

学校代码：10378

密级：

分类号：F832.5



安徽财经大学  
ANHUI UNIVERSITY OF FINANCE & ECONOMICS

## 硕士学位论文

# 基于深度学习的股指波动率预测与 VaR 测度实证研究

学 号： 3202000151

学生姓名： 刘 滢

学位类别： 经济学硕士

专业名称： 数量经济学

研究方向： 金融计量经济分析

导师姓名： 朱海龙

二〇二三年六月

School code : 10378

Security:

Classification: F832.5



# Empirical research on stock index volatility prediction and VaR measurement based on deep learning

Student ID: 3202000151

Name: Liu Ying

Degree category: Master of Economics

The professional name: Quantitative Economics

Research direction: Financial Economics Analysis

Tutor's name: Zhu Hailong

June, 2023

## 摘要

金融风险管理是防范风险、化解风险、保持市场稳定与快速发展的重要内容，波动率预测是风险管理的重要工具和核心变量。因此，在波动率预测和风险管理问题上，在处理与拟合现阶段金融系统较为复杂的非线性结构方面，本文在利用传统计量模型的基础上希冀尝试寻找新的方法。针对这一问题，本文拟将深度学习算法与计量经济学模型有机地融合，在该框架下开展基于深度学习的波动率预测和 VaR 测度的建模与分析，从而实现在计量经济学方法下提供有力的金融解释力度，在深度学习算法下对数据的非线性部分进行有效地拟合，进而达到提高预测的准确性的目的。

文章首先对金融市场波动率预测和 VaR 测度的历史和研究现状作出综述；随后，介绍了相关理论基础以及模型构建；其次，使用单一的 GARCH 模型和 LSTM 模型以及二者的 7 种融合模型对波动率进行建模，通过损失函数对波动率预测效果进行评价，并将综合表现最优的模型预测结果作为后续 VaR 测度模型的输入变量；最后，在神经网络分位数回归（QRNN）模型的基础上考虑滞后内生变量，建立神经网络分位数自回归（QARNN）模型进行 VaR 测度，并通过返回测试对比、评价模型表现。

主要的研究结果：（1）在波动率预测中，通过损失函数横向比较 GARCH 模型在不同残差分布假设下的波动率预测能力，当假设残差服从于 GED 分布时，模型预测精确性高于其他 GARCH 族模型；通过损失函数横向对比单一模型的表现，LSTM 模型的预测精度高于 GARCH 模型；在量价数据的基础上增加 GARCH 模型参数作为 LSTM 模型输入得到的融合模型，通过损失函数对比融合模型和单一模型的表现，融合具有更好的预测性能；在融合模型中，结合三种 GARCH 模型和 LSTM 模型的 L-G-GT-GGED 融合模型表现效果最优；（2）在 VaR 测度中，通过返回测试对比模型 VaR 测度的准确程度，代表传统计量模型的分位数回归（QR）模型和分位数自回归（QAR）模型均在测试集内出现低估损失的情况；QRNN 模型的模型预测性能优于 QR 模型和 QAR 模型，失败率接近期望失败率，对 VaR 测度较为合理；QARNN 模型整体表现最佳，失败率低于期望失败率，该模型不仅考虑了波动率对 VaR 的关键影响，而且考虑了滞后内生变量的非线性效应，使得模型准确性显著提升。

综上所述，在对上证指数的波动率预测中，传统计量模型与深度学习算法相结合的 LSTM-GARCH 融合模型相较于单一模型都有着更为优异的表现，同时为后续 VaR 测度提供了更准确的波动率输入变量；在 VaR 测度方面，综合考虑了滞后内生变量和融合了传统计量模型的 QARNN 模型在很大程度上改进了传统分位数回归模型在 VaR 测度上的预测性能。

**关键词：**波动率；VaR；LSTM-GARCH 模型；QARNN 模型

## ABSTRACT

Financial risk management is an important part of preventing risks, resolving risks, maintaining market stability and rapid development, and volatility forecasting is an important tool and core variable of risk management. Therefore, in terms of volatility forecasting and risk management, in dealing with and fitting the complex nonlinear structure of the current financial system, this paper hopes to find new methods based on traditional econometric models. To solve this problem, this article aims to combine deep learning with organic standard model algorithms. In this case, the volatility of VaR prediction is modeled and analyzed based on deep learning. Improve prediction accuracy.

Firstly, this paper summarizes the history and research status of financial market volatility prediction and VaR measurement; Then, it introduces the relevant theoretical basis and model construction; Secondly, we use the single GARCH model, LSTM model and their seven fusion models to model the volatility, evaluate the volatility prediction effect through the loss function, and take the model prediction result with the best comprehensive performance as the input variable of the subsequent VaR measurement model; Finally, on the basis of the Quantile Regression Neural Network (QRNN) model, considering the lag endogenous variables, the Quantile Autoregression Neural Network (QARNN) model is established to measure VaR, and the performance of the model is compared and evaluated through the return test.

Main research results: (1) In the volatility prediction, we compare the volatility fitting prediction ability of GARCH model under different residual distribution assumptions. Based on the results of multiple loss functions, it is proved that the accuracy of the model prediction can be improved when the residual is assumed to be subject to GED; Compared with the performance of single model through loss function, the prediction accuracy of LSTM model is higher than that of GARCH model; After using GARCH model to obtain model parameters through rolling prediction, GARCH model parameters are added on the basis of volume and price data as input of LSTM model, and the performance of the fusion model obtained is generally better than that of single model; Among the fusion models, the fusion model that integrates three GARCH models and LSTM models has the best performance. (2) In the VaR measure, the Quantile Regression (QR) model and Quantile Autoregressive (QAR) model representing the traditional econometric model are underestimated in the test set by returning the accuracy of the VaR measure of the comparison model; The model effect

of QRNN model is better than that of QR model and QAR model, the failure rate is close to the expected failure rate, and the VaR measure is more reasonable; The QARNN model has the best overall performance, and the failure rate is lower than the expected failure rate. This model not only takes into account the key impact of volatility on VaR, but also takes into account the nonlinear effect of lag endogenous variables, which significantly improves the accuracy of the model.

In summary, in the volatility prediction of the Shanghai Composite Index, the LSTM-GARCH model combining traditional econometric model and deep learning algorithm has better performance than a single model, and provides more accurate volatility input variables for VaR measurement. In terms of VaR measures, the QARNN model, which comprehensively considers lagging endogenous variables and integrates traditional econometric models, greatly improves the prediction performance of traditional quantile regression models on VaR measures.

**KEYWORDS:** volatility; VaR; LSTM-GARCH model; QARNN model

# 目 录

第一章 绪论.....	1
第一节 研究背景与意义.....	1
一、研究背景.....	1
二、研究意义.....	2
第二节 文献综述.....	3
一、波动率预测的研究现状.....	4
二、VaR 风险测度的研究现状.....	7
三、简要评述.....	9
第三节 研究思路及方法.....	10
一、研究思路.....	10
二、研究方法.....	11
第四节 创新与不足.....	12
一、可能的创新点.....	12
二、不足之处.....	13
第二章 金融市场波动率及金融风险的基本理论 .....	14
第一节 金融市场波动率的基本理论 .....	14
一、波动率的理论基础.....	14
二、波动率的分类.....	14
三、波动率的性质 .....	16
第二节 金融风险的基本理论 .....	17
一、金融风险及其分类.....	17
二、金融风险度量的传统方法.....	18
三、金融风险度量的 VaR 方法.....	19
第三章 波动率预测与 VaR 测度模型的理论基础 .....	22
第一节 LSTM-GARCH 模型 .....	22
一、GARCH 模型.....	22
二、LSTM 模型 .....	25
三、LSTM-GARCH 模型.....	26
第二节 神经网络分位数回归模型 .....	27
一、分位数回归模型.....	27
二、分位数自回归模型.....	28
三、神经网络分位数回归模型.....	28
第四章 基于 LSTM-GARCH 混合模型的波动率预测 .....	30
第一节 基于 LSTM-GARCH 模型的波动率预测方法 .....	30

一、基于 LSTM-GARCH 的波动率预测模型构建 .....	30
二、波动率预测评价指标 .....	33
第二节 特征选取及统计分析 .....	34
一、特征选取 .....	34
二、统计分析与检验 .....	35
第三节 基于 LSTM-GARCH 模型的波动率预测 .....	37
第四节 模型预测效果评价 .....	39
第五节 本章小结 .....	41
第五章 基于 QARNN 模型的 VaR 测度 .....	43
第一节 基于 QARNN 模型的 VaR 测度方法 .....	43
一、基于 QARNN 的 VaR 测度模型构建 .....	43
二、VaR 返回测试评价指标 .....	44
第二节 数据选取与统计检验 .....	45
一、数据选取 .....	45
二、统计检验 .....	46
第三节 基于 QARNN 的 VaR 风险测度研究 .....	47
第四节 返回测试 .....	48
第五节 本章小结 .....	49
第六章 结论与展望 .....	51
第一节 研究结论 .....	51
第二节 研究展望 .....	52
参考文献 .....	53

# 第一章 绪论

## 第一节 研究背景与意义

### 一、研究背景

随着布雷顿森林体制的崩塌,以及不可阻挡的全球一体化进程,这为我国的金融机构和个体投资者提供了许多机会,但同时也增加了其面临的金融风险。最近几年,世界范围内和局部地区的金融危机时有发生,致使人们愈发认识到:金融市场的整体结构越发复杂,内部也存在日渐明显的不确定性。因此,外汇、债券、商品、衍生品等金融工具的不断发展和完善,使得金融风险管理已成为防范风险、化解风险、保持市场稳定、快速发展的重要内容。

市场风险是金融风险中的重要构成部分。金融市场风险主要来自于金融资产收益率的不确定性。随着金融市场的日渐繁荣、金融衍生工具种类的增加,金融系统中的不确定性因素也随之增多,导致金融市场风险大幅度增大。金融资产的收益率呈现出尖峰厚尾、非线性和异质性的特征。

因此,金融市场风险度量日渐成为研究的焦点。在金融风险管理领域,在险价值 VaR (Value at Risk) 是 JP Morgan 在 20 世纪 90 年代开发的用于度量资产所面临市场风险的一种风险测量方法,其相关理论与方法是金融风险计量的主流方法。VaR 更为具体的定义为:一项金融资产或资产组合在一定时间内和一定置信水平下所面临的最大损失。更严格地说, VaR 描述了在一定时间间隔内资产损益预期分布的分位数。它可将无形的风险量化为具体的损失额,有助于投资者和金融机构更加直观地监测风险大小并对风险进行把控,进而达到合理配置资本金、实现最大化收益的目的。

VaR 定义相对简单,但金融市场风险计量是一个更为复杂和前沿的问题。基于过去的研究,金融市场的资产收益率时间序列建模主要存在以下两点问题,一方面,与正态分布的假设相悖,呈现尖峰厚尾的分布特征;另一方面,具有波动聚集性、非线性和异质性。因此如何正确拟合收益率序列,准确刻画收益率分布,成为精确计量金融市场风险的研究重点,同时也是本文力求解决的问题。

实现更为精确的金融风险计量离不开对金融资产收益率的波动性的建模与预测。也就是说,金融资产收益的波动率是影响 VaR 值的决定因素和测度 VaR 值的有力工具,是金融风险测度的核心变量,同时也是资产定价理论、投资组合构建以及衍生产品避险策略设计等领域的关键影响因素,因此,针对于波动率的建模研究始终是金融



风险管理的活跃领域。

波动率与价格不同,我们无法从市场中直接观测并获取到波动率的相关数据,因此衍生出了一些相对直观的方法对波动率进行简单地估算,如采用样本方差估计量来作为波动率的量化值。此类方法计算简便,但预测精度较低,也无法对金融资产收益率的波动率进行较为准确地刻画。为了改进波动率的预测效果,国内外学者们展开了大量的研究并提出了多种方法用以预测波动率。其中,极具代表性的是 Bollerslev 提出的广义自回归条件异方差 (GARCH) 模型,该模型是波动率研究中的一大重要进展,可以有效地描述时间序列数据残差的异方差特性。自从 GARCH 模型被提出以后,它就被众多研究者进行了理论扩充和方法拓展,到现在为止, GARCH 族模型相关研究已较为完善,诸多学者将 GARCH 方法应用于波动率预测和研究并取得许多进展。但现存的问题是, GARCH 模型是线性模型,而我国股票市场由于存在过多的噪声交易,而且其波动率不仅受到国家宏观经济政策的影响,而且受投资者个体、机构投资者和企业自身的决策等众多因素的影响与制约,多重影响因素导致了我国金融市场的波动性呈现出十分复杂的、非线性的特点,这使得传统的线性波动率预测模型的预测性能有所降低。近年来,机器学习、人工智能等相关技术的日臻完善在模拟非线性结构方面有着愈发优异的效果,为准确预测波动率带来新的可能性。最有代表意义的是基于神经网络的一类深度学习模型,它的优点在于:一方面,是可以有效地处理大规模、高维度的数据,大大提高了拟合效率;另一方面,由于深度学习算法具备优良的非线性拟合能力和高效的自主学习能力,能够传统线性模型的基础上提高在预测复杂数据方面的准确率。因此,在最近几年的研究中,已有许多学者将深度学习算法应用到了对时间序列数据的预测中。

如今,大数据和人工智能的发展在很大程度上提高了数据采集、清洗和分析的速度,使金融机构和个人投资者能够更有效、更快捷地对风险敞口进行测量和管理,然而深度学习相关算法也存在着缺乏金融理论阐释机制的问题。因此,本文拟将经典的计量经济学模型和新兴的深度学习算法有机地结合起来。二者结合生成的融合模型不仅具备了经典计量经济学理论的可解释能力,而且还具备了深度学习对复杂金融市场非线性结构的拟合能力,从而为更加精准地进行波动率预测和 VaR 度量奠定了坚实的基础。

## 二、研究意义

本文选取“基于深度学习的股指波动率预测与 VaR 测度实证研究”作为研究议题,考虑到波动率预测在金融计量分析中占据重要地位,是测度 VaR 值的核心变量和有力工具。本文将波动率预测加入到金融风险计量框架中,拟通过将传统 GARCH

模型与 LSTM 神经网络相结合以提高波动率的预测精确度,为后续金融风险量化提供更加准确的波动率输入变量;并在 Taylor 提出的神经网络分位数回归模型基础上进行改进,在此基础上考虑滞后内生变量的影响,构建神经网络分位数自回归模型,并给出相应的 VaR 值测度模型与方法。本文所具有的研究价值,可以概括为两个方面:一是理论价值,二是实践意义。

### (一) 理论意义

在波动率预测的模型建立方面,本文以过往文献为基础,构建 LSTM-GARCH 模型并将其应用于股指波动率预测,一方面 GARCH 模型具有较为强大的金融解释力度,相较于神经网络模型具备更好的说服力;另一方面, LSTM 模型擅长捕捉时间维度的波动率相关特征,且结合我国股票市场的长记忆性特征考虑,应用该模型能够有效地提取股票指数的波动信息,提高预测精度。

在 VaR 值测度的模型建立方面,本文以 Taylor 提出的神经网络分位数回归模型为基础,进一步改进提出了基于神经网络的分位数自回归模型。一方面,是从非线性的角度拓展经典的分位数自回归,非线性模型相较于线性模型在处理非线性数据时具有更好的拟合效果和预测能力,结合金融产品收益率序列存在非线性特征考虑,应用该模型将具有较好地适应性;另一方面,在神经网络分位数回归的基础上考虑了滞后内生变量的影响,同时,将分位数自回归与神经网络相融合,以改善神经网络分位数回归模型在时序数据自相关性刻画上有所欠缺的缺点。

### (二) 应用价值

在应用研究方面,本文选取“金融风险计量”这一热点问题,在传统计量理论和新兴神经网络框架的基础上展开相关研究,以波动率的准确预测为辅助风险度量的重要工具,进而展开 VaR 测度分析,以期更好地揭示金融风险变动规律、进行金融风险的控制与管理。

总的来说,本文将波动率预测引入金融风险计量框架,在一定程度上丰富了金融风险测度方法的研究视角拓宽了现有金融风险计量方法的思路,同时,将传统计量模型与神经网络模型相结合,为波动率预测和 VaR 值测度提供了新的研究方法,在理论研究和实践应用中都具有一定意义。

## 第二节 文献综述

在中国知网资源总库检索“波动率预测”和“VaR”两个关键词,从图 1-1 可以看出,学者对于 VaR 的探索起步于 20 世纪 60 年代左右,相较于波动率预测的研究起步更早,二者的相关研究发展在之后的 20 年间均较为缓慢;波动率预测和 VaR 风险

计量的相关文献分别于 20 世纪 90 年代和 20 世纪 80 年代开始逐渐增加,学者于二者的探索日渐成为金融领域的热点问题,并在探索过程中逐步补充、完善相关理论;近十年,关于波动率预测和 VaR 风险计量的文献数量达到巅峰,并始终保持较高的热度,相关理论取得了较为显著的进展,运用相关理论进行的实证研究也有了比较可观的进展。

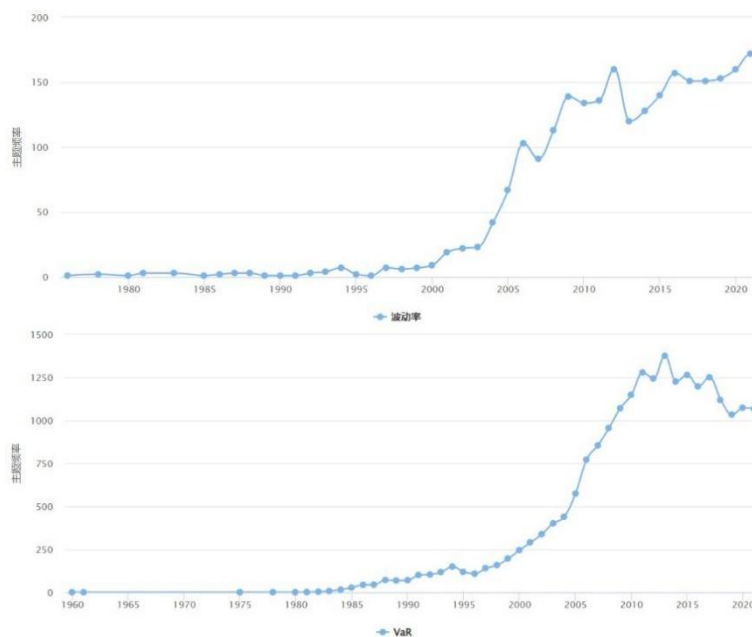


图 1-1 关于波动率和 VaR 的学术关注度

资料来源:图中数据由作者通过检索知网资源总库可得。考虑到近两年文章收录可能存在滞后,因此文章只选取至 2021 年。

## 一、波动率预测的研究现状

### (一) 基于传统计量模型和时间序列模型的方法

Engle<sup>[1]</sup> (1982) 通过对英国通胀指数的波动性集聚现象的研究,提出了条件异方差自回归 (Autoregressive Conditional Heteroskedasticity Model, ARCH) 的方法,有效地克服了传统计量经济学中因二个假定 (方差恒为常数) 所引发的误差,这个模型是获得 2003 年诺贝尔经济学奖的计量经济学成果之一。Bollerslev<sup>[2]</sup> (1986) 在 ARCH 模型的基础上,提出了广义条件异方差自回归 (Generalized AutoRegressive Conditional Heteroskedasticity, GARCH) 模型。在某种意义上,该模型可以较好地克服多重共线性问题。在后续的工作中,学者们不断对 GARCH 模型进行研究与发展,以解决研究过程中的特定具体问题。研究相同大小的正负收益率对波动率造成不同程度的冲击,如门限 ARCH (Zakoian<sup>[3]</sup>, 1990) 模型、EGARCH (Nelson<sup>[4]</sup>, 1991) 模型、GJR-GARCH (Glosten<sup>[5]</sup>, 1993) 模型;研究股票收益率与波动率相关性的 GARCH-M 模型 (Engle<sup>[6]</sup>,

1987)；研究波动率长记忆特征的 IGARCH (Engle<sup>[7][8]</sup>, 1986) 模型。

目前, 国内 GARCH 族模型的相关研究已取得诸多可观的理论和实践成果。如田华和曹家和<sup>[9]</sup> (2003) 以上海股市与深圳股市为研究对象, 使用 GARCH-M 模型对二者的波动率进行了实证研究, 证实了两个股市之间存在着非对称性现象; 黄海南等<sup>[10]</sup> (2007) 多种 GARCH 族模型, 在三种不同的残差分布假定下的波动率预测性能进行对比。实证结果显示, GARCH 模型对上海股市的收益率波动性有较好的预测效果, 综合来看, 残差假定为 GED 分布的 GJR (1,1) 模型预测性能最为突出, 而 GARCH 族模型的预测表现效果则因市场而有所差异; 魏宇等<sup>[11]</sup> (2007) 通过实证分析得出, 将非对称因素引入到 GARCH 模型中, 其在衡量收益率的波动时具有更高的精度; 赵伟雄等<sup>[12]</sup> (2010) 研究表明, 中国铜货市场不存在非对称特征, 但是其周期性的长记忆特征较为明显, 在对其进行分析的过程中使用了 4 种 GARCH 族模型来比较, 结果表明, 残差服从 t 分布 FIGARCH 模型对于股票价格的波动率具有最佳的预测性能。进一步地, 后续有学者将 GARCH 模型与其他模型相结合, 如杨继平等<sup>[13]</sup> (2014) 采用考虑金融波动变结构的马尔可夫结构转换 GARCH 模型度量我国股市的波动性, 结果表明该类模型能够更好地刻画股票波动特征, 波动率预测结果更加精确; 王天一等<sup>[14]</sup> (2015) 将 GAS 模型的思想引入到 GARCH 模型中并与之结合, 拓展了传统 GARCH 的预测模式。

## (二) 基于深度学习的方法

在深度学习算法占据一席之地的人工智能发展趋势下, 目前, 很多研究者把深度学习引入到 NLP、语音识别等应用研究领域, 并取得了一系列成果。Lecun 等<sup>[15]</sup> (2015) 全面回顾了深度学习的现阶段的研究成果和应用方面的发展现状。其最为显著的优势在于不需要人为地进行训练与参数选择, 模型能够从大量的数据中自行提取出相应的特性, 这一优点使得其在现实中众多领域的应用日益广泛。特别是在市场表现日渐复杂的金融领域, 面对金融市场的海量数据, 复杂的非线性关系, 以及难以精确预测等诸多问题, 以及中国股市的收益率和波动性表现出的长记忆性 (李亚静等<sup>[16]</sup>, 2003)。而深度学习算法则擅长处理这种复杂的非线性关系, 因此, 将深度学习算法应用到波动率预测中是十分有意义的。

卷积神经网络 (Convolutional Neural Network, CNN)、循环神经网络 (Recurrent Neural Network, RNN)、长短期记忆网络 (Long Short-Term Memory, LSTM) 是当前较为主流的几种神经网络模型。CNN 的应用领域以图像处理居多, RNN 主要应用于处理时间序列数据, LSTM 模型则是以 RNN 为基础进行了进一步的改良, 它具有对长时间的历史数据进行部分记忆的能力, 该算法的实现是在神经网络结构中添加不同控制门机制的基础上进行。

RNN 模型在理论上可以存储并使用任何长度的数据,但是,当 RNN 模型在进行训练时,由于“梯度消失”和“梯度爆炸”的问题十分突出,导致 RNN 丧失了对历史信息的长期记忆及使用能力。这不但影响训练效率,还会对输出结果的精度造成直接影响,甚至导致模型训练的失败。Hochreiter 等<sup>[17]</sup>(1997)利用长短期记忆单元对 RNN 模型进行优化,不仅可以克服了上述 RNN 所面临的问题,而且还可以解决有效历史信息会因新输入特征的影响而不能长期维持的缺陷。在此基础上,Graves 等<sup>[18]</sup>(2013)改进了 LSTM,并对比了 LSTM 与传统 RNN 模型的准确度,结果表明 LSTM 的预测效果更好。

具体来说,运用深度学习算法对波动率进行建模的有关研究主要有以下几种:李亚静等<sup>[16]</sup>(2003)的实证研究结果表明上证指数的波动率表现出明显的持续性和长期记忆特征,因此更适用于通过 LSTM 模型进行数据的处理与训练;陈卫华<sup>[19]</sup>(2018)使用 LSTM 模型和传统计量经济学模型分别预测上证综指波动率,结果表明,在不同的损失函数中,LSTM 模型的预测准确度均为最高;张蕾等<sup>[20]</sup>(2021)在 LSTM 的输入变量中添加了代表市场情绪指标的 VIX 指数,经实证分析研究表明,该模型比原模型具有更优的预测性能;刘广应等<sup>[21]</sup>(2021)将 LSTM 神经网络模型与波动率高频信息结合,构建了 LSTM-RV 模型实现了波动率的动态预测;Horvath 等<sup>[22]</sup>(2021)将深度神经网络引入到隐含波动率的建模当中,能够快速实现对隐含波动率的预测校准

### (三) 基于深度学习方法与传统计量模型相结合的方法

在变化多端的金融市场中,GARCH 族模型的建模与实际应用受到了严格假设条件的制约,因此,如何将 GARCH 类模型与其他新兴方法有机地融合已成为一个热门课题。与此同时,大多数的深度学习都是以数据为基础的建模方式,虽然模型的总体精度得到了很大的提升,但它也有一个显著的缺点,那就是削弱了解释变量的显著性,无法对金融市场中的相关性统计特性进行有效的刻画,从而大大降低了它的金融解释力度。为使得深度学习算法与金融相关理论进行更加紧密有机的整合,诸多学者利用经典的计量经济学方法和深度学习的融合建模框架。该方法综合了两类模型的优势,是目前金融时间序列分析的一大热点。在波动率预测中,利用融合模型进行预测的研究工作主要包括:Kim 等<sup>[23]</sup>(2018)以 KOSPI 200 的波动率作为研究对象,构建了 LSTM 与 GARCH、EGARCH、EWMA 三种模型的混合模型 GEW-LSTM,该模型相较于其他单一模型有更小的均方误差(MSE)和平均绝对误差(MAE),具有更好的股市波动预测性能;Yan 等<sup>[24]</sup>(2020)提出了一种新的混合方法来预测铜价波动,不仅将深度神经网络与经典的 GARCH 相结合,还将 LSTM 与传统的人工神经网络(Artificial Neural Network, ANN)相结合,实验结果表明,GARCH 模型可以作为信

息特征显著提高神经网络模型的预测能力, LSTM 与 ANN 的融合是构建神经网络结构以提高预测性能的有效途径; 操玮等<sup>[25]</sup> (2020) 构建了 LSTM 与 GARCH 族混合模型预测人民币汇率波动率, 实验结果表明, 混合模型的表现优于单一模型, 单一 LSTM 模型表现优于单一 GARCH 模型; Verma<sup>[26]</sup> (2021) 提出 GARCH-GJR-LSTM 混合模型预测原油波动率; Zeng 等<sup>[27]</sup> (2022) 将深度学习方法和广义自回归条件异方差混合数据采样模型 (GARCH-MIDAS) 相结合, 用 GARCH-MIDAS 模型处理宏观经济变量与股市波动率之间的异频问题, 并对短期波动率进行预测, 最后将预测的短期波动率作为深度学习模型的输入指标, 对股市已实现波动率进行预测。

## 二、VaR 风险测度的研究现状

VaR 的概念最早起源于 20 世纪 80 年代, 是指在一定持有期内, 一定置信水平下, 某种金融资产或组合可能遭受的最大损失, 它的数值与损失序列的分位数相对应。(RueyS 和 Tsay<sup>[28]</sup>, 2000) RiskMetric 是美国摩根银行于 1996 年开发的一种金融风险管理模型, 该模型首次说明 VaR 的计算公式和使用范围。巴塞尔委员会也明确提出了“商业银行的资本金配置应以 VaR 值为标准”的规定。该政策旨在激励各大商业银行和金融机构等金融组织采用 VaR 方法计量其金融风险, 以达到合理配置资本金和提高商业银行风险管理能力的目的。

到目前为止, 在理论和应用两个层面上, 有诸多学者都在对 VaR 进行了深入的探讨。VaR 主流的分类方法是由 Engle 和 Manganeli<sup>[29]</sup> (2004) 提出的, 将 VaR 的计量方法归纳为三类: 参数法、半参数法和非参数法。

### (一) 参数法

参数法主要是对金融资产的收益率序列分布进行假定, 并利用对收益率的波动率进行估计, 从而对该金融资产的 VaR 值进行度量, 比如: 美国摩根银行<sup>[30]</sup> (1996) 提出的 RiskMetrics 方法; Nelson<sup>[31]</sup> (1991) 提出运用指数 GARCH 模型测度 VaR 值的思想; 陈泽忠等<sup>[32]</sup> (2000) 运用 EGARCH 模型, 研究发现在中国股票市场上, 波动率对收益率冲击的反应具有非对称效果, 并且收益率与波动性之间存在着明显的正向相关关系; 江涛<sup>[33]</sup> (2010) 利用 GARCH 模型对上海股票市场进行了 VaR 的测算, 结果显示, 与常规的传统方法相比, 文章提出的 VaR 算法具有更高的效率; Xiao 等<sup>[34]</sup> (2012) 提出采用两步法对线性 GARCH 模型进行分位数回归估计, 研究证实两步法比条件分位数估计方法的效果更好。

### (二) 半参数法

半参数方法主要依据分位数回归或极值理论, 直接估计收益率序列的分位数。如: Koenker 和 Bassett<sup>[35]</sup> (1978) 提出的分位数回归模型在 VaR 风险测度中占据着极为

关键的地位；Engle 和 Manganelli<sup>[29]</sup>（2004）在分位数回归模型的基础上，提出了条件自回归 VaR 模型（CAViaR）；王新宇等<sup>[36]</sup>（2010）通过测算 VaR 值，将分位数回归理论与传统最小二乘回归模型的运用结果进行比较，得出结论：在低显著性水平的条件下，分位数回归模型适用于对分布呈现出厚尾特征的收益率数据进行 VaR 估计；王传美等<sup>[37]</sup>（2019）通过对比 GARCH 类模型和分位数回归对创业板市场隔夜风险的测度结果，研究比较发现分位数回归具有更优良的预测结果；闫昌荣<sup>[38]</sup>（2012）提出了基于流动性调整的 CAViaR，它可以直观地反映一种资产的流动性对其未来的风险所产生的冲击，并据此测算出经流动性调整的 VaR 值，从而更好地描述中国股票市场的流动性风险的特征。Diebold 等<sup>[39]</sup>（2000）首次将极值理论（Extreme Value Theory, EVT）应用于金融风险管理，准确地测量了金融资产收益率的极限分位数和尾部概率，从而极大地提高了对极端金融风险的管理水平；王艺馨等<sup>[40]</sup>（2012）利用极值理论 POT（Peak Over Threshold）模型，并非拟合损益的整体分布，而是仅使用上证指数日收益率的尾部数据，拟合尾部分布，进而计算出 VaR 风险价值；Yu 等<sup>[41]</sup>（2013）将多重分形波动率模型（Multifractal Volatility, MFV）与极值理论有机地融合起来，提出一种新的 VaR 值测算方法，并将该方法与传统 GARCH 模型的性能作比较，实证结果表明基于多重分形波动率模型和 EVT 方法的 VaR 值度量在高风险水平上表现优于 GARCH 类模型；许启发等<sup>[42]</sup>（2015）针对具有更加显著的极值特征的金融产品，对比分析了传统 GARCH 模型和与机制理论相结合的 GARCH 模型在度量极端风险方面的优劣，并通过实证分析发现：基于 GARCH-EVT 模型的风险测度结果更为准确；Ausín 等<sup>[43]</sup>（2014）对 GARCH 模型中的参数分布假设进行了放宽，并将贝叶斯半参数的思想引入 GARCH 模型中，以应用于估计和测度 VaR 值。胡宗义等<sup>[44]</sup>（2019）构建条件自回归 Expectile（CARE）模型对我国上证指数和深证成指的 VaR 进行度量，结果表明半参数 CARE 模型和 GARCH 模型在度量不同置信水平下的 VaR 值时各有优劣。

### （三）非参数法

非参数的 VaR 值度量方法主要以历史模拟法或者蒙特卡洛模拟法等方法为代表。如：黄晓<sup>[45]</sup>（2009）经实证分析得出结论：在样本数据匮乏的前提下，一般历史模拟法比混合历史模拟法对 VaR 值的度量结果更准确；吕永健等<sup>[46]</sup>（2019）在传统历史模拟法和滤波历史模拟法（FHS, Filtered Historical Simulation）的基础上构造了 BHW（Bootstrapped Hull and White）历史模拟法，相较于传统历史模拟法，该方法在 VaR 预测上有更优异的表现；Wang<sup>[47]</sup>（2002）利用蒙特卡洛模拟法推导出了增量 VaR 的精确分布，表明增量 VaR 的推断依赖于仿真方法；肖佳文等<sup>[48]</sup>（2016）提出一种新的非参数 VaR 度量方法，即将金融市场状态进行分类，并使用核密度估计法拟合分

类后的几种子分布,进而得到损益的总体分布以测算 VaR 值;Taylor<sup>[49]</sup>(2000)提出的 QRNN 模型是近年来应用较为频繁的一种非参数方法,它将神经网络的非线性拟合能力和分位数回归的金融解释力度结合在一起,因此该融合模型具有极好的拟合与预测性能。该方法不但能对被解释变量的整个条件分布进行全面的建模分析,而且还能对金融市场数据的各种复杂的非线性结构进行拟合。

### 三、简要评述

国内外学者针对波动率预测和VaR值计量的研究上取得了一定的成果,这些研究对于我们了解波动率和金融风险计量的基本方法提供了理论指导,可以得出如下结论:

(1) 波动率预测的研究起步较早,但将神经网络应用于波动率预测是较为新兴的领域。早期相关学者针对于波动率时序的不同特征,着力于在GARCH模型的基础上进行延伸、拓展,如Threshold ARCH、EGARCH、GJR-GARCH等传统计量模型。近些年来,随着人工智能的兴起与发展,深度学习算法在非线性的特征处理方面表现出优异的性能。许多学者考虑到金融市场的收益率序列存在非线性特点,进而将深度学习相关算法引入到波动率预测的研究工作中。尤其是具备长记忆性的LSTM模型,诸多学者的研究表明该模型在波动率预测方面的表现要优于传统模型。

(2) VaR测度方法通常在传统参数法、半参数和非参数法的框架下展开,参数法主要以GARCH模型及其扩展模型为代表,半参数法主要依据分位数回归或极值理论,非参数法的传统模型主要以历史模拟法为代表。近些年来,有越来越多的学者将深度学习算法用以改进传统非参数方法,特别是以Taylor提出的QRNN模型为代表,为VaR测度提供了深度学习算法与分位数回归模型结合的新思路。



### 第三节 研究思路及方法

#### 一、研究思路

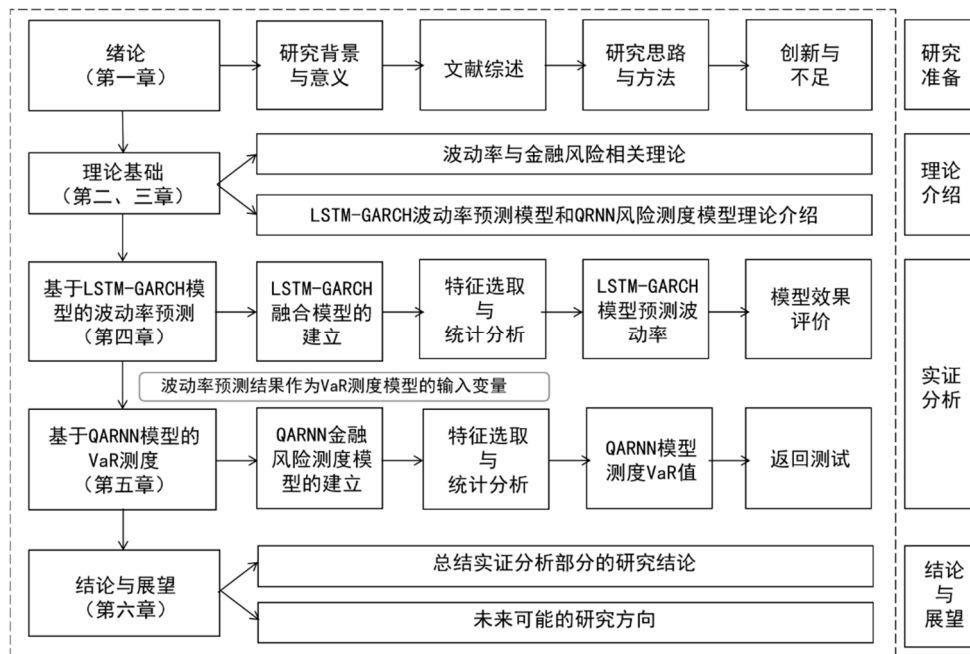


图 1-2 文章研究思路

文章将传统计量模型与深度学习算法结合,在融合模型的基础上进行波动率预测与 VaR 测度的实证分析,并比较模型结果。研究的路线如图 1-2 所示,首先,分析国内外有关波动率预测和 VaR 测度的相关文献,并进行总结。其次,界定了波动率与 VaR 的金融概念,并对相关的计算方法进行了梳理。在整理了传统 GARCH 模型和分位数回归模型的相关理论的基础上,引入深度学习算法,构建了基于 LSTM-GARCH 融合模型的波动率预测模型和基于 QARNN 模型的 VaR 测度方法。再次,基于以上建立的模型进行实证分析,运用 LSTM-GARCH 融合模型预测波动率,并将融合模型的预测值和单一模型的预测值与波动率真实值进行比较,并分析融合模型的改进效果。将预测性能最优的模型的波动率预测值作为 QARNN 模型的输入变量,进一步地对 VaR 值进行预测。最后,根据相关理论与实证结果,给出了相关的结论与展望。

本文共包括六个部分,第一部分为绪论,第二部分为金融市场波动率及金融风险的基本理论,第三部分为波动率预测与 VaR 测度模型的构建,第四部分为基于 LSTM-GARCH 混合模型的波动率预测,第五部分为基于神经网络分位数自回归的 VaR 测度,第六部分为研究结论与展望。

第一部分主要介绍本文的研究背景和研究意义,回顾了金融市场波动率预测的相关文献和金融资产 VaR 度量方法,在此基础上做出总结,并对其进行了简要地阐述,最后对本文的总体框架以及可能的创新之处进行阐述,同时也说明了本文存在的一些

不足。

第二部分是金融市场波动率及金融风险的基本理论，主要介绍本文的理论基础。首先，对波动率的基本概念、分类以及性质进行阐述，其中波动率按照其计算方法和应用的不同可归为隐含波动率、历史波动率等类，波动率分类不同对后续波动率预测以及 VaR 值估计都会存在一定影响，因此，了解波动率相关理论知识并确定其计算形式是至关重要的。随后，对金融风险的概念、分类及度量方法进行详细介绍，并着重说明金融风险度量方法中的 VaR 方法，为后续的模型构建奠定理论基础。

第三部分主要利用深度学习算法和传统计量模型相结合，构建股指波动率预测与 VaR 测度模型。波动率作为影响 VaR 值估计的重要因素，传统的方法通常是使用 GARCH 模型对其进行预测，虽然传统计量模型具有较好的金融解释力度，但我国股票市场存在较为明显的非线性特征，传统计量方法在拟合非线性变量方面的表现难以收到令人满意的成效。因此，考虑到波动率存在长记忆性和非线性，本文将长短期记忆模型（LSTM）与传统的 GARCH 模型相结合，构建 LSTM-GARCH 混合模型对波动率进行预测；VaR 本身即为分位数的概念，因此考虑将分位数自回归与神经网络相结合，并将 LSTM-GARCH 模型预测的波动率作为输入变量，构建神经网络分位数回归（QARNN）模型。融合模型既考虑到深度学习算法在模拟非线性特征上的优势，又继承了传统计量模型的可解释性，能够有效提高预测模型准确性。

第四部分主要是基于 LSTM-GARCH 混合模型的波动率预测。本文选取上证综指为研究对象，构建 4 种单一模型和 7 种 LSTM-GARCH 混合模型进行股指波动率的拟合与预测，并选取 MSE、MAE 等 7 种损失函数作为评价指标，将混合模型与传统 GARCH 模型、单一 LSTM 模型预测的波动率结果进行模型效果比较、评价。

第五部分主要基于神经网络分位数自回归的 VaR 测度。首先，将第三章 LSTM-GARCH 混合模型预测的波动率结果和滞后 3 期的 VaR 作为输入变量，构建分位数回归（QR）、分位数自回归（QAR）、神经网络分位数回归（QRNN）、神经网络分位数自回归（QARNN）共 4 种模型；随后通过上述 4 个模型对 VaR 值进行度量，与此同时，可以通过网格搜索和交叉验证（GridSearchCV）的方法寻求模型的最佳参数；最后，通过返回测试，如失败天数、失败率、似然比检验的 P 值和相对误差率等指标，评价 4 种模型的 VaR 风险测度效果。

第六部分是本文的研究结论与展望。总结归纳本文在波动率预测和 VaR 度量两方面的研究成果，并指出波动率预测和金融风险度量未来可能的研究的方向和问题。

## 二、研究方法

### （一）文献研究法

本文在大量查阅文献和收集资料的基础上,本文对国内外有关波动率预测和 VaR 测度的研究进行了归纳与总结,了解相关波动率和金融风险的概念、性质以及相关预测方法。并在研究、总结现有的理论和方法的基础上,本文将经典计量经济学模型和深度学习算法有机结合在一起,并将融合模型应用到实际的波动率预测与 VaR 测度研究中,明确研究思路与实证分析框架,为本文后续研究夯实理论根基。

## (二) 实证分析法

本文选取 2012-2022 年上证综指开盘价、收盘价等时间序列数据,首先利用单一的 GARCH 模型和 LSTM 神经网络以及二者的融合模型预测上证指数的波动率,并选取多种损失函数作为评价指标,对比、分析单一模型和融合模型的预测性能。其次,在现有的神经网络分位数回归(QRNN)的基础上,考虑了滞后内生变量,将波动率和滞后三期收益率作为输入变量,构建神经网络分位数回归(QARNN)模型进行 VaR 测度。最后,在返回测试中,运用多种评价指标来评价分析各种 VaR 测度模型的表现效果。

## (三) 比较分析法

本文在波动率预测中选取多种损失函数,对单一 GARCH 模型和 LSTM 模型以及多种 LSTM-GARCH 混合模型的波动率预测精度进行比较。该方法不仅包括相同模型、不同损失函数的横向比较,还包括不同模型、相同损失函数的纵向比较,以取得更加全面、准确的模型评价结果,选取最佳波动率预测模型为后续 VaR 测度模型提供波动率作为输入变量。

# 第四节 创新与不足

## 一、可能的创新点

在方法上,本文利用深度学习算法和传统计量模型相结合,构建股指波动率预测与 VaR 测度模型。

在波动率预测方面,GARCH 模型是应用较多的传统计量模型,具有较好的金融解释力度,但受到其自身条件限制,在拟合时间序列数据的非线性特征方面表现欠佳,而指数的波动率往往呈现非线性的特征。再者,考虑到波动率自身的长记忆性特征,本文首先通过上证指数的日度对数收益率滚动估计 GARCH 模型的参数,提取波动率的相关特征,并将估计出的参数与量价数据作为 LSTM 模型的输入变量,以此构建 LSTM-GARCH 混合模型,并选取多种损失函数作为模型评价指标,进行单一模型与融合模型的预测精度对比,通过将经典计量模型和深度学习算法融合的方式来提高模型整体对波动率预测的准确性。在 VaR 测度方面,本文在 Taylor 提出的神经网络分

位数回归模型 (QRNN)<sup>[42]</sup>的基础上,将神经网络引入分位数自回归构建神经网络分位数自回归 (QARNN) 模型,该模型相较于 QRNN 模型,考虑了滞后内生变量的影响,即在 QRNN 模型以波动率为输入变量的基础上增加了滞后三期的 VaR 值作为输入特征。并将该模型的 VaR 测度结果与分位数回归 (QR) 和分位数自回归 (QAR) 两种传统计量模型和 QRNN 模型三种模型进行对比,尝试在传统计量模型和 QRNN 模型的基础上提高 VaR 金融风险计量的准度,金融风险管理提供新思路。

## 二、不足之处

数据方面,从理论上讲,通过高频数据预测已实现波动率,再进一步测度 VaR 值的方法会更加精确。但受到数据资源获取的限制,本文仅取得到上证指数的日度数据,只能通过对日度波动率的预测,进而实现对金融风险进行量化。

风险量化方法方面, VaR 方法存在不满足次可加性和不足以表达尾部风险问题,在提供了一种量化风险的简单思路的同时也损失了一定的准确性。

模型算法方面,深度学习本身发展时间不是很长,还存在很大的发展空间,且相关算法更新速度快,还需要投入更多的时间、精力去进行更深一步的学习。

## 第二章 金融市场波动率及金融风险的基本理论

### 第一节 金融市场波动率的基本理论

#### 一、波动率的理论基础

从统计意义上来说,波动率是指在连续复利计息的条件下,一个特定时期内某一特定金融资产或投资组合收益率的标准差。在金融领域,波动率是金融资产价格的波动程度的一种度量,它反映了标的资产收益率的不确定性。一般而言,当金融市场中的股票、债券等金融产品出现较大的波动性时,意味着此时会出现较大的投资回报不确定性;相反,随着市场的波动性变得越来越低,投资回报率也变得越来越低。从实质上来说,波动性指的是一种对资产收益率不确定性的量化方式,从而可以反映出金融产品所面临的风险水平的高低。

从经济学的角度来说,产生波动率的主要原因可以归纳为以下三个方面:①系统性风险,即宏观经济的运行所造成的冲击;②非系统性风险,即企业受到某个特定事件带来的冲击;③投资者心理状态或对未来预期的变化对股票价格所产生的作用。综合以上原因可以得到,股票市场的波动率是一个极为复杂的变量,它是由多种因素共同影响而形成的。

#### 二、波动率的分类

按照计算方法与应用的不同,可将波动率划分为如下类别:

(1) 历史波动率法。历史波动率是通过某一特定金融资产的历史交易信息进行计算,进而求得该资产的波动率。历史波动率有以下几类:

① 样本标准差。在某一确定的历史时期内,用收益率的标准差来衡量这一时期内资产价格的波动率,计算方式如下:

$$u_i = \ln \left( \frac{S_i}{S_{i-1}} \right)$$

$$\sigma_n^2 = \frac{\sum_{i=1}^m (u_i - \bar{u})^2}{m - 1}$$

$$\bar{u} = \frac{\sum_{i=1}^m u_i}{m}$$

其中,  $S_i$  是第  $i$  时刻的资产收盘价,  $u_i$  是第  $i$  时刻资产相对于前一天的对数收益率,  $m$  是观测区间的天数,  $\sigma_n$  为第  $n$  天对数收益率的标准差。

② 极差波动率。以上所述的样本标准差波动率计算方法,是一种常用的统计学算

法, 后来诸多研究者对此进行了进一步的完善, 并给出了极差波动率的概念, 主要有以下几种形式:

i. Parkinson 法 (1980), 使用在一段时间间隔中的每天的最高价和最低价两个交易数据, 使用极端差值来计算波动率。这种方法的优点是它可以使用更短的时间来收敛到实际的波动率, 缺点是它没有将隔夜、价格存在漂移等特性考虑进去, 而且它不一定是一个无偏估计, 而是属于一种经验性的研究, 其具体的计算模型如下:

$$\sigma = \sqrt{\frac{1}{4m \ln 2} \sum_{t=1}^T (\ln(H_t/L_t))^2}$$

ii. Garman-Klass 法 (1980), 使用交易区间内每日的量价数据来计算波动率。这种方法的主要具有以下两方面的优点: 一是以较快的速度逼近于真实的波动率, 二是运用估计量除以调整因子来校准误差, 从而获得方差的无偏估计量。劣势是它没有将隔夜、价格存在漂移等特性考虑进去, 计算方法如下:

$$\sigma = \sqrt{\frac{1}{m} \sum_{t=1}^T (\ln(H_t/L_t))^2 - \frac{1}{m} \sum_{t=1}^T (2 \ln 2 - 1) \ln(c_t/c_{t-1})}$$

iii. 基于以上模型和方法, Rogers 和 Satchell (1991) 认为, 针对于价格存在漂移的现象, 一个新的波动率计算方法如下:

$$\sigma = \sqrt{\frac{1}{m} \sum_{t=1}^T \ln(H_t/c_t) \ln(H_t/O_t) - \ln(L_t/c_t) \ln(L_t/O_t)}$$

iv. Garman Klass 和 Yang Zhang 在研究了隔夜因素对波动率的影响后, 得到了如下波动率模型:

$$\sigma = \sqrt{\frac{1}{m} \sum_{t=1}^T \left( \frac{1}{2} \ln(H_t/L_t) \right)^2 - (2 \ln 2 - 1) \left( \ln \left( \frac{H_t}{L_t} \right) \right)^2 + \ln(L_t/c_t)}$$

v. Yang Zhang 借鉴先前学者的研究成果, 弥补上述模型存在的一些缺陷, 提出了波动率预测新方法:

$$\begin{aligned} \sigma &= \sqrt{\sigma_{\text{隔夜}}^2 + k \sigma_{\text{当天}}^2 + (1-k) \sigma_{RS}^2} \\ \sigma_{\text{隔夜}}^2 &= \frac{1}{m-1} \sum_{t=1}^m (\ln(O_t/O_{t-1}) - \overline{\ln(O_t/O_{t-1})})^2 \\ \sigma_{\text{当天}}^2 &= \frac{1}{m-1} \sum_{t=1}^m (\ln(c_t/O_{t-1}) - \overline{\ln(c_t/O_{t-1})})^2 \\ \sigma_{RS}^2 &= \sqrt{\frac{1}{m} \sum_{t=1}^m (\ln(H_t/c_t) \ln(H_t/O_t) + \ln(L_t/c_t) \ln(L_t/O_t))} \end{aligned}$$

通过以上分析, 我们发现每一个新模型都是在原有模型的基础上加以完善和改进的结果。应当指出的是, 从现有的理论与研究来看, 以上几类模型在实证分析研究中

并没有什么绝对的优势或劣势，我们应该依据自身研究的需求来选择合适的模型。

③已实现波动率（Realized Volatility, RV）。已实现波动率利用了数学中积分的思想，其具体计算公式如下：

$$RV_t = \sum_{i=1}^n r_{t,k}^2$$

已实现波动率的计算方法是 $t$ 日短时高频收益率的平方值加总， $n$ 为日内总时长与时间间隔之比。

（2）隐含波动率法。在期权定价中，很多时候我们无从获取历史波动信息，此时可以通过对未来的预期进行折算，得到当前的波动率。在期权市场上，理论上可以采用 Black-Scholes 期权定价模型，可以从一个期权的价格中，反向推导出该期权的隐含波动率。

$$c = Se^{-qt}N(d_1) - Ke^{-rt}N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r - q + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

其中， $c$ 为欧式看涨期权价格， $S$ 为标的资产价格， $r$ 为无风险利率， $K$ 为执行价格。

### 三、波动率的性质

构建一个理论模型，不能脱离实际观察和对实践的总结。从学者们对各个国家金融市场的观察研究来看，除了无法被直接观测这一独特的性质之外，波动率还具有以下特性。（Christopher 和 William<sup>[50]</sup>，1990）

（1）尖峰厚尾。以往建立的波动率预测模型，一般都是假定金融资产的收益率序列分布服从正态分布，即偏度为 0，峰度为 3 的模型设定。然而，大量的经验分析和实证研究显示，我国金融资产收益的波动率分布呈现出尖峰厚尾的特征，意味着有更多数据分布在均值和尾部附近。

（2）波动率聚集性。波动聚集性指的是金融时间序列会出现大波动聚集在一些时间段发生、小波动聚集在一些时间段发生的现象，波动之间存在正相关关系。金融市场中的大部分数据都具备这种特点。

（3）长记忆性。许多学者研究了具有长期记忆性的波动率时间序列数据，发现其绝对或幂指的衰减速率很缓慢，并且具有很强的自相关性。这主要体现在：一个外部冲击事件在接下来一段时间内会对收益率的波动产生持续性的作用，而这种作用会随着时间的推移而逐渐减弱，并且会在将来的波动变化中继续存在。

（4）非对称性。波动率的非对称性是由市场上的两类信息所引起的，即积极消息和消极消息的不对称效应。一般来说，与好的正面信息相比，波动性更容易受到超

乎预期的负面信息的影响。而在资产收益率为正值的情况下,收益率的波动性比较小,价格表现也比较平稳;而在收益率为负值时,收益率的波动程度就会加大。从某种意义上说,这也是体现了金融投资者存在风险规避的特征。

(5) 跳跃性。在股票市场中,一旦有重要的事情或者关键性的新闻发生,就会引起股市时序数据出现剧烈的变动。相对于正常的微小波动,这种价格震荡程度是非常大的,呈现出波动率的跳跃性。

(6) 溢出效应。各种金融资产的波动率彼此之间存在着交互作用。一种资产价格的变动将会以某种方式传递给其它的金融产品,从而影响到其它的金融产品价格。

## 第二节 金融风险的基本理论

### 一、金融风险及其分类

对风险的界定可分为有广义和狭义两类。一般意义上的风险是把风险看成是一种不确定因素,会产生亏损、盈利和无亏损三种后果,因此,其重点在于收益的不确定因素;狭义上的风险则是认为,风险只能带来损失,因此它主要强调损失的不确定性。在实际生活中,特别是在金融领域中,金融风险通常表现为多种风险因素的叠加。

本文的研究问题是金融风险计量,以金融风险为主要研究对象。中国金融市场的风险类型主要包括市场风险、信用风险、操作风险及系统性风险。

市场风险就是因为市场价格的变化,使得一项投资的公允价值或者未来的现金流量产生了变化,这种变化有可能在未来造成某种损失。市场风险主要包括外汇风险和利率风险。外汇风险也被称为“汇率风险”,是国际金融体系中最难以管理的一种外在风险。通常,影响汇率变化的因素有:国际收支,通货膨胀,利率,市场预期利率,国家宏观经济政策,货币当局的介入等。按汇率风险可分为交易风险、会计风险和经济风险。

信用风险是一种从金融机构建立之日起就始终伴随着的历史悠久的风险,它经常会出现在买卖两方履行合同义务的环节中,一般情况下会因为一方违约、信用等级发生变化或者因标的资产价格的变化而产生信用风险。尽管在现实生活中难以对信用风险进行有效的辨识与定量,但学术界还是发展了如 Merton 模型、KMV 模型等具有代表性的信用风险量化模型。

操作风险是由于人员操作存在失误或不当行为、机构或企业内部管理不规范等多方面人为因素而造成损失的风险。与其他风险相比,操作风险的涵盖的范围要宽泛得多,导致其产生的因素也要复杂得多,并且获得相关数据的困难程度更大,亦或是无法定量数据,因此,对操作风险进行度量的难度要远远高于以上两种风险。

系统性风险指的是金融机构所面临的风险的总和,如果发生系统性风险,那么就



整个金融体系都将蒙受巨大的损失，它不但会对整个股票市场的正常运作造成一定的影响，还会与其它金融风险发生共振，进而引起更大的灾难性危机。

当前，针对不同类型的金融风险，从理论上分析不同类型风险之间的关系，包括完善不同类型风险的度量方法，准确刻画不同类型风险之间的关系等问题仍是当前金融领域研究的热点。在金融风险的量化研究中，研究者和投资者对市场风险的重视程度是最高的，这是由于，金融市场风险不但是金融市场中的每个参与者都会遭遇的一种风险，还是其它几种类型的金融风险发生的根本因素。因此，在金融风险之中，金融市场风险处于十分特殊的位置。通常意义下的金融风险度量，指的是金融市场风险的度量。

## 二、金融风险度量的传统方法

要对金融市场参与者所面对的市场风险进行有效的控制与管理，就必须对其进行测度。传统的测量风险的方式主要可归类为如下几种：

(1) 简单算术法。

在简单算数的方法中，是利用价格的差异来对风险进行定量。价差率是对单一证券的波动和风险的一个度量，价差率值的大小与一支股票所面临的风险水平的高低是成正比例的，它的计算公式如下：

$$\text{价差率} = \frac{2(H_t - L_t)}{(H_t + L_t)},$$

其中 $H_t$ 表示相应时间段内股票的最高价，表示 $L_t$ 相应时间段内股票的最低价。

(2) 波动性分析。

①单个资产的标准差。在多数研究中，资产收益率 $r$ 通常被作为随机变量来处理。令 $E(r)$ 表示收益率的期望值， $\sigma^2(r)$ 表示收益率的方差，根据收益率随机变量为离散型或连续型的两种不同情况，标准差存在下述两种计算方法：

$$E(r) = \sum_{i=1}^n r_i p_i$$

$$\sigma^2(r) = \sum_{i=1}^n (r_i - E(r))^2 p_i$$

其中， $r$ 为离散型随机变量， $r_i p_i (i = 1, 2, 3, \dots, n)$ 表示收益率取值与其对应概率的乘积。

$$E(r) = \int_{-\infty}^{+\infty} r f(r) dr$$

$$\sigma(r) = \int_{-\infty}^{+\infty} |r - E(r)| f(r) dr$$

其中， $r$ 为连续型随机变量， $f(r)$ 为其对应的概率密度函数。

②资产组合的标准差。假设某一资产组合由 $m$ 个资产构成，每种资产对应权重为 $w_i (i = 1, 2, 3, \dots, m)$ ，且 $\sum_{i=1}^m w_i = 1$ ， $\sigma_i$ 表示资产 $i$ 的标准差，则有：

$$\sigma^2 = \sum_{i=1}^m \sum_{j=1}^m w_i w_j \text{cov}(r_i, r_j),$$

其中,  $\text{cov}(r_i, r_j)(i, j = 1, 2, 3, \dots, m)$  为资产  $i$  与资产  $j$  的协方差。

当  $r_i$  和  $r_j$  为离散型随机变量时, 有:

$$\begin{aligned} \text{cov}(r_i, r_j) &= E\left((r_i - E(r_i))(r_j - E(r_j))\right) \\ &= \sum_{k=1}^n (r_{ik} - E(r_{ik}))(r_{jk} - E(r_{jk})) \end{aligned}$$

当  $r_i$  和  $r_j$  为连续型随机变量时, 有:

$$\text{cov}(r_i, r_j) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (r_i - E(r_i))(r_j - E(r_j)) f(r_i, r_j) dr_i dr_j$$

根据  $\text{cov}$ 、 $\sigma$ 、 $\rho$  三者之间的关系, 可推出如下关系式:

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^m \sum_{j=1}^m w_i w_j \rho_{ij} \sigma_i \sigma_j \\ &= \sum_{i=1}^m w_i^2 \sigma_i^2 + \sum_{i=1}^m \sum_{j=1, j \neq i}^m w_i w_j \rho_{ij} \sigma_i \sigma_j \end{aligned}$$

以矩阵形式表达上式可得:

$$\sigma^2 = [W_1 \quad \dots \quad W_m] \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1m} \\ \vdots & & & & \vdots \\ \sigma_{m1} & \sigma_{m2} & \sigma_{m3} & \dots & \sigma_m^2 \end{bmatrix} \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix},$$

其中,  $\sigma_{ij}$  表示资产  $i$  的收益率  $r_i$  与资产  $j$  的收益率  $r_j$  的协方差。

(3) 下方风险度量方法。在现实的风险控制过程中, 投资者更多的是考虑到了投资组合的收益小于期望收益的情形, 基于该情形而得到的就是下方风险度量方法, 其中 LPM 法和 VaR 法是最具代表意义的。在本文中也选取 VaR 方法对风险进行量化。

### 三、金融风险度量的 VaR 方法

VaR 又称为在险价值, 为 JP Morgan 在 20 世纪 90 年代开发的一种风险测量方法, 主要应用于度量资产所面临的市场风险, 其相关的理论和方法是目前量化金融风险的重要手段。VaR 方法与传统的风险度量方法相比, 具有可比性、全面性、直观性、简单性和易于理解等优点。

VaR 以概率统计理论为基础, 采用合理的、科学的、理性的数学模型, 借助现代的计算机技术, 对市场数据展开了研究和计算。VaR 不但可以较为精确地度量出市场

的金融风险,全面地反映出市场各个方面的风险状态,能够对金融风险进行较为完整的测度,还具有简单的可操作性,便于金融机构对其面临的金融风险进行有效的测度,所以 VaR 的应用场景十分广泛,被有关金融机构和监管部门普遍使用于日常的风险控制与管理。

VaR 指的是一项金融资产或资产组合在一定持有期 $I$ 内和一定显著性水平 $\alpha$ 下所面临的 $\text{最大损失}$ 。其数学定义为:

$$P(Lost \leq VaR(\alpha)) = 1 - \alpha,$$

其中,  $Lost$ 为持有期 $I$ 内资产的亏损值。

用 VaR 来测度的风险是在一般的股市波动性条件下发生的,因此 VaR 值的计算方法以资产组合面临的风险敞口为依据,其计算方法简单、成本低廉、具有一定的可比性。这种金融风险量化方法不需要单独考虑利率、汇率等多种影响资产价格波动的因素对损失的影响,而是综合考量各种风险因子对损失的影响,并对其进行统一的定量。因此, VaR 风险度量方法既可以自适于复杂多变的金融市场,又与世界各地日益一体化的发展潮流相一致,成为目前金融风险度量方法的主流工具。

在 VaR 度量中,如下几种方式是比较具有代表性的。

(1) 德尔塔—正态分布法。在这种模型中,假定了资产收益率服从正态分布。那么,该资产的 VaR 值就是收益率标准差和对应的显著性水平下分位数的乘积:

$$VaR = Z_{\alpha} \cdot \sigma \cdot \sqrt{\Delta t},$$

其中,  $Z_{\alpha}$ 是标准正态分布在显著性水平 $\alpha$ 下的分位数,通常情况下 $\alpha$ 取 0.1 或 0.05。 $\sigma$ 为收益率的标准差,  $\Delta t$ 为持有期。

虽然这种新的计算方式极大程度上简化了风险量化的计算,但是在现实的金融市场中,大部分的收益率分布都呈现出尖峰厚尾的特征。而该方法由于假定了资产收益率服从正态分布,所以其对 VaR 值的估计存在很大的偏差,从而导致了 VaR 的估计值偏低,因此存在低估风险的缺陷。

(2) 历史模拟法。这种分析方法的思路是:将来资产收益率的变化是历史的再现。历史模拟方法是以实证为基础,用重建金融资产的历史收益率序列来模拟这一类资产的收益与亏损的分布,将收益率的历史分布替代为其真实分布,并通过在给定置信度下的分位数来计算出 VaR 的数值。

这种方法具有以下优势:它不需要假设资产的收益率分布,也不需要考虑投资组合中每个金融资产之间的相关性。该方法的缺点是,这种模型假设了市场因子的变动与历史的变动规律一致,从而造成了模型中的“鬼影效应”(Ghost Effect)的产生,另外,该方法对数据的要求较高,必须通过对历史数据进行大量的模拟才能得到准确的预测。

(3) Monte Carlo 模拟法。该方法无需对金融资产的收益率分布进行假定,而是

运用几何布朗运动（GBM）模型来刻画资产价格变化。Monte Carlo 模拟法的优势是可以根据不同场景，通过人工设置假设数据，以此来弥补样本数量匮乏的缺陷，并且能通过假设收益服从不同的分布进行多次拟合，与实际金融市场中的收益分布更为符合。

## 第三章 波动率预测与 VaR 测度模型的理论基础

### 第一节 LSTM-GARCH 模型

#### 一、GARCH 模型

##### (一) 标准 GARCH 模型

对金融资产的波动率序列进行建模,对于资产定价和金融风险管理等都具有非常关键的作用。早期,Markowitz 提出用收益率的方差来估计波动率,但是因为波动率自身存在集聚效应,且真实的波动性并不是一成不变的,因此这种模型会与实际情况产生一定的偏差。Engle<sup>[1]</sup>(1982)提出了自回归条件异方差模型(AutoRegressive Conditional Heteroskedasticity Model, ARCH),该模型的假设条件进行了放宽,不再包括方差为常数这一假定,与现实中收益率的异方差特性相一致。因此,这种新的模型能够反映出序列中方差和残差之间的动态关系。ARCH(q)模型可表示如下:

$$\begin{aligned}\varepsilon_t &= \sigma_t v_t & v_t &\sim N(0,1) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2,\end{aligned}$$

其中 $\varepsilon_t$ 为资产收益率的扰动项,  $\sigma_t^2$ 为 $\varepsilon_t$ 的条件方差即波动率的平方。

虽然 ARCH 模型的结构相对比较简洁,然而在实际应用中,建立 ARCH 模型所需要估计的参数数量很多,而且在高阶情形下,其精度也会受到影响。此外,为了确保方差为正数,要求参数是恒为正数的,因此,采用这种方法的要求比较苛刻,并且与其约束条件与现实情形存在着一定的偏差,这些缺陷都在很大程度上制约了 ARCH 方法在实际分析问题中的应用。

Bollerslev<sup>[2]</sup>(1996)在 ARCH 模型的基础上进一步提出了广义自回归条件异方差模型(Generalized AutoRegressive Conditional Heteroskedasticity Model, GARCH),该模型是将条件异方差转化为 ARMA 过程,增加了 p 阶滞后和 q 阶残差滞后。与 ARCH 模型相比,其具有更少的约束条件和待估参数,极大地改善了模型的精度。GARCH(p, q)模型如下:

$$\begin{cases} y_t = x_t \gamma + \varepsilon_t \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases}, \quad (3.1)$$

其中, p 表示 GARCH 的阶数, q 表示 ARCH 的阶数。使得 GARCH(p, q)平稳的条件是 $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$ 。且随着 $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j$ 逐渐逼近于 1,衰减程度也逐渐减慢。

当  $p = 1, q = 1$  时, 就构成了 GARCH(1,1)模型。且当  $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j$  逐渐逼近于 1, 则说明这一时期的数据存在波动聚集性。这种模型不仅简单明了, 而且可以很好地体现出序列的波动聚集性等特征, 因此被普遍用于当前对金融市场的波动性进行预测。标准化的 GARCH(1,1)模型形式如下所示:

$$\begin{cases} y_t = x_t \gamma + \varepsilon_t \\ \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \end{cases}$$

## (二) 不同残差分布假设下的 GARCH 模型

在早期研究中, 对于进行 GARCH 模型的构建时, 大多数研究者都是采用假定收益序列服从于正态分布的方法。尽管此种方法相对简单易行, 但与实际的收益率序列的真实分布之间还有很大的偏差: 在实际的金融市场中, 金融产品的收益率不服从正态分布, 反而相对于正态分布具有更厚的尾部和更高的峰度, 并且呈现出偏态分布的特点。所以, 将金融产品的收益率序列假设为正态分布, 并用以实证研究是存在不合理之处的。

(1) GARCH(1,1)-t 分布。t 分布又可称之为学生氏分布, 在金融领域的研究中, 除正态分布外, t 分布也是常用的分布之一。William Gosset 首先引入了 t 分布有关的理论, 并在此基础上以 Student 为笔名发表了与 t 分布有关的文章。随后, 有其它的研究者不断对 t 分布进行改进与扩展。t 分布的概率密度函数表达式如下:

$$f(y, v) = \frac{\Gamma\left(\frac{v+1}{2}\right) \left(1 + \frac{y^2}{v}\right)^{-\frac{v+1}{2}}}{(v\pi)^{\frac{1}{2}} \Gamma\left(\frac{v}{2}\right)},$$

其中,  $\Gamma$  为 Gamma 函数,  $v$  为自由度。当自由度  $v$  趋向于正无穷时, t 分布趋向于标准正态分布的概率密度函数。总的来说, t 分布相对于标准正态分布来说, 分布特征呈现尖峰厚尾的特点。

GARCH(1,1)-t 分布的具体模型如下:

$$\begin{aligned} y_t &= x_t \gamma + \mu_t \\ \mu_t &= v_t \sigma_t, v_t \sim t(n) \\ \sigma_t^2 &= \omega + \alpha \mu_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned}$$

(2) GARCH(1,1)-GED 分布。广义误差分布 (Generalized Error Distribution, GED) 是一种常见的厚尾分布, 由 JP Morgan 的 Risk Metrics 提出, 在金融市场上被广泛地应用于基于 GARCH 模型的 VaR 风险度量方法。其对数似然函数的数学表达形式见下式:

$$\ln L(\theta) = -\frac{T}{2} \ln \left( \frac{\Gamma(1/r)^3}{\Gamma(3/r)(r/2)^2} \right) - \frac{1}{2} \sum_{t=1}^T \ln \sigma_t^2 - \sum_{t=1}^T \left( \frac{\Gamma(3/r)(y_t - x_t \gamma)^2}{\sigma_t^2 \Gamma(1/r)} \right)^{r/2}$$

GED 分布的概率密度函数如下:

$$f(x, n) = \frac{n}{2^{\frac{n+1}{2}} \times \lambda \Gamma(1/n)} \times \exp \left( -\frac{1}{2} \left| \frac{x}{\lambda} \right|^n \right),$$

$$\lambda = \sqrt{\frac{2^{2n} \Gamma(1/n)}{\Gamma(3/n)}}$$

其中,  $n$  为自由度,  $\Gamma(1/n)$  表示参数为  $1/n$  的 Gamma 分布,  $\lambda$  为尾部厚度参数, 即决定分布的尾部厚度。当  $\lambda > 2$  时, GED 分布呈现厚尾特征; 当  $\lambda = 2$  时, GED 分布趋近于正态分布; 当  $\lambda < 2$  时, GED 分布呈现瘦尾特征。

GARCH(1,1)-GED 分布的具体模型如下:

$$\begin{aligned} y_t &= x_t \gamma + \mu_t \\ \mu_t &\sim GED(\sigma_t^2, \lambda) \\ \sigma_t^2 &= \omega + \alpha \mu_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned}$$

### (三) GARCH 拓展模型

(1) EGARCH 模型。在实际的金融市场交易中, 投资者对股票收益的期望越低, 则股票收益率的波动越大, 反之当投资者对股票收益的期望越高, 则股票收益率的波动越小。这被称为杠杆效应。然而, 无论 ARCH 模型还是 GARCH 模型, 它们都无法模拟出收益率变动的非对称特征。基于此问题, 在 1991 年 Nelson 针对 GARCH 模型仅能体现波动性聚集性的缺陷, 又推出了一种新的指数 GARCH 模型 (EGARCH), 该方法的基本结构是:

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \left( \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} - E \left( \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right) \right| + \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right) + \sum_{i=1}^p \beta_i \ln \sigma_{t-i}^2$$

在 EGARCH 模型中, 标准差呈现为  $\sigma_t$  为对数形式  $\ln \sigma_t^2$ , 确保其不是负值; 模型中的参数  $\gamma_i$  为杠杆系数, 用以衡量波动率自身存在的杠杆效应特性。

(2) TGARCH 模型。另一种是一个与 EGARCH 模型相似的门限 GARCH (TGARCH) 模型, 它可以用来反映在金融市场中的广泛存在的杠杆效应。TGARCH (1,1) 模型是其最简单的形式, 并且具有如下的条件方差方程:

$$\sigma_t^2 = y_0 + y_1 u_{t-1}^2 + y_1^* u_{t-1}^2 d_{t-1} + y_2 \sigma_{t-1}^2,$$

式中的  $u_t$  为第  $t$  日的资产收益率,  $d_{t-1}$  为虚拟变量, 具体地,  $d_{t-1} = \begin{cases} 1, & u_t < 0 \\ 0, & u_t \geq 0 \end{cases}$ 。

(3) GIR-GARCH 模型。该模型同样能够刻画出波动的非对称性特征, 模型如下:

$$\sigma_t^2 = \omega + \sum_{i=1}^q (\alpha_i \varepsilon_{t-i}^2 + \gamma_i S_{t-1} \varepsilon_{t-1}^2) + \sum_{j=1}^p \beta_j \sigma_{t-1}^2$$

该模型将市场的冲击分成正负两各部分, 并加入到模型中, 并以各部分的系数来

判断其非对称效应。

## 二、LSTM 模型

长短期记忆（Long Short Term Memory, LSTM）网络是一类与梯度学习算法相融合的较为准确的网络结构。这是目前应用较多的 RNN 神经网络的一种变形，与普通的 RNN 网络结构在实质上相较几乎没有任何区别，LSTM 只不过是采用了一个不同的用于计算隐藏层状态的函数。传统 RNN 模型的隐藏层仅具有一个对短时的信息更灵敏的状态，而 LSTM 的隐藏层添加了一个对长时的信息更具存储能力的状态，这种状态就类似于一个“处理器”能够判断对历史数据选择保存还是遗忘，即单元状态（Cell State）。这种改进能够很好地克服 RNN 存在的长期依赖性。

RNN 模型都是由重复的简单神经网络模块构成的。LSTM 同样是相同结构的重复，但是与 RNN 结构不同，相对来说更具复杂性。图 2-1 所示 LSTM 神经网络的结构图，在  $t$  时刻该模型共有 3 个输入变量： $h_{t-1}$ 、 $c_{t-1}$ 、 $x_t$ 。其中  $h_{t-1}$  代表模型上一时刻的输出值， $c_{t-1}$  代表模型上一时刻的单元状态， $x_t$  代表模型在当前时刻的输入。

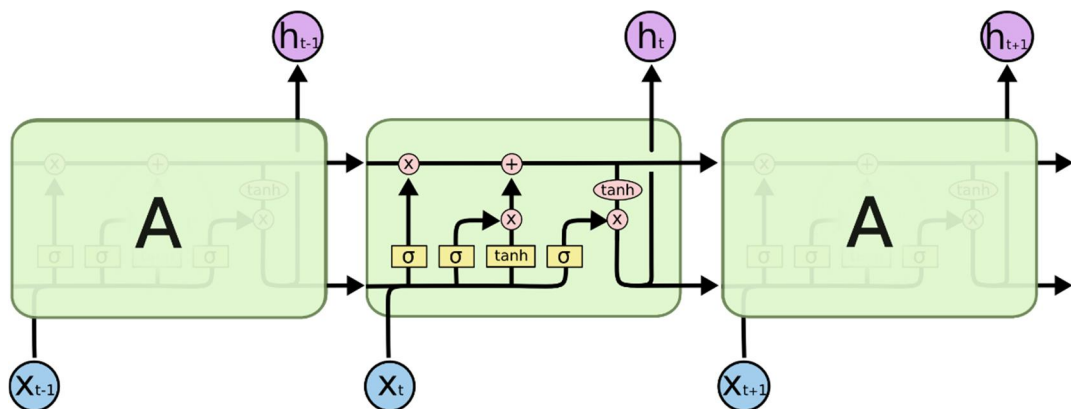


图 2-1 LSTM 神经网络结构图

LSTM 神经网络在学习过程中，是利用一个“门”的结构完成了对所学信息的记忆与遗忘。在门结构中，一个向量被输入，之后利用激活函数，将该向量转换为每个元素都在 0 到 1 之间的向量，最终将转换后的向量输出。门的控制功能是通过将这个输出向量与控制的向量相乘实现的，因为输出向量中的每一个在 0 到 1 之间的实数都代表赋予信息一定的权重，例如，0 代表“不保留任何信息”，1 代表“保留所有信息”，0 至 1 之间的权重代表“保留部分信息”。

LSTM 神经网络结构中存在三个门：输入门、遗忘门和输出门。门结构能够控制历史信息的更新与遗忘。

遗忘门是指在一个神经网络的建模过程中，对神经元在上一节点上的状态  $c_{t-1}$



的信息进行处理, 决定存储到当前时刻的单元状态 $c_t$ 中的信息和决定遗忘上一节点的多少信息。该门会输入向量 $h_{t-1}$ 和 $x_t$ , 通过 Sigmoid 激活函数进行运算, 输出一个所有元素都介于 0 和 1 之间的向量 $f_t$ , 并将其作为权重赋予到单元状态 $c_{t-1}$ 中的每一个元素, 从而达到对信息状态的控制。 $f_t$ 的数值愈趋向 1, 则表示前一时间节点中含有更多有用的信息;  $f_t$ 的数值越是靠近 0, 说明目前绝大多数的信息都是无效的, 因此赋予更小的权重, 即达到被“遗忘”的效果。计算过程可以用下式表示。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f),$$

其中 $W_f$ 和 $b_f$ 分别表示模型遗忘门的权重和偏置项,  $\sigma$ 表示 Sigmoid 激活函数。

输入门是用来确定在一个神经网络的建模过程中, 现时间节点的信息输入 $x_t$ 存储到当前时刻的单元状态 $c_t$ 中的信息量。输入门的通过下式实现。

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i),$$

其中 $W_i$ 和 $b_i$ 分别表示模型输入门的权重和偏置。

输出门是用于决定当前时刻的单元状态 $c_t$ 输出到当前时刻的输出值 $h_t$ 中的信息量。

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \cdot \tanh(c_t)$$

LSTM 神经网络采用门控制机制实现了对历史信息的遗忘与保留, 因此与 RNN 相比, 在处理时间序列数据方面, 特别是具有长记忆性的波动率序列, 具有更好的应用效果。

### 三、LSTM-GARCH 模型

将融合模型用于时间序列预测时, 已有的基础研究理论大多为先对数据集进行划分, 再利用多个模型对分割后的数据集进行拟合, 得到多个模型的预测结果后再进行加总, 从而得到融合模型的预测结果。常见的融合模型构建方法主要包括 HP 滤波分解法、MA 滤波分解法等。

HP 滤波法 (Hodrick Prescott) 是一种对时序数据进行分类的主要手段。这种滤波方法可以把一个时间序列分为两个部分, 分别是线性的趋势项和非线性的周期项。趋势数据序列具有显著的线性特点, 且整体趋势较为平滑。而在周期性数据中, 则表现出更明显的非线性和更强的波动性。在此基础上, 利用多种模型方法对时间序列的趋势项和周期项进行拟合, 并将二者相加得到的数据进行综合作为融合模型的最终结果,

从而达到对时间序列进行精确预测的目的。

MA 滤波法 (Moving average filter) 利用了时序中包含高波动和低波动两种不同特征波动性的特点, 将其按照某种方法进行了分离。其主要思路为: 当一组平稳的时间序列数据服从正态分布的情况下, 是严平稳的, 通常采用峰度来判断其是否服从正态分布。一般情况下, 当峰度为 3 时, 时间序列服从正态分布, 意味着时间序列的波动性较小, 为低波动成分; 如果峰度值太大或太小时, 那么该序列要么是低峰度要么是高峰度, 也就是说, 该序列为高波动成分。最后, 分别对高低波动率数据进行预测, 并对所得数据进行求和, 从而得出了融合方法的组合预测结果。

## 第二节 神经网络分位数回归模型

### 一、分位数回归模型

基于均值回归的概念, Koenker 和 Bassett<sup>[35]</sup> (1978) 引入了分位数回归 (Quantile Regression, QR) 这一新的理论。传统的最小二乘法主要关注的是分布的均值, 而忽视了对分布尾部数据特征的研究。与最小二乘方法相比, 分位数回归可以更为全面地捕捉条件分布的特征, 从而更好地刻画样本尾部的状态变化规律。在金融领域中, 学者往往更加注重对金融数据的尾部分布进行深入的研究, 因而, 分位数回归被大量地运用到了金融时序的研究中。近年来, 分位数回归的数学建模与分析方法得到了快速发展, 并获得了丰富的实证研究结果, 这不但使人们对经典的回归分析有了更深层次的认识, 同时也极大地拓展了分位数回归分析的理论内涵与应用范围。

分位数回归是以最小二乘法为理论依据进行线性扩展, 是建立在解释变量和被解释变量的条件分位数之间的函数关系。

$$Quant_{\tau}[Y/X = x] = \inf\{y: F(y/X = x) \geq \tau\}$$

通过上述公式, 可将分位点拓展为分位数回归模型:

$$y_t = x_t^T \beta(\tau), t = 1, 2, \dots, n$$

其中,  $\tau(0 < \tau < 1)$  是分位点;  $x_t = (1, x_{t,1}, x_{t,2}, \dots, x_{t,k})'$  是自变量构成的矩阵;  $\beta(\tau) = (\beta_0(\tau), \beta_1(\tau), \dots, \beta_k(\tau))'$  是回归系数向量;  $\varepsilon_t(\tau)$  是误差项。

上式中的  $\beta(\tau)$  可以通过线性规划的方法求解, 下式即为分位数回归的目标函数:

$$\min_{\beta} \sum_{t=1}^n \rho_{\tau}(y_t - x_t^T \beta), \quad (3.2)$$

其中  $\rho_{\tau}(u) = \begin{cases} \tau u, & u \geq 0 \\ (1 - \tau)u, & u < 0 \end{cases}, 0 < \tau < 1$ 。

分位点回归是建立在非对称形式的绝对值残差极小化的基础上进行模型拟合的。当残差不服从正态分布时,与最小二乘法相比,分位数回归具有更准确的估算精度和更高的鲁棒性。总体而言,与普通最小二乘法相比较,分位数回归方法有以下优势:第一、能较好地描述解释变量对被解释变量整个条件分布的影响,较好地避免了离群值对估计结果的影响;第二,不需要对模型中的随机扰动项进行任何分布上的假定,从而使得模型整体的稳定性得到改善。

在初始阶段,由于分位数回归模型估计参数的难度较大,因此其最开始的应用范围很窄。直到后续有学者在研究非线性分位数回归时,引入了内点方法,极大地减少了估算参数的困难,从而使分位数回归在金融、教育等许多领域得到了广泛的运用。

## 二、分位数自回归模型

在均值回归方法中,由于对模型残差的假设是服从正态分布,所以其计算结果仅能体现被解释变量的平均水平。Koenker 等<sup>[51]</sup>(2006)提出的分位数自回归(Quantile Autoregression, QAR)能够在不同分位数条件下模拟被解释变量和解释变量之间的映射关系,并且适用于处理具有异方差性的模型,在时间序列数据建模中起到了尤为关键的作用,在模拟具有尖峰厚尾特征的金融数据方面具有广泛的应用。

QAR 模型的具体表达形式如下:

$$Q_{y_t}(\tau | y_{t-1}, \dots, y_{t-p}) = \alpha_0(\tau) + \alpha_1(\tau)y_{t-1} + \dots + \alpha_p(\tau)y_{t-p} = z_t'\theta(\tau),$$

其中,  $\alpha_i(\tau)(i = 0, 1, \dots, p)$  为模型的待估参数, 可以通过最小化如下的目标函数来求得。

$$\arg \min_{\theta \in R^{p+1}} \sum_{t=1}^T \rho_{\tau}(y_t - z_t'\theta)$$

其中,  $\rho_{\tau}$  为非对称损失函数, 具体表达式如下:

$$\rho_{\tau}(u) = \begin{cases} \tau u, & u \geq 0 \\ (1 - \tau)u, & u < 0 \end{cases}$$

QAR 方法打破了传统的基于均值的回归方法仅考虑样本平均水平的限制, 而是在不同分位点下的寻求目标函数最小化的参数解, 进而求解在不同水平下的模型问题。这种方法不但可以得到更为全面、精确的参数估计值, 并且其估计结果对于异常数据和非对称分布也有较好的拟合能力。

## 三、神经网络分位数回归模型

Taylor<sup>[49]</sup>(2000)正在早年间神经网络理论和技术发展的基础上, 首次提出了神经网络分位数回归(Quantile Regression Neural Network, QRNN)模型。QRNN 模型

一种基于人工神经网络和分位数回归的混合建模方法,该模型将二者的优势结合起来,既能够很好地拟合时间序列数据中的非线性成分,又能完整地刻画出解释变量在不同分位点下的分布。

QRNN 模型在结构上与普通神经网络的结构无差异,同样包括了输入层、隐藏层及输出层。与大部分神经网络选取 MSE 等指标作为损失函数的思路不同,QRNN 模型考虑到与分位数回归模型的结合,选取分位数回归的目标函数作为 QRNN 模型的损失函数,因此 QRNN 模型也具备分位数回归的优良特性,可以完整地反映出在各种分位数条件下被解释变量和解释变量之间的映射关系,刻画整体分布。具体地,在分位点 $\tau$ 下的解释变量 $Q_{y_i}(\tau|x_i)$ 是通过下列函数 $f(x_i, w_i(\tau), v_i(\tau))$ 计算得到的:

$$f(x_i, w_i(\tau), v_i(\tau)) = h_2\left\{\sum_{k=1}^m v_{i,k}(\tau)h_1\left[\sum_{j=1}^r w_{i,j,k}(\tau)x_i\right]\right\} \quad (3.3)$$

其中  $h_1, h_2$  为激活函数,通常为  $\tanh$  或  $\text{sigmoid}$  函数。 $w_{i,j,k}(\tau)$ 表示在 $\tau$ 分为点处输出层到隐藏层的权重,  $v_{i,k}(\tau)$ 表示在 $\tau$ 分为点处隐藏层到输出层的权重,  $x_i$ 为模型的输入变量,  $x_i$ 在本文模型中为上证综指的波动率 $\sigma_i$ , 具体函数形式如下式:

$$f(\sigma_i, w_i(\tau), v_i(\tau)) = h_2\left\{\sum_{k=1}^m v_{i,k}(\tau)h_1\left[\sum_{j=1}^r w_{i,j,k}(\tau)\sigma_i\right]\right\}$$

模型中未知权重向量 $w(\tau), v(\tau)$ 的求解,可以通过最小化目标函数求得,目标函数如式:

$$\min_{w,v} \left( \sum_{t|y_t < f(x,w,v)} (1-\tau)|y_t - f(x,w,v)| + \sum_{t|y_t \geq f(x,w,v)} \tau|y_t - f(x,w,v)| \right)$$

$w(\tau)$ 和 $v(\tau)$ 可以通过梯度下降的方法确定下来,最终也能计算出条件分位数下  $Y$  的值。 $\hat{w}(\tau)$ 和  $\hat{v}(\tau)$ 分别表示 $w(\tau)$ 和  $v(\tau)$ 的估计值。

$$VaR = \hat{Q}_Y(\tau|x) = f(x, \hat{w}(\tau), \hat{v}(\tau))$$

QRNN 模型将分位数回归和神经网络的优点相结合。一方面,分位数回归模型可以更精确地刻画出金融资产收益率分布的非对称性和尖峰厚尾特征;另一方面,将神经网络应用到金融系统的时序数据拟合中,可以更好地刻画出金融数据复杂的非线性结构。

## 第四章 基于 LSTM-GARCH 混合模型的波动率预测

本章首先在前述的模型理论基础上,通过引入 GARCH 族模型的参数和其他解释变量作为 LSTM 神经网络模型的输入变量,构建新的 LSTM-GARCH 融合模型,并阐述了建模步骤以及模型效果评价指标;其次,选取量价数据和 GARCH 族模型参数作为输入特征,并进行基本的统计分析与检验;再次,分别通过单一 GARCH 模型和 LSTM 模型以及对波动率进行建模,并在验证集上比较模型预测效果;最后通过多种损失函数对模型效果进行分析,将预测准确度最高的模型波动率输出作为后续 VaR 测度的输入变量,提供更精确的金融风险量化结果。

### 第一节 基于 LSTM-GARCH 模型的波动率预测方法

#### 一、基于 LSTM-GARCH 的波动率预测模型构建

由于神经网络和与 GARCH 族模型在波动率的预测方面有着各自的优势和不足,并且与 GARCH 族模型的参数输入能够更好地帮助 LSTM 模型获取波动率的时序特征。为此,本文拟通过引入 GARCH 族模型的参数和其他解释变量作为 LSTM 神经网络模型的输入变量,构建新的融合模型,来提升波动率的预测精度。

在 GARCH 模型中,GARCH 项系数表示波动的持续性,ARCH 项系数表示波动率冲击的大小。若将不同残差分布的 GARCH 模型的参数都加入到 LSTM 神经网络模型中,则该混合模型将能够充分挖掘出参数中蕴含的波动率信息,从而使模型预测准确率在单一 LSTM 模型的基础上得到显著提升。因此,本文首先运用 GARCH 族模型进行建模,将其参数作为 LSTM 模型的输入变量,使神经网络模型能够对时间序列中的波动率信息进行有效地挖掘,捕捉到更多波动率的相关特征,从而提高模型的预测性能。

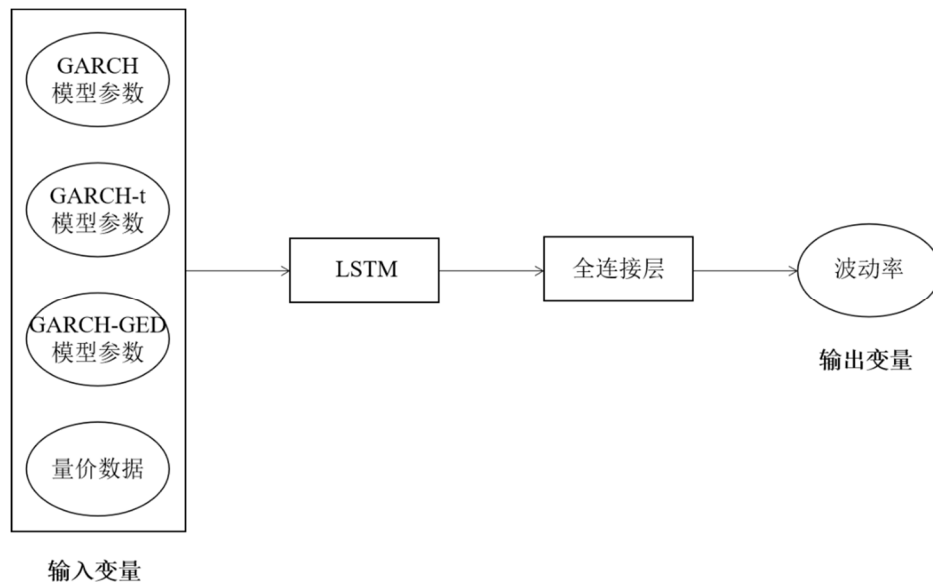


图 4-1 LSTM-GARCH 模型结构图

如图 4-1 所示的 LSTM-GARCH 模型结构图，首先将上证指数的量价数据、GARCH(1,1)模型参数、GARCH(1,1)-t 分布模型参数和 GARCH(1,1)-GED 模型参数作为输入变量输入到 LSTM 模型中进行训练，而 LSTM 模型输出结果后经过一个全连接层，最后得到上证指数的波动率作为输出变量。

根据输入的 GARCH 模型参数不同，构建出的融合模型也有所差异。本文共建立 7 种 LSTM-GARCH 模型，具体模型说明如表 4-1 所示。

表 4-1 LSTM-GARCH 融合模型说明

模型简称	说明
GARCH	GARCH 模型
GT	GARCH(1,1)-t 分布模型
GGED	GARCH(1,1)-GED 分布模型
LSTM	LSTM 模型
L-G	LSTM 模型加入 GARCH 三个参数
L-GT	LSTM 模型加入 GT 三个参数
L-GGED	LSTM 模型加入 GGED 三个参数
L-G-GT	LSTM 模型加入 GARCH 三个参数、GT 三个参数
L-G-GED	LSTM 模型加入 GARCH 三个参数、GGED 三个参数
L-GT-GGED	LSTM 模型加入 GT 三个参数、GGED 三个参数
L-G-GT-GGED	LSTM 模型加入 GARCH 三个参数、GT 三个参数、GGED 三个参数

利用 LSTM-GARCH 混合模型预测波动率的具体步骤如下：

第一步，量价数据选取与归一化。选取研究对象每日的开盘价、最低价、最高价、

收盘价作为量价数据这一类输入变量，并用最大最小标准化法对量价数据进行规整。该标准化方法的如式 4.1 所示：

$$x_i^* = \frac{x_i - x_{max}}{x_{max} - x_{min}} \tag{4.1}$$

其中， $x_i$ 为待规整的变量， $x_i^*$ 为经过 $x_i$ 归一化处理的数值， $x_{max}$ 代表输入变量中的最大的数值， $x_{min}$ 代表输入变量中的最小的数值。通过归一化处理，能较好地避免数据存在两极分化的现象，有利于后续的建模分析。

第二步，GARCH 模型参数的获取。GARCH 族模型参数的估计是滑动预测的过程，以 15 天作为窗口长度，t 日的模型参数是根据 t-16 日到 t-1 日的收益率数据预测得出，t+1 日的模型参数是根据 t-15 日到 t 日的收益率数据预测得出，以此类推，窗口向后滑动估计模型参数。滑动预测过程如图 4-2 所示。

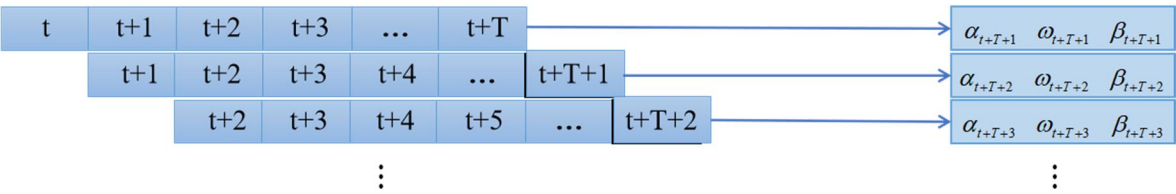


图 4-2 GARCH 模型参数滚动预测过程图

其中 T 表示时间步长， $\omega$ 、 $\alpha$ 和 $\beta$ 均为式 3.1 中的 GARCH 模型参数。经滑动预测得到的参数蕴含着波动率的相关信息，将作为后续 LSTM-GARCH 波动率预测模型的输入变量。

第三步，搭建 LSTM 神经网络框架。利用 Tensorflow 搭建神经网络框架，并利用网格搜索和交叉验证的方法决定神经网络最佳的迭代次数、网络层数等参数设置。

第四步，构建 LSTM-GARCH 混合模型。

表 4-2 LSTM-GARCH 模型参数设置

输入变量	GARCH	GT	G GED	LS TM	L- G	L- GT	L- GGED	L-G- GT	L-G- GGED	L-GT- GGED	L-G-GT- GGED
收益率	√	√	√								
量价数据				√	√	√	√	√	√	√	√
GARCH 系数					√			√	√		√
GT 系数						√		√		√	√
GGED 系数							√		√	√	√

注：表中 √ 意为选择行中表头数据作为输入变量。

对于单一 LSTM 模型和 LSTM-GARCH 融合模型的预测，LSTM 模型的输入变量为量价数据，LSTM-GARCH 模型的输入变量为量价数据和 GARCH 模型参数，并需要保持单一 LSTM 模型和七种 LSTM-GARCH 融合模型的输入变量长度一致。模

型具体输入变量设置如表 4-2 所示。

第五步，模型训练。对 GARCH 模型和神经网络相关模型分别进行训练，得到在不同模型下波动率的预测结果。

## 二、波动率预测评价指标

由于单一的误差评估指标有其局限，若只使用一个指标来评估模型效果，则会产生较大的偏差，且很难对模型的预测结果做出综合全面的评价。为此，为了能更全面地衡量不同模型的训练效果，本文引入了 7 个误差评估指标，分别为均方误差(MSE)，平均绝对值误差(MAE)，均方根误差(RMSE)，经异方差调整的 MSE(HMSE)，经异方差调整的 MAE(HMAE)、高斯准极大似然函数误差(QLIKE)和对数损失函数误差(R2LOG)。其中损失函数的计算如表 4-3 所示。

表 4-3 损失函数

损失函数	计算公式
MSE	$\frac{1}{T} \sum_{t=1}^T (RV_t - \hat{\sigma}_t^2)^2$
MAE	$\frac{1}{T} \sum_{t=1}^T  RV_t - \hat{\sigma}_t^2 $
RMSE	$\sqrt{\frac{1}{T} \sum_{t=1}^T (RV_t - \hat{\sigma}_t^2)^2}$
HMSE	$\frac{1}{T} \sum_{t=1}^T \left(1 - \frac{\hat{\sigma}_t^2}{RV_t}\right)^2$
HMAE	$\frac{1}{T} \sum_{t=1}^T \left 1 - \frac{\hat{\sigma}_t^2}{RV_t}\right $
QLIKE	$\frac{1}{T} \sum_{t=1}^T \left(\ln \hat{\sigma}_t^2 - \frac{RV_t}{\hat{\sigma}_t^2}\right)$
R2LOG	$\frac{1}{T} \sum_{t=1}^T \left(\ln \frac{RV_t}{\hat{\sigma}_t^2}\right)^2$

这七个指标评价标准都是计算值越小代表模型的预测结果越好。相较于 RMSE 指标，MAE 指标在样本有离群点时的具有更强的稳健性。然而，MAE 也会因为梯度基本一致的特点而导致训练效果不好。与其它三种误差指标相比，MAPE 是通过相对的



百分数提供了波动率的真实值和预测值之间的比较,因此更容易衡量波动率预测值的准确性。不过,对于应用到神经网络模型中进行训练的误差损失函数来说,MSE 的学习速率会随预测值与真实值之间的误差的大小而改变,当预测误差较小时,学习速率则会变得更慢,反之当预测误差较大时,学习速率则会变得更快,因此,学习效果会更好。因此本文选择 MSE 指标应用到神经网络模型的训练中,将剩余指标作为模型预测效果评价指标。

## 第二节 特征选取及统计分析

### 一、特征选取

本章实证研究对象是上证综指及其波动率。上证综合指数是于 1991 年 7 月 15 日正式发布。本文研究对象选择该支综合指数的原因是其股票覆盖面广、行业分布均衡,能够较为全面地反映我国股市的整体走势。本章波动率预测的时间范围选取为 2012 年 1 月 4 日到 2018 年 12 月 28 日,模型特征选取量价数据和 GARCH 模型参数两类数据。实证研究基于 Python3.6 软件进行。

#### (一) 量价数据选取

量价数据特征包括开盘价、收盘价、最高价、最低价共计四个数据。数据来自于 Wind 金融终端。上证指数 2012 年 1 月 4 日到 2018 年 12 月 28 日的收盘价变化趋势如图 4-3 所示。从图中可以看出:上证综指的收盘价整体呈现上升趋势,在 2015 年至 2016 年上证指数出现了较大幅度的上涨,波动幅度较为剧烈。

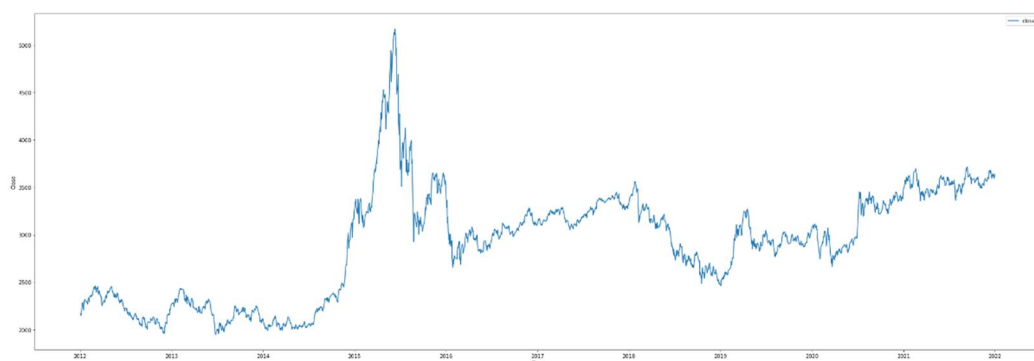


图 4-3 上证综指收盘价趋势图

#### (二) GARCH 族模型参数估计

本文使用的 GARCH 族模型包括 GARCH 模型、GARCH-t 模型和 GARCH-GED 模型三种,共计 9 个参数,包括 GARCH 模型的 $\omega$ 、 $\alpha$ 和 $\beta$ 三个参数,GARCH-t 模型的 $\omega$ 、 $\alpha$ 和 $\beta$ 三个参数,GARCH-GED 模型的 $\omega$ 、 $\alpha$ 和 $\beta$ 三个参数。

GARCH 族模型参数的估计是滑动预测的过程,以 15 天作为窗口长度,将 15 天的上证指数收益率数据作为模型的输入,拟合得到的 GARCH 模型的三个参数作为第 16 天的特征向量。即  $t$  日的模型参数是根据  $t-16$  日到  $t-1$  日的收益率数据预测得出, $t+1$  日的模型参数是根据  $t-15$  日到  $t$  日的收益率数据预测得出,以此类推,窗口向后滑动估计模型参数。部分交易日的 GARCH 模型参数估计的部分结果如表 4-4 所示。

表 4-4 GARCH 模型部分参数

日期	omega_g	alpha_g	beta_g	omega_t	alpha_t	beta_t
2012/2/1	0.818	5.7E-16	0.746	2.599	0.184	8.26E-11
2012/2/2	1.558	9.5E-17	0.503	0.689	4.4E-15	0.779
2012/2/3	1.571	1.6E-16	0.503	2.582	0.165	1.1E-20
2012/2/6	1.969	0.352	0.878	1.974	0.352	1.3E-19
2012/2/7	1.529	0.454	0.551	1.534	0.453	0.932
2012/2/8	1.298	1.2E-16	0.503	2.461	0.029	5.9E-17
2012/2/9	4.9E-14	3.8E-14	0.983	1.990	0.351	6.4E-12
2012/2/10	0.092	1.0E-15	0.961	1.2E-13	5.6E-15	0.962

注:表中的 omega\_g、alpha\_g 和 beta\_g 分别表示 GARCH 模型中的  $\omega$ 、 $\alpha$  和  $\beta$  参数,omega\_t、alpha\_t 和 beta\_t 分别表示 GT 模型中的  $\omega$ 、 $\alpha$  和  $\beta$  参数;由于页面显示限制,仅展示 GARCH 模型和 GT 模型的参数滚动预测结果,未呈现 GGED 模型的参数估计结果。

## 二、统计分析与检验

### (一) 相关统计特征

量价数据的描述性统计结果如表 4-5 所示。

表 4-5 量价数据描述性统计

	开盘价	最高价	最低价	收盘价
均值	2913.411	2936.001	2891.683	2916.912
标准差	571.041	578.040	562.026	571.631
最小值	1935.522	1959.165	1849.654	1950.013
中位数	2987.024	3007.353	2968.035	2987.124
最大值	5174.427	5178.195	5103.496	5166.352

通过表 4-3 可得知,四类量价数据的最大值和最小值差距较大,且存在一定的波动性,因此通过归一化对量价数据进行标准化处理是十分有必要的。

### (二) 相关统计检验

在对时间序列数据进行模型构建前，须对序列进行平稳性检验。最典型的检验方式就是单位根 ADF 检验，该检验的原假设（H0）是所研究的时间序列存在单位根，如果 ADF 检验的统计量小于显著的水平下所相应的值，则可以拒绝原假设，这说明该时间序列不存在单位根，序列是平稳的。

表 4-6 上证指数 ADF 检验

	ADF 值	不同置信水平下的统计值			P 值
		1%	5%	10%	
上证指数	-9.508	-3.433	-2.862	-2.567	3.3e-16

表 4-6 中，上证指数对数收益率序列的 ADF 检验统计量的值为-9.508，小于 1% 显著性水平下的值-3.435，而且 p 趋近 0。综合表格中的数据分析，则能够拒绝原假设，表明上证指数收益率序列平稳。

使用 GARCH 族模型之前还需要对时间序列进行 ARCH 效应检验，即检验是否存在序列相关性。上证指数日对数收益率平方序列如图 4-4 所示，从这张图中我们可以较为直观地看到，序列呈现出明显的波动性聚集效应。

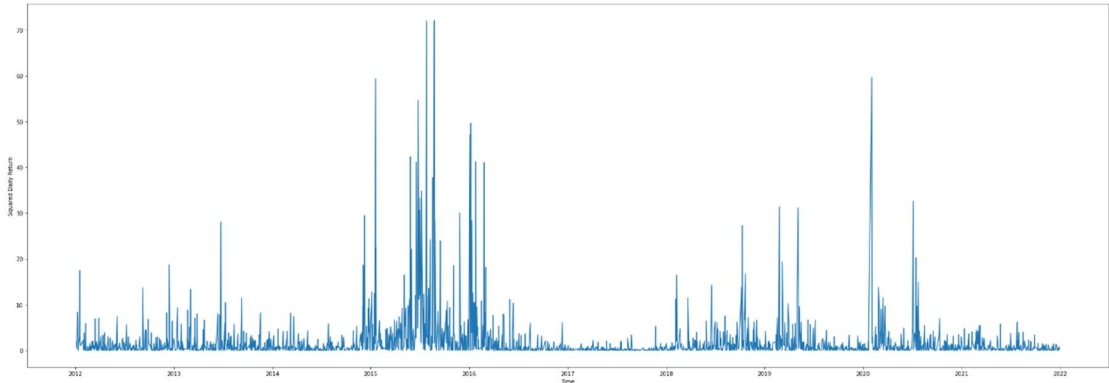


图 4-4 上证指数日对数收益率平方序列图

检验 ARCH 效应通常采用 Ljung-Box 检验。接受原假设（H0）表明时间序列是白噪声序列，拒绝原假设则认为序列存在相关性。本文设置滞后阶数为 1 到 12，Ljung-Box 检验的 LB 统计量的值和 p 值在不同阶数下的变化如图 4-5 所示。图 4-5 中显示 p 值趋近于 0，远小于临界值 0.05，可以拒绝上证指数日对数收益率平方序列是白噪声的原假设，说明该序列存在 ARCH 效应。经过 ARCH 效应检验，证明了上证指数日对数收益率序列可用 GARCH 族模型进行建模。

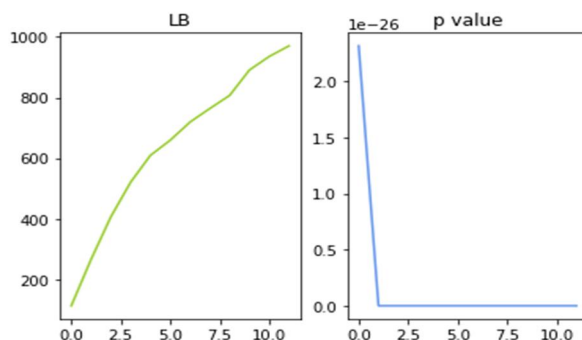


图 4-5Ljung-Box 检验的 LB 统计量和 p 值

### 第三节 基于 LSTM-GARCH 模型的波动率预测

在本章第一节,通过 LSTM 模型融合 GARCH 族模型,构建了 LSTM-GARCH 模型并给出建模步骤;在本章第二节选取了模型特征,并进行了必要的统计性检验,证明上证指数的收益率序列适合用 GARCH 模型来进行建模。在上述基础上,本节将 LSTM-GARCH 模型应用于波动率预测的实证分析,并比较单一 GARCH 模型和 LSTM 模型与七种 LSTM-GARCH 融合模型的波动率预测精度。

对于 GARCH 族模型的参数估计与波动率预测,本文使用 Python 的 arch 库来实现。采用滚动预测的方式,设置窗口长度为 15,第一个训练集是从 2012 年 1 月 2 日到 2012 年 1 月 20 日,共 15 个样本收益率,以此预测 2012 年 1 月 23 日的波动率与 GARCH 模型参数,随后窗口继续向后滑动 1 天,预测 2012 年 1 月 24 日的波动率与 GARCH 模型参数,以此类推,直至预测完成 2018 年 12 月 28 日的相关数据,训练结束。这样就保证了 GARCH 族模型预测出的参数长度能够覆盖深度学习模型训练集与测试集,以将 GARCH 模型的参数作为融合模型的输入变量,并保持与量价数据的长度一致。具体地,2012 年 1 月 23 日至 2017 年 12 月 29 日区间内的数据作为实验集,2018 年 1 月 1 日至 2018 年 12 月 28 日的数据作为测试集。GARCH 模型的数据集划分如表 4-7 所示。

表 4-7 GARCH 模型数据集划分

GARCH 族模型	实验集	测试集
时间区间	2012 年 1 月 4 日-2017 年 12 月 29 日	2018 年 1 月 2 日-2018 年 12 月 28 日
数据量	1458	243

对于深度学习相关的 LSTM 模型和 LSTM-GARCH 融合模型,本文均利用 Tensorflow 搭建神经网络框架,将 2012 年 1 月 23 日至 2017 年 12 月 29 日区间内的数据作为实验集,其中共计 1458 条数据按照 8:2 的比例划分为训练集和验证集,训练集中的 1166 条数据用于深度学习模型的学习与训练,验证集中的 292 条数据用于

验证模型表现效果；2018 年 1 月 1 日至 2018 年 12 月 28 日的数据作为测试集。深度学习模型的测试集与 GARCH 模型的测试集相同是为了比较分析两类模型的预测精度。具体数据集划分如表 4-8 所示

表 4-8 深度学习相关模型数据集划分

深度学习相关模型	训练集	验证集	测试集
时间区间	2012 年 1 月 4 日-2017 年 12 月 29 日	2018 年 1 月 2 日-2018 年 12 月 28 日	
数据量	1458*80%=1166	1458*20%=292	243

数据集划分完成后，可进行 LSTM 模型训练。在进行训练前，需要提前对模型参数进行设置。考虑到神经网络模型的参数设置能够在很大程度上影响模型的训练表现，选择一组好的参数能够提高模型的训练性能和效果。因此，本文通过网格搜索和交叉验证（GridSearchCV）的方法得到 LSTM 神经网络模型最佳参数设置如表 4-9 所示。得到最佳参数后，即可开始利用最佳参数对 LSTM 模型和 LSTM-GARCH 融合模型的训练。

表 4-9 LSTM 模型最佳参数

参数	epochs	batch_size	optimizer
最佳设置	49	32	Adam

经过三种 GARCH 模型、LSTM 模型和 7 种 LSTM-GARCH 模型的训练和预测，十一种模型在测试集上的波动率的预测结果和真实波动率的对比如图 4-5 所示。其中，绿色实线代表的是各个模型在测试集上的波动率预测预测结果，蓝色实线代表的则是测试区间的真实波动率。

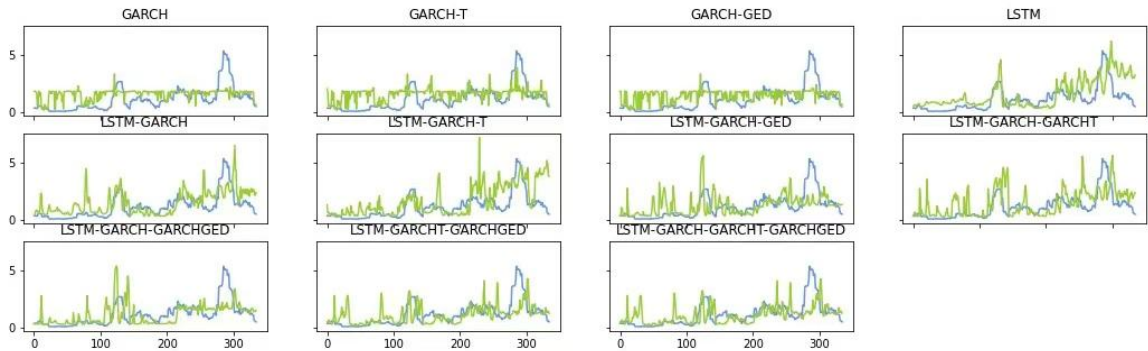


图 4-6 单一模型和 LSTM-GARCH 混合模型的波动率预测对比图

总体上看，根据图 4-6 可以很明显地看到，三种 GARCH 族模型在测试集中，预测的波动率与实际波动率相比，仅能模拟出大致趋势，对波动率的低波动部分预测准确性更高，对高波动部分预测性能较差，这是可能是由于 GARCH 族模型在实质上属于一种线性模型，因此，它的预测结果具有比较平稳的滞后性，这就会造成在非线性特征较为显著的区间，GARCH 族模型的预测偏差。而 LSTM 模型和 LSTM-GARCH

融合模型由于其本身为非线性模型，能够更好地捕捉到波动率的非线性特性，因此，在波动率的高波动部分具有更好的预测性能，使得模型整体的拟合效果比单一的 GARCH 模型的拟合效果更好，从图 4-5 中也可以证实这一点：LSTM 模型和七种 LSTM-GARCH 融合模型的预测效果要远好于三种 GARCH 族模型。综合来看，LSTM-GARCH 融合模型的预测性能整体上更好，远优于单一 GARCH 模型和 LSTM 模型的预测结果。

对比 GARCH 族模型。三种模型在低波动部分预测能力相当，直观上看来预测波动率与真实值存在偏离，仅能预测总体趋势，具体的波动预测并不准确，均表现出平滑性；三种模型中只有 GT 模型在高波动部分表现出一定的预测能力，表现相对较好。

对比 LSTM 融合一种 GARCH 模型的 L-G 模型、L-GT 模型和 L-GGED 模型。三个模型在低波动预测部分表现出了较好的预测能力，可以直观看出神经网络算法对非线性序列的拟合优势。L-GGED 模型在后半段的高波动部分表现略逊色于其他两个模型，但 L-G 模型和 L-GT 模型在高波动部分相较于低波动部分重合度更低。

对比 LSTM 融合两种 GARCH 模型的 L-G-GT 模型、L-GGED 模型和 L-GT-GGED 模型，唯有 L-GT-GGED 模型未能较好预测出后半部分的大波动，但整体上来说三个模型在前半段的低波动部分和后半段高波动部分均有良好表现，拟合程度较高。

对比 GARCH 族模型和 LSTM 模型两类的单一模型。LSTM 模型在前半段的低波动部分与真实波动率值的重合度更高，且预测高波动率方面比单一 GARCH 模型有更好的表现，显然 LSTM 模型良好的非线性拟合能力和长记忆性在波动率预测上预测具备优势。

对比单一模型和融合模型。整体上来看 LSTM-GARCH 融合模型的预测准确度高于一 GARCH 模型和单一 LSTM 模型，且在低波动和高波动部分的预测表现均有所提高。但是七种融合模型的具体差异与优劣判断还需借助多种损失函数做进一步的评定。

## 第四节 模型预测效果评价

十一种模型在测试集上的损失函数如表 4-9 所示。分析表中数据可知：

(1) 三种 GARCH 族模型在测试集上的七种损失函数的值相较于 LSTM 模型和 LSTM-GARCH 模型普遍偏大，因此 GARCH 族模型的预测精度低于其他两类模型。具体分析 GARCH 族模型的表现：GARCH 模型和 GT 模型的损失函数的值相差不大，模型表现能力相仿，整体表现最差；而 GGED 模型在测试集上的七种损失函数中，有六种损失函数的值在 GARCH 族模型中均为最小。说明相较于 GARCH 模型和 GT

模型，GGED 模型对上证指数波动率的预测效果更好，在 GARCH 族模型中预测性能最佳。

(2) LSTM 模型和 LSTM-GARCH 融合模型在测试集上的七种损失函数的值均远小于 GARCH 族模型。说明在拟合非线性特征明显的波动率数据上，非线性模型比线性的 GARCH 模型有更好的表现，证实深度学习模型捕捉非线性特性的能力可以在时间序列中发挥出很好的预测作用。

(3) LSTM-GARCH 融合模型的七种损失函数的值普遍小于单一的 LSTM 模型，说明 LSTM-GARCH 融合模型对 LSTM 模型的预测性能有一定程度上的改进，GARCH 模型参数蕴含的信息有助于 LSTM 模型更好地学习波动率相关特征，证明本文进行模型融合是有效的。

(4) 具体对比七种 LSTM-GARCH 融合模型，整体上七种损失函数值的差距较小，表明融合模型在预测上证指数波动率方面皆有较好的表现；融合三种 GARCH 模型的 L-G-GT-GGED 模型的 MAE、HMAE、HMSE 和 R2LOG 值最小，总体上表现最优；L-GT-GGED 模型的 HSE 和 RMSE 的值最小，表现次之。

因此，本文选取模型整体表现最佳的 L-G-GT-GGED 融合模型进一步地对模型表现做出分析。

表 4-9 单一模型和融合模型的波动率预测效果损失函数对比

模型	MAE	MSE	RMSE	HMAE	HMSE	QLIKE	R2LOG
GARCH	0.847	1.173	1.083	2.917	37.833	1.119	1.523
GT	0.846	1.158	1.076	2.861	36.424	1.111	1.495
GGED	0.780	1.259	1.029	2.513	30.063	1.079	1.335
LSTM	0.864	1.341	1.158	1.748	9.018	1.072	1.077
L-G	0.796	1.143	1.069	1.723	9.034	1.060	1.061
L-GT	0.881	1.617	1.272	1.624	8.503	<b>1.033</b>	0.985
L-GGED	0.751	1.234	1.111	1.476	9.847	1.055	0.932
L-G-GT	0.846	1.423	1.193	1.561	8.805	1.043	0.967
L-G-GGED	0.701	1.101	1.049	1.191	6.974	1.110	0.885
L-GT-GGED	0.686	<b>1.081</b>	<b>1.040</b>	1.452	7.687	1.086	0.909
L-G-GT-GGED	<b>0.604</b>	1.107	1.056	<b>1.073</b>	<b>5.133</b>	1.071	<b>0.795</b>



注：表中黑色加粗字体表示在每种损失函数下函数值最小的模型。

由于本文旨在分析融合模型相较于单一模型的优越性能，因此将进一步地比较融合模型相对单一 GARCH 和 LSTM 模型的提升幅度与改进效果，选取表现最佳的 L-G-GT-GGED 模型分别与 GARCH 模型中表现最佳 GGED 模型和 LSTM 模型进行比较。

对比 L-G-GT-GGED 模型与 GGED 模型，显著降低的指标有：经异方差调整的 MAE、经异方差调整的 MSE 和对数损失函数误差 R2LOG，分别降低了 57.3%、82.9% 和 40.4%；平均绝对误差 MAE 和均方误差 MSE 存在小幅度降低，分别为 22.6% 和 12.1%；高斯准极大似然函数误差 QLIKE 差距微小，仅降低 0.7%，而均方根误差 RMSE 上升了 2.6%。整体上来看，L-G-GT-GGED 模型的波动率预测性能明显优于 GGED 模型，表明神经网络的加入显著提高了模型的预测能力。

对比 L-G-GT-GGED 模型与 LSTM 模型，显著降低的指标有：平均绝对误差 MAE、经异方差调整的 MAE、经异方差调整的 MSE 和对数损失函数误差 R2LOG，分别降低了 30.1%、38.6%、43.1% 和 26.2%；均方误差 MSE 和均方根误差 RMSE 分别小幅度降低 17.4%、8.8%；高斯准极大似然函数误差 QLIKE 改善不显著，仅降低 0.1%。整体上来看，L-G-GT-GGED 模型的波动率预测性能明显优于 LSTM 模型，表明混合模型相较于单一的神经网络模型不仅在理论上增强了模型的可解释力度，还在实证中显著改善模型的预测性能。

## 第五节 本章小结

本章主要内容是对上证指数的波动率预测，目的是为后续 VaR 测度更为准确的波动率输入变量。首先，本章基于正态分布、t 分布、GED 分布三种残差分布假设条件下建立 GARCH 模型，利用滚动预测的方式预测上证指数的波动率，并将含有波动率信息的 GARCH 模型参数作为 LSTM 模型的输入变量，构建 LSTM-GARCH 融合模型预测波动率，根据输入参数的不同，LSTM-GARCH 模型共计分为七种：L-G、L-GT、L-GGED、L-G-GT、L-G-GGED、L-GT-GGED、L-G-GT-GGED，上述融合模型既具备 GARCH 模型对波动率信息的捕捉能力，同时又具有 LSTM 深度学习模型对非线性结构的拟合能力。最后，本章通过七种常用的误差评价指标（MSE、MAE、RMSE、HMSE、HMAE、QLIKE、R2LOG）评价四个单一模型和七个融合模型的整体预测性能。

实证结果表明：

通过损失函数横向比较 GARCH 模型在不同残差分布假设下的波动率拟合能



力, 当假设残差服从于 GED 时, 模型预测精确性高于其他 GARCH 族模型; 通过损失函数横向对比单一模型的表现, LSTM 模型的预测精度高于 GARCH 模型; 在量价数据的基础上增加 GARCH 模型滚动预测得到的参数作为 LSTM 模型输入得到的融合模型, 通过损失函数对比融合模型和单一模型的表现, 融合具有更好的预测性能; 在融合模型中, 融合三种 GARCH 模型和 LSTM 模型的 L-G-GT-GGED 融合模型表现效果最优。

具体比较 L-G-GT-GGED 融合模型对 LSTM 模型和 GARCH 模型的改进: 对比 L-G-GT-GGED 模型与 GGED 模型, HMAE、HMSE 和 R2LOG 三个指标的函数值显著降低, MAE 和 MSE 存在小幅度降低, QLIKE 差距微小, 而 RMSE 值存在上升。整体上来看, L-G-GT-GGED 模型的波动率预测性能明显优于 GGED 模型, 表明神经网络的加入显著提高了模型的预测能力; 对比 L-G-GT-GGED 模型与 LSTM 模型, MAE、HMAE、HMSE 和 R2LOG 四个指标存在显著的降低, MSE 和 RMSE 存在小幅度的降低, QLIKE 值的改善并不显著。整体上来看, L-G-GT-GGED 模型的波动率预测性能明显优于 LSTM 模型, 表明混合模型相较于单一的神经网络模型不仅在理论上增强了模型的可解释力度, 还在实证中显著改善模型的预测性能。

## 第五章 基于 QARNN 模型的 VaR 测度

首先,考虑到 VaR 本身存在自相关性,本章在第三章提及的神经网络分位数回归模型的基础上,加入滞后内生变量作为神经网络模型的输入变量,构建神经网络分位数自回归(QARNN)模型,并介绍了该模型测度 VaR 的基本步骤和返回测试方法;其次,选取上章 L-G-GT-GGED 模型预测的波动率和上证指数滞后三期的 VaR 值作为输入特征,并对上证指数的收益率进行基本的统计性检验;再次,建立 QR 模型、QAR 模型、QRNN 模型和 QARNN 模型共计 4 种模型,分别在置信水平为 95%和 99%的条件下对上证指数的 VaR 值进行测度;最后,在返回测试中,通过多种评价指标对 4 种 VaR 测度模型在回测区间上的表现进行对比分析。

### 第一节 基于 QARNN 模型的 VaR 测度方法

#### 一、基于 QARNN 的 VaR 测度模型构建

受 Taylor 提出的 QRNN 模型启发,本文在其模型式 3.3 的基础上考虑滞后内生变量的影响。将分位数自回归模型与神经网络相结合中,建立神经网络分位数自回归(QARNN)模型。模型具体表达形式如下所示:

$$Q_{y_i}(\tau) = f(x_i, \dots, x_{i-p}, Q_{y_{i-1}}(\tau), \dots, Q_{y_{i-p}}(\tau); w_i(\tau); v_i(\tau)),$$

其中,  $x_{i-j}(j = 0, \dots, p)$  为解释变量,  $Q_{y_{i-j}}(\tau)(j = 0, \dots, p)$  为滞后  $i$  期的被解释变量。具体在本文中,模型解释变量为上证指数的波动率  $\sigma_i$ , 被解释变量为 VaR 值。

$$VaR_t = f(\sigma_i, VaR_{t-1}, VaR_{t-2}, VaR_{t-3}; w_i; v_i)$$

其中,  $VaR_t$  即为收益率序列在分位点  $\tau$  处的分位数  $Q_{y_i}(\tau)$ 。

对于模型中未知权重向量  $w(\tau), v(\tau)$  的求解,大多数神经网络模型会选择 MSE 等指标作为损失函数进行训练。但在本章构建的 QARNN 模型中,损失函数则选择为分位数回归的目标函数式 3.2。具体到本文的波动率预测模型中,权重向量  $w(\tau), v(\tau)$  的求解可以通过以下最小化目标函数求得。

$$\min_{w, v} \left( \sum_{t|y_t < VaR_t} (1 - \tau) |y_t - VaR_t| + \sum_{t|y_t \geq VaR_t} \tau |y_t - VaR_t| \right)$$

从上述公式可以看出, QARNN 模型是基于 Taylor 的 QRNN 模型提出。在将波动率作为神经网络的输入变量的基础上,增加了滞后 3 期的收益率分位数(即 VaR 值)作为输入变量,考虑了滞后内生变量的非线性效应。具体神经网络结构如图 5-1

所示。

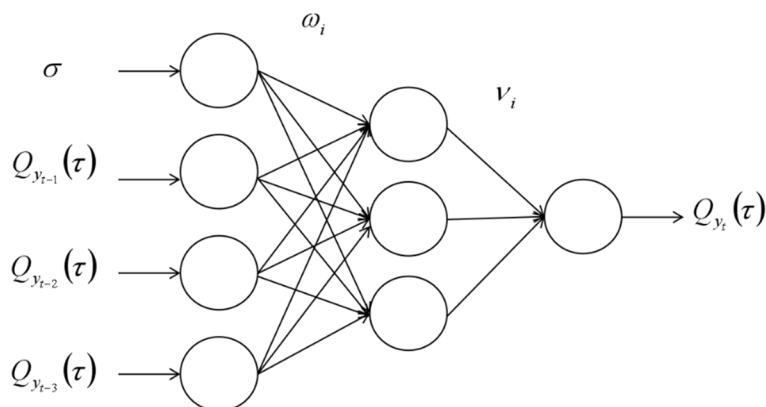


图 5-1 基于 QARNN 的 VaR 测度模型结构图

## 二、VaR 返回测试评价指标

返回测试是检验 VaR 建模精度的一种特有手段。VaR 风险度量模型必须将数据的预测范围和回测范围之间进行分割，类似于机器学习中划分的训练集与测试集，将一部分历史的数据作为预测区间，用不同的建模方法对这些收益率数据进行 VaR 值的预测，得到的预测结果即为回测区间。随后，对于回测区间的预测值和真实值进行返回测试，本文选择多种评价指标，对 VaR 值的预测精度进行评估，并与实际损失进行对比。为便于对每一种 VaR 风险度量模型进行预测效果评估，下文将简要说明在本文中使用到的几种 VaR 返回测试方法。

首先，我们将实际亏损大于 VaR 预测值的情况，称为发生了一次“VaR 违背事件”，时间序列  $\{I_{t+1}\}_{t=1}^T$ ，定义如下：

$$I_{t+1} = \begin{cases} 1, & y_{t+1} < -\text{VaR}_{t+1} \\ 0, & y_{t+1} \geq -\text{VaR}_{t+1} \end{cases}$$

VaR 金融风险计量模型的表现将通过下列三种评价指标来进行对比分析：

(1) 失败天数  $T_1$ ：失败天数为时间序列中值为 1 的次数，表示损失真实值超过 VaR 预测值的天数。

(2) 失败率  $\pi$ ：失败率为损失真实值超过 VaR 预测值的天数  $T_1$  占回测区间总天数  $T$  的比例。

$$\pi = 100\% \cdot \left( \frac{T_1}{T} \right)$$

(3) Kupiec 似然比检验：Kupiec (1995) 提出了失败频率检验法来检验 VaR 测度模型的预测精确度。

该方法中的似然比检验统计量 LR 服从自由度为 1 的  $\chi^2$  分布。给定显著性水平为  $\tau$ ，失败率的期望值即为显著性水平  $\pi^* = \tau$ 。一个恰当的 VaR 风险测度意味着  $\pi = \frac{T_1}{T} = \pi^*$ ，当  $\pi < \pi^*$  时，表示模型对金融风险的估计值过高，反之则表示低估金融风险。因

此, 似然比检验法的主要思想是检验模型 VaR 预测失败率 $\pi$ 是否显著不同于预设的显著性水平, 即提出原假设 $H_0: \pi = \tau$ ;  $H_1: \pi \neq \tau$ 。似然比检验统计量的形式为:

$$LR = -2\ln \frac{L(\tau)}{L(\pi)} \sim \chi_1^2$$

其中,

$$L(\tau) = \prod_{t=1}^n (1 - \tau)^{1-I_{t+1}} \tau^{I_{t+1}} = (1 - \tau)^{T_0} \tau^{T_1}, L(\pi) = \left(1 - \frac{T_1}{T}\right)^{T_0} \left(\frac{T_1}{T}\right)^{T_1}$$

## 第二节 数据选取与统计检验

### 一、数据选取

本章仍然选取上证指数为研究对象, 日收益率采用对数收益率。样本区间为 2019 年 1 月 1 日至 2022 年 12 月 30 日, 按照 8:2 的比例划分训练集和测试集。

表 5-1 VaR 测度模型数据集划分

VaR 测度模型	测试集	测试集
时间区间	2019 年 1 月 1 日至 2022 年 12 月 30 日	
数据量	730*80%=584	730*20%=146

基于传统计量模型和 Taylor 提出的 QRNN 模型, 本章共建 4 种模型对上证指数的 VaR 值进行测度, 模型分别为: 分位数回归模型(QR)、分位数自回归模型(QAR)、神经网络分位数回归模型(QRNN)和神经网络分位数自回归模型(QARNN)。根据模型选择不同, 输入变量也有所不同。具体参数设置如表 5-1 所示。

表 5-2 VaR 测度模型参数设置

输入变量	QR	QAR	QRNN	QARNN
$\sigma$	√		√	√
$VaR_{t-3}$		√		√
$VaR_{t-2}$		√		√
$VaR_{t-1}$		√		√

注: 表中 √ 意为选择行中表头数据作为输入变量。

QR 模型和 QRNN 模型的输入变量皆为 L-G-GT-GGED 融合模型预测得出的波动率, QAR 模型的输入变量为滞后 3 期的 VaR 测度值, QARNN 模型的输入变量为滞后 3 期的 VaR 测度值和 L-G-GT-GGED 融合模型预测得出的波动率。

## 二、统计检验

通过 2019 年 1 月 1 日至 2022 年 12 月 30 日上证指数日度收盘价数据计算日对数收益率序列，计算方式如下所示：

$$R_t = \ln \frac{P_{t+1}}{P_t}$$

其中  $P_t$  为上证指数  $t$  日收盘价， $P_{t+1}$  为上证指数  $t$  日后一日的收盘价， $R_t$  为上证指数日对数收益率序列，对数形式有较为良好的统计特征。结果如图 4-2 所示。

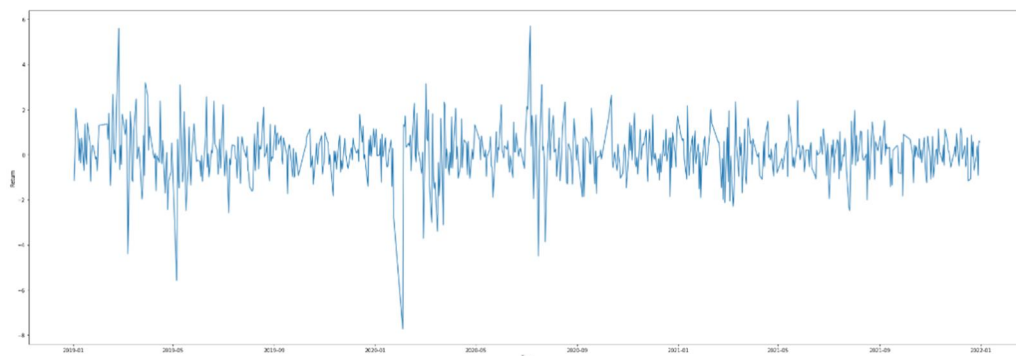


图 5-2 上证指数日对数收益率

根据表 5-2 的统计量显示，上证指数的收益率均值接近 0；偏度小于 0，表明收益率序列的分布呈现左偏；峰度大于 3，即大于正态分布的峰度，表明呈现尖峰的特征；J-B 检验原假设为序列服从正态分布，上证指数的 J-B 统计量值较大且  $p$  值近似为 0，拒绝原假设，表明收益率序列不服从正态分布。

表 5-3 上证指数对数收益率描述性统计

均值	标准差	中位数	偏度	峰度	J-B 统计量
0.00021	0.01374	0.00062	-1.00899	6.99199	1804.07709

为进一步揭示上证指数收益率序列的分布特征，本文在 J-B 统计量基础上，对序列进行正态性检验，作出上证指数收益率序列的 Q-Q 图如图 5-3 所示。

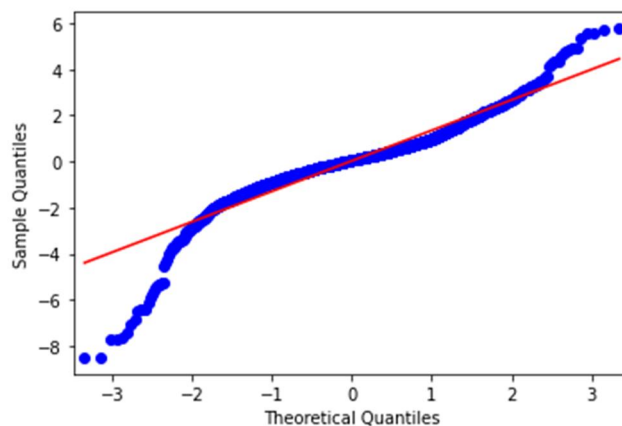


图 5-3 上证指数收益率 Q-Q 图

图 5-3 中红线表示正态分布，蓝点表示上证指数日对数收益率序列的分布。由图可以看到：Q-Q 图中间部分蓝点基本与红色直线偏离不大，但左侧略微偏离，右侧略微偏低；蓝点两端偏离红色直线的幅度较大，且左端偏离的程度更大。因此，上证指数的日对数收益率不服从正态分布。上述特征说明与标准正态分布相比，上证指数的收益率序列有更高的峰度和明显的肥尾特征，分布呈现为尖峰厚尾特征。

综合 J-B 统计量的检验结果和 Q-Q 图的分析考虑，上证指数的收益率序列不服从正态分布，且具有尖峰厚尾的分布特征。

### 第三节 基于 QARNN 的 VaR 风险测度研究

下面做出 QR、QAR、QRNN 和 QARNN 模型的 VaR 预测图，并与原收益率序列进行比较。图 5-4 和图 5-5 分别为置信水平 95%和置信水平 99%情形下的四种模型在测试集的 VaR 对比预测结果。图中红色圆点代表测试集收益率的实际值，深绿色实线代表测试集模型预测的 VaR 值。

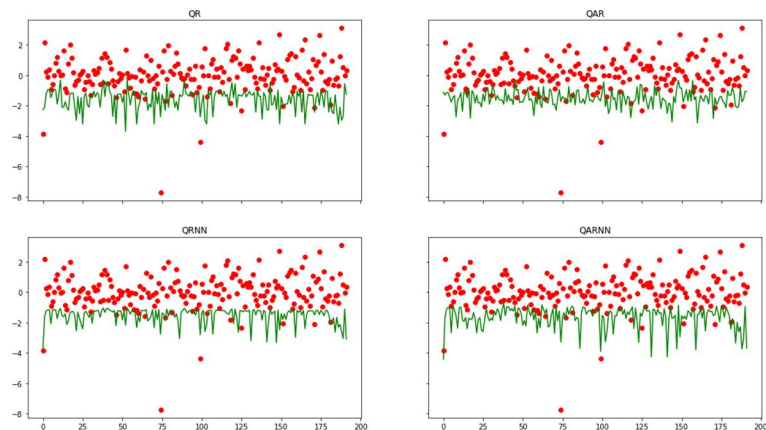


图 5-4 95%置信水平下的 VaR 值预测

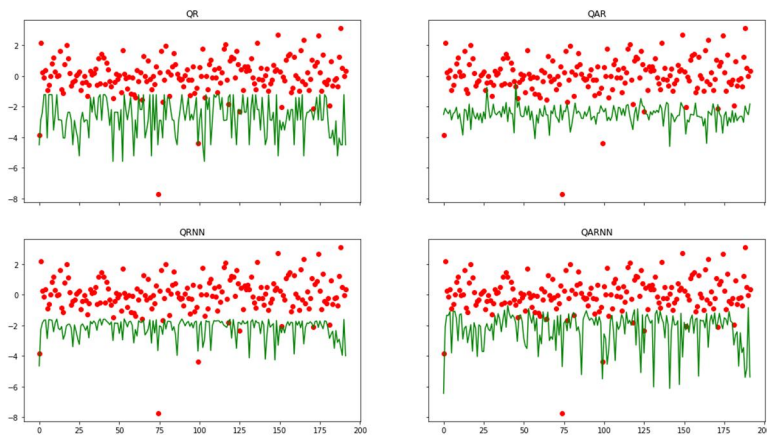


图 5-5 99%置信水平下的 VaR 值预测

由图 5-4 可知，置信水平为 95%时，四个模型预测的 VaR 值绝大部分都在真实

值以下，均在 0 值下方波动，且波动效果明显。QAR 模型预测的 VaR 值波动最小。QARNN 模型的 VaR 预测值波动最为剧烈，表明融合神经网络并加入波动率变量后较好地避免了对风险的低估。

由图 5-5 可知，置信水平为 99% 时，四个模型预测的 VaR 值绝大部分都在真实值以下，仅有个别极端值点落在 VaR 预测值之外。QAR 模型和 QRNN 模型预测的 VaR 值波动较大，QR 模型和 QARNN 模型预测的 VaR 值波动较小。从图中可以直观看出 QARNN 模型对极端值有较好的预测效果。

对比图 5-4 和图 5-5 可以看出，相较于置信水平为 95% 的情况，置信水平为 99% 时模型对极端风险的预测效果更好，且综合两种置信水平下的模型表现，QARNN 模型具有更好的 VaR 值测度效果，但对于在不同置信水平下的不同模型的效果还需通过返回测试进行更为精准的比较。

#### 第四节 返回测试

选择在置信水平 95% 和置信水平 99% 的情形下，四种模型对于上证指数的返回测试结果如下表 5-4、5-5 所示。表中的失败次数意味着返回测试区间内 VaR 值大于实际损失的天数之和；失败率表示未通过检验的天数与总天数的比值，即失败次数占回测区间长度的百分比；LR 值为 Kupiec 似然比检验中的检验统计量，对应的 LR 值越小，表明失败率越接近给定的显著性水平；在 95% 的置信水平和 99% 的置信水平下，LR 值分别小于 3.81 和 6.63 即意味着通过检验。

表 5-4 置信水平 95% 下的上证指数 VaR 模型返回测试结果

模型	失败次数	失败率	LR
QR	15	7.8%	1.194
QAR	13	6.7%	0.497
QRNN	11	5.7%	0.089
QARNN	9	4.6%	0.017

表 5-5 置信水平 99% 下的上证指数 VaR 模型返回测试结果

模型	失败次数	失败率	LR
QR	6	3.1%	2.432
QAR	4	2.1%	0.753
QRNN	3	1.6%	0.234
QARNN	1	0.5%	0.127

表 5-4 和 5-5 给出了四种模型在不同置信水平下的失败次数、失败率和 LR 值，分析检验结果可得到如下结论：

第一，在 95%和 99%两种置信水平下，模型的表现均较为稳定，拟合失败率均很低，表明模型的准确性可接受。从失败次数和失败率来看，99%置信水平下的值皆低于 95%置信水平下的值；从 LR 值来看，99%置信水平下的值高于 95%置信水平下的值，说明 99%置信水平下的对风险测度的要求更为极端，它并非只关注常态水平下的风险损失，更为重要的是反映了标的资产在未来一段时间内所可能面临的尾部极端风险。

第二，95%和 99%的置信度水平意味着两种情况下期望失败率分别为 5%和 1%。在期望失败率为 5%的情况下，QR 模型和 QAR 模型的失败率分别为 7.8%和 6.7%，均显著高于期望失败率；QRNN 模型的失败率为 5.7%，略高于期望失败率；QARNN 模型的失败率为 4.6%，低于期望失败率。在期望失败率为 1%的情况下，QR 模型、QAR 模型和 QRNN 模型的失败率分别为 3.1%、2.1%和 1.6%均略高于期望失败率；QARNN 模型的失败率为 0.5%，显著低于期望失败率。通过上述分析可得知，代表传统计量模型的 QR 模型和 QAR 模型均在测试集内出现低估的情况；QRNN 模型的模型效果比 QR 模型和 QAR 模型要好，失败率接近期望失败率，对 VaR 测度较为合理；QARNN 模型整体表现最佳，失败率低于期望失败率，该模型不仅考虑了波动率对 VaR 的关键影响，而且考虑了滞后内生变量的非线性效应，使得模型准确性显著提升。

## 第五节 本章小结

本章的主要内容是 VaR 金融风险度量，首先借鉴 Taylor 提出的 QRNN 模型，在此基础上加入滞后内生变量建立 QARNN 模型，并将上章 LSTM-GARCH 融合模型预测得出的波动率结果作为神经网络的输入变量。其次，选取上证指数的波动率（由第三章 L-G-GT-GED 模型预测得出）和滞后三期的 VaR 值作为模型输入特征，并对上证指数进行基本的统计性检验。再次，在 95%和 99%两种置信水平下使用 QR 模型、QAR 模型、QRNN 模型和 QARNN 模型分别测算 VaR 值，最后，利用失败天数、失败率和 LR 统计值来比较四种模型对的 VaR 测度效果。

通过对上证指数的统计检验可以得出：上证指数的对数收益率不服从正态分布，序列存在尖峰厚尾、偏态分布的特征。同时，该序列存在波动聚集性，表明波动随时间变化而变化，大波动聚集在一些时间段发生、小波动聚集在一些时间段发生的现象，波动之间存在正相关关系，呈现出波动聚集的特征。

通过返回测试对比模型 VaR 测度的准确程度可得出：代表传统计量模型的 QR 模型和 QAR 模型的失败率较高，意味着低估了损失；QRNN 模型的模型预测性能优



于 QR 模型和 QAR 模型,失败率略高于期望失败率,对 VaR 测度较为合理; QARNN 模型整体表现最佳,失败率低于期望失败率且与期望失败率最为接近,该模型不仅考虑了波动率对 VaR 的关键影响,而且考虑了滞后内生变量的非线性效应,使得模型准确性显著提升。

## 第六章 结论与展望

### 第一节 研究结论

本文以上证指数为研究对象,在传统计量理论和新兴神经网络框架的基础上展开相关研究。

在波动率预测方面,本文基于正态分布、t 分布、GED 分布三种残差分布假设条件下建立 GARCH 模型,利用滚动预测的方式预测上证指数的波动率,并将含有波动率信息的 GARCH 模型参数作为 LSTM 模型的输入变量,构建 LSTM-GARCH 融合模型预测波动率,根据输入参数的不同,LSTM-GARCH 模型共计分为七种:L-G、L-GT、L-GGED、L-G-GT、L-G-GGED、L-GT-GGED、L-G-GT-GGED,并比较融合模型与单一模型的拟合效果,为 VaR 测度提供更准确的波动率输入变量。

在 VaR 金融风险度量方面,本文借鉴 Taylor 提出的 QRNN 模型,在此基础上加入滞后内生变量建立 QARNN 模型,并将融合模型预测得出的波动率结果作为神经网络的输入变量,从不同置信水平分别测算 VaR 值,且将 QARNN 模型与 QR 模型、QAR 模型、QRNN 模型比较,利用失败率和 Kupiec 检验中的似然统计量 LR 来对比分析四种模型估计 VaR 的准确程度。

通过本文四、五章的实证分析,主要得出以下结论:

上证指数的收益率序列存在尖峰厚尾、偏态分布的特征。同时,该序列存在波动聚集性,表明波动随时间变化而变化,大波动聚集在一些时间段发生、小波动聚集在一些时间段发生的现象,波动之间存在正相关关系,呈现出波动聚集的特征。

在波动率预测模型中,通过损失函数横向比较 GARCH 模型在不同残差分布假设下的波动率拟合能力,当假设残差服从于 GED 时,模型预测精确性高于其他 GARCH 族模型;通过损失函数横向对比单一模型的表现,LSTM 模型的预测精度高于 GARCH 模型;在量价数据的基础上增加 GARCH 模型滚动预测得到的参数作为 LSTM 模型输入得到的融合模型,通过损失函数对比融合模型和单一模型的表现,融合具有更好的预测性能;在融合模型中,融合三种 GARCH 模型和 LSTM 模型的 L-G-GT-GGED 融合模型表现效果最优。

在 VaR 测度中,通过返回测试对比模型 VaR 测度的准确程度,代表传统计量模型的 QR 模型和 QAR 模型均在测试集内出现低估损失的情况;QRNN 模型的模型预测性能优于 QR 模型和 QAR 模型,失败率接近期望失败率,对 VaR 测度较为合理;QARNN 模型整体表现最佳,失败率低于期望失败率,该模型不仅考虑了波动率对 VaR 的关键影响,而且考虑了滞后内生变量的非线性效应,使得模型准确性显著提升。

综上所述,在对上证指数的波动率预测中,传统计量模型与深度学习算法相结

合的 LSTM-GARCH 融合模型相较于单一 GARCH 模型和单一 LSTM 模型都有着更为优异的表现,同时为后续 VaR 测度提供了更准确的波动率输入变量;在 VaR 测度方面,综合考虑了滞后内生变量和融合了传统计量模型的 QARNN 模型在很大程度上改进了传统分位数回归模型在 VaR 测度上的预测性能。

## 第二节 研究展望

本文以上证指数为研究对象,建立了 LSTM-GARCH 融合模型预测波动率,该模型相较于单一 GARCH 模型和单一 LSTM 模型准确率显著提高;随后在 QRNN 模型的基础上加入滞后内生变量,建立 QARNN 模型估计 VaR 值,该模型大大降低了返回测试中的失败率,为 VaR 的准确测度提供了新的方法,该模型对金融风险的控制与管理也有较好的实用意义。然而,本文中也还存在很多不足之处,对于波动率预测和金融风险定量分析也有待进行更多的探索。

数据方面。本文利用上证指数日收盘价计算收益率,并在此基础上展开波动率预测和 VaR 值测度的研究。随着研究的深入,可以进一步地利用日内高频数据预测日实现波动率(Realized Volatility, RV)对金融市场风险进行量化。通过对一天内多个时间段的波动率刻画,利用已实现波动率进一步地去测度每日的 VaR 值,实现对金融风险更加准确的计量。

风险量化方法方面,除在险价值 VaR 外,损失期望值 ES 也是行之有效的一种风险度量方法。“损失期望值”顾名思义,是通过对超过 VaR 值的损失求均值的方法来量化尾部风险,能够有效解决 VaR 方法存在的不满足次可加性和不足以表达尾部风险问题,对风险的准确度量提供了另一种可能性。

## 参考文献

- [1] Engle R. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation[J]. *Econometrica*, 1982, 50(50).
- [2] Bollerslev T. Generalized autoregressive conditional heteroscedasticity[J]. *Journal of Econometrics*, 1986, 31(2): 307-327.
- [3] Zakoian J. Threshold heteroskedastic models[J]. *Journal of Economic Dynamics & Control*, 2004,18(5):931-955.
- [4] Nelson D. Conditional heteroskedasticity in asset returns: A new approach[J]. *Econometrica*, 1991,59:347- 370.
- [5] Glosten, Jagannathan, Runkle. On the Relation Between Expected Value and the Volatility of the Nominal Excess Return on Stocks[J]. *Journal of Finance*, 1993,48: 1779- 1801.
- [6] Engle R. Estimating time varying risk premia in the term structure: The ARCH\_Model [J]. *Econometrica*, 1987,55:391-407.
- [7] Engle R, Bollerslev T. Modeling the persistence of conditional variances[J]. *Econometric Reviews*, 1986,5:1-50.
- [8] Engle R, Kroner K. Multivariate simultaneous generalized ARCH[J]. *Econometric Theory*, 2000,11(1):122-150.
- [9] 田华, 曹家和. 中国股票市场报酬与波动的 GARCH-M 模型[J].系统工程理论与实践, 2003(8):81-86.
- [10] 黄海南, 钟伟. GARCH 类模型波动率预测评价[J].中国管理科学,2007(06):13-19.
- [11] 魏宇, 余怒涛. 中国股票市场的波动率预测模型及其 SPA 检验[J].金融研究,2007(07):138-150.
- [12] 赵伟雄, 崔海蓉, 何建敏. GARCH 类模型波动率预测效果评价——以沪铜期货为例 [J].西安电子科技大学学报(社会科学版),2010,20(04):27-32.
- [13] 杨继平, 袁璐, 张春会. 基于结构转换非参数 GARCH 模型的 VaR 估计[J].管理科学学报,2014,17(02):69-80.
- [14] 王天一, 黄卓. Realized GAS-GARCH 及其在 VaR 预测中的应用[J].管理科学学报,2015,18(05):79-86.
- [15] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015,521(7553).
- [16] 李亚静, 何跃, 朱宏泉. 中国股市收益率与波动性长记忆性的实证研究[J].系统工程理论与实践,2003(01):9-15.
- [17] Hochreiter S, Schmidhuber J. Long short-term memory [J]. *Neural computation*, 1997,

- 9(8): 1735-1780.
- [18] Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural network [C]. *IEEE*, 2013
- [19] 陈卫华. 基于深度学习的上证综指波动率预测效果比较研究[J]. 统计与信息论坛, 2018, 33(05): 99-106.
- [20] 张蕾, 孙尚红, 王月. 基于深度学习 LSTM 模型的汇率预测[J]. 统计与决策, 2021, 37(13): 158-162.
- [21] 刘广应, 吴鸿超, 孔新兵. 深度学习 LSTM 模型与 VaR 风险管理[J]. 统计与决策, 2021(8): 136-140.
- [22] Horvath B, Muguruza A, Tomas M. Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models[J]. *Quantitative Finance*, 2021, 21(1).
- [23] Kim H, Won C. Forecasting the Volatility of Stock Price Index: A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models[J]. *Expert Systems with Applications*, 2018, 103(aug.): 25-37.
- [24] Yan H, Jian N, Liu W. A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction[J]. *Physica A: Statistical Mechanics and its Applications*, 2020, 557.
- [25] 操玮, 任思儒. 基于 LSTM 与 GARCH 族混合模型的人民币汇率波动预测研究[J]. 计算机应用研究, 2020, 37(S1): 79-82.
- [26] Verma S. Forecasting volatility of crude oil futures using a GARCH-RNN hybrid approach[J]. *Intelligent Systems in Accounting, Finance and Management*, 2021, 28(2).
- [27] Zeng H, Shao B, Bian G, Dai H, Zhou F. A hybrid deep learning approach by integrating extreme gradient boosting-long short-term memory with generalized autoregressive conditional heteroscedasticity family models for natural gas load volatility prediction[J]. *Energy Science & Engineering*, 2022, 10(7).
- [28] Ruey S. Tsay. 金融时间序列分析[M]. 机械工业出版社, 2009.
- [29] Manganelli S, Engle R. CAViaR: Conditional Value at Risk by Quantile Regression[J]. *NBER Working Papers*, 1999.
- [30] Morgan J. Riskmetrics: Technical document[R]. *Working Paper, Morgan Guaranty Trust Company of New York*, 1996.
- [31] Nelson, Daniel B. Conditional Heteroskedasticity in Asset Returns: A New Approach[J]. *Modelling Stock Market Volatility*, 1991, 59(2): 347-0.
- [32] 陈泽忠, 杨启智, 胡金泉. 中国股票市场的波动性研究——EGARCH-M 模型的应用[J]. 决策借鉴, 2000, 13(5): 24-27.

- [33]江涛. 基于 GARCH 与半参数法 VaR 模型的证券市场风险的度量和分析:来自中国上海股票市场的经验证据[J]. 金融研究, 2010(6):9.
- [34]Xiao Z, Koenker R. Conditional Quantile Estimation for Generalized Autoregressive Conditional Heteroscedasticity Models[J]. *Journal of the American Statistical Association*, 2012, 104: 1696-1712.
- [35]Koenker R, Bassett G. Regression quantiles[J]. *Econometrica*, 1978, 46(1):211-244.
- [36]王新宇, 赵绍娟. 基于分位数回归模型的沪深股市风险测量研究[J]. 中国矿业大学学报, 2008, 37(3):416-421.
- [37]王传美,张炼,贺素香,吴海英.基于 GARCH 类模型和分位数回归的创业板市场隔夜风险度量[J].武汉金融,2019(02):32-36.
- [38]闫昌荣. 基于流动性调整 CAViaR 模型的风险度量方法[J]. 数量经济技术经济研究, 2012, 29(3):11.
- [39]Diebold F, Schuermann T, Stroughair J. Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management[J]. *Center for Financial Institutions Working Papers*, 1998.
- [40]王艺馨, 周勇. 极端情况下对我国股市风险的实证研究[J]. 中国管理科学, 2012, 20(003):20-27.
- [41]Yu W, Wang C, Yu L. Measuring daily Value-at-Risk of SSE index: A new approach based on multifractal analysis and extreme value theory[J]. *Physica A: Statistical Mechanics and its Applications*, 2013, 392( 9):2163-2174.
- [42]许启发, 陈士俊,蒋翠侠. 基于 GARCH-EVT 模型的证券投资基金动态风险测度[J]. 合肥工业大学学报:自然科学版, 2015(38):1003.
- [43]MC Ausín, Galeano P, Ghosh P. A semiparametric Bayesian approach to the analysis of financial time series with applications to value at risk estimation[J]. *European Journal of Operational Research*, 2014, 232.
- [44]胡宗义,李毅,万闯,唐建阳.基于半参数 CARE 模型的金融市场 VaR 度量[J].统计与信息论坛,2019,34(04):19-24.
- [45]黄晓. 混合历史模拟法在金融风险度量中的应用[J]. 南方金融, 2009.
- [46]吕永健, 符廷奎, 胡颖毅. 基于拔靴滤波历史模拟法的黄金市场 VaR 测度研究[J]. 中国管理科学, 2019, 27(7):10.
- [47]Wang Z. The Properties of Incremental VaR in Monte Carlo Simulations[J]. *Journal of Risk Finance*, 2002, 3(3):14-23.
- [48]肖佳文,杨政.混合分布的 VaR 非参数估计:对期货市场的实证分析[J]. 系统工程学报,2016,31(04):471-480.
- [49]Taylor J. A quantile regression neural network approach to estimating the conditional

- density of multiperiod returns[J]. *Journal of Forecasting*, 2000, 19(4):299-311.
- [50] Christopher G, William D. Persistence in Variance, Structural Change, and the GARCH Model[J]. *Journal of Business & Economic Statistics*, 1990, 225-234.
- [51] Koenker, Roger W, Xiao Z. Quantile Autoregression[J]. *Journal of the American Statistical Association*, 2006, 980-990.