	147
第 6 章 总结与研究展望	149
6.1 总结	149
6.2 研究展望	150
6.2.1 多元模型与二维函数型模型的研究	151
6.2.2 混合模型的研究	152
6.2.3 集成风险模型的研究	154
参考文献	155
附录	179
后记	181

第 1 章 导论

1.1 研究背景和意义

金融是现代经济活动最为活跃和最重要的组成部分，伴随经济全球化和金融一体化的进程，金融更加广泛而深入地渗透到社会经济活动的方方面面。同时，金融也是创新意识最为重要的领域。金融活动的实践、理论与方法相互推动，表现出惊人的发展速度并取得了大量突破性的发展成果。对金融市场的复杂行为进行合理的计量和统计，建立观测市场变化的有效指标和手段，是对金融市场进行研究的重要领域。

从市场运行角度而言，近年来金融市场飞速发展，金融市场的运行呈现出复杂化特征，不断拓展认识和理解金融市场的视角，尝试利用新的思维方式和分析工具来研究金融问题，是一项与时俱进、具有挑战性和实践意义深刻的工作。首先，了解金融市场运行的现状、方式和现象需要用合理的数据分析工具，将经济现象得以客观、充分而准确地呈现。其次，探索金融市场运行的规律，需要理论和分析工具的高度融合。再次，利用已有的认知保障金融市场的健康发展、维护金融稳定是政府、企业和社会共同面临的挑战和任务。

从风险防范角度而言,金融风险管理对于现代社会具有重要意义,也是金融学研究的一个重要领域。随着全球经济一体化以及各国金融市场国际化程度不断加强,国际、国内金融市场波动增大、关联性增强,同时,金融资产的风险结构变得愈加复杂。2008年新一轮的金融危机蔓延全球,金融风险的重要性也因此得到重新认识。而2015年中国股票市场经历的剧烈震荡再次警示金融风险的重要影响。无论对于金融市场的参与者还是监管者,理解金融资产价格变化的特征、充分认识金融市场存在的风险、从而展开有效的风险管理具有重要意义。对风险进行准确测度是金融风险管理的关键环节。有效进行金融风险管理从而防范和规避金融风险的重要前提,是合理地测度风险。合理测度风险不仅要求准确衡量风险水平、刻画风险特征,还要求对风险具有一定的预测能力。

此外,由于通信和计算机技术使高频数据得到普遍使用,基于高频数据分析的高频交易也快速发展。金融市场的数据分析在实现了从低频数据分析向高频数据的跨越后,进一步发展到超高频数据的应用阶段。超高频数据与基于超高频数据的高频交易使金融资产价格形成的机制发生了显著变化。当前,高维大型金融数据成为现代金融市场的重要信息,在金融市场信息传播迅速、交易高频化的背景下,传统分析方法不断受到挑战。探索如何有效利用高频金融数据对金融资产的价格特征进行描述、如何对金融风险进行准确的衡量、如何基于高维大型数据对金融市场的风险进行合理测度,对于新信息环境下金融风险的识别和防范十分重要。

函数型数据分析方法 (functional data analysis, FDA) 起源于 20 世纪 50 年代的心理学研究。在发展的初期, 较多地应用于生物统计、海洋、医学等领域。近年来, 函数型方法发展迅速, 逐渐被应用到经济相关的研究中并取得了巨大进展。特别地, 金融市场的迅速发展对数据处理和分析提出了新的要求。而函数型数据分析这项分析工具为应对和解决这些问题提供了一种全新的视角和研究工具。

函数型数据分析作为金融领域中一种较新的数据研究方法, 在处理高维问题和适用灵活性方面具有显著优势, 金融数据的分析是其应用的重要领域。相较于传统方法, 函数型数据分析方法的主要优势有: 首先, 无须事先假设数据所服从的分布, 根据观测值本身进行建模, 可以较好地弥补现有研究依赖分布假定的缺陷; 其次, 充分考虑资产之间的相关性, 并且借助多个存在相关性信息实现更有效的分析; 再次, 函数型主成分分析的降维思想在尽可能提取重要信息的同时较好地解决了高维大型数据分析的“维数诅咒”问题。

综上, 为更好地把握金融市场的变化规律、有效管理金融风险, 本书将函数型数据分析的方法引入到金融市场的相关的研究中, 以函数型思维方式为切入点, 分别从金融市场的函数型表达、金融市场的风险以及风险管理三条线索研究金融市场中的函数型数据分析。

1.2 国内外研究现状

函数型数据分析将数据的观测值作为包含了更多信息的曲线进行

分析。与传统数据分析相比, FDA 具有许多优越性, 通过对曲线性质的分析能够挖掘出更多重要信息, 其在医学、气象学、生物统计学、海洋研究已经得到了广泛应用 (Ramray 和 Silverman, 2005), 但在经济学和金融学方面的应用相对较少。

在金融市场中, 资产持续地流动, 资产的价格及其他相关指标以高频率不断地记录, 因此, 金融市场产生的数据大部分可以被视为函数型数据。金融领域是该方法应用的重要领域。国外关于 FDA 在金融中的应用研究近年来逐渐发展, 出现了许多有价值的研究成果。但是, 函数型数据在金融应用中的研究仍然处于起步阶段, 需要更深入地研究和进一步地完善。

一、国外文献

国外学者在函数型数据分析的经济和金融应用方面已经展开了一系列的研究。在经济学应用方面, Ramsay 和 Ramsey (2001) 将函数型数据分析用于研究非耐用消费品指数的动态特征。Matzner 和 Villa (2004) 使用函数型主成分分析探索来收益率曲线的特征。而 Wang (2008) 则将其应用于网上商品拍卖价格的分析中。

由于金融市场的交易是连续不断进行的且市场上的相关的量化信息, 如金融资产的价格、交易量等, 能够以较高的频率提取, 金融市场上存在丰富的可观测信息成为“天然”的函数型数据。函数型数据分析在金融市场中的应用成为一个非常有吸引力的研究方向。

在金融应用方面, Anderson 和 Newbold (2002) 首次基于函数型

数据分析金融资产收益,用函数自回归方法估计了瑞士—法郎—美元汇率时变的日内逆分布函数的自回归模型,研究结果显示,相较于传统的统计方法 FDA 在金融数据的分析中更有优势。随后,函数型数据分析被广泛应用于分析不同标的金融数据上。

Salvastore 和 Damiana (2005) 使用对 2000 年 1 月 3 日至 2002 年 12 月 30 日的 MIB30 指数进行了函数型数据主成分分析,研究结果显示,函数型数据分析是对金融序列进行分析的一项有利工具。

Müller, Stadtmüller 和 Yao (2006) 首次提出了函数型方差过程的概念,从理论角度证明了实际数据转化为函数型数据进行分析的原理和过程,为方差过程分析提供了新思路与新方法。

Sebastian 和 Eddie (2007) 基于传统函数型主成分分析方法提出了基于向量自回归 (VAR) 方法的函数型主成分分析,对纽约交易所 2002 年 1 月至 2006 年 12 月轻质原油期货合约价格进行了分析。

Kai 等 (2009) 首次使用函数型主成分分析方法分析隐含波动率,通过引入函数共同主成分分析模型对隐含波动率进行了研究。Ingrassia 等 (2009) 也将函数型数据分析方法应用于对金融时间序列数据的分析,以 MBI300 指数为样本进行实证分析并提出了基于载荷因子聚类的股票指数构建方法。Alva, Juan 和 Esther (2009) 基于函数型数据主成分分析方法,利用 AR(1) 模型对日内波动率的预测进行了改进。Müller, Sen 和 Stadtmüller (2011) 提出了使用函数型数据分析方法估计股票的日内波动率曲线,并建立了使用日内上午交易信息预测日内下午股票波动的预测模型。Wang, Sun 和 Li (2014) 将函数型数据的主

成分分析方法直接应用于中国上证 50 指数 2011 年月收益率的分析中, Fearghal, Finbarr 和 Mark (2015) 基于跳跃扩散模型, 利用函数型数据分析方法对国际商品市场 2007 年至 2013 年 WTI 期货合约的隐含波动率进行了研究。

二、国内文献

函数型数据在中国的研究起步较晚, 研究领域也相对单一。国内现有关于函数型数据的研究多为基础性描述研究或直接的应用研究, 对函数型数据分析方法的拓展及更深层次的研究还较少。张崇岐等 (2006) 讨论了将函数型数据进行平滑曲线展示的方法, 并对方法应用过程中的潜在问题进行了说明。严明义 (2007) 首次系统地讨论了函数型数据分析统计意义上的基本思想和方法。近年来, 国内文献在函数型数据分析的经济、金融领域的应用方面取得了巨大进步。

在经济学研究方面, 靳刘蕊 (2008) 对函数型数据的聚类分析、主成分分析等方法进行综述, 同时利用该方法对国内的宏观数据进行了分析。孟凤银 (2011) 对比分析了函数型数据分析中的主成分分析和多元统计分析的差别, 并利用函数型主成分分析方法对中国财政指数情况进行了研究。不同于主成分分析和聚类分析, 剡亮亮 (2013) 从泛函特征而非多元统计延伸的视角, 介绍了函数型数据的主微分分析方法并对全国银行间同业拆借利率进行了分析。

在金融学研究方面, 毛娟 (2008) 使用美国股票市场 S&P500 指

数在 2003 年 9 月 23 日至 2004 年 6 月 2 日的指数期权的隐含波动率使用函数型数据方法进行研究。岳敏和朱建平 (2009) 及王诚 (2014) 运用函数型主成分分析方法, 分别以 2007 年沪市股票和 2013 年 A 股股票为样本, 把股票的月收益率视为连续的函数数据, 应用函数型主成分方法来提取主成分因子以刻画股票收益率的主要变动因素随时间变化的特征。曲爱丽 (2009) 选取我国沪深两市的五只蝶式权证作为研究对象, 对 2006 年 5 月 22 日至 2006 年 11 月 15 日权证收盘价格数据的周收益率进行了分析。朱建平等 (2010) 介绍了函数型数据聚类及其在金融时间序列分析中进行应用的相关问题。郭均鹏、孙钦堂和李汶华 (2012) 使用函数型数据分析的区间主成分分析法对 Shibor 收益率进行实证研究。徐佳 (2008)、胡梦荻 (2014)、陈丽琼 (2014)、张蕾蕾 (2014) 及田和鹭 (2015) 等利用函数型数据分析方法分别针对中国的证券投资基金、股票、期货和期权进行了研究。

三、文献评述

综上所述, 国内外研究者近年来对函数型数据的金融应用有所关注, 国外文献方面已经有较为深入的研究。然而, 函数型数据在国内金融研究方面的应用还处于起步阶段, 虽然近两年发展较快, 但以直接将函数型方法应用到实证分析中的研究方式为主, 并没有对相对的理论和方法进行拓展。函数型数据分析在国内金融研究中的应用具有较大的发展空间。

函数型数据还处在发展的初始阶段，它与其他数据类型的界限是相对模糊的。因此，在考虑函数型数据的金融应用时，更重要的是用合理的方式思考数据、使用数据，帮助分析实际的问题。由于金融数据的高维特性及函数型数据分析在降维处理中的优越性，在金融高频数据的处理中，这一方法受到格外地关注并被一些学者提倡应用于金融高维数据的分析。在研究不断推进和深入的过程中，函数型数据分析的思想在金融数据的统计分析中的应用被证明是一项重要的分析工具和值得探索的重要方向。

从函数型数据分析的国内外研究现状来看，存在以下几个方面的问题值得深入思考与研究：

首先，从函数型数据分析方法引入金融研究的总体框架来看，当前研究多针对金融市场价格变化的某一个方面进行研究，或是针对某一种具体的模型进行研究，而将研究对象和模型紧密结合、进行系统性研究的很少。尤其是进行多种途径综合分析的实证研究不多。

其次，对于具体的函数型分析过程，准确认识函数型数据和金融数据的共性与个性以及其统计特征，是进行金融市场分析的基础。但当前的研究多为直接将函数型数据的一般分析方法套用到金融数据上，缺乏对其适用性和经济含义进行合理的解释。

再次，无论是理论还是应用，国内的研究都存在一定程度的滞后，在实际应用中有算法改进、方法拓展或创新性结论的研究较少。

1.3 研究思路与研究方法

函数型数据分析从函数视角探索数据,将观测数据视为一个整体,将其表示成平滑的曲线或连续的函数 $x_i(t)$, $i=1, 2, \dots, n$ 。其中, n 代表观测曲线条数或观测对象个数, t 是观测点,可以是时间或者时间以外的其他变量。对于所有的观测对象,并不要求它们的观测点个数或进行观测的次数相等。函数型数据分析的出发点就是将函数数据作为一个单独的对象(通常是一条曲线)而非一系列观测点的序列,强调的是数据的内在结构而非其外在形式。本书利用函数型数据分析的方法,以函数型数据的视角为切入点,分别从金融市场的资产价格、资产价格的波动率以及金融风险管理三个方面讨论了函数型数据的金融应用。

研究的基本思路为:以函数型分析方法为主线,以金融市场的价格变化、波动率和风险为线索,对函数型视角下的资产价格变化特征进行考察。采用理论分析和实证分析相结合的研究方法。具体而言,将函数型主成分分析与聚类分析相结合,构建基于特征聚类的投资组合;将函数型主成分分析与时间序列模型相结合应用于波动率的研究,通过期望最大化(expectation-maximization, EM)算法实现;将函数型数据的主成分分析、分位数回归与时间序列模型相结合应用于风险的研究,通过最小极优化(majorize-minimize, MM)算法实现。

首先,基于函数型波动率过程建立刻画股票市场价格波动的计量模型,考察中国股票市场资产价格的变化特征。其次,将分位数回归引入到函数型数据分析中,利用其对分位数信息的详细刻画研究极端波动时股票价格波动的方式,即考察中国股票市场波动过程的风险建立基于函数型数据的多元风险模型。再次,在分析中,将资产价格变化的特征从横向和纵向两个维度进行分析,横向方面比较不同股票、不同分位数水平的波动率和风险价值特征,纵向方向考察不同风险测度随时间发生的变化,反映不同风险测度的动态变化过程。同时,结合具体的市场背景,分析不同阶段股票市场波动风险的表现和内在风险结构的变化。

1.4 特色与创新之处

本书从函数型数据分析的理论基础和一般分析方法出发,详细讨论了函数型数据分析应用于金融市场分析面临的问题、解决方法及存在的问题,并基于中国股票市场的实际数据进行了讨论。首先,本书围绕金融市场分析的中心和关注焦点,对函数型数据分析思想的经典文献和新近研究成果进行了较为系统的梳理。其次,本书展示了如何具体地将函数型数据分析方法应用到实际的金融市场分析中,并对一般方法进行了扩展,提出了新的研究方法和研究视角。再次,本书利用中国股票市场的实际数据提供了大量基于函数型数据视角的实证

证据。

本书对函数型数据在金融市场中的应用进行了较为全面的分析。从金融市场出发,提供了一系列重要问题的具体应用和解决方案。更重要的是,提供了从函数视角出发,分析金融市场行为的思维方式。基于函数型数据分析这一新的视角,为更好地认识和理解中国金融市场的变化规律和风险提供了新的分析工具和实证证据。研究结合了函数型数据分析方法、时间序列分析和统计算法中的重要思想和方法,并将其应用到刻画资产价格变化和衡量风险的研究中。

本书主要的创新之处如下:

第一,研究视角的创新。在研究视角上,从函数型数据分析方法出发对金融资产的价格变化进行了系统研究,包括对资产价格变化及其风险的刻画。一方面,利用函数型数据的思想通过对曲线特征进行建模的角度衡量金融市场的风险,为研究金融风险测度提供了新的视角。另一方面,区别于传统的函数型数据分析,将其拓展至分位数即分布函数尾部特征的研究,拓展了函数型数据分析方法的应用领域。

第二,研究方法的创新。利用所提取的“函数型”数据信息,考虑不同资产之间的相关性,建立了基于函数型数据的多元风险模型。在数据分析上,提高了数据的使用效率,增加了用于估计的信息量,尤其在分析极值问题中避免了由于尾部数据稀缺性而造成的估计偏误,提高了尾部信息估计的有效性。