

分类号: O211.4

单位代码: 10335

密 级: 无

学 号: 11635035

浙江大学

博士学位论文



中文论文题目: 基于厚尾时间序列的非参假设检验

英文论文题目: Nonparametric Hypothesis Testing for
Heavy-tailed Time Series

申请人姓名: 周沫

指导教师: 张荣茂教授

专业名称: 概率论与数理统计

研究方向: 时间序列分析

所在学院: 数学科学学院

论文提交日期: 二〇二一年十月

基于厚尾时间序列的非参假设检验



论文作者签名: 周沫

指导教师签名: 张勇

论文评阅人 1: 陈坤 教授 西南财经大学
评阅人 2: 黎德元 教授 复旦大学
评阅人 3: 王文胜 教授 杭州电子科技大学
评阅人 4: 隐名评阅人
评阅人 5: 隐名评阅人

答辩委员会主席: 梁汉营 教授 同济大学
委员 1: 梁汉营 教授 同济大学
委员 2: 苏中根 教授 浙江大学
委员 3: 王江峰 教授 浙江工商大学
委员 4: 庞天晓 教授 浙江大学
委员 5: 张朋 教授 浙江大学

答辩日期: 二〇二一年十一月

Nonparametric Hypothesis Testing for Heavy-tailed Time Series



Author's signature: 周沐

Supervisor's signature: 张英

External Reviewers:

Kun Chen Prof. Southwestern University of Finance and Economics

Deyuan Li Prof. Fudan University

Wensheng Wang Prof. Hangzhou Dianzi University

Anonymous Reviewer

Anonymous Reviewer

Examining Committee Chairperson:

Hanying Liang Prof. Tongji University

Examining Committee Members:

Hanying Liang Prof. Tongji University

Zhonggen Su Prof. Zhejiang University

Jiangfeng Wang Prof. Zhejiang Gongshang University

Tianxiao Pang Prof. Zhejiang University

Peng Zhang Prof. Zhejiang University

Date of oral defence: November 2021

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得浙江大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：周沫

签字日期：2021 年 12 月 24 日

学位论文版权使用授权书

本学位论文作者完全了解浙江大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权浙江大学可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名：周沫

导师签名：

张英

签字日期：2021 年 12 月 24 日

签字日期：2021 年 12 月 24 日

致 谢

九年的求是之路即将告一段落，在此论文完成之际，谨向所有帮助过我的老师、同学、朋友及家人表示衷心的感谢。

我最想感谢的是我的导师张荣茂老师，从我本科第三年开始接触科研到如今的博士毕业论文，我的点滴成长都凝聚了张老师的心血。依稀记得第一次见到张老师是在他的金融风险管理课堂，张老师深入浅出的授课让我第一次接触到时间序列在金融领域的应用。张老师对我本科毕业论文的悉心指导，让我有幸能深入学习时间序列的相关知识，并且尝试自己动手去解决实际问题。在整个博士求学生涯中，张老师的悉心指导和严格要求让我打下了扎实的基础。忙碌的张老师总是坚持每周抽出时间来主持讨论班，和我们探讨前沿问题，并跟进我们的研究进度；每次遇到困难，就算是在假期和休息时间，张老师都会及时地提供指导和帮助。与张老师无论是科研还是生活中的每次交流都让我受益匪浅，张老师对科研敏锐的直觉、精辟独到的想法以及扎实的理论基础让人钦佩，我非常荣幸能够成为张老师的学生。

感谢美国乔治亚州立大学的彭亮教授。依稀记得第一次见到彭老师是在他来浙江大学做学术报告的期间，在和彭老师一个多小时的交流中，他给予了我很多学术和职业规划上的建议。在有幸和彭老师的三次合作中，彭老师在学术上给予了我莫大的帮助和悉心的指导，尤其是在模拟研究方面。无论多忙，彭老师都会及时跟进项目的进展，迅速回复邮件与我们一起探讨问题。彭老师认真负责的态度深深地感染了我，使我养成了很好的科研习惯。感谢这几年来彭老师在科研上对我的帮助。

感谢北方民族大学的马耀兰老师。在我的博士生涯中，无论我碰到什么问题，马老师都第一时间和我一起探讨和解决。我记得每一个我们探讨问题到深夜的情形，每一次我遇到困难做不出来时马老师对我的鼓励。在数次的合作和交流中，马老师和我也建立了坚固的友谊。感谢马老师在我的求学生涯中的支持和鼓励。

同时，我也要感谢张立新教授，苏中根教授，庞天晓教授，张朋教授等统计研究所的老师们对我的指导和帮助。

感谢廖桂丽师姐，刘其梦师姐，习代青师姐和许佳攀师姐在科研上对我的帮助；感谢一起朝夕相处，并肩作战的池昌雄，吴潇然，胡日成，梁浩杰，他们的陪伴让我的科研之路再也不孤单；也感谢很多未提到的同学和朋友们，能和他们相识，是我一生的宝贵财富。

一路走来，我要特别感谢我的母亲一直以来对我的支持和鼓励，在我快要坚持不下去的时候给我勇气和动力，她的默默付出让我可以义无反顾，安心地做自己喜欢做的事情；同时，也要感谢我的男朋友沈阳同学，从本科以来就一直支持我选择读博这条路，有了他的支持和陪伴，我才能心无旁骛地专心科研。最后，感谢我的两只猫咪傻脸和卷福，它们的陪伴让我不再孤单。此文也献给我在天堂的外公。

周沫

二零二一年十一月于求是园

摘 要

重尾现象及其理论在许多科学领域都起着非常重要的作用,例如保险和经济学。本文主要研究了基于重尾时间序列的非参数假设检验。我们讨论了当误差四阶矩不存在时,对 ARMA-GARCH 模型的经验似然检验;以及当误差具有无限方差时的 Portmanteau 型单位根检验。本论文由以下部分组成。

第一部分,我们考虑了当误差四阶矩不存在时,ARMA-GARCH 模型中误差零均值的检验问题。对于重尾 ARMA-GARCH 模型,经典的拟极大似然估计(QMLE)不能保证估计量的渐近正态性成立,所以一般经常使用自加权拟极大指数似然估计(SWQMELE)。但是,SWQMELE 要求误差的中位数为 0,而经典的 ARMA-GARCH 模型通常假设均值为 0。为了保证 SWQMELE 的假设成立,我们在对模型做相应的变换时可能会破坏 ARMA-GARCH 模型的结构。因此,在误差的中位数为 0、绝对矩为 1 的假设下,我们提出了一个可以有效检验误差零均值的经验似然检验。在常规的假设下,我们建立了该检验统计量在原假设和局部备择假设下的极限分布。模拟研究及对美国房价指数和金融收益率序列的应用证实了该检验在有限样本下具有良好的数值效果。

第二部分,为了提高检验对异方差和重尾的稳定性,我们假设误差均值为 0、方差为 1,通过使用 GARCH 结构而无需估计 GARCH 模型来对误差的中位数是否为 0 进行检验。此外,我们也给出了改进过后的检验统计量在原假设和局部备择假设下的极限分布。该检验的有效性也被模拟研究所证实。对金融收益率序列的数据分析表明,在使用 SWQMELE 时,应该注意其条件是否满足。

第三部分,为了提高零均值误差检验的稳定性,我们再次利用第二部分的思想,即利用 GARCH 结构而不估计 GARCH 模型,在误差的中位数为 0、二阶矩为 1 的条件下,结合基于中位数的估计方法和随机加权自助抽样法来检验误差零均值。我们给出了在原假设和局部备择假设下检验的渐近理论。对几种外汇数据的实证分析表明,在拟合模型之前进行假设检验很有必要。

最后,以上研究是针对误差四阶矩不存在且方差有限的 ARMA-GARCH 模型。为了进一步研究方差无限的时间序列,我们考虑了带有重尾线性噪声的自回归模型的单位根检验,即噪声项具有 $\varepsilon_t = \sum_{j=0}^{\infty} d_j \eta_{t-j}$ 的形式,其中 $\{\eta_t\}$ 独立同分布且属于 α -stable 分布吸引域, $\alpha \in (0, 2)$ 。我们提出了基于样本协方差的 Portmanteau 型

检验统计量，该检验统计量有非参的形式，且易于实现。我们证明了该检验统计量的极限分布为标准 **stable** 分布的泛函。模拟研究证明了我们提出的统计量在 **size** 和 **power** 上表现优于传统的单位根检验统计量。对世界原油价格和金融商业票据利率的应用表明我们提出的统计量有良好的表现。

关键词：重尾；非参数；ARMA-GARCH 模型；经验似然；拟极大似然估计；自加权拟极大指数似然估计；加权估计；零中位数；零均值；单位根检验；线性过程；**stable** 分布

Abstract

The phenomena of heavy tails exist in many area of science, such as insurance and economics. This thesis mainly focuses on the nonparametric hypothesis testing for heavy-tailed time series. We study the hypothesis testing for errors using empirical likelihood method in the procedure of fitting an ARMA-GARCH model when the errors have infinite fourth moment, as well as the Portmanteau-type test for unit root when errors have infinite variance. This thesis consists of following parts.

First, in the procedure of fitting an ARMA-GARCH model, the self-weighted quasi maximum exponential likelihood estimation (SWQMELE) is often employed instead of the quasi maximum likelihood estimation (QMLE) to reduce the moment constraints. However, SWQMELE requires the errors to have zero median instead of zero mean which is assumed by QMLE. Because changing zero mean to zero median destroys the ARMA-GARCH structure and has a serious effect on skewed data, we consider an efficient empirical likelihood test for zero mean of errors in the application of SWQMELE to ensure that the model still concerns conditional mean. We derive the limit distributions of the test under both null hypothesis and local alternative hypothesis under mild assumptions. A simulation study confirms the good finite sample performance before applying the test to the US housing price indexes and financial returns for the study of comovement.

Second, to improve the robustness against heteroscedasticity and heavy tails, we employ the empirical likelihood test for testing zero median of errors by using the GARCH structure to reduce the moment effect without estimating the GARCH model. The limit distribution of the test under null hypothesis and local alternative hypothesis are developed as well. The effectiveness of the proposed test is confirmed by simulation study. The data analysis shows that some financial returns do not have zero median of errors, which cautions the use of the SWQMELE.

Third, to improve the robustness of the test for zero mean of the errors, we also consider to use a random weighted bootstrap method combined with the idea of using the GARCH structure to reduce the moment effect without estimating the GARCH model. Asymptotic theory for our proposed test has been developed under both null hypothesis and local al-

ternative hypothesis. Empirical analysis on several exchange rates shows the hypothesis testing before fitting the model is necessary.

Finally, the above study focuses on ARMA-GARCH model whose errors have infinite fourth moment and finite variance. To further study the time series with infinite-variance errors, we focus on the test for the existence of the unit-root of an autoregressive model under a heavy-tailed linear noise, where the noise $\varepsilon_t = \sum_{j=0}^{\infty} d_j \eta_{t-j}$ and $\{\eta_t\}$ is a sequence of i.i.d. innovations belonging to the domain of attraction of an α -stable distribution for some $\alpha \in (0, 2)$. We propose a Portmanteau-type statistic based on the sample covariance which is nonparametric and easy to implement. Under certain regular assumptions, the limit distribution of the statistics is shown to be a functional of a standard stable distribution. The good finite sample studies show that the new test outperforms conventional unit root tests in size and power. Applications to the daily world crude oil price and 3-month AA financial commercial paper rate are also given to illustrate the performance of the new test.

Keywords: heavy tail; nonparametric; ARMA-GARCH model; empirical likelihood; quasi-maximum likelihood estimation; self-weighted quasi maximum exponential likelihood estimation; weighted estimation; zero median; zero mean; unit root test; linear process; stable distribution

文中部分缩写及符号说明

i.i.d.	独立同分布
$ \cdot $	绝对值函数
$\ \cdot\ $	欧式范数
$\exp\{\cdot\}$	指数函数
$\mathbb{I}(A)$	集合 A 的示性函数
EX	随机变量 X 的期望
$\text{Var}X$	随机变量 X 的方差
$\mathbb{R}$	实数集合
$\mathbb{R}_0$	正实数集合
$D[0, 1]$	在 $[0, 1]$ 上右连续且有左极限的函数空间
$\sim$	服从分布
$N(\cdot, \cdot)$	正态分布
$\text{Beta}(\cdot, \cdot)$	Beta 分布
$\text{Laplace}(\cdot, \cdot)$	Laplace 分布
$\text{Uniform}(\cdot, \cdot)$	均匀分布
$\text{Pareto}(\cdot, \cdot)$	Pareto 分布
$\xrightarrow{d}$	依分布收敛
$\xrightarrow{p}$	依概率收敛
$\Rightarrow$	弱收敛
$a_n = O(b_n)$	$\lim_{n \rightarrow \infty} a_n/b_n < \infty$
$a_n = o(b_n)$	$\lim_{n \rightarrow \infty} a_n/b_n = 0$
$a_n = O_p(b_n)$	a_n/b_n 依概率有界
$a_n = o_p(b_n)$	a_n/b_n 依概率收敛到 0

图目录

3.1	S&P500 和 Microsoft Stock (MSFT) 每日收盘价的时间序列图	20
3.2	S&P500 和 Microsoft Stock (MSFT) 每日收盘价的对数收益率 ($\times 100$) . .	20
3.3	对 S&P 500 和 MSFT 拟合的 η_t 和 $ \eta_t $ 的自协方差函数图	22
3.4	对 S&P 500 和 MSFT 拟合的 η_t 的 Hill 估计量和其 90% 置信区间 . . .	22
4.1	HKD/USD, USD/EUR, CAD/USD 和 MXN/USD 外汇的 Hill 估计量 . .	39
4.2	RIG, S&P 500, MSFT 和 DJI 的 Hill 估计量	40
5.1	HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 INR/USD 外汇的时间序列图	58
5.2	HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 INR/USD 外汇的 Hill 估计量	59
6.1	世界原油价格和 3 个月 AA 金融商业票据利率 (DCPF3M) 的时间序 列图	75
6.2	原油价格和 3 个月 AA 金融商业票据利率 (DCPF3M) 的 Hill 估计量 .	76

表目录

3.1	原假设下 profile 经验似然检验的 size	18
3.2	备择假设下 profile 经验似然检验的 power	18
3.3	对美国房价指数 (HPI) 拟合 AR(3)-GARCH(1,1) 模型的 SWQMELE .	19
3.4	美国房价指数 (HPI) 的 profile 经验似然检验	19
4.1	$\delta = 0.5$ 时, ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型经验似然检验的 size	36
4.2	$\delta = 0.3$ 时, ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型经验似然检验的检验 power	37
4.3	$\delta = 0.4$ 时, ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型经验似然检验的检验 power	38
4.4	HKD/USD, USD/EUR, CAD/USD 和 MXN/USD 外汇经验似然检验的 p 值	40
4.5	RIG, S&P 500, MSFT 和 DJI 经验似然检验的 p 值	41
5.1	ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型检验的 size .	57
5.2	ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型检验的 power	58
5.3	对 HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 IN-R/USD 外汇的对数收益率拟合的 ARMA 模型和检验的 p 值	60
6.1	$\alpha = 1.5$ 时, 不同 M 下 T_n 的经验分位数	68
6.2	$\alpha = 1$ 时, 不同 M 下 T_n 的经验分位数	69
6.3	$\alpha = 0.5$ 时, 不同 M 下 T_n 的经验分位数	70
6.4	$\alpha = 1.5$ 和 $c = 40$ 时 Portmanteau 型检验、PP、DF 和 ADF 检验的 power	71
6.5	$\alpha = 1$ 和 $c = 40$ 时 Portmanteau 型检验、PP、DF 和 ADF 检验的 power	72
6.6	$\alpha = 0.5$ 和 $c = 40$ 时 Portmanteau 型检验、PP、DF 和 ADF 检验的 power	72
6.7	$\alpha = 1.5$ 时 Portmanteau 型检验的 size	72
6.8	$\alpha = 1$ 时 Portmanteau 型检验的 size	73
6.9	$\alpha = 0.5$ 时 Portmanteau 型检验的 size	73
6.10	Wild 自助法下, $\alpha = 0.5$ 和 $c = 40$ 时 Portmanteau 型检验的 power . . .	74

6.11 原油价格和 3 个月 AA 金融商业票据利率 (DCPF3M) 在 4 种检验统 计量下的值	76
--	----

目 次

致谢	I
摘要	III
Abstract	V
文中部分缩写及符号说明	VII
图目录	VIII
表目录	X
目录	
1 引言	1
1.1 选题背景与意义	1
1.2 研究方法简介	4
1.3 文章结构	5
2 预备知识	7
2.1 经验似然理论	7
2.1.1 非参极大似然	7
2.1.2 经验似然检验	7
2.1.3 均值的经验似然检验	8
2.1.4 基于估计方程的经验似然检验	9
2.1.5 Profile 经验似然检验	9
2.1.6 经验似然检验在时间序列上的应用	10
2.2 重尾理论	11
2.2.1 Stable 分布	11
2.2.2 吸引域	12
2.2.3 Hill 估计量	13
3 ARMA-GARCH 模型误差零均值的经验似然检验	14
3.1 引言	14
3.2 方法和理论结果	14
3.3 模拟研究	17
3.4 实证分析	18
3.4.1 美国房价指数	18

3.4.2 金融收益率序列	19
3.5 结论	23
3.6 定理 3.1 和 3.2 的证明	23
4 ARMA-GARCH 模型误差零中位数的经验似然检验	31
4.1 引言	31
4.2 方法和理论结果	32
4.3 模拟研究	35
4.4 实证分析	38
4.4.1 外汇序列	38
4.4.2 股票序列	40
4.5 结论	41
4.6 定理 4.1 和 4.2 的证明	41
5 基于随机加权自助法的 ARMA-GARCH 模型误差零均值检验	54
5.1 引言	54
5.2 方法和理论结果	54
5.3 模拟研究	57
5.4 实证分析：汇率序列	58
5.5 结论	60
5.6 定理 5.1 和 5.2 的证明	60
6 重尾噪声下的 Portmanteau 型单位根检验	64
6.1 引言	64
6.2 方法和理论结果	64
6.3 Wild 自助抽样算法	66
6.4 模拟研究	67
6.5 实证分析	74
6.6 结论	76
6.7 定理 6.1 的证明	76
6.7.1 S_{n1} 的收敛性	77
6.7.2 S_{n2} 的收敛性	80
6.7.3 S_{n3} 的收敛性	80
6.7.4 T_n 的收敛性	86
7 总结与展望	89
参考文献	91

附录	99
作者简历及在学期间所取得的科研成果	113

1 引言

1.1 选题背景与意义

重尾现象一直以来在许多科学领域都很常见，比如在保险和金融领域。该现象主要是由于序列中取极值的概率相比一般分布（比如正态分布和指数分布）要大，其分布函数往往呈现“尖峰厚尾”的形态。重尾现象的分析是极值理论的一个分支，后者主要研究一些极端罕见事件的发生，如百年一遇的洪水和地震等。与经典的统计建模和推断不同的是，重尾现象的统计分析很大程度上受到尾部形态的影响，其分析和研究主要基于规则变化函数、概率测度的弱收敛、随机测度和点过程；其应用范围十分广泛，比如在金融中，资产回报往往是重尾的，进行资产管理和风险度量时需要使用重尾理论；在保险中，再保险领域往往更加关注极值的存在。因此，当系统的行为被极值所支配时，如何对其进行统计推断也是近几年来学者十分关注的问题。

在许多经济和金融领域中，序列经常存在偏斜性、时间相关性和波动的持久性。为了刻画这些特点，一个常用的时间序列模型是著名的自回归滑动平均-广义自回归条件异方差（ARMA-GARCH）模型，表示为：

$$\begin{cases} y_t = \varphi + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \psi_j \varepsilon_{t-j} + \varepsilon_t, \varepsilon_t = \eta_t \sqrt{h_t}, \\ h_t = \alpha_0 + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \end{cases} \quad (1.1)$$

这里 $\alpha_0 > 0$, $\alpha_i \geq 0$ ($i = 1, \dots, r$), $\beta_j \geq 0$ ($j = 1, \dots, s$), $\{\eta_t\}$ 是 i.i.d. 的随机变量且均值为 0 方差为 1。基于 $E\eta_t = 0$ 和 $E\eta_t^2 = 1$, 参数模型 (1.1) 很好地刻画了 $\{X_t\}$ 的条件均值和条件标准差。自从 Engle (1982) 和 Bollerslev (1986) 以来，对 GARCH 模型和 ARMA-GARCH 模型的理论性质的研究和统计推断受到了极大的关注。比如，Basrak et al. (2002) 研究了 GARCH 模型的尾部行为和样本自相关函数；Mikosch and Starica (2000)、Berkas et al. (2003)、Chan et al. (2012) 和 Zhang et al. (2019) 考虑了 GARCH(1,1) 模型尾部指数的估计。拟合模型 (1.1) 常用的统计方法是著名的拟极大似然估计（QMLE）法，参考 Ling and Li (1997a)、Jeantheau (1998)、Berkas et al. (2003)、Ling and McAleer (2003)、Hall and Yao (2003) 和 Francq and Zakoian (2004) 等。众所周知，当 $E\varepsilon_t^4 < \infty$ 和 $E\eta_t^4 < \infty$ 成立时，即当序列存在一个较轻的尾部时，QMLE 收敛到一个正态分布。

实际中， $\sum_{i=1}^r \alpha_i + \sum_{j=1}^s \beta_j$ 经常接近于 1，表明 $E\varepsilon_t^4 < \infty$ 经常不成立，即模型 (1.1) 中误差经常是重尾的。如何在重尾情况下对模型 (1.1) 进行统计推断，是近年来一直受到研究者们备受关注的的问题。为了解决这个问题，Ling (2007) 曾提出了

一种自加权的局部 QMLE, 当 $E\eta_t^4 < \infty$ 和 $E|\varepsilon_t|^\tau < \infty$ 对某个 $\tau > 0$ 成立时, 该估计量可以有正态极限。当模型 (1.1) 是一个纯 GARCH 模型时, Hall and Yao (2003) 证明了 QMLE 在 $E\eta_t^4 = \infty$ 下收敛到一个 stable 分布。然而, 由于 stable 随机变量的分布函数没有显性的表达式, 使得给定置信 (显著) 水平下的置信 (拒绝) 域的确确定变得非常复杂。因此, 我们往往希望在 $E\eta_t^4 = \infty$ 下, 所得的估计量依然能够收敛于正态分布, 以便进行统计推断。一个常用的方法是使用基于中位数的统计推断, 绝对偏差估计 (LADE) 是其中最为常见的方法。该方法可以提高在四阶矩不存在的情况下估计的稳定性, 并且在一定的条件下具有渐近正态性。局部 LAD 的方法曾被 Peng and Yao (2003) 和 Li and Li (2005, 2008) 等研究过; 全局 LAD 则被 Horváth and Liese (2004) 和 Berkes and Horváth (2004) 等研究过。Peng and Yao (2003) 提出了针对 GARCH 序列的 LADE, 并推导出了该估计量的渐近正态性。他们假设了 η_t^2 的中位数为 1 而不是方差为 1。但由于 LAD 的目标函数是不可导且非凸的, 此方法在使用时会过于复杂, 计算成本可能会过大。最近 Zhu and Ling (2011) 提出了一个自加权拟极大指数似然估计 (SWQMELE) 来解决重尾情况下模型估计的问题, 通过假设 $E|\eta_t| = 1$ 和 η_t 的中位数为 0, 而不是 (1.1) 中的 $E\eta_t^2 = 1$, 无论 $E\eta_t^4 = \infty$ 还是 $E\eta_t^4 < \infty$, 都能确保估计量收敛到正态分布。

然而, 当 η_t 不满足 SWQMELE 所需的假设: $E|\eta_t| = 1$ 和 η_t 的中位数为 0 时, 为了使得模型满足假设, 我们需要对模型做一些变换。若 $E|\eta_t| = d > 0$, 其中 d 未知, 通过定义

$$\eta_t^* = \eta_t/d, h_t^* = d^2 h_t, \alpha_i^* = d^2 \alpha_i, \text{ 其中 } i = 0, 1, \dots, r, \beta_j^* = \beta_j, \text{ 其中 } j = 1, \dots, s, \quad (1.2)$$

模型 (1.1) 相当于

$$y_t = \varphi + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \psi_j \varepsilon_{t-j} + \varepsilon_t, \varepsilon_t = \eta_t^* \sqrt{h_t^*}, h_t^* = \alpha_0^* + \sum_{i=1}^r \alpha_i^* \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j^* h_{t-j}^*. \quad (1.3)$$

因此, 为了使得 $E|\eta_t| = 1$ 成立, 只需要对参数 α_i 进行简单的尺度变换。更重要的是, 这种尺度转换一般不会改变常用的风险度量的统计推断, 例如 conditional Value-at-Risk (CVaR)、 y_t 在给定 $\mathcal{F}_{t-1}$ ($\mathcal{F}_t$ 为 $\{y_s : s \leq t\}$ 生成的 σ 域) 下的 conditional Expected Shortfall (CES)、 η_t 的 VaR 和 ES 及多个序列之间的联动性。例如, 由模型 (1.1) 和 (1.3) 可得:

$$\begin{aligned} P(y_t \leq y | \mathcal{F}_{t-1}) &= P(\eta_t \leq (y_t - \varphi - \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \psi_j \varepsilon_{t-j}) / \sqrt{h_t} | \mathcal{F}_{t-1}) \\ &= P(\eta_t^* \leq (y_t - \varphi - \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \psi_j \varepsilon_{t-j}) / \sqrt{h_t^*} | \mathcal{F}_{t-1}). \end{aligned}$$

也就是说, 给定 $\mathcal{F}_{t-1}$ 下, 根据 (1.1) 计算出的 y_t 的 CVaR 与根据 (1.3) 计算出的 CVaR

是一样的。然而,当模型 (1.1) 中 η_t 的中位数不为 0 时,为了使得 η_t 的中位数为 0 来满足 SWQMELE 的条件,我们需要进行复杂的位移变换,但是变换后 ARMA-GARCH 结构将不再成立。比如,如果 η_t 的中位数为 $m \neq 0$,那么为了让中位数为 0,需要做如下变换:

$$\varepsilon_t = m\sqrt{h_t} + (\eta_t - m)\sqrt{h_t}.$$

然而由于 $m\sqrt{h_t}$ 的存在,此模型不再具有零均值的 GARCH 误差,这样 SWQMELE 便不再适用。更多相关的信息,可参见 Fan et al. (2014)。因此,当应用 SWQMELE 来估计重尾情况下的 ARMA-GARCH 模型时,检验其假设是否满足具有重要意义。这将是本论文重点考虑的一个关键问题,我们将在本论文的第 3-5 章对此问题进行详细的探讨。

此外,非平稳性在金融和经济领域中也很有见,刻画这种现象最常用的模型之一是单位根模型。具体地说,对于一个 AR(1) 过程:

$$Y_t = \beta Y_{t-1} + \varepsilon_t, t = 1, \dots, n. \quad (1.4)$$

当 $\beta = 1$ 时,模型 (1.4) 为一个 AR(1) 单位根过程,也常被称为随机游走过程。单位根模型具有不可预测性,如果忽略单位根的存在就进行下一步的统计分析,经常会导致伪回归等问题。因此,对时间序列进行统计分析时,我们首先都会检验单位根是否存在。最经典的单位根检验方法是 Dickey and Fuller (1979) 提出的 DF 检验。但当 β 接近 1 且 ε_t 具有序列相关性时,基于最小二乘法的 DF 检验在有限样本下 power 很低。因此,研究人员们提出了一些改进方法来提高检验的 power,参考 Said and Dickey (1984) 提出的针对 ARMA 误差的 augmented Dickey-Fuller (ADF) 检验,和 Phillips and Perron (1988) 提出的 PP 检验(即 DF 检验的非参数版本)。最近, Jansson and Moreira (2006) 以及 Jansson (2008) 在不同形式的趋势项下使用了 Gaussian power envelope 进行单位根检验。频域方法,如冯-诺依曼方差比(VN)检验(参考 Sargan and Bhargava (1983) 和 Cai and Shintani (2006)), Fan and Gencay (2010) 的小波谱方法也被研究过。Zhang and Chan (2018) 提出了一种 Portmanteau 型检验统计量,该统计量通过利用序列的部分阶数的自相关系数,来提高序列相关情况下检验的 power,因此具有稳健性且易于使用。为了推导该 Portmanteau 型检验统计量的渐近分布, Zhang and Chan (2018) 要求噪声的四阶矩存在,即噪声项 ε_t 有较轻的尾部。然而,众所周知,重尾现象在许多经济和金融序列中经常存在,因此该假设在实际中很难满足。

与轻尾的情况相比,关于厚尾数据单位根检验的研究相对还是不够完善,部分相关文献如下: Chan and Tran (1989) 考虑了当误差为 i.i.d. 且属于 α -stable 分布吸引域 ($\alpha \in (0, 2)$) 的单位根检验。他们建立了在 $\beta = 1$ 下的普通最小二乘 (OLS) 估

计量的强一致性, 并证明了此情况下的 DF 检验的极限分布为 α -stable 过程的泛函。Knight (1989, 1991) 考虑了当噪声有无限方差时, 自回归参数 β 的 OLS 和 M-估计量的渐近行为, 他们得出 M-估计量的收敛速度更快。Ahn et al. (2001) 推导出了当噪声 ε_t 属于对称的 α -stable 分布的吸引域时, DF 检验、Lagrange 乘法、Durbin-Watson 检验和 Phillips 修正检验的极限分布。最近, Samarakoon and Knight (2009) 拓展了 Knight (1989, 1991) 的结果, 对带有方差无穷噪声的 p 阶自回归模型 (AR(p)), 考虑了基于 M-估计量的 DF 检验。Chan and Zhang (2010) 建立了当 ε_t 为重尾 GARCH(1,1) 噪声且方差无限时 DF 检验的极限分布。Arvanitis (2017) 得出了当噪声为一个属于 α -stable 分布吸引域的 i.i.d. 序列和一个正序列 (该正序列的无穷大截尾 α 阶矩在无穷点变化缓慢) 的乘积时, DF 检验的极限分布。然而, 他们并没有给出一个在实践中可以近似得出检验统计量经验分布的算法。在由线性过程驱动的渐近单位根过程的背景下, 其中该线性过程是一个重尾分布和一个轻尾分布的线性组合, Georgiev et al. (2017) 研究了重尾噪声对包括 ADF 和 PP 检验在内的常用的单位根检验在理论和实践中的影响。此外, 他们提出了一种基于 Eicker-White 标准误差的 ADF 检验的变形来提高传统单位根检验的 power。Cavaliere et al. (2018) 研究了当 ε_t 是一个由 i.i.d. 无限方差噪声驱动的线性过程时, ADF 统计量的渐近分布。Zhang and Chan (2020) 进一步考虑了当 ε_t 是由重尾 GARCH(1,1) 模型驱动的线性过程时, ADF 检验的渐近分布。然而, DF 检验对于序列相关的噪声通常来说 power 很低, 而 ADF 检验需要对差分过程进行参数化回归。这促使我们想要找到一种简单而直接的方法来检验重尾情况下, 尤其是无限方差噪声下单位根是否存在。本论文的第 6 章将对这个重要的检验问题进行探讨。

1.2 研究方法简介

本文中主要用到的方法之一是经验似然法, 该方法是由 Owen (1988, 1990) 提出的, 是一种结合了参数方法和非参数方法优点的数据驱动方法。与自助抽样法 (bootstrap) 和刀切法 (jackknife) 一样, 经验似然方法不需要假设数据服从一个特定的分布族; 与参数似然法类似, 经验似然法可以自动确定置信区间/区域的形状, 它直接融入了通过约束条件或先验分布表达的信息; 它也可以延伸到有偏抽样和删失数据的情况, 而且具有良好的渐近 power。经验似然法可以被当做是一种不进行再抽样的自助抽样法, 或者无需参数假设的似然法。在一定的正则条件下, 经验似然比统计量渐近地服从一个卡方分布, 所获得的置信区间/区域的形状由数据自动决定。由于其有效性, 经验似然法已被应用于各个领域来提供强有力的检验和准确的区间估计, 其中包括 Owen (1991) 和 Kolaczyk (1994) 在一般回归问题中的应用, Chuang and Chan (2002) 在非平稳 AR 模型中的应用, Chan and Ling (2006) 在 GARCH 模型中的

应用, Liu et al. (2008) 和 Ciuperca and Salloum (2015) 在变点检测中的应用, Fan and Huang (2005) 在变系数部分线性模型中的应用, Chen et al. (2012) 在门限 AR 模型中的应用, Li et al. (2012) 在 AR-ARCH 模型中对 LAD 估计的应用, 以及 Zhang et al. (2019) 在 GARCH 类型模型的尾指数推断中的应用。更多关于经验似然法的内容可以参考 Owen (2001)。

此外, 本文中还将使用的另一个方法是随机加权自助抽样法 (random weighted bootstrap method)。随机加权自助抽样法改进于 Wu (1986) 提出的传统的 wild 自助抽样法, 最初由 Jin et al. (2001) 提出。它是一种简单的重抽样方法, 通过多次反复扰动目标函数来进行统计推断。随机权重可由标准指数分布、伯努利分布等生成。到目前为止, 此方法已被广泛用于回归分析和时间序列模型的统计推断, 来解决估计量的渐近协方差矩阵不能被直接估计的问题。更多关于随机加权自助抽样法的内容可以参考 Chen et al. (2008, 2010)、Li et al. (2014, 2015)、Zhu and Ling (2015) 和 Zhu (2016)。

最后, 我们在单位根检验中用到了 Portmanteau 型检验统计量, 此类检验统计量针对的是原假设明确但备择假设比较松散的假设检验, 它对各种偏离原假设的情况都有很好的检验 power。使用这种检验统计量可以避免在各种特定的偏离原假设的情况下分别讨论的繁琐性。最著名的该类型统计量是由 Box and Pierce (1970) 以及 Ljung and Box (1978) 提出的主要用于检验 GARCH 模型中残差的 i.i.d. 假设 (即模型 (1.1) 中的 η_t), 它已成为对 GARCH 模型进行检验的常用工具。其他关于此类型统计量的例子, 可以参考 Li and Mak (1994)、Ling and Li (1997b)、Carbon and Francq (2011)、Hong and Lee (2003)、Escanciano (2008) 和 Ling and Tong (2011)。

1.3 文章结构

在第 2 章, 我们将给出本文所需要的预备知识, 包括经验似然理论和重尾理论。

在第 3 章, 我们考虑了用经验似然方法来检验拟合 SWQMELE 于 ARMA-GARCH 模型时假设是否满足。也就是说, 假设模型 (1.1) 中 $E|\eta_t| = 1$ 、 η_t 的中位数为 0 和 $E\eta_t^2 < \infty$ 成立, 我们用 profile 经验似然方法来检验下面的问题:

$$H_0 : E\eta_t = 0 \quad \text{v.s.} \quad H_1 : E\eta_t \neq 0.$$

在常规的假设条件下, 我们推导出了该检验统计量在原假设和局部备择假设下的极限分布。此外, 我们也给出了此方法的模拟研究, 并在实证分析部分对美国房价指数和金融收益率序列进行了检验。

在第 4 章, 我们通过使用 GARCH 结构而无需估计 GARCH 模型来提高检验的稳健性。我们继续考虑模型 (1.1), 但假设 $E\eta_t = 0$ 且 $E\eta_t^2 = 1$, 令 η_t 的中位数为 d ,

我们用 **profile** 经验似然方法检验：

$$H_0 : d = 0 \quad \text{v.s.} \quad H_1 : d \neq 0.$$

该检验对波动性和重尾稳健，且具有很好的 **power**。我们给出了在原假设和局部备择假设下该检验统计量的渐近理论。此外，详细的模拟研究和对金融收益率序列的数据分析表明，该检验统计量在有限样本下具有良好的数值效果。

在第 5 章，我们阐述了如何结合基于中位数的推断方法和随机加权自助抽样法来检验误差零均值。我们和第 3 章一样考虑检验：

$$H_0 : E\eta_t = 0 \quad \text{v.s.} \quad H_1 : E\eta_t \neq 0.$$

和第 3 章不同的是，我们应用了第 4 章的思想，即通过使用 **GARCH** 结构而无需估计 **GARCH** 模型，在 η_t 的中位数为 0 和 $E\eta_t^2 = 1$ 的假设下，检验 $E\eta_t$ 是否为 0。我们首先使用 **LAD** 方法估计了 **ARMA** 模型，估计量在当 $E(|\varepsilon_t|^\delta) < \infty$ 对某个 $\delta > 0$ 成立时渐近收敛于正态分布。接下来，我们考虑了用随机加权自助抽样法来检验原假设，并给出了相应的渐近理论。此外，对几种外汇数据的应用也表明，拟合模型 (1.1) 之前进行假设检验很有必要。

在第 6 章，我们考虑了 **Portmanteau** 型检验统计量来检验重尾噪声下单位根是否存在。我们将 **Zhang and Chan (2018)** 中对于有限方差情况下的 **Portmanteau** 型单位根检验拓展到了无限方差的情况。假设模型 (1.4) 中的 ε_t 是一个重尾线性过程，其中 $\varepsilon_t = \sum_{i=0}^{\infty} d_i \eta_{t-i}$ ， $\{\eta_t\}$ 是属于一个 α -stable 分布吸引域的 i.i.d. 噪声。我们证明了该检验统计量的极限分布是一个 **stable** 过程的泛函。我们的方法不仅保留了 **Zhang and Chan (2018)** 中检验方法易于实施的特性，而且相对于传统的检验统计量 **PP**、**DF** 和 **ADF** 检验，有更精确的 **size** 和更大的 **power**。详细的模拟研究表明，我们的方法具有良好的数值效果。在实证分析部分，我们也对世界原油价格和金融商业票据利率进行了检验。

在第 7 章，我们给出了本文的总结，并且对将来可能的拓展工作做了展望。

本文收录了作者五年来所撰写的部分论文，发表和投稿的详细情况可参见文末的附表。在写作过程中，本文也得到了 **NSFC (No.11771390/11371318)**、**USyd-ZJU** 合作研究项目和中央高校基础研究基金的支持。限于作者自身的学术水平，文中难免会有不当和谬误之处，也有很多值得深入思考和拓展的问题，恳请各位专家和读者不吝批评和指正。

2 预备知识

2.1 经验似然理论

2.1.1 非参极大似然

对一个随机变量 $X \in \mathbb{R}$, 其分布函数定义为 $F(x) = P(X \leq x)$, 其中 $-\infty < x < \infty$ 。我们用 $F(x-)$ 来表示 $P(X < x)$, 因此 $P(X = x) = F(x) - F(x-)$ 。下面我们给出经验分布函数和其非参似然函数的定义。

定义 2.1: 令 $X_1, \dots, X_n \in \mathbb{R}$, $X_1, \dots, X_n$ 的经验分布函数为

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(x_i \leq x), -\infty < x < \infty.$$

定义 2.2: 若 $X_1, \dots, X_n \in \mathbb{R}$ 独立且有共同的分布函数 F_0 , 若 F 为任意一个分布函数, 则其非参似然函数定义为:

$$L(F) = \prod_{i=1}^n (F(X_i) - F(X_i-)) = \prod_{i=1}^n p_i.$$

定理 2.1 证明了经验分布函数 F_n 可以使得非参似然函数 $L(F)$ 达到最大, 因此 F_n 是 F 的非参极大似然估计。

定理 2.1: 若 $X_1, \dots, X_n \in \mathbb{R}$ 为独立的随机变量, 有共同的分布函数 F_0 , 记 F_n 为其经验分布函数, F 为任意一个分布函数, 若 $F \neq F_n$, 则有 $L(F) < L(F_n)$ 。

记参数 $\theta = T(F)$, 其真值为 $\theta_0 = T(F_0)$, 这里 T 是分布的实值函数。 θ 的非参极大似然估计为 $\hat{\theta} = T(F_n)$ 。比如, 若 θ 为均值, 则 θ_0 的非参极大似然估计为 F_n 的均值, 即 $\bar{X} = (1/n) \sum_{i=1}^n X_i$ 。

2.1.2 经验似然检验

类似参数似然比方法, 我们可以对分布 F 定义如下非参似然比函数:

$$R(F) = \frac{L(F)}{L(F_n)},$$

其中 $F_n, L(F)$ 如定义 2.1-2.2 中定义, 通过简单的计算我们可以得出 $R(F) = \prod_{i=1}^n np_i$ 。则参数 θ 的似然比函数可以定义为:

$$R(\theta) = \sup\{R(F) | T(F) = \theta, F \in \mathcal{F}\},$$

$\mathcal{F}$ 为 F 所属的分布的集合。一般情况下, 我们取 $\mathcal{F}$ 为 $\mathbb{R}$ 上所有分布的集合或者某些分布的集合。经验似然假设检验在 $R(\theta) < r_0$ 时拒绝原假设 $H_0: T(F_0) = \theta_0$, r_0 为某个阈值, 相应的经验似然置信区间有如下的形式:

$$\{\theta | R(\theta) \geq r_0\}.$$

2.1.3 均值的经验似然检验

若我们对分布 F 的均值 μ 感兴趣, 则均值的经验似然比函数可以写为

$$R(\mu) = \sup \left\{ \prod_{i=1}^n (np_i) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i X_i = \mu \right\}.$$

Owen (1988) 指出, 若 μ 在 $X_1, \dots, X_n$ 的凸包内, 则上式的右侧存在唯一解。 $R(\mu)$ 的精确表达式可以由拉格朗日乘子方法得到: 当

$$p_i = p_i(\mu) = \frac{1}{n(1 + \lambda(X_i - \mu))}$$

时, $\prod_{i=1}^n (np_i)$ 在约束 $p_i \geq 0$ 、 $\sum_{i=1}^n p_i = 1$ 和 $\sum_{i=1}^n p_i X_i = \mu$ 下达到最大值。这里 $\lambda = \lambda(\mu)$ 为拉格朗日乘子, 且为下式的解:

$$\sum_{i=1}^n \frac{X_i - \mu}{1 + \lambda(X_i - \mu)} = 0.$$

因为 F_n 可以使 $\prod_{i=1}^n p_i$ 达到最大 (定理 2.1), 因此 $R(\mu)$ 在 $\hat{\mu} = \bar{X}$ 时达到最大, 且

$$R(\mu) = \prod_{i=1}^n \frac{1}{1 + \lambda(X_i - \mu)}.$$

经验似然比统计量则为 $W(\mu) = -2 \log R(\mu)$, 即

$$W(\mu) = 2 \sum_{i=1}^n \log(1 + \lambda(X_i - \mu)).$$

类似参数似然比方法中的 Wilks 定理, Owen (1988) 给出了如下的经验似然定理 (ELT)。

定理 2.2: 若 $X_1, \dots, X_n \in \mathbb{R}$ 为独立的随机变量且有共同的分布 F_0 , 令 $\mu_0 = E(X_i)$, 假设 $0 < \text{Var}(X_i) < \infty$, 当 $n \rightarrow \infty$ 时, 则有 $-2 \log(R(\mu_0))$ 依分布收敛到 χ_1^2 , χ_1^2 表示自由度为 1 的卡方分布。

根据定理 2.2, 当 $-2 \log(R(\mu_0)) > \chi_{1,\alpha}^2$ 时, 我们在显著性水平 α 上拒绝原假设, $\chi_{1,\alpha}^2$ 为自由度为 1 的卡方分布的 α 分位数。

Owen (1988) 给出了经验似然检验的覆盖精确度, 即当 $n \rightarrow \infty$ 时,

$$P(-2\log(R(\mu_0)) \leq \chi_{1,1-\alpha}^2) - (1 - \alpha) \rightarrow 0.$$

经验似然检验的 **power** 可以通过 R 在非参极大似然估计 $\bar{X}$ 上的曲率来评估。对于大样本来说, 当 μ 接近 $\bar{X}$ 时, $-2\log(R(\mu)) = \frac{n(\mu - \bar{X})^2}{2\sigma_0^2}$, 其中 $\sigma_0^2 = \text{Var}(X_i)$ 。曲率 (的绝对值) 越大, 在给定置信水平上的置信区间就越短, **power** 就越大。Owen (2001) 中给出:

$$-2\log(R(\mu_0 + \frac{\tau\sigma_0}{\sqrt{n}})) \xrightarrow{d} \chi_1^2(\tau^2),$$

这里 τ^2 为非中心参数, 这表明经验似然检验有和参数似然检验类似的 **power**。

2.1.4 基于估计方程的经验似然检验

Qin and Lawless (1994) 将经典的经验似然结果扩展到基于估计方程来构造经验似然比。给定一个随机变量 $X \in \mathbb{R}^d$ 、参数 $\theta \in \mathbb{R}^p$ 和向量值的函数 $m(X, \theta) \in \mathbb{R}^s$ 且

$$E(m(X, \theta)) = 0,$$

其中 $s \geq p$, 一般情况下, 真值 θ_0 可以通过解估计方程:

$$\frac{1}{n} \sum_{i=1}^n m(X_i, \hat{\theta}) = 0$$

来估计。类似 Owen (1988, 1990), 我们可以定义经验似然比函数为:

$$R(\mu) = \sup \left\{ \prod_{i=1}^n (np_i) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i m(X_i, \theta) = 0 \right\},$$

其相应的 ELT 有

定理 2.3: 若 $X_1, \dots, X_n \in \mathbb{R}^d$ 为独立的随机变量且有共同的分布 F_0 , 对于 $\theta \in \Theta \subseteq \mathbb{R}^p$, $X \in \mathbb{R}^d$ 和 $m(X, \theta) \in \mathbb{R}^s$, 令 $\theta_0 \in \Theta$ 使得 $\text{Var}(m(X_i, \theta_0))$ 有限且秩 q 满足 $q > 0$ 。若 θ_0 满足 $E(m(X, \theta_0)) = 0$, 则当 $n \rightarrow \infty$ 时, $-2\log(R(\theta_0)) \xrightarrow{d} \chi_q^2$, χ_q^2 表示自由度为 q 的卡方分布。

2.1.5 Profile 经验似然检验

很多时候, 我们会面对估计方程个数比参数个数更多的情形, 即我们有 $p+q$ 个估计等式和 q 个未知参数, 我们可以选择只关注 q 个估计等式而忽略其余 p 个, 此方法我们称之为 **profile** 经验似然检验。

Qin and Lawless (1994) 给出了 **profile** 经验似然检验的 ELT, 更多关于 **profile** 经验似然的内容可以参考 Owen (2001) 的第 3.10 章。

定理 2.4: 令 $X_i \in \mathbb{R}^d$ 为 i.i.d. 的随机向量, 假设 $\theta_0 \in \mathbb{R}^p$ 由 $E(m(X, \theta)) = 0$ 唯一确定, 此处 $m(X, \theta) \in \mathbb{R}^{p+q}$, $q \geq 0$ 。记 $\tilde{\theta} = \arg \max_{\theta} R(\mu)$, 其中

$$R(\mu) = \sup \left\{ \prod_{i=1}^n (np_i) | p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i m(X_i, \theta) = 0 \right\}.$$

假设存在 θ_0 的邻域 Θ 和函数 $M(x)$, 其中 $E(M(X)) < \infty$, 使得下列成立:

- (a) $E\left(\frac{\partial m(X, \theta_0)}{\partial \theta}\right)$ 秩为 p ,
- (b) $E(m(X, \theta_0)m(X, \theta'_0))$ 正定,
- (c) $\frac{\partial m(x, \theta)}{\partial \theta}$ 对 $\theta \in \Theta$ 连续,
- (d) $\frac{\partial^2 m(x, \theta)}{\partial \theta \partial \theta'}$ 对 $\theta \in \Theta$ 连续,
- (e) $\|m(x, \theta)\|^3 \leq M(x)$ 对 $\theta \in \Theta$ 成立,
- (f) $\left\| \frac{\partial m(x, \theta)}{\partial \theta} \right\| \leq M(x)$ 对 $\theta \in \Theta$ 成立,
- (g) $\left\| \frac{\partial^2 m(x, \theta)}{\partial \theta \partial \theta'} \right\| \leq M(x)$ 对 $\theta \in \Theta$ 成立。

则当 $n \rightarrow \infty$ 时,

$$-2 \log(R(\theta_0)/R(\tilde{\theta})) \xrightarrow{d} \chi_p^2,$$

且

$$-2 \log(R(\tilde{\theta})) \xrightarrow{d} \chi_q^2.$$

2.1.6 经验似然检验在时间序列上的应用

当经验似然检验应用于时间序列中时, 我们一般基于条件似然来生成估计方程, 在一些宽松的条件下, 利用对偶似然 (dual likelihood) 方法来进行统计推断。

由 Mykland (1995) 所提出的对偶似然方法将拉格朗日乘子 λ 作为参数。对每个固定参数值 θ , 对 θ 的检验相当于对 $\lambda = 0$ 的检验。当数据为 i.i.d. 时, 对偶似然和经验似然是一致的。在一些宽松的条件下, 对偶似然统计量也依分布收敛于卡方分布。更多关于对偶似然及其在时间序列中的应用可以参考 Owen (2001) 的第 8 章。

2.2 重尾理论

2.2.1 Stable 分布

定义 2.3: 分布函数 F 是 stable 分布, 如果对任意 $a_1, a_2 > 0$ 和任意 b_1, b_2 , 存在常数 $a > 0$ 和 b 使得

$$F(a_1x + b_1)F(a_2x + b_2) = F(ax + b). \quad (2.1)$$

若取 $b_1 = b_2 = 0$, 则 F 的特征函数 f 有如下的形式:

$$f(t/a_1)f(t/a_2) = f(t/a)e^{-ibt}.$$

Stable 分布在重尾理论中起着非常重要的作用。比如, 若 $X_1, \dots, X_n$ 为强混合平稳序列, B_n 和 A_n 为常数序列, 在一定的宽松条件下,

$$Z_n = \frac{X_1 + \dots + X_n}{B_n} - A_n \quad (2.2)$$

往往都依分布收敛于一个 stable 分布, 参考 Ibragimov and Linnik (1971) 的定理 18.1.1。

为了引入 stable 分布的相关性质, 我们首先给出缓变函数的定义。缓变函数是一类规则变化函数, 后者对于处理重尾和吸引域来说不可或缺。粗略地说, 规则变化函数是那些渐近表现类似幂函数的函数。我们给出如下定义:

定义 2.4: 可测函数 $U: \mathbb{R}_0 \rightarrow \mathbb{R}_0$ 在 ∞ 是以指标 ρ 规则变化的 (记为 $U \in RV_\rho$), 若对任意 $x > 0$:

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho.$$

我们称 ρ 为变化指数。

若 $\rho = 0$, 我们称 U 为缓变函数。

Ibragimov and Linnik (1971) 在第 2 章中给出了如下结果:

定理 2.5: 对于 i.i.d. 分布的随机变量序列 $X_i, i = 1, \dots, n$, 记如 (2.2) 中定义的 Z_n 分布的弱极限为分布函数 F 。 F 为 stable 分布的充分必要条件为:

除非 F 是退化分布, (2.2) 中的常数 B_n 有形式 $B_n = n^{1/\alpha}h(n)$, 其中 $0 < \alpha \leq 2$, $h(n)$ 为一个缓变函数。

下面我们给出 stable 分布的指数的定义。

定理 2.6: 分布 F 是 stable 分布的充分必要条件是它的特征函数有以下形式:

$$\log f(t) = i\gamma t - c|t|^\alpha \left(1 - i\beta \frac{t}{|t|} \omega(t, \alpha) \right),$$

其中 α, β, γ, c 为常数 ($C \geq 0, 0 < \alpha \leq 2, |\beta| \leq 1$) 并且

$$\begin{aligned} \omega(t, \alpha) &= \tan\left(\frac{1}{2}\pi\alpha\right), \alpha \neq 1, \\ &= 2\pi^{-1} \log |t|, \alpha = 1. \end{aligned}$$

我们称 α 为 stable 分布 F 的指数, 也就是我们通常所说的尾指数。Ibragimov and Linnik (1971) 给出了如下结果:

若 $X_1, \dots, X_n$ 为 i.i.d. 的随机变量, 并且使得 Z_n 收敛到指数为 α 的 stable law, 则 $B_n = n^{1/\alpha} h(n)$, 其中 $h(n)$ 为一个缓变函数。

2.2.2 吸引域

若 $X_1, \dots, X_n$ 为独立随机变量序列, 有相同的分布函数 $F(x)$ 。定义

$$S_n = X_1 + X_2 + \dots + X_n, \quad (2.3)$$

再令 $G(x)$ 为一个非退化的分布。我们称 $F(x)$ 在 $G(x)$ 的吸引域内, 当且仅当存在常数 $a_n > 0$ 和 b_n , 使得 $\frac{S_n}{a_n} - nb_n$ 的分布收敛到 $G(x)$ 。更多关于吸引域的内容可以参见 Feller (1971) 的第 XVII 章。

对于 stable law 的吸引域, 我们有以下性质:

定理 2.7: 一个分布 $F(x)$ 属于指数为 α ($0 < \alpha < 2$) 的 stable law 吸引域的充分必要条件为: 当 $|x| \rightarrow \infty$ 时,

$$\begin{aligned} F(x) &= \frac{c_1 + o(1)}{(-x)^\alpha} h(-x), \quad x < 0, \\ F(x) &= 1 - \frac{c_2 + o(1)}{x^\alpha} h(x), \quad x > 0, \end{aligned}$$

其中函数 $h(x)$ 为一个缓变函数, c_1, c_2 为常数, $c_1, c_2 \geq 0, c_1 + c_2 > 0$ 且和 stable law 等式 (2.1) 有关。

对于在指数为 α ($0 < \alpha < 2$) 的 stable law 吸引域内的分布函数, 其矩有以下性质:

定理 2.8: 若分布 $F(x)$ 属于指数为 α ($0 < \alpha < 2$) 的 stable law 吸引域, 则对任意 δ ($0 \leq \delta < \alpha$),

$$\int_{-\infty}^{\infty} |x|^\delta dF(x) < \infty,$$

即分布 F 的 δ 阶矩存在。任意 $\delta > \alpha$ 阶矩都不存在。

2.2.3 Hill 估计量

接下来,我们简单介绍一下经典的尾指数估计量: **Hill 估计量**。为了简便性,我们假设观测量 $X_1, \dots, X_n$ 非负。对于 $1 \leq i \leq n$, 我们记 $X_1, \dots, X_n$ 的第 i 个最大值为 $X_{(i)}$, 即

$$X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}.$$

Hill 估计量基于前 k 个次序统计量来估计 $1/\alpha$:

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \log \frac{X_{(i)}}{X_{(k+1)}}.$$

De Haan and Ferreira (2006) 证明了如下的渐近性质:

定理 2.9: 假设 $X_1, \dots, X_n$ 为独立随机变量序列, 有相同的分布函数 $F(x)$ 。若 $F(x)$ 属于指数为 $\gamma = 1/\alpha$ 的极值分布 $G(x)$ 的吸引域, 则当 $n \rightarrow \infty$, $k = k(n) \rightarrow \infty$ 且 $k/n \rightarrow 0$ 时,

$$H_{k,n} \xrightarrow{p} 1/\alpha.$$

注 2.1: 定理 2.9 中, 指数为 $\gamma = 1/\alpha$ 的极值分布 $G(x)$ 定义为

$$G(x) = \exp \left\{ - \left(1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right)^{-1/\gamma} \right\},$$

其中 $\mu \in \mathbb{R}$ 为位置参数, $\sigma > 0$ 为尺度参数。Stable 分布是此分布的一种特殊情况。更多关于极值分布的内容可以参考 De Haan and Ferreira (2006)。

关于相依随机变量的 Hill 估计量的一致性, 可以参考 Resnick and Starica (1995)、Rootzen et al. (1990) 和 Resnick and Starica (1997) 等。

3 ARMA-GARCH 模型误差零均值的经验似然检验

3.1 引言

本章我们考虑如下标准 ARMA-GARCH 模型：

$$\begin{cases} y_t = \varphi + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \psi_j \varepsilon_{t-j} + \varepsilon_t, \varepsilon_t = \eta_t \sqrt{h_t}, \\ h_t = \alpha_0 + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \end{cases} \quad (3.1)$$

这里 $\alpha_0 > 0$, $\alpha_i \geq 0$ ($i = 1, \dots, r$), $\beta_j \geq 0$ ($j = 1, \dots, s$)。对模型 (3.1) 来说, 经典的估计方法 QMLE 在 $E\varepsilon_t^4 < \infty$ 和 $E\eta_t^4 < \infty$ 时收敛到正态分布, 但在实际中, 此假设经常不成立, 模型 (3.1) 中误差经常存在重尾的现象。因此, 在本章节, 我们考虑 $E\eta_t^2 < \infty$ 但 $E\eta_t^4$ 可能不存在的情况。在此情况下, Zhu and Ling (2011) 提出的 SWQMELE 方法经常被用来拟合 ARMA-GARCH 模型, 通过假设 $E|\eta_t| = 1$ 和 η_t 的中位数为 0, 而不是 QMLE 中常用的 $E\eta_t^2 = 1$, 无论 $E\eta_t^4 = \infty$ 还是 $E\eta_t^4 < \infty$, 都能确保估计量收敛到正态分布。

正如第 1 章中所介绍, 当 SWQMELE 的假设不满足时, 为了使得模型 (3.1) 满足假设, 我们需要对模型做一些变换。若 η_t 的中位数不为 0, 此时涉及到的复杂的位移变换会使得变换后 ARMA-GARCH 结构不再成立。因此, 当应用 SWQMELE 来拟合重尾情况下的 ARMA-GARCH 模型时, 检验其假设是否满足具有重要意义。

此外, 在研究美国各州房价指数的波动性时, Zimmer (2012, 2015) 曾用 QMLE 对房价指数拟合 AR-GARCH 模型。然而, Huang et al. (2019) 使用 Hill 估计 (Hill (1975)) 证实了 $E\eta_t^4$ 和 $E\varepsilon_t^4$ 都可能不存在, 在接下来应用 SWQMELE 进行拟合时也发现 η_t 的均值和中位数的估计均接近于 0。此现象也促使了我们在本章提出, 在使用 SWQMELE 拟合 ARMA-GARCH 模型时利用经验似然方法来检验 $E\eta_t = 0$ 。

本章的结构如下。第 3.2 节介绍了经验似然检验及其渐近结果。第 3.3 节和第 3.4 节分别介绍了模拟研究和对美国房价指数及几个金融收益率序列的实证分析。第 3.5 节给出了我们的结论。证明和理论细节均在第 3.6 节中。

3.2 方法和理论结果

考虑 ARMA(p, q)-GARCH(r, s) 模型 (3.1), 令 $\theta = (\gamma', \delta')'$ 表示未知参数, 其真值为 θ_0 , 其中 $\gamma = (\varphi, \phi_1, \dots, \phi_p, \psi_1, \dots, \psi_q)'$, $\delta = (\alpha_0, \alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s)'$ 。令 $\Theta = \Theta_\gamma \times \Theta_\delta$ 表示参数空间, 其中 $\Theta_\gamma \subset \mathbb{R}^{p+q+1}$, $\Theta_\delta \subset \mathbb{R}_0^{r+s+1}$, $\mathbb{R} = (-\infty, \infty)$ 且 $\mathbb{R}_0 = [0, \infty)$ 。

给定观测值 $\{y_n, \dots, y_1\}$ 和初值 $\{y_0, y_{-1}, \dots\}$, 我们把模型 (3.1) 写成参数形式:

$$\begin{cases} \varepsilon_t(\gamma) = y_t - \varphi - \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \psi_j \varepsilon_{t-j}(\gamma), \eta_t(\theta) = \frac{\varepsilon_t(\gamma)}{\sqrt{h_t(\theta)}}, \\ h_t(\theta) = \alpha_0 + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2(\gamma) + \sum_{j=1}^s \beta_j h_{t-j}(\theta). \end{cases} \quad (3.2)$$

记 $\eta_t(\theta_0) = \eta_t$, $\varepsilon_t(\gamma_0) = \varepsilon_t$, $h_t(\theta_0) = h_t$, 再记 $\phi(z) = 1 - \sum_{i=1}^p \phi_i z^i$, $\psi(z) = 1 + \sum_{i=1}^q \psi_i z^i$, $\alpha(z) = \sum_{i=1}^r \alpha_i z^i$ 和 $\beta(z) = 1 - \sum_{i=1}^s \beta_i z^i$. 类似 Ling (2007) 和 Zhu and Ling (2011), 我们对 $i \leq 0$ 取 $y_i = 0$, 并且做出以下假设:

假设 3.1: 假设 Θ_γ 和 Θ_δ 是紧的, 且 θ_0 是 Θ 的一个内点. 对任意一个 $\theta \in \Theta$, 当 $|z| \leq 1$ 时, $\phi(z) \neq 0$; $\phi(z)$ 和 $\psi(z)$ 无共同的根且 $\phi_p \neq 0$ 或 $\psi_q \neq 0$.

假设 3.2: 对任意一个 $\theta \in \Theta$, $\alpha(z)$ 和 $\beta(z)$ 无共同根, $\alpha(1) \neq 1$, $\alpha_r + \beta_s \neq 0$, 且 $\sum_{i=1}^s \beta_i < 1$.

假设 3.3: $\{\eta_t\}$ 是 i.i.d. 的随机变量。

假设 3.4: 对任意 $\rho \in (0, 1)$, $E[(w_t + w_t^2)\xi_{\rho,t-1}^4] < \infty$, 其中 $\xi_{\rho,t} = 1 + \sum_{i=0}^{\infty} \rho^i |y_{t-i}|$.

假设 3.5: η_t 的中位数为 0, 且 $E|\eta_t| = 1$, $E\eta_t^2 < \infty$, 其有一个连续的密度函数 $g(x)$, 满足 $g(0) > 0$ 和 $\sup_{x \in \mathbb{R}} g(x) < \infty$.

假设 3.6: $E|\varepsilon_t|^\iota < \infty$ 对某个 $\iota > 0$ 成立。

假设 3.1 保证了模型 (3.2) 的 ARMA 部分的平稳性、可逆性和可识别性, 在此条件下有

$$\psi^{-1}(z) = \sum_{i=0}^{\infty} a_\psi(i) z^i \text{ 和 } \phi(z)\psi^{-1}(z) = \sum_{i=0}^{\infty} a_\gamma(i) z^i,$$

其中 $\sup_{\theta \in \Theta_\gamma} a_\psi(i) = O(\rho^i)$ 和 $\sup_{\theta \in \Theta_\gamma} a_\gamma(i) = O(\rho^i)$ 对某个 $0 < \rho < 1$ 成立。

假设 3.2 是模型 (3.2) 的 GARCH 部分的可识别条件。在此条件下, 我们有

$$\beta^{-1}(z) = \sum_{i=0}^{\infty} a_\beta(i) z^i \text{ 和 } \alpha(z)\beta^{-1}(z) = \sum_{i=1}^{\infty} a_\delta(i) z^i,$$

这里 $\sup_{\theta \in \Theta_\delta} a_\beta(i) = O(\rho^i)$ 和 $\sup_{\theta \in \Theta_\delta} a_\delta(i) = O(\rho^i)$ 对某个 $0 < \rho < 1$ 成立, 参考 Ling (1999) 的引理 2.1. 假设 3.6 保证了 $\{\varepsilon_t\}$ 的平稳性。

假设 3.4 中的权重 w_t 在 Zhu and Ling (2011) 中被用来减少 ε_t 的矩条件。Zhu and Ling (2011) 和 Pan et al. (2007) 给出了推荐的权重选择, 我们将其用在了下文即将介绍的模拟研究和实证分析中。

假设 3.5 保证了即使 η_t 的四阶矩不存在, SWQMELE 也能渐近收敛于正态分布。更加确切地说, 我们临时假设 η_t 服从标准指数分布, 通过最小化如下的加权负似然函数来得到 SWQMELE:

$$L_{sn}(\boldsymbol{\theta}) = \frac{1}{n} \sum_{t=1}^n w_t l_t(\boldsymbol{\theta}), \quad \text{其中 } l_t(\boldsymbol{\theta}) = \log \sqrt{h_t(\boldsymbol{\theta})} + \frac{|\varepsilon_t(\gamma)|}{\sqrt{h_t(\boldsymbol{\theta})}}.$$

正如引言中所说, 如果对模型进行复杂的位移变换来满足 SWQMELE 的假设条件, 可能会使得 ARMA-GARCH 结构不再成立, 尤其是对偏斜数据有很大的影响。因此, 我们需要在假设 3.1-3.6 下进行如下检验:

$$H_0 : E\eta_t = 0 \quad \text{v.s.} \quad H_1 : E\eta_t \neq 0. \quad (3.3)$$

当我们接受原假设时, ARMA-GARCH 的结构仍然成立。

为了得到经验似然检验统计量, 我们根据 Qin and Lawless (1994) 中的想法来得到估计方程, 即:

$$\begin{aligned} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \frac{1}{2h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \boldsymbol{\theta}} \text{sgn}(\eta_t(\boldsymbol{\theta})) - \frac{|\varepsilon_t(\gamma)|}{2\sqrt{h_t(\boldsymbol{\theta})}} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ &= \frac{1}{2h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (1 - |\eta_t(\boldsymbol{\theta})|) + \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \boldsymbol{\theta}} \text{sgn}(\eta_t(\boldsymbol{\theta})), \end{aligned}$$

这里 sgn 表示符号函数。记 $\mathbf{D}_{t,1}(\boldsymbol{\theta}) = w_t \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, $D_{t,2}(\boldsymbol{\theta}, \mu) = w_t(\eta_t(\boldsymbol{\theta}) - \mu)$, 其中 $\mu = \mu(\boldsymbol{\theta}) = E\eta_t(\boldsymbol{\theta})$ 且真值为 μ_0 , 通过解以下的估计方程可以同时估计 $\boldsymbol{\theta}$ 和 μ :

$$\sum_{t=1}^n \mathbf{D}_{t,1}(\boldsymbol{\theta}) = 0 \quad \text{和} \quad \sum_{t=1}^n D_{t,2}(\boldsymbol{\theta}, \mu) = 0.$$

相应的经验似然函数就可以定义为:

$$L(\boldsymbol{\theta}, \mu) = \sup \left\{ \prod_{t=1}^n (np_t) : p_1 \geq 0, \dots, p_n \geq 0, \sum_{t=1}^n p_t = 1, \sum_{t=1}^n p_t \mathbf{D}_t(\boldsymbol{\theta}, \mu) = 0 \right\},$$

其中 $\mathbf{D}_t(\boldsymbol{\theta}, \mu) = (\mathbf{D}'_{t,1}(\boldsymbol{\theta}), D_{t,2}(\boldsymbol{\theta}, \mu))'$ 。用拉格朗日乘子方法, 我们有对数似然函数:

$$l(\boldsymbol{\theta}, \mu) = -2 \log(L(\boldsymbol{\theta}, \mu)) = 2 \sum_{t=1}^n \log \{1 + \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, \mu)\}, \quad (3.4)$$

其中 $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\theta}, \mu)$ 满足

$$\sum_{t=1}^n \frac{\mathbf{D}_t(\boldsymbol{\theta}, \mu)}{1 + \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, \mu)} = 0.$$

因为只对 μ 感兴趣, 我们采用 profile 经验似然比 $l_p(\mu) = \min_{\boldsymbol{\theta} \in \Theta} l(\boldsymbol{\theta}, \mu)$ 。接下来的定理证明了我们所提出的经验似然检验具有 Wilks 定理。

定理 3.1: 若模型 (3.1) 满足假设 3.1-3.6, 在 $H_0 : E\eta_t = 0$ 下, 当 $n \rightarrow \infty$ 时, $l_p(0)$ 依分布收敛到自由度为 1 的卡方分布。

基于上述的定理 3.1, 当 $l_p(0) > \chi_{1,a}^2$ 时, 我们在置信水平 $1 - a$ 下拒绝 $H_0 : E\eta_t = 0$, 这里 $\chi_{1,a}^2$ 表示自由度为 1 的卡方分布的 a 分位数。

接下来的定理给出了我们提出的检验的 power。

定理 3.2: 若模型 (3.1) 满足假设 3.1-3.6, 如果 $\mu_0 = M/\sqrt{n}$ 对于某个常数 M 成立, 则当 $n \rightarrow \infty$ 时, $l_p(0)$ 依分布收敛到一个自由度为 1 的非中心卡方分布, 非中心参数为 $M^2(Ew_t)^2$ 。

3.3 模拟研究

这节我们用模拟研究来验证, 在 3.2 节中我们提出的经验似然检验在有限样本下具有良好的数值效果。我们的数据由 AR(1)-GARCH(1,1) 模型生成:

$$y_t = \varphi + \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = \eta_t \sqrt{h_t}, h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}.$$

这里, 我们取 $\theta_0 = (0.0797, -0.0465, 0.0347, 0.1572, 0.8057)$, $n = 200, 500, 1000, 2500$ 和 5000 , 权重 w_t 选为

$$w_t = \left(\max \left\{ 1, \frac{1}{C} \sum_{k=1}^{\infty} \frac{1}{k^9} |y_{t-k}| I\{|y_{t-k}| > C\} \right\} \right)^{-4}, \quad (3.5)$$

其中 C 是 $\{|y_t|\}_{t=1}^n$ 的 95% 或 90% 分位数, 这也和 Ling (2007) 和 Pan et al. (2007) 推荐的当 $E|\varepsilon_t| < \infty$ 时选取的权重一样。为了在显著性水平 10% 和 5% 上计算我们所提出的经验似然检验的 size, 我们考虑 $\eta_t = \tilde{\eta}_t / E|\tilde{\eta}_t|$, 其中 $\tilde{\eta}_t \sim \text{Laplace}(0, 1)$ 或 $\tilde{\eta}_t \sim N(0, 1)$ 。在这两种情况下 η_t 是对称的, 所以具有零均值和零中位数。表 3.1 给出了基于 10000 次重复所计算出的 size。我们观察到:

- 经验似然检验的 size 对两种 C 的选择都很稳定,
- 在正态分布情况下 size 的精确度不如在 Laplace 分布情况下 size 的精确度,
- 样本量越大, size 的精确度越高。

表 3.1 $H_0 : E\eta_t = 0$ 下 profile 经验似然检验的 size, 这里 (3.5) 里的 C 选为 $\{|y_t|\}_{t=1}^n$ 的 95% 分位数 (左边) 和 90% 分位数 (右边)

n	5%	10%	5%	10%	5%	10%	5%	10%
	Laplace	Laplace	Normal	Normal	Laplace	Laplace	Normal	Normal
200	0.0644	0.1181	0.0419	0.0797	0.0644	0.1160	0.0418	0.0801
500	0.0418	0.0843	0.0320	0.0654	0.0431	0.0849	0.0306	0.0683
1000	0.0455	0.0900	0.0328	0.0766	0.0458	0.0915	0.0333	0.0750
2500	0.0425	0.0879	0.0387	0.0803	0.0419	0.0856	0.0392	0.0811
5000	0.0456	0.0946	0.0436	0.0939	0.0455	0.0941	0.0418	0.0914

接下来,我们验证了所提出的经验似然检验在显著性水平 10% 和 5% 上的 power。我们考虑了标准的 Beta (1.2, 1) 分布和 Beta (2, 2.7) 分布, 即 $E|\eta_t| = 1$ 且中位数为 0。通过从上述两个 Beta 分布里重复生成 100,000 次随机样本, 我们发现 Beta (1.2, 1) 分布的均值在 -0.0659 周围波动, Beta (2, 2.7) 分布的均值在 0.0660 周围波动, 即原假设 H_0 不成立。表 3.2 给出了我们提出的检验在这两种 Beta 分布下的 power, 从中可以看出:

- 我们的经验似然检验具有很大的 power,
- power 随着样本量的增大而提高,
- power 对于两种 C 的选择很稳定,
- Beta (2, 2.7) 分布的 power 大于 Beta (1.2, 1), 因为前者偏离原假设更远。

表 3.2 对 $H_0 : E\eta_t = 0$ 的 profile 经验似然检验的 power, 这里 (3.5) 里的 C 选为 $\{|y_t|\}_{t=1}^n$ 的 95% 分位数 (左边) 和 90% 分位数 (右边)

n	5%	10%	5%	10%	5%	10%	5%	L10%
	Beta(1.2,1)	Beta(1.2,1)	Beta(2,2.7)	Beta(2,2.7)	Beta(1.2,1)	Beta(1.2,1)	Beta(2,2.7)	Beta(2,2.7)
200	0.1249	0.1914	0.0977	0.1657	0.1192	0.1851	0.0986	0.1660
500	0.1746	0.2712	0.1783	0.2846	0.1687	0.2674	0.1755	0.2734
1000	0.3364	0.4705	0.3830	0.5090	0.3247	0.4584	0.3678	0.5030
2500	0.7408	0.8347	0.8083	0.8857	0.7274	0.8243	0.7974	0.8776

综上所述, 我们提出的经验似然检验有精确的 size 和良好的 power, 并且对 C 的选择表现稳定。

3.4 实证分析

3.4.1 美国房价指数

本节将提出的经验似然检验应用于 1975 年至 2018 年美国加利福尼亚州 (CA)、佛罗里达州 (FL)、内华达州 (NV) 和亚利桑那州 (ZA) 的州级房价指数 (HPI) 的

季度百分比变化序列，这些指数是由美国联邦住房金融局估计和公布的，每个州的观测值总数均为 176。为了研究住房市场的联动性和相关性，用 AR-GARCH 模型来拟合是个不错的选择。最近，Huang et al. (2019) 对这些序列应用了 SWQMELE，发现 AR(3)-GARCH(1,1) 模型可以很好地拟合这四个序列。他们用了经典的 Hill 尾指数估计方法 (Hill (1975)) 估计了 η_t 和 ε_t 的尾指数后发现， $E\eta_t^4 < \infty$ 和 $E\varepsilon_t^4 < \infty$ 的假设均不成立，参考表 3.3 中的估计结果。虽然 η_t 的均值和中位数的估计均接近于 0 (见表 3.4)，但 Huang et al. (2019) 并没有正式地检验 $H_0 : E\eta_t = 0$ 是否成立。在这里，我们应用了本章中的 profile 经验似然检验，对加利福尼亚州、佛罗里达州、内华达州和亚利桑那州分别给出了 0.267、0.480、0.637 和 0.830 的 p 值，见表 3.4。因此，我们不能拒绝原假设 $H_0 : E\eta_t = 0$ ，即拟合的 AR(3)-GARCH(1,1) 模型仍然成立，且 Huang et al. (2019) 中用 copula 和 SWQMELE 来研究房价指数的联动性在方法上没有问题。

表 3.3 对美国房价指数 (HPI) 拟合 AR(3)-GARCH(1,1) 模型的 SWQMELE

State	$\hat{\varphi}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$
CA	0.3067	0.7274	-0.0069	0.1220	0.3327	0.3305	0.1252
FL	0.2662	0.5008	0.0344	0.3174	0.0666	0.1380	0.6706
NV	0.2703	0.4475	0.0094	0.3231	0.0821	0.2810	0.5647
AZ	0.4155	0.3253	0.1169	0.2565	0.0333	0.3037	0.5810

表 3.4 美国房价指数 (HPI) 的 profile 经验似然检验

State	$\hat{\eta}_t$ 均值	$\hat{\eta}_t$ 中位数	$ \hat{\eta}_t $ 均值	$\hat{l}_p(\mu_0)$	p 值
CA	-0.0944	-0.0068	1.0161	1.2332	0.267
FL	-0.0661	-0.0217	0.9888	0.4995	0.480
NV	0.0245	-0.0230	0.9951	0.2231	0.637
AZ	0.0028	-0.0454	0.9863	0.0459	0.830

3.4.2 金融收益率序列

这节我们研究了 S&P500 和 Microsoft Stock (MSFT) 从 2009 年 8 月 3 日到 2019 年 7 月 29 日的每日收盘价，每个序列均有 2514 个观测值。图 3.1 和 3.2 分别给出了收盘价和对数收益率的曲线。

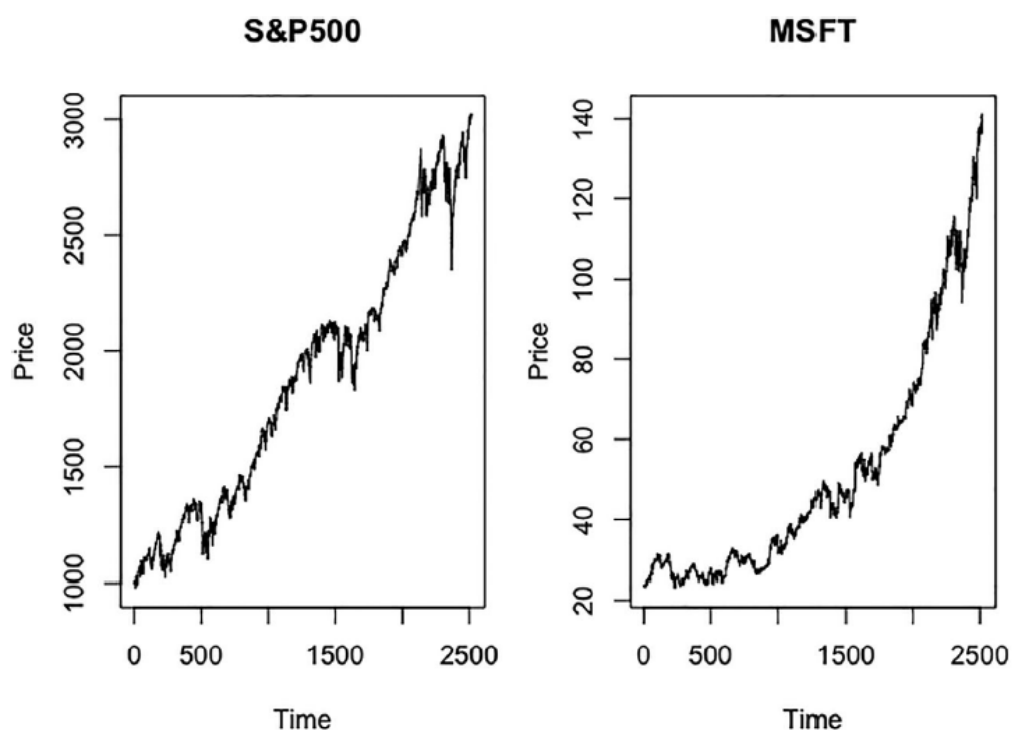


图 3.1 2009 年 8 月 3 日到 2019 年 7 月 29 日 S&P50 和 Microsoft Stock (MSFT) 每日收盘价的时间序列图

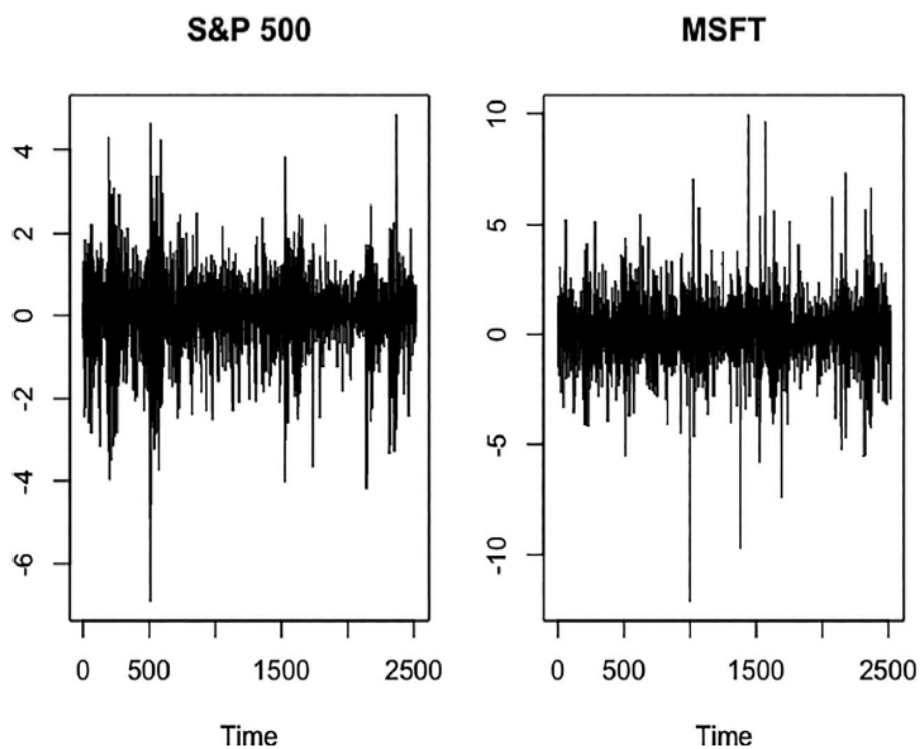


图 3.2 2009 年 8 月 3 日到 2019 年 7 月 29 日 S&P50 和 Microsoft Stock (MSFT) 每日收盘价的对数收益率 ($\times 100$)

我们对 $100 \times (\log p_t - \log p_{t-1})$ (p_t 为收盘价) 的序列拟合 ARMA-GARCH 模型。我们先用 R 软件包 “forecast” 里的函数 “auto.arima” 和赤池信息量准则 (AIC) 来选择 ARMA 部分的阶数。最终, 我们用 ARMA(2,2)-GARCH(1,1) 模型来拟合 S&P500 的对数收益率, 用 ARMA(1,1)-GARCH(1,1) 来拟合 MSFT 的对数收益率。其次, 我们用 R 软件包里的 “fGarch” 函数和 QMLE 方法来拟合模型: 对 S&P500 拟合的 ARMA(2,2)-GARCH(1,1) 模型为

$$\begin{cases} y_t = 0.0054 - 0.0287y_{t-1} + 0.9469y_{t-2} + 0.0052\varepsilon_{t-1} - 0.9784\varepsilon_{t-2} + \varepsilon_t, \\ \quad (0.0015) \quad (0.0118) \quad (0.0115) \quad (0.0071) \quad (0.0072) \\ \varepsilon_t = \eta_t \sqrt{h_t}, \quad h_t = 0.0345 + 0.1588\varepsilon_{t-1}^2 + 0.8042h_{t-1}, \\ \quad (0.0084) \quad (0.0244) \quad (0.0247) \end{cases}$$

对 MSFT 拟合的 ARMA(1,1)-GARCH(1,1) 模型为

$$\begin{cases} y_t = 0.0217 + 0.7716y_{t-1} - 0.8083\varepsilon_{t-1} + \varepsilon_t, \\ \quad (0.0105) \quad (0.0974) \quad (0.0952) \\ \varepsilon_t = \eta_t \sqrt{h_t}, \quad h_t = 0.3391 + 0.1206\varepsilon_{t-1}^2 + 0.7216h_{t-1}. \\ \quad (0.1353) \quad (0.0372) \quad (0.0846) \end{cases}$$

这里, 括号里的数字表示标准误差。

图 3.3 给出了拟合的 η_t 和 $|\eta_t|$ 的自相关函数, 证实了上述拟合的模型对两个对数收益率都是合理的。我们用 R 软件包 “evmix” 画出了拟合的 η_t 的 Hill 估计和其 90% 置信区间, 见图 3.4, 这表明 S&P500 的 QMLE 可能有正态极限, 但 MSFT 不一定有。接下来, 我们检验了 SWQMELE 是否适用。

应用 SWQMELE 后我们发现, 对 S&P500 和 MSFT 来说, η_t 的均值分别是 -0.0201 和 0.0029, η_t 的中位数分别是 0.0335 和 -0.0083, $|\eta_t|$ 的均值分别是 1.0048 和 1.0014。我们所提出的 profile 经验似然检验对 S&P500 和 MSFT 分别给出了 $\hat{l}_p(0) = 16.1283$ 和 0.6071, 对应的 p 值分别为 5.919×10^{-5} 和 0.436。因此, 尽管估计的 η_t 的均值和中位数都很接近 0, 原假设 $H_0: E\eta_t = 0$ 对于 S&P 500 来说是不成立的, 但对于 MSFT 来说是成立的。以上表明, MSFT 的对数收益率具有重尾的特性, 且均值和中位数均为 0, 因此应用 SWQMELE 是合理的。反之, 我们的检验表明 S&P500 的对数收益率不满足零均值和零中位数的假设, 如果对其应用 SWQMELE, 得出的模型将不再具有 ARMA-GARCH 结构。此外, 由 Hill 估计量可以看出此序列不存在重尾效应, 所以用 QMLE 来进行拟合可能是个更好的选择。

综上所述, 在实际中应用 SWQMELE 时, 检验 ARMA-GARCH 模型的误差是否有零均值和零中位数很有必要。

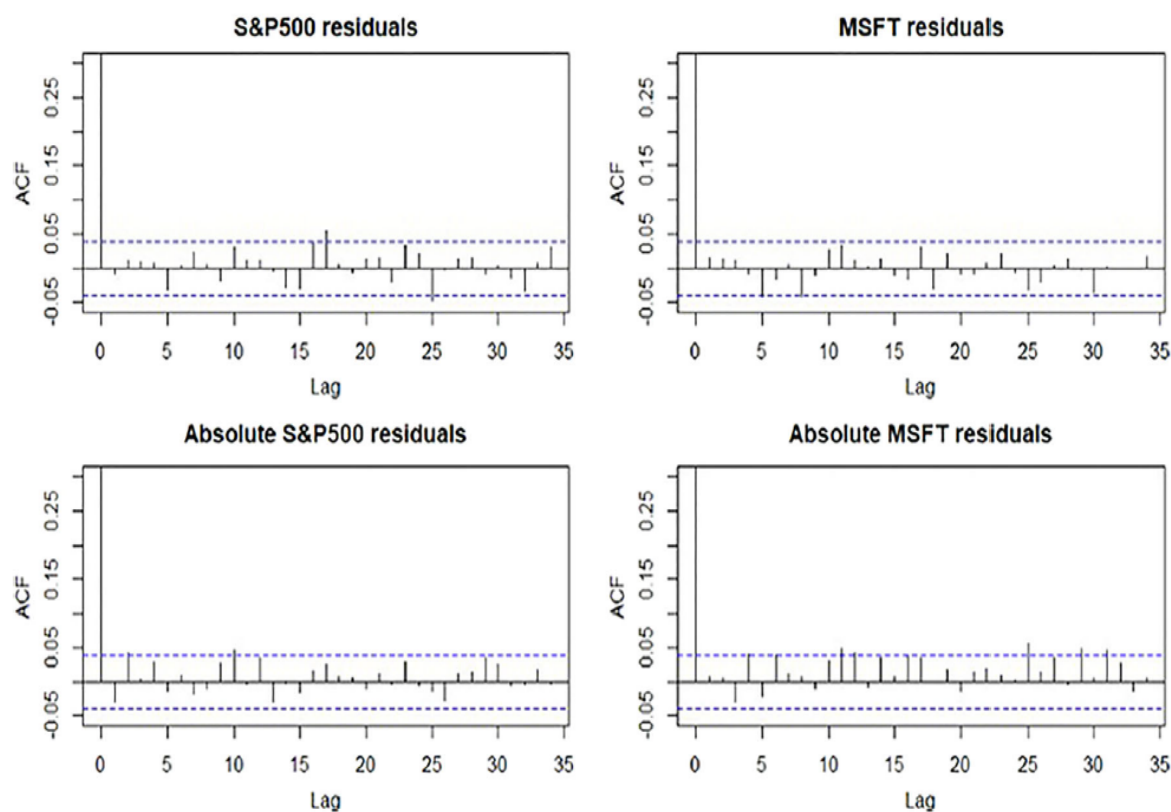


图 3.3 对 S&P 500 和 MSFT 拟合的 η_t 和 $|\eta_t|$ 的自协方差函数图

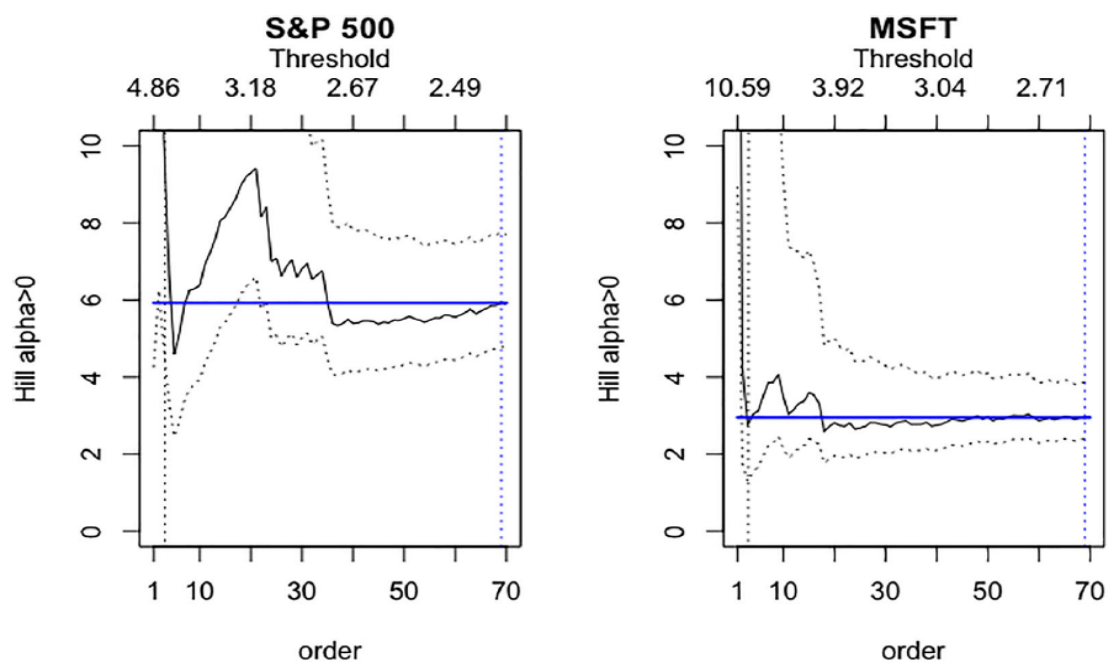


图 3.4 对 S&P 500 和 MSFT 拟合的 η_t 的 Hill 估计量及其 90% 置信区间 (黑色虚线)

3.5 结论

在对 ARMA-GARCH 模型进行统计推断时, 一般经常使用的 QMLE 方法, 需要误差和序列本身的四阶矩存在来确保估计量收敛到正态极限。SWQMELE 显著地减少了矩条件, 但要求误差具有零中位数而不是原始 ARMA-GARCH 模型中要求的零均值, 而为了确保零中位数成立而进行的变换可能会破坏 ARMA-GARCH 结构。本章提出了一种有效的 profile 经验似然方法来检验使用 SWQMELE 时, ARMA-GARCH 模型中的误差是否有零均值。实证研究表明, SWQMELE 不适用于 S&P500 对数收益率, 但对美国住房价格指数和 MFST 对数收益率来说仍然是一个合适的估计方法。

3.6 定理 3.1 和 3.2 的证明

首先, 我们给出一些引理, 引理 3.2-3.3 的证明将在本文末的附录中给出。我们用 μ_0 表示 $\mu = E\eta_t$ 的真值, 其在原假设 $H_0: E\eta_t = 0$ 下为 0。

引理 3.1: 在定理 3.1 的条件下, 存在常数 $\rho \in (0, 1)$, $C > 0$ 和 θ_0 的邻域 Θ_0 , 使得对任意常数 $\iota_1 \in (0, 1)$, 我们有:

- (a) $\sup_{\Theta_0} |\varepsilon_t(\gamma)| \leq C\xi_{\rho t-1},$
- (b) $\sup_{\Theta_0} \left\| \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma} \right\| \leq C\xi_{\rho t-1},$
- (c) $\sup_{\Theta_0} \left\| \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma \partial \gamma'} \right\| \leq C\xi_{\rho t-1},$
- (d) $\sup_{\Theta_0} \left| \frac{\partial^3 \varepsilon_t(\gamma)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} \right| \leq C\xi_{\rho t-1},$ 其中 $i, j, k = 1, \dots, p+q+1,$
- (e) $1 \leq \sup_{\Theta_0} \frac{h_t(\theta)}{\alpha_0} \leq C\xi_{\rho t-1}^2,$ 其中 $\underline{\alpha}_0 = \inf_{\Theta_0} \alpha_0,$
- (f) $\sup_{\Theta_0} \left\| \frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \delta} \right\| \leq C\xi_{\rho t-1}^{\iota_1},$
- (g) $\sup_{\Theta_0} \left\| \frac{1}{h_t(\theta)} \frac{\partial^2 h_t(\theta)}{\partial \delta \partial \delta'} \right\| \leq C\xi_{\rho t-1}^{\iota_1},$
- (h) $\sup_{\Theta_0} \left\| \frac{1}{h_t(\theta)} \frac{\partial^3 h_t(\theta)}{\partial \delta_i \partial \delta_j \partial \delta_k} \right\| \leq C\xi_{\rho t-1}^{\iota_1},$ 其中 $i, j, k = 1, \dots, r+s+1,$
- (i) $\sup_{\Theta_0} \left\| \frac{1}{\sqrt{h_t(\theta)}} \frac{\partial h_t(\theta)}{\partial \gamma} \right\| \leq C\xi_{\rho t-1},$
- (j) $\sup_{\Theta_0} \left\| \frac{1}{\sqrt{h_t(\theta)}} \frac{\partial^2 h_t(\theta)}{\partial \gamma \partial \gamma'} \right\| \leq C\xi_{\rho t-1},$
- (k) $\sup_{\Theta_0} \left| \frac{1}{\sqrt{h_t(\theta)}} \frac{\partial^3 h_t(\theta)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} \right| \leq C\xi_{\rho t-1},$ 其中 $i, j, k = 1, \dots, p+q+1,$

$$\begin{aligned}
(l) \quad & \sup_{\Theta_0} \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\gamma}'} \right\| \leq C \xi_{\rho t-1}, \\
(m) \quad & \sup_{\Theta_0} \left| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^3 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j \partial \gamma_k} \right| \leq C \xi_{\rho t-1}, \text{ 其中 } i, j = 1, \dots, p+q+1, k = 1, \dots, r+s+1, \\
(n) \quad & \sup_{\Theta_0} \left| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^3 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_j \partial \gamma_k} \right| \leq C \xi_{\rho t-1}, \text{ 其中 } i = 1, \dots, p+q+1, j, k = 1, \dots, r+s+1.
\end{aligned}$$

证明：参考 Ling (2007) 或 Chan and Ling (2006)。

□

引理 3.2：在引理 3.1 的条件下，我们有

$$\begin{aligned}
(a) \quad & E \sup_{\Theta_0} \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|^2 < \infty, \\
(b) \quad & E \sup_{\Theta_0} \|\mathbf{P}_t(\boldsymbol{\theta}, \mu_0)\| < \infty, \text{ 其中 } \mathbf{P}_t = \frac{\partial \mathbf{D}_t(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\omega}'} \text{ 和 } \boldsymbol{\omega} = (\boldsymbol{\theta}', \mu)'.
\end{aligned}$$

证明：由于该引理证明较长，其证明详见本文末的附录。

□

引理 3.3：令 $V_0 = \{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \frac{M}{\sqrt{n}}\}$ ，其中常数 $M > 0$ 。在引理 3.1 的条件下，我们有

$$\begin{aligned}
(a) \quad & \max_{1 \leq t \leq n} \sup_{V_0} \|\mathbf{P}_t(\boldsymbol{\theta}, \mu_0)\| = o_p(n), \\
(b) \quad & \max_{1 \leq t \leq n} \sup_{V_0} \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\| = o_p(\sqrt{n}), \\
(c) \quad & \frac{1}{n} \sum_{t=1}^n \mathbf{P}_t(\boldsymbol{\theta}, \mu_0) = \boldsymbol{\Psi} + o_p(1) \text{ 对所有的 } \boldsymbol{\theta} \in V_0 \text{ 一致成立}, \\
(d) \quad & \frac{1}{n} \sum_{t=1}^n [\mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \mathbf{D}_t'(\boldsymbol{\theta}, \mu_0)] = \boldsymbol{\Omega} + o_p(1) \text{ 对所有的 } \boldsymbol{\theta} \in V_0 \text{ 一致成立},
\end{aligned}$$

$$\begin{aligned}
\text{其中 } \boldsymbol{\Psi} &= \begin{pmatrix} \frac{1}{4} E\left\{ \frac{w_t}{h_t^2(\boldsymbol{\theta}_0)} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right\} & \mathbf{0} \\ E\left\{ \frac{w_t}{\sqrt{h_t(\boldsymbol{\theta}_0)}} \frac{\partial \varepsilon_t(\gamma_0)}{\partial \boldsymbol{\theta}'} \right\} & -E w_t \end{pmatrix}, \quad \boldsymbol{\Omega} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}, \\
\Omega_{11} &= E\left\{ \frac{w_t^2}{h_t(\boldsymbol{\theta}_0)} \frac{\partial \varepsilon_t(\gamma_0)}{\partial \boldsymbol{\theta}} \frac{\partial \varepsilon_t(\gamma_0)}{\partial \boldsymbol{\theta}'} \right\} + \frac{E \eta_t^2 - 1}{4} E\left\{ \frac{w_t^2}{h_t^2(\boldsymbol{\theta}_0)} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right\}, \\
\Omega_{12} &= -\frac{E(\eta_t^2 \text{sgn}(\eta_t))}{2} E\left\{ \frac{w_t^2}{h_t(\boldsymbol{\theta}_0)} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \right\}, \\
\Omega_{21} &= -\frac{E(\eta_t^2 \text{sgn}(\eta_t))}{2} E\left\{ \frac{w_t^2}{h_t(\boldsymbol{\theta}_0)} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right\}, \quad \Omega_{22} = E(w_t^2 \eta_t^2).
\end{aligned}$$

证明：由于该引理证明较长，其证明详见本文末的附录。

□

引理 3.4: 在引理 3.1 的条件下, 我们有

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}_0, \mu_0) \xrightarrow{d} N(0, \boldsymbol{\Omega}),$$

其中 $\boldsymbol{\Omega}$ 如引理 3.3 中定义。

证明: 因为 $\mathcal{F}_t$ 是由 $\{y_m, m \leq t\}$ 生成的 σ 域, 我们有

$$\begin{aligned} E(D_{t,1}(\boldsymbol{\theta}_0)|\mathcal{F}_{t-1}) &= E\left(\frac{w_t}{2h_t(\boldsymbol{\theta}_0)} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} (1 - |\eta_t(\boldsymbol{\theta}_0)|) \right. \\ &\quad \left. + \frac{w_t}{\sqrt{h_t(\boldsymbol{\theta}_0)}} \frac{\partial \varepsilon_t(\gamma_0)}{\partial \boldsymbol{\theta}} \text{sgn}(\eta_t(\boldsymbol{\theta}_0)) \middle| \mathcal{F}_{t-1}\right) \\ &= 0, \quad \text{且} \end{aligned}$$

$$E(D_{t,2}(\boldsymbol{\theta}_0, \mu_0)|\mathcal{F}_{t-1}) = E(w_t \eta_t | \mathcal{F}_{t-1}) = w_t E \eta_t = 0.$$

因此, $D_t(\boldsymbol{\theta}_0, \mu_0)$ 是一个鞅差序列, 引理 3.3 的结论由鞅差序列的中心极限定理可得, 参考 Hall and Heyde (1980) 的第 58 页。□

定理 3.1 的证明: 令 $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \frac{\boldsymbol{\nu}}{\sqrt{n}}$, 其中 $\boldsymbol{\nu}$ 是一个 $(p + q + r + s + 2)$ 维向量。

记 $\mathbf{g}(\boldsymbol{\theta}, \mu_0, \boldsymbol{\lambda}) = \frac{1}{n} \sum_{t=1}^n \frac{D_t(\boldsymbol{\theta}, \mu_0)}{1 + \boldsymbol{\lambda}' D_t(\boldsymbol{\theta}, \mu_0)}$ 和 $\gamma_t(\boldsymbol{\theta}, \mu_0) = \boldsymbol{\lambda}' D_t(\boldsymbol{\theta}, \mu_0)$, 其中 $\boldsymbol{\lambda}$ 是等式 $\mathbf{g}(\boldsymbol{\theta}, \mu_0, \boldsymbol{\lambda}) = \mathbf{0}$ 的一个解。我们首先证明 $\max_{1 \leq t \leq n} |\gamma_t(\boldsymbol{\theta}, \mu_0)| = o_p(1)$ 。

令 $\boldsymbol{\lambda} = \rho \mathbf{r}$, 其中 $\|\mathbf{r}\| = 1$ 。注意到

$$\begin{aligned} 0 = \|\mathbf{g}(\boldsymbol{\theta}, \mu_0, \rho \mathbf{r})\| &\geq \|\mathbf{r}' \mathbf{g}(\boldsymbol{\theta}, \mu_0, \rho \mathbf{r})\| \\ &= \left| \frac{1}{n} \sum_{t=1}^n \mathbf{r}' D_t(\boldsymbol{\theta}, \mu_0) - \frac{1}{n} \rho \sum_{t=1}^n \frac{\mathbf{r}' D_t(\boldsymbol{\theta}, \mu_0) D_t'(\boldsymbol{\theta}, \mu_0) \mathbf{r}}{1 + \rho \mathbf{r}' D_t(\boldsymbol{\theta}, \mu_0)} \right| \\ &\geq -\frac{1}{n} \left| \sum_{t=1}^n \mathbf{r}' D_t(\boldsymbol{\theta}, \mu_0) \right| + \frac{\rho \mathbf{r}' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \mathbf{r}}{1 + \rho Z_n(\boldsymbol{\theta}, \mu_0)}, \end{aligned}$$

其中, $\mathbf{S}_n(\boldsymbol{\theta}, \mu_0) = \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \mathbf{D}_t'(\boldsymbol{\theta}, \mu_0)$, $Z_n(\boldsymbol{\theta}, \mu_0) = \max_{1 \leq t \leq n} \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|$ 。因此,

$$\begin{aligned} \frac{\rho \mathbf{r}' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \mathbf{r}}{1 + \rho Z_n(\boldsymbol{\theta}, \mu_0)} &\leq \frac{1}{n} \left| \sum_{t=1}^n \mathbf{r}' \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right| \\ &\leq \frac{1}{n} \left| \sum_{t=1}^n \mathbf{r}' \mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) \right| + \frac{1}{n} \left| \sum_{t=1}^n \mathbf{r}' \mathbf{P}_t(\boldsymbol{\theta}^*, \mu_0) \right| \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \\ &\leq \frac{1}{n} \left| \sum_{t=1}^n \mathbf{r}' \mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) \right| + \left(\frac{1}{n} \sum_{t=1}^n \|\mathbf{P}_t(\boldsymbol{\theta}^*, \mu_0)\| \right) \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \\ &= O_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

对任意 $\boldsymbol{\theta} \in V_0$ 一致成立, 其中最后一个等式由引理 3.4 得出的 $\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) = O_p(1)$ 和附录中引理 3.2 的证明可得。因此,

$$\frac{\rho \mathbf{r}' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \mathbf{r}}{1 + \rho Z_n(\boldsymbol{\theta}, \mu_0)} = O_p\left(\frac{1}{\sqrt{n}}\right).$$

由引理 3 (d) 可得, $\mathbf{r}' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \mathbf{r} \geq a + o_p(1)$ 对 $\boldsymbol{\theta} \in V_0$ 一致成立, 其中 a 是 $\boldsymbol{\Omega}$ 最小的特征值。因此,

$$\rho = \|\boldsymbol{\lambda}\| = O_p\left(\frac{1}{\sqrt{n}}\right), \quad (3.6)$$

对 $\boldsymbol{\theta} \in V_0$ 一致成立。于是, 由引理 3 (b), 我们有

$$\max_{1 \leq t \leq n} |\gamma_t(\boldsymbol{\theta}, \mu_0)| \leq \|\boldsymbol{\lambda}\| \max_{1 \leq t \leq n} \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\| = o_p(1), \quad (3.7)$$

对 $\boldsymbol{\theta} \in V_0$ 一致成立。

此外, 泰勒展开和引理 3 (c) 表明:

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) &= \frac{1}{n} \sum_{t=1}^n (\mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) + \mathbf{P}_t(\boldsymbol{\theta}^*, \mu_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + o_p\left(\frac{1}{\sqrt{n}}\right)) \\ &= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) + \frac{\boldsymbol{\nu}}{n\sqrt{n}} \sum_{t=1}^n \mathbf{P}_t(\boldsymbol{\theta}^*, \mu_0) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= O_p\left(\frac{1}{\sqrt{n}}\right) + \frac{\boldsymbol{\nu}}{\sqrt{n}} (\boldsymbol{\Psi} + o_p(1)) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= O_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= O_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

在 V_0 内一致成立, 其中 $\boldsymbol{\theta}^*$ 位于 $\boldsymbol{\theta}_0$ 和 $\boldsymbol{\theta}$ 中间。由 (3.7), 我们有

$$\begin{aligned}
\mathbf{0} = \mathbf{g}(\boldsymbol{\theta}, \mu_0, \boldsymbol{\lambda}) &= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \boldsymbol{\lambda} + \frac{1}{n} \sum_{t=1}^n \frac{\mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \gamma_t^2(\boldsymbol{\theta}, \mu_0)}{1 + \gamma_t(\boldsymbol{\theta}, \mu_0)} \\
&\leq \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \boldsymbol{\lambda} \\
&\quad + \frac{1}{(1 - \max_{1 \leq t \leq n} |\gamma_t(\boldsymbol{\theta}, \mu_0)|)} \frac{1}{n} \sum_{t=1}^n \left(\mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \gamma_t^2(\boldsymbol{\theta}, \mu_0) \right) \\
&\leq \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \boldsymbol{\lambda} + \frac{\max_{1 \leq t \leq n} |\gamma_t(\boldsymbol{\theta}, \mu_0)|^2}{(1 - \max_{1 \leq t \leq n} |\gamma_t(\boldsymbol{\theta}, \mu_0)|)} \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) \\
&= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \boldsymbol{\lambda} + o_p(1) O_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \boldsymbol{\lambda} + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

由引理 3.3 (d), $\mathbf{S}_n^{-1}(\boldsymbol{\theta}, \mu_0) \geq C$ 依概率成立, 因此,

$$\boldsymbol{\lambda} = \mathbf{S}_n^{-1}(\boldsymbol{\theta}, \mu_0) \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) + \mathbf{L}_n, \text{ 其中 } \|\mathbf{L}_n\| = o_p\left(\frac{1}{\sqrt{n}}\right).$$

由泰勒展开和 $o_p(|\gamma_t(\boldsymbol{\theta}, \mu_0)|^2) = o_p(1)$, 我们有

$$\ln(1 + \gamma_t(\boldsymbol{\theta}, \mu_0)) = \gamma_t(\boldsymbol{\theta}, \mu_0) - \frac{\gamma_t^2(\boldsymbol{\theta}, \mu_0)}{2} + R_t(\boldsymbol{\theta}, \mu_0),$$

其中, $R_t(\boldsymbol{\theta}, \mu_0)$ 为冗余项。因此, 我们可以得到:

$$\begin{aligned}
l(\boldsymbol{\theta}, \mu_0) &= -2 \ln L(\boldsymbol{\theta}, \mu_0) = 2 \sum_{t=1}^n \ln\{1 + \gamma_t(\boldsymbol{\theta}, \mu_0)\} \\
&= 2 \boldsymbol{\lambda}' \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) - n \boldsymbol{\lambda}' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \boldsymbol{\lambda} + 2 \sum_{t=1}^n R_t(\boldsymbol{\theta}, \mu_0) \\
&= 2n \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right)' \mathbf{S}_n^{-1}(\boldsymbol{\theta}, \mu_0) \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) + 2 \mathbf{L}_n' \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) \\
&\quad - n \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right)' \mathbf{S}_n^{-1}(\boldsymbol{\theta}, \mu_0) \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) - \mathbf{L}_n' \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) \\
&\quad - n \mathbf{L}_n' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \mathbf{L}_n - \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right)' \mathbf{L}_n + 2 \sum_{t=1}^n R_t(\boldsymbol{\theta}, \mu_0) \\
&= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right)' \mathbf{S}_n^{-1}(\boldsymbol{\theta}, \mu_0) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) - n \mathbf{L}_n' \mathbf{S}_n(\boldsymbol{\theta}, \mu_0) \mathbf{L}_n \\
&\quad + 2 \sum_{t=1}^n R_t(\boldsymbol{\theta}, \mu_0).
\end{aligned}$$

首先, 由 $\|\mathbf{L}_n\| = o_p(\frac{1}{\sqrt{n}})$ 可推出 $n\mathbf{L}'_n\mathbf{S}_n(\boldsymbol{\theta}, \mu_0)\mathbf{L}_n = o_p(1)$ 。其次, 由引理 3.3 (d) 可得:

$$\begin{aligned}
& \sum_{t=1}^n \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|^3 \\
& \leq \max_{1 \leq t \leq n} \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\| \left(\sum_{t=1}^n \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|^2 \right) = o_p(\sqrt{n}) \left(\sum_{t=1}^n \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|^2 \right) \\
& = o_p(\sqrt{n}) \cdot n \cdot \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}'_t(\boldsymbol{\theta}, \mu_0) \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) = o_p(\sqrt{n}) \cdot n \cdot \text{tr} \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}'_t(\boldsymbol{\theta}, \mu_0) \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) \\
& = o_p(\sqrt{n}) \cdot n \cdot \text{tr} \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \mathbf{D}'_t(\boldsymbol{\theta}, \mu_0) \right) = o_p(\sqrt{n}) \cdot n \cdot \text{tr}(\boldsymbol{\Omega} + o_p(1)) \\
& = o_p(\sqrt{n}) \cdot n \cdot O_p(1) = o_p(n\sqrt{n}),
\end{aligned}$$

这里 tr 表示矩阵的迹。因此, 我们有 $\|\boldsymbol{\lambda}\|^3 = O_p(n^{-\frac{3}{2}})$, 且

$$\sum_{t=1}^n |\gamma_t(\boldsymbol{\theta}, \mu_0)|^3 \leq \sum_{t=1}^n \|\boldsymbol{\lambda}\|^3 \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|^3 = O_p(n^{-\frac{3}{2}}) o_p(n\sqrt{n}) = o_p(1)。$$

所以, $R_t(\boldsymbol{\theta}, \mu_0)$ 满足

$$\left| 2 \sum_{t=1}^n R_t(\boldsymbol{\theta}, \mu_0) \right| \leq 2B \sum_{t=1}^n |\gamma_t(\boldsymbol{\theta}, \mu_0)|^3 = o_p(1)。 \quad (3.8)$$

结合引理 3.3 (d), 我们有

$$\begin{aligned}
l(\boldsymbol{\theta}, \mu_0) &= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right)' \mathbf{S}_n^{-1}(\boldsymbol{\theta}, \mu_0) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) + o_p(1) \\
&= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right)' \boldsymbol{\Omega}^{-1} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0) \right) + o_p(1),
\end{aligned}$$

对 $\boldsymbol{\theta} \in V_0$ 一致成立。因此,

$$l(\boldsymbol{\theta}_0, \mu_0) = \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) \right)' \boldsymbol{\Omega}^{-1} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}_0, \mu_0) \right) + o_p(1)。$$

$$\text{定义 } \boldsymbol{\Delta}_n(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, \mu_0), \text{ 则}$$

$$\begin{aligned}
l(\boldsymbol{\theta}, \mu_0) - l(\boldsymbol{\theta}_0, \mu_0) &= (\boldsymbol{\Delta}_n(\boldsymbol{\theta}) - \boldsymbol{\Delta}_n(\boldsymbol{\theta}_0))' \boldsymbol{\Omega}^{-1} \boldsymbol{\Delta}_n(\boldsymbol{\theta}_0) + \boldsymbol{\Delta}'_n(\boldsymbol{\theta}_0) \boldsymbol{\Omega}^{-1} (\boldsymbol{\Delta}_n(\boldsymbol{\theta}) - \boldsymbol{\Delta}_n(\boldsymbol{\theta}_0)) \\
&\quad + (\boldsymbol{\Delta}_n(\boldsymbol{\theta}) - \boldsymbol{\Delta}_n(\boldsymbol{\theta}_0))' \boldsymbol{\Omega}^{-1} (\boldsymbol{\Delta}_n(\boldsymbol{\theta}) - \boldsymbol{\Delta}_n(\boldsymbol{\theta}_0)) + o_p(1),
\end{aligned}$$

对所有 $\boldsymbol{\theta} \in V_0$ 一致成立, 并且有

$$\boldsymbol{\Delta}_n(\boldsymbol{\theta}) - \boldsymbol{\Delta}_n(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left[(\mathbf{D}_{t,1}(\boldsymbol{\theta}) - \mathbf{D}_{t,1}(\boldsymbol{\theta}_0))', \mathbf{D}_{t,2}(\boldsymbol{\theta}, \mu_0) - \mathbf{D}_{t,2}(\boldsymbol{\theta}_0, \mu_0) \right]'.$$

令 $\Psi_{11} = \frac{1}{4}E\left\{\frac{w_t}{h_t^2(\theta_0)}\frac{\partial h_t(\theta_0)}{\partial \theta}\frac{\partial h_t(\theta_0)}{\partial \theta'}\right\}$ 为 Ψ 左上角的分块, 则由引理 3.3 (c) 和泰勒展开, 我们有

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=1}^n(D_{t,1}(\theta) - D_{t,1}(\theta_0)) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n\left(\frac{\partial D_{t,1}(\theta_0)}{\partial \theta'}(\theta - \theta_0) + o_p\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^n\left(\frac{\partial D_{t,1}(\theta_0)}{\partial \theta'}\frac{\nu}{\sqrt{n}} + o_p\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \nu\left(\frac{1}{n}\sum_{t=1}^n\frac{\partial D_{t,1}(\theta_0)}{\partial \theta'}\right) + o_p(1) \\ &= \Psi_{11}\nu + o_p(1),\end{aligned}$$

对 $\theta \in V_0$ 一致成立。类似地, 令 $\Psi_{21} = E\left\{\frac{w_t}{\sqrt{h_t(\theta_0)}}\frac{\partial \varepsilon_t(\gamma_0)}{\partial \theta'}\right\}$, 即 Ψ 的左下角分块, 我们有

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=1}^n(D_{t,2}(\theta, \mu_0) - D_{t,2}(\theta_0, \mu_0)) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n\left(\frac{\partial D_{t,2}(\theta_0)}{\partial \theta'}(\theta - \theta_0) + o_p\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^n\left(\frac{\partial D_{t,2}(\theta_0, \mu_0)}{\partial \theta'}\frac{\nu}{\sqrt{n}} + o_p\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^n\left(\frac{\partial D_{t,2}(\theta_0, \mu_0)}{\partial \theta'}(\theta - \theta_0) + o_p\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \nu\left(\frac{1}{n}\sum_{t=1}^n\frac{\partial D_{t,2}(\theta_0, \mu_0)}{\partial \theta'}\right) + o_p(1) \\ &= \Psi_{21}\nu + o_p(1),\end{aligned}$$

对 $\theta \in V_0$ 一致成立。

令 $A = (\Psi_{11}, \Psi_{21})$, 我们有

$$\Delta_n(\theta) - \Delta_n(\theta_0) = (\Psi'_{11}, \Psi'_{21})'\nu + o_p(1) = A'\nu + o_p(1),$$

对 $\theta \in V_0$ 一致成立。所以,

$$\begin{aligned}l(\theta, \mu_0) - l(\theta_0, \mu_0) \\ = \nu' A \Omega^{-1} \Delta_n(\theta_0) + \Delta_n'(\theta_0) \Omega^{-1} A' \nu + \nu' A' \Omega^{-1} \nu A' + o_p(1),\end{aligned}\quad (3.9)$$

对 $\theta \in V_0$ 一致成立。类似 Qin and Lawless (1994) 中引理 1 的证明, 我们知道使得 (3.9) 达到最小值的 $\hat{\theta} = \theta_0 + \frac{\nu}{\sqrt{n}}$ 一定在 V_0 内, 即 $\hat{\nu} = -(A \Omega^{-1} A')^{-1} A \Omega^{-1} \Delta_n(\theta_0) + o_p(1)$ 。因此,

$$l(\hat{\theta}, \mu_0) = [\Omega^{-\frac{1}{2}} \Delta_n(\theta_0)]'[I - \Omega^{-\frac{1}{2}} A' (A \Omega^{-1} A')^{-1} A \Omega^{-\frac{1}{2}}][\Omega^{-\frac{1}{2}} \Delta_n(\theta_0)] + o_p(1).$$

由引理 3.4, $\Omega^{-\frac{1}{2}}\Delta_n(\theta_0)$ 依分布收敛到一个多维标准正态分布, 且

$$\begin{aligned} \text{tr}(\mathbf{I} - \Omega^{-\frac{1}{2}}\mathbf{A}'(\mathbf{A}\Omega^{-1}\mathbf{A}')^{-1}\mathbf{A}\Omega^{-\frac{1}{2}}) &= \text{tr}(\mathbf{I}) - \text{tr}(\Omega^{-\frac{1}{2}}\mathbf{A}'(\mathbf{A}\Omega^{-1}\mathbf{A}')^{-1}\mathbf{A}\Omega^{-\frac{1}{2}}) \\ &= p + q + s + r + 3 - \text{rank}(\Omega^{-\frac{1}{2}}\mathbf{A}') = 1, \end{aligned}$$

因此, $l(\hat{\theta}, \mu_0) \xrightarrow{d} \chi_1^2$. □

定理 3.2 的证明: 类似定理 3.1 的证明, 我们有

$$l(\hat{\theta}, 0) = [\Omega^{-\frac{1}{2}}\Delta_n^*(\theta_0)]'[\mathbf{I} - \Omega^{-\frac{1}{2}}\mathbf{A}'(\mathbf{A}\Omega^{-1}\mathbf{A}')^{-1}\mathbf{A}\Omega^{-\frac{1}{2}}][\Omega^{-\frac{1}{2}}\Delta_n^*(\theta_0)] + o_p(1), \quad (3.10)$$

其中 $\Delta_n^*(\theta_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\theta_0, 0)$.

因为 $\mu_0 = M/\sqrt{n}$, 我们有

$$\begin{aligned} \Delta_n^*(\theta_0) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\theta_0, 0) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n (D'_{t,1}(\theta_0), D_{t,2}(\theta_0, 0)) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left[(D'_{t,1}(\theta_0), w_t(\eta_t - \mu_0))' + (\mathbf{0}, \frac{M}{\sqrt{n}}w_t)' \right] \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n (D'_{t,1}(\theta_0), w_t(\eta_t - \mu_0))' + \frac{1}{n} \sum_{t=1}^n (\mathbf{0}, Mw_t)' \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\theta_0, \mu_0) + \frac{1}{n} \sum_{t=1}^n (\mathbf{0}, Mw_t)', \end{aligned}$$

其中 $\mathbf{0}$ 是一个 $(p + q + s + r + 2)$ 维向量。由引理 3.4, 我们有 $\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\theta_0, \mu_0) \xrightarrow{d} N(\mathbf{0}, \Omega)$ 。

由平稳序列的弱大数定律, $\frac{M}{n} \sum_{t=1}^n w_t \xrightarrow{p} MEw_t$ 。因此, $\Omega^{-\frac{1}{2}}\Delta_n^*(\theta_0)$ 依分布收敛到一个均值为 $-(\mathbf{0}', MEw_t)'$ 、方差为 $\mathbf{I}$ 的多维正态分布。

由 (3.10), 当 $n \rightarrow \infty$ 时, $l(\hat{\theta}, 0)$ 收敛到一个非中心的自由度为 1 的卡方分布, 非中心参数为 $M^2(Ew_t)^2$. □

4 ARMA-GARCH 模型误差零中位数的经验似然检验

4.1 引言

在第 3 章中, 我们探究了如何在应用 SWQMELE 拟合 ARMA-GARCH 模型时对误差的零均值进行检验。考虑一个 $\text{ARMA}(r, s)\text{-GARCH}(p, q)$, 定义为:

$$\begin{cases} X_t = \mu + \sum_{i=1}^r \phi_i X_{t-i} + \sum_{j=1}^s \psi_j \varepsilon_{t-j} + \varepsilon_t, \\ \varepsilon_t = \sigma_t \eta_t, \sigma_t^2 = w_1 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \end{cases} \quad (4.1)$$

第 3 章中所提出的 **profile** 经验似然检验要求事先给定模型的阶数 r 、 s 、 p 和 q , 或者其可以被精确的估计; 并且该检验需要对 $(p+q+r+s+2)$ 维的冗余参数空间进行搜索, 来找到能够使得对数似然函数达到最小的冗余参数值。然而, 在实际中, 我们很难识别模型的阶数, 一般用来选择阶数的方法是 AIC 或贝叶斯信息准则 (BIC), 然而 AIC 容易高估, BIC 容易低估, 参考 Burnham and Anderson (2004); 并且当阶数较大时, 对冗余参数空间的搜索在计算上效率较低。因此, 我们需要找到一个对参数选择不敏感、更稳健的方法来进行检验。

为了找到一个更稳健的检验方法, 本章我们考虑了在重尾情况下, 对 ARMA-GARCH 模型中 η_t 的中位数进行检验。由于在模型 (4.1) 中, η_t 的零中位数相当于 ε_t 的零中位数, 这样我们可以在 p 和 q 未知以及不估计未知参数 $w_1, \{a_i\}_{i=1}^p$ 和 $\{b_j\}_{j=1}^q$ 的情况下, 对 ε_t 的零中位数进行检验, 该检验对波动和重尾具有稳健的特性。为了估计 ε_t 的中位数, 且保证估计量具有正态极限, 同时避免估计模型的 GARCH 部分, 我们不能使用最小二乘法来估计 ARMA 模型中的参数 $\theta = (\mu, \phi_1, \dots, \phi_r, \psi_1, \dots, \psi_s)'$, 因为 Zhang and Ling (2015) 证明了: 当 $E\varepsilon_t^4 = \infty$ 时, 其估计量有一个非正态极限。本文我们将采用 Huang et al. (2020) 提出的加权最小二乘法来估计 θ , 然后通过加权中位数来估计 ε_t 的中位数。尽管我们没有估计 GARCH 模型, 但由于所采用的权重函数可以弱化 σ_t 的作用, 降低矩条件的要求, 即使在 $E\varepsilon_t^2 = \infty$ 的情况下, 我们方法中 ε_t 中位数的估计量也渐近收敛于正态分布。然而, 此估计量的渐近方差非常复杂, 因此我们提出应用经验似然检验方法来加以解决。

本章结构如下。第 4.2 节介绍了方法和理论结果。第 4.3 节和第 4.4 节分别给出了模拟研究和实证分析。第 4.5 节给出了结论。定理的证明将在第 4.6 节给出。

4.2 方法和理论结果

令 $\mathcal{F}_t$ 表示由 $\{\varepsilon_s : s \leq t\}$ 生成的 σ -域, $\boldsymbol{\theta} = (\mu, \phi_1, \dots, \phi_r, \psi_1, \dots, \psi_s)'$, $\boldsymbol{\delta} = (w_1, a_1, \dots, a_p, b_1, \dots, b_q)'$, 且 $\boldsymbol{\gamma} = (\boldsymbol{\theta}', \boldsymbol{\delta}')'$, $\boldsymbol{\gamma}_0$ 表示真值。再令 $\eta_t(\boldsymbol{\gamma}_0) = \eta_t, \varepsilon_t(\boldsymbol{\theta}_0) = \varepsilon_t$ 和 $\sigma_t(\boldsymbol{\gamma}_0) = \sigma_t$ 。定义 $\phi(z) = 1 - \sum_{i=1}^r \phi_i z^i$, $\psi(z) = 1 + \sum_{j=1}^s \psi_j z^j$, $a(z) = \sum_{i=1}^p a_i z^i$ 和 $b(z) = 1 - \sum_{j=1}^q b_j z^j$ 。

给定观测值 $\{X_1, \dots, X_n\}$ 和初始值 $\{X_0, X_{-1}, \dots\}$, 我们在模拟研究和实证分析中取初始值为 0。我们将模型 (4.1) 写为参数形式:

$$\begin{cases} \varepsilon_t(\boldsymbol{\theta}) = X_t - \mu - \sum_{i=1}^r \phi_i X_{t-i} - \sum_{j=1}^s \psi_j \varepsilon_{t-j}(\boldsymbol{\theta}), \\ \sigma_t^2(\boldsymbol{\gamma}) = w_1 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2(\boldsymbol{\theta}) + \sum_{j=1}^q b_j \sigma_{t-j}^2(\boldsymbol{\gamma}), \quad \eta_t(\boldsymbol{\gamma}) = \frac{\varepsilon_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\gamma})}. \end{cases}$$

要检验 ε_t 是否有零中位数, 我们不能直接用最小二乘法来估计 $\boldsymbol{\theta}$, 因为 Zhang and Ling (2015) 证明了: 当 $E\varepsilon_t^4 = \infty$ 时, 最小二乘估计没有正态极限。反之, 我们提出用加权最小二乘法, 即最小化 $\sum_{t=1}^n w_{t-1}^{-2} \varepsilon_t^2(\boldsymbol{\theta})$, 相当于解估计方程:

$$\sum_{t=1}^n w_{t-1}^{-2} \varepsilon_t(\boldsymbol{\theta}) \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0,$$

其中 w_t 是 $\mathcal{F}_t$ 可测的, 其具体意义将在下文中定义。因为 $\varepsilon_t(\boldsymbol{\theta}) = \sigma_t(\boldsymbol{\gamma})\eta_t(\boldsymbol{\gamma})$, 以上的估计方法使用了一个 w_{t-1} 来弱化 $\varepsilon_t(\boldsymbol{\theta})$ 中 σ_t 的作用, 另一个 w_{t-1} 是用来弱化 $\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ 的作用。因此, 我们提出的加权最小二乘估计 $\hat{\boldsymbol{\theta}}$ 在 $E\varepsilon_t^2 = \infty$ 下也有正态极限。接下来, 我们使用 $\hat{\boldsymbol{\theta}}$ 来估计 ε_t 的中位数, 即解估计方程:

$$\sum_{t=1}^n w_{t-1}^{-1} \text{sgn}(\varepsilon_t(\hat{\boldsymbol{\theta}}) - d) = 0,$$

其中 sgn 为符号函数。类似上文所述, w_{t-1} 减少了估计量 $\hat{\boldsymbol{\theta}}$ 的矩效应。因此, 所得到的中位数估计在 $E\varepsilon_t^2 = \infty$ 时也有正态极限, 但它的渐近方差非常复杂。为了避免估计渐近方差, 我们使用了 Qin and Lawless (1994) 提出的基于估计方程构造经验似然比的方法。

令

$$D_{t,1}(\boldsymbol{\theta}) = w_{t-1}^{-2} \varepsilon_t(\boldsymbol{\theta}) \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \quad D_{t,2}(\boldsymbol{\theta}, d) = w_{t-1}^{-1} \text{sgn}(\varepsilon_t(\boldsymbol{\theta}) - d),$$

和

$$D_t(\boldsymbol{\theta}, d) = (D_{t,1}'(\boldsymbol{\theta}), D_{t,2}(\boldsymbol{\theta}, d))',$$

其中 d 为 $\varepsilon_t(\boldsymbol{\theta})$ 的中位数。我们定义对 $\boldsymbol{\theta}$ 和 d 的经验似然函数为:

$$L(\boldsymbol{\theta}, d) = \sup \left\{ \prod_{t=1}^n (np_t) : p_1 \geq 0, \dots, p_n \geq 0, \sum_{t=1}^n p_t = 1, \sum_{t=1}^n p_t D_t(\boldsymbol{\theta}, d) = \mathbf{0} \right\}.$$

使用拉格朗日乘子方法, 我们得到 $p_t = \frac{1}{n\{1 + \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, d)\}}$ 和对数经验似然比

$$l(\boldsymbol{\theta}, d) = -2 \log(L(\boldsymbol{\theta}, d)) = 2 \sum_{t=1}^n \log\{1 + \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, d)\},$$

其中 $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\theta}, d)$ 满足

$$\sum_{t=1}^n \frac{\mathbf{D}_t(\boldsymbol{\theta}, d)}{1 + \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, d)} = 0.$$

因为只对 d 感兴趣, 所以我们考虑 profile 经验似然比 $l_p(d) = \min_{\boldsymbol{\theta} \in \Theta} l(\boldsymbol{\theta}, d)$ 。

为了证明所提出的经验似然检验具有 Wilks 定理, 我们引入以下几个条件。

首先, 令

$$\mathbf{A}_t = \begin{pmatrix} a_1 \eta_t^2 + b_1 & b_2 & \cdots & b_{q-1} & b_q & a_2 & a_3 & \cdots & a_p \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ \eta_t^2 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

γ 为随机矩阵 $\{\mathbf{A}_t\}$ 的 Lyapunov 指数, 即对 $(p+q-1) \times (p+q-1)$ 矩阵空间内的任意范数 $\|\cdot\|$,

$$\gamma = \inf \left\{ \frac{1}{n} E(\ln \|\mathbf{A}_1 \cdots \mathbf{A}_n\|), n \in \mathbb{N} \right\},$$

其中 $\|\mathbf{A}_t\| = \sup_{|x|=1} |\mathbf{A}_t \mathbf{x}|$ 。

假设 4.1: $\boldsymbol{\theta}_0$ 为 Θ 的内点。对每个 $\boldsymbol{\theta} \in \Theta$, 当 $|z| \leq 1$ 时, $\phi(z) \neq 0$ 且 $\psi(z) \neq 0$, $\phi(z)$ 和 $\psi(z)$ 无共同根, $\phi_r \neq 0$ 或 $\psi_s \neq 0$ 。

假设 4.2: $w_1 > 0$, $\gamma < 0$, 且 $E(\ln(\max(|\eta_1|, 1))) < \infty$ 。

假设 4.3: 对任意 $\rho \in (0, 1)$, $E(w_{t-1}^{-4} \xi_{\rho,t}^4) < \infty$, 其中 $\xi_{\rho,t} = 1 + \sum_{i=0}^{\infty} \rho^i |X_{t-i}|$, $w_t > 0$ 且 $\mathcal{F}_t$ 可测, $\inf_{t \geq 1} w_t > 0$, $\{w_t\}$ 为一个平稳序列。

假设 4.4: $\{\eta_t\}$ 为 i.i.d. 随机变量, 均值为 0, 方差为 1。

假设 4.5: $\{\eta_t\}$ 有连续密度函数 $g(x)$, 满足 $g(0) > 0$ 和 $\sup_{x \in \mathbb{R}} g(x) < \infty$ 。

假设 4.1 和 4.2 保证了 (4.1) 的第一个和第二个等式均存在唯一严平稳因果解（更多关于 GARCH 模型的内容，参考 Basrak et al. (2002) 的定理 3.1）。在假设 4.3 下，我们可以使用权重来减少 σ_t 的作用。假设 4.4 表明我们对条件均值和条件标准差考虑经典的 ARMA-GARCH 模型。假设 4.5 为中位数估计的常规假设。

定理 4.1: 在假设 4.1-4.5 和原假设 $H_0 : d_0 = 0$ 下，当 $n \rightarrow \infty$ 时， $l_p(0)$ 依分布收敛到自由度为 1 的卡方分布。

对于检验 (4.1) 中的 ε_t 是否有零中位数，上述定理表明，当 $l_p(0) \geq \chi_{1,1-\alpha}^2$ 时，在显著性水平 α 上拒绝原假设，其中 $\chi_{1,1-\alpha}^2$ 表示自由度为 1 的卡方分布的 $(1 - \alpha)$ 分位数。

为了研究所提出的检验的 power，我们考虑如下的局部备择假设：

$$H_a : d_0 = \frac{M_1}{\sqrt{n}}, M_1 \text{ 为常数。} \quad (4.2)$$

下面的定理证明了我们所提出的检验具有良好的 power。当 $|M_1| \rightarrow \infty$ 时，power 趋向于 1。

定理 4.2: 若对模型 (4.1) 有假设 4.1-4.5 成立，在备择假设 (4.2) 下， $l_p(0)$ 依分布收敛到自由度为 1 的非中心卡方分布，非中心参数为

$$(\mathbf{0}', 2g(0)M_1E(w_{t-1}^{-1})E(\sigma_t^{-1}))\Lambda^{-1}(\mathbf{0}', 2g(0)M_1E(w_{t-1}^{-1})E(\sigma_t^{-1})),$$

其中 $\Lambda = E(D_1(\theta_0, 0)D_1'(\theta_0, 0))$ 。

关于权重函数的选取，由 Ling (2007) 和 Zhu and Ling (2011, 2015) 可知，选择权重函数的主要思路是对假设 4.3 中定义的 $\xi_{\rho,t}$ 施加一个上界。有很多的权重函数可供选择，例如 Ling (2007) 中所提出的。我们注意到，当 t 充分大时，对任意给定的 $\rho \in (0, 1)$ 和 $h \in (0, 1)$ ， $\sum_{i=0}^{\infty} e^{\log(h)\log^2(i+1)} < \infty$ 和 $e^{\log(h)\log^2(t+1)} \geq \rho^t$ 成立。因此，我们可以使用 $\sum_{i=0}^{t-1} e^{\log(h)\log^2(i+1)}|X_{t-i}|$ 来给 $\xi_{\rho,t}$ 施加一个上界，从而控制 σ_{t+1} 的作用。因此，可以采用以下的权重函数：

$$w_t(h) = \max(C, \sum_{i=0}^{t-1} e^{\log(h)\log^2(i+1)}|X_{t-i}|), \text{ 其中 } h \in (0, 1) \text{ 且 } t = 1, \dots, n, \quad (4.3)$$

其中 C 为 $\{|X_t|\}_{t=1}^n$ 的 90% 分位数且 $w_0(h) = 1$ 。类似 He et al. (2020)，我们可以证明：若上述权重函数中样本分位数 C 由其他相似的分位数代替（记为 $\tilde{C}$ ），那么假设 4.3 也是满足的；并且用 C 和 $\tilde{C}$ 的估计量都有相同的渐近分布。类似核密度估计，

按照 **size** 来选择 h 时需要计算经验似然比的导数的 Edgeworth 展开, 因此将会十分复杂。尽管如此, 下面的模拟结果仍然证实了使用 $h = 0.1, 0.2, 0.3, 0.4, 0.5$ 和 0.6 都有准确的 **size** 和很好的 **power**。

4.3 模拟研究

本节我们探究了在有限样本下, 所提出的经验似然检验在 **size** 和 **power** 上的表现。

我们从 ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型中重复生成了 5000 次随机样本, 样本量 $n = 1000$ 和 2000 。模型参数选为 $\mu = 0.1, \phi_1 = 0.5, \psi_1 = 0.2, w_1 = 0.1, a_1 = 0.2$ 或 $0.1, b_1 = 0.8$ 。我们考虑 $\eta_t = (V - E(V))/\sqrt{\text{Var}(V)}$, 其中

$$V = \frac{I(U < \delta)V_1}{1/(\alpha_1 - 1)} - \frac{I(U \geq \delta)V_2}{1/(\alpha_2 - 1)},$$

$U \sim \text{Uniform}(0, 1)$, $V_1 \sim \text{Pareto}(1, \alpha_1)$ 和 $V_2 \sim \text{Pareto}(1, \alpha_2)$ 独立。显然, $E\varepsilon_t^2$ 当 $a_1 = 0.1$ 时有限, 当 $a_1 = 0.2$ 时无限。通过简单的计算, 我们可以得出:

$$E(V) = 2\delta - 1, E(V^2) = \frac{2\delta(\alpha_1 - 1)}{\alpha_1 - 2} + \frac{2(1 - \delta)(\alpha_2 - 1)}{\alpha_2 - 2},$$

$$P(\eta_t < 0) = 1 - \delta + \delta P(V_1 < \frac{2\delta - 1}{\alpha_1 - 1}) = 1 - \delta(1 + \frac{2\delta - 1}{\alpha_1 - 1})^{-\alpha_1}, \text{ 对 } 2\delta - 1 \geq 0,$$

和

$$P(\eta_t < 0) = \delta P(V_2 > \frac{1 - 2\delta}{\alpha_2 - 1}) = \delta(1 + \frac{1 - 2\delta}{\alpha_2 - 1})^{-\alpha_2}, \text{ 对 } 2\delta - 1 < 0.$$

因此, η_t 的右尾指数为 α_1 , 左尾指数为 α_2 , 若 $\delta = 0.5$ 则有零中位数, 当 $\alpha_1 \neq \alpha_2$ 时不对称。经过基本的计算可知, 当 $\delta \neq 1/2, \alpha_1 > 2$ 和 $\alpha_2 > 2$ 时, $P(\eta_t < 0) \neq 1/2$, 即若 $\delta \neq 1/2, \eta_t$ 中位数不为 0。注意, 若 η_t 有标准的偏斜 t 分布, 均值为 0, 方差为 1, 则它有相同的左右尾指数, 在中位数为 0 时是对称的。因此, 我们所构造的 η_t 的分布在对尾部和对称性建模时比常用的偏斜 t 分布更灵活。

我们取 $\alpha_1 = 2.2$ 或 $4.5, \alpha_2 = 2.2, \delta = 0.3, 0.4$ 或 0.5 。因此, $\delta = 0.5$ 时我们对检验的 **size** 进行评估, 其他情况下我们对检验的 **power** 进行评估。我们使用 R 软件包 “emplik” 来计算经验似然函数, R 函数 “optim” 来计算 profile 经验似然函数。我们使用 (4.3) 中的权重函数 $w_t(h)$, 其中 $h = 0.1, 0.2, 0.3, 0.4, 0.5$ 和 0.6 。表 4.1 和表 4.2-4.3 分别给出了显著性水平为 10%、5% 和 1% 时, profile 经验似然检验的经验 **size** 和经验 **power**。为了节省空间, 我们只给出 $h = 0.1$ 和 0.3 时的结果, 其他 h 下的结果均类似。

表 4.1 $\delta = 0.5$ 时, ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型经验似然检验的 size

a_1	α_1	n	显著性水平	ARMA(1,0)-GARCH(1,1)		ARMA(1,1)-GARCH(1,1)	
				$h = 0.1$	$h = 0.3$	$h = 0.1$	$h = 0.3$
0.2	2.2	1000	10%	0.1282	0.1292	0.1448	0.1458
			5%	0.0738	0.0752	0.0824	0.0868
			1%	0.0164	0.0174	0.0250	0.0264
		2000	10%	0.1194	0.1162	0.1326	0.1322
			5%	0.0608	0.0604	0.0744	0.0742
			1%	0.0148	0.0128	0.0212	0.0202
	4.5	1000	10%	0.1160	0.1154	0.1168	0.1188
			5%	0.0586	0.0590	0.0626	0.0664
			1%	0.0122	0.0136	0.0194	0.0170
		2000	10%	0.1012	0.1040	0.1120	0.1156
			5%	0.0530	0.0520	0.0630	0.0612
			1%	0.0126	0.0098	0.0156	0.0144
0.1	2.2	1000	10%	0.1276	0.1298	0.1408	0.1424
			5%	0.0730	0.0708	0.0804	0.0856
			1%	0.0164	0.0186	0.0230	0.0262
		2000	10%	0.1194	0.1148	0.1378	0.1378
			5%	0.0606	0.0580	0.0756	0.0654
			1%	0.0128	0.0132	0.0208	0.0200
	4.5	1000	10%	0.1144	0.1146	0.1164	0.1206
			5%	0.0590	0.0608	0.0654	0.0668
			1%	0.0128	0.0128	0.0172	0.0170
		2000	10%	0.1002	0.1022	0.1112	0.1134
			5%	0.0530	0.0492	0.0624	0.0594
			1%	0.0116	0.0098	0.0150	0.0156

表 4.1 表明:

- 当样本量变大时, size 的精确性也随之提升,
- $\alpha_1 = 4.5$ 时的 size 比 $\alpha_1 = 2.2$ 时的 size 更精确,
- size 对 h 的选择稳健,
- ARMA (1,0) 模型的 size 比 ARMA (1,1) 模型的 size 更精确,
- η_t 的重尾程度而不是 ε_t 的重尾程度对 size 有影响。

表 4.2 $\delta = 0.3$ 时, ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型经验似然检验的检验 power

a_1	α_1	n	显著性水平	ARMA(1,0)-GARCH(1,1)		ARMA(1,1)-GARCH(1,1)	
				$h = 0.1$	$h = 0.3$	$h = 0.1$	$h = 0.3$
0.2	2.2	1000	10%	0.8368	0.8344	0.8358	0.8322
			5%	0.7672	0.7584	0.7728	0.7648
			1%	0.5960	0.5918	0.5908	0.5880
		2000	10%	0.9394	0.9374	0.9478	0.9448
			5%	0.9082	0.9038	0.9164	0.9160
			1%	0.8110	0.8130	0.8218	0.8202
	4.5	1000	10%	0.9514	0.9532	0.9498	0.9474
			5%	0.9178	0.9152	0.9086	0.9078
			1%	0.7806	0.7792	0.7696	0.7670
		2000	10%	0.9986	0.9990	0.9976	0.9986
			5%	0.9960	0.9970	0.9950	0.9952
			1%	0.9812	0.9822	0.9804	0.9820
0.1	2.2	1000	10%	0.8422	0.8382	0.8460	0.8382
			5%	0.7786	0.7654	0.7658	0.7734
			1%	0.6118	0.6000	0.6102	0.5988
		2000	10%	0.9426	0.9362	0.9508	0.9466
			5%	0.9114	0.9042	0.9200	0.9176
			1%	0.8252	0.8172	0.8318	0.8234
	4.5	1000	10%	0.9590	0.9536	0.9584	0.9500
			5%	0.9282	0.9146	0.9226	0.9120
			1%	0.8040	0.7920	0.7958	0.7810
		2000	10%	0.9992	0.9990	0.9984	0.9984
			5%	0.9978	0.9962	0.9964	0.9962
			1%	0.9848	0.9854	0.9860	0.9838

表 4.3 $\delta = 0.4$ 时, ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型经验似然检验的检验 power

a_1	α_1	n	显著性水平	ARMA(1,0)-GARCH(1,1)		ARMA(1,1)-GARCH(1,1)	
				$h = 0.1$	$h = 0.3$	$h = 0.1$	$h = 0.3$
0.2	2.2	1000	10%	0.3926	0.3926	0.4004	0.4040
			5%	0.2930	0.2958	0.3048	0.3038
			1%	0.1424	0.1406	0.1400	0.1426
		2000	10%	0.5632	0.5576	0.5658	0.5656
			5%	0.4472	0.4462	0.4594	0.4586
			1%	0.2402	0.2432	0.2658	0.2664
		1000	10%	0.4326	0.4380	0.4420	0.4398
			5%	0.3184	0.3264	0.3144	0.3274
			1%	0.1424	0.1492	0.1358	0.1414
0.2	4.5	2000	10%	0.6854	0.6834	0.6870	0.6816
			5%	0.5656	0.5722	0.5774	0.5806
			1%	0.3178	0.3318	0.3346	0.3372
		1000	10%	0.4022	0.3996	0.4108	0.4086
			5%	0.3020	0.2996	0.3098	0.3040
			1%	0.1506	0.1462	0.1484	0.1468
		2000	10%	0.5708	0.5610	0.5686	0.5642
			5%	0.4606	0.4488	0.4680	0.4578
			1%	0.2550	0.2446	0.2658	0.2630
0.1	2.2	1000	10%	0.4544	0.4474	0.4594	0.4508
			5%	0.3418	0.3324	0.3448	0.3380
			1%	0.1602	0.1570	0.1498	0.1480
		2000	10%	0.7048	0.6954	0.7094	0.6942
			5%	0.5984	0.5826	0.6028	0.5830
			1%	0.3486	0.3444	0.3676	0.3512

此外, 我们从表 4.2和 4.3得出:

- 当样本量变大时, 检验的 power 随之提升,
- $\delta = 0.3$ 时的检验比 $\delta = 0.4$ 时的检验有更大的 power, 因为前者误差的中位数比后者离 0 更远,
- $\alpha_1 = 4.5$ 时的检验比 $\alpha_1 = 2.2$ 时的检验 power 更好,
- η_t 的重尾程度而不是 ε_t 的重尾程度对 power 有影响, 因为我们使用了权重函数来减少 σ_t 的作用。

4.4 实证分析

4.4.1 外汇序列

本节研究了 2009 年 10 月 5 日至 2019 年 10 月 4 日的每日 HKD/USD、USD/EUR、CAD/USD 和 MXN/USD 外汇序列。我们用 $\{X_t\}$ 表示每个序列的对数收益率 ($\times 100$)。

我们首先通过 Hill (1975) 中的 Hill 估计量 $\hat{\alpha}(k)$ 来估计 $\{X_t\}$ 的尾部指数, 定义如下:

$$\hat{\alpha}(k) = \left[\frac{1}{k} \sum_{i=1}^k \log \left(\frac{X_{(n-i+1)}}{X_{(n-k)}} \right) \right]^{-1},$$

其中 $X_{(t)}$ 是 X_t 的升序统计量。我们在图 4.1 中画出了 Hill 估计量与不同的 k 之间的关系, 可以看出 HKD/USD、CAD/USD 和 MXN/USD 外汇的尾部指数小于 4, 即 $E(X_t^4) = \infty$ 。因此, 当我们用 QMLE 来拟合 ARMA-GARCH 模型时, 估计量可能没有正态极限。为了探索使用 SWQMELE 来拟合 ARMA-GARCH 模型的可能性, 我们必须检验模型是否具有零中位数。

为了进行这样的检验, 我们首先使用 R 软件包 “forecast” 中的 “auto.arima” 函数和 AIC 准则来选择 ARMA 模型合适的阶数。最终, 我们用 ARMA(2,2) 模型来拟合 HKD/USD 外汇, 用 ARMA(1,1) 模型来拟合 USD/EUR 外汇, 用 ARMA(5,4) 模型来拟合 CAD/USD 外汇, 用 ARMA(3,2) 模型来拟合 MXN/USD 外汇。接下来, 我们用 (4.3) 中的权重函数 $w_t(h)$ 在 $h = 0.1, 0.2, 0.3, 0.4$ 和 0.5 的情况下, 计算经验似然检验的 p 值。表 4.4 表明, 我们对 MXN/USD 外汇拒绝零中位数的原假设, 而对 HKD/USD 和 USD/EUR 外汇来说, 原假设成立。因此, 在使用 SWQMELE 对 MXN/USD 外汇拟合 ARMA-GARCH 模型时应该谨慎。

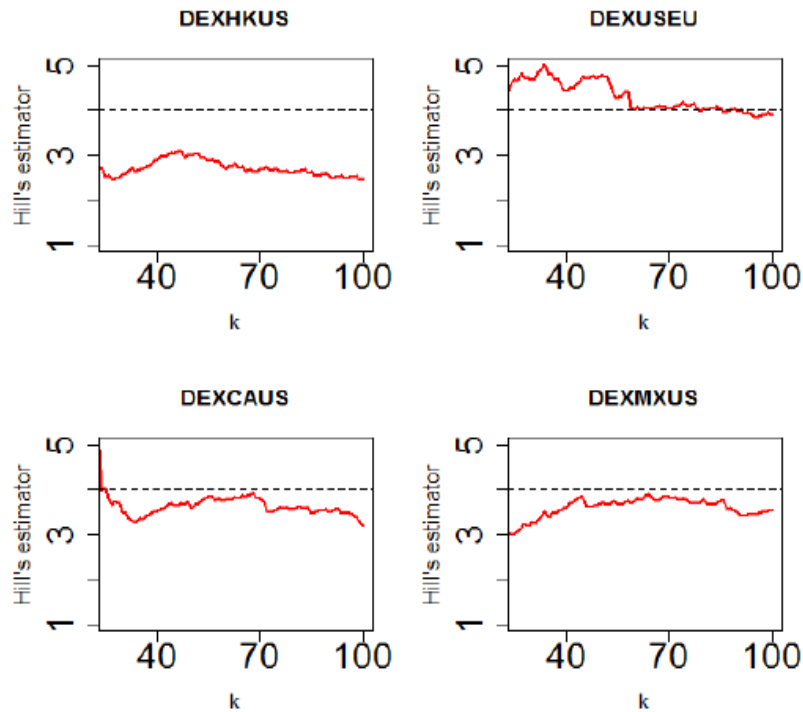


图 4.1 HKD/USD, USD/EUR, CAD/USD 和 MXN/USD 外汇的 Hill 估计量

表 4.4 HKD/USD, USD/EUR, CAD/USD 和 MXN/USD 外汇经验似然检验的 p 值

外汇	ARMA 模型	$h = 0.1$	$h = 0.2$	$h = 0.3$	$h = 0.4$	$h = 0.5$
HKD/USD	ARMA(2,2)	0.1754	0.2971	0.1935	0.2555	0.6919
USD/EUR	ARMA(1,1)	0.6431	0.6178	0.8581	0.8583	0.9567
CAD/USD	ARMA(5,4)	0.2100	0.1267	0.1044	0.0841	0.0153
MXN/USD	ARMA(3,2)	0.0134	0.0157	0.0158	0.0107	0.0109

4.4.2 股票序列

本节考虑了从 2009 年 10 月 28 日到 2019 年 7 月 29 日以下股票和指数的每日收盘价：Transocean (RIG)、S&P 500、Microsoft Stock (MSFT) 和 Dow Jones Index (DJI)。与第 4.4.1 节一样，我们首先在图 4.2 中画出了每个序列的 Hill 估计量，从图中可以看出 S&P 500 和 MSFT 的尾部指数均小于 4。因此，我们需要在采用 SWQMELE 来拟合 ARMA-GARCH 模型前确认模型有零中位数。

如前所述，我们首先用“auto.arima”和 AIC 准则选择 ARMA 模型的阶数，然后计算在 $h = 0.1, 0.2, 0.3, 0.4$ 和 0.5 下误差零中位数检验的 p 值。表 4.5 表明，我们在使用 SWQMELE 对 S&P500 和 DJI 进行 ARMA-GARCH 模型拟合时也应该比较谨慎，因为误差的中位数可能不为 0。

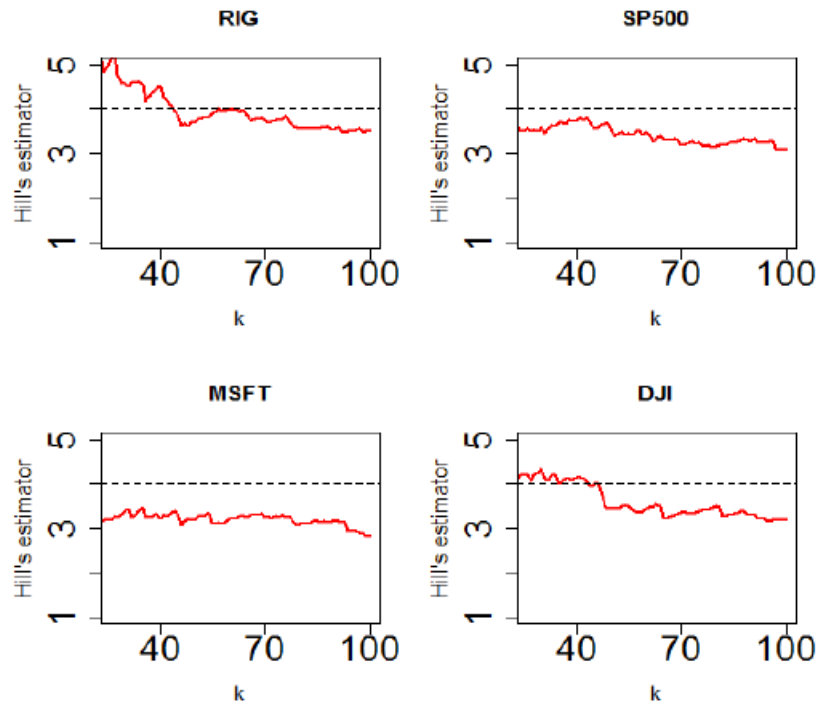


图 4.2 RIG, S&P 500, MSFT 和 DJI 的 Hill 估计量

表 4.5 RIG, S&P 500, MSFT 和 DJI 经验似然检验的 p 值

股票/指数	ARMA 模型	$h = 0.1$	$h = 0.2$	$h = 0.3$	$h = 0.4$	$h = 0.5$
RIG	ARMA(1,0)	0.9743	0.9013	0.9756	0.9811	0.7316
S&P 500	ARMA(3,1)	0.0937	0.1037	0.0996	0.0761	0.0551
MSFT	ARMA(1,1)	0.9624	0.7056	0.9750	0.8013	0.4259
DJI	ARMA(3,1)	0.0044	0.0675	0.2006	0.0216	0.0169

4.5 结论

为了找到一个稳健的方法来检验应用 SWQMELE 时 ARMA-GARCH 模型满足其假设, 本章在不估计 GARCH 模型和其渐近方差的情况下, 探究了在重尾情况下对误差零中位数的经验似然检验。模拟研究证实了该检验在有限样本下有良好的 size 和 power。实际数据分析表明, 在对模型应用 SWQMELE 前进行检验十分重要。

4.6 定理 4.1 和 4.2 的证明

在证明定理 4.1 和 4.2 前, 我们需要一些引理。引理 4.1 给出了使用权重函数减少矩效应的相关上界。引理 4.2 证明了 $D_t(\theta, 0)$ 的样本均值和方差可以被 $D_t(\theta_0, 0)$ 的样本均值和方差逼近。引理 4.3 得出了 $\sum_{t=1}^n D_t(\theta_0, 0)$ 的渐近正态性。对于冗余参数的处理, 引理 4.4 给出了对 $\sum_{t=1}^n D_{t,2}(\theta, 0)$ 的逼近。

引理 4.1: 在定理 4.1 的条件下, 存在常数 $\rho \in (0, 1)$, 常数 $C > 0$ 和 θ_0 的邻域 Θ_0 满足

$$\sup_{\Theta_0} |\varepsilon_t(\theta)| \leq C\xi_{\rho,t-1}, \quad \sup_{\Theta_0} \left\| \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \right\| \leq C\xi_{\rho,t-1}, \quad \text{and} \quad \sup_{\Theta_0} \left\| \frac{\partial^2 \varepsilon_t(\theta)}{\partial \theta \partial \theta'} \right\| \leq C\xi_{\rho,t-1},$$

其中 $\xi_{\rho,t}$ 如假设 4.3 中定义。

证明: 参考 Ling (2007)。 □

引理 4.2: 令 $\mathcal{B}_0 = \{\theta : \|\theta - \theta_0\| \leq \frac{C_0}{\sqrt{n}}\}$, 其中 $C_0 > 0$ 。在定理 4.1 的条件下, 我们有

$$(i) \quad \max_{1 \leq t \leq n} \sup_{\mathcal{B}_0} \|D_t(\theta, 0)\| = o_p(\sqrt{n}),$$

$$(ii) \quad \frac{1}{n} \sum_{t=1}^n D_t(\theta, 0) = \frac{1}{n} \sum_{t=1}^n D_t(\theta_0, 0) + O_p(n^{-1/2}) \text{ 对 } \theta \in \mathcal{B}_0 \text{ 一致成立,}$$

$$(iii) \quad \frac{1}{n} \sum_{t=1}^n \{D_t(\theta, 0) D_t'(\theta, 0)\} = E\{D_1(\theta_0, 0) D_1'(\theta_0, 0)\} + o_p(1) \text{ 对 } \theta \in \mathcal{B}_0 \text{ 一致成立。}$$

证明：(i) 我们首先证明

$$\max_{1 \leq t \leq n} \sup_{\mathcal{B}_0} \|D_{t,1}(\boldsymbol{\theta})\| = o_p(\sqrt{n}). \quad (4.4)$$

由引理 4.1, 对充分大的 n 有以下成立:

$$\begin{aligned} \sup_{\mathcal{B}_0} \|D_{t,1}(\boldsymbol{\theta})\| &= \sup_{\mathcal{B}_0} \left\| w_{t-1}^{-2} \varepsilon_t(\boldsymbol{\theta}) \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| \\ &\leq \sup_{\mathcal{B}_0} w_{t-1}^{-2} \|\varepsilon_t(\boldsymbol{\theta})\| \left\| \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| \leq C^2 w_{t-1}^{-2} \xi_{\rho,t-1}^2 =: V(t). \end{aligned}$$

由切贝雪夫不等式, 对任意 $\epsilon > 0$, 当 $n \rightarrow \infty$ 时,

$$\begin{aligned} &P\left(\max_{1 \leq t \leq n} V(t) > \sqrt{n}\epsilon\right) \\ &\leq \frac{1}{n\epsilon^2} \sum_{t=1}^n E\left\{V^2(t) I(V(t) > \sqrt{n}\epsilon)\right\} \\ &\leq \frac{1}{\epsilon^2} \max_{1 \leq t \leq n} E\left\{V^2(t) I(V(t) > \sqrt{n}\epsilon)\right\} \rightarrow 0, \end{aligned}$$

其中最后一个不等式可由

$$E(V^2(t)) = C^4 E(w_{t-1}^{-4} \xi_{\rho,t-1}^4) < \infty,$$

和假设 4.3 得到。因此, (4.4) 成立。

因为

$$\sup_{\mathcal{B}_0} |D_{t,2}(\boldsymbol{\theta}, 0)| = \sup_{\mathcal{B}_0} |w_{t-1}^{-1} \text{sgn}(\varepsilon_t(\boldsymbol{\theta}))| \leq w_{t-1}^{-1} \sup_{\mathcal{B}_0} |\text{sgn}(\varepsilon_t(\boldsymbol{\theta})) - \text{sgn}(\varepsilon_t)| + w_{t-1}^{-1} \leq 3w_{t-1}^{-1},$$

由假设 4.3 可得:

$$\max_{1 \leq t \leq n} \sup_{\mathcal{B}_0} |D_{t,2}(\boldsymbol{\theta}, 0)| = o_p(\sqrt{n}). \quad (4.5)$$

因此, 由 (4.4) 和 (4.5), 我们有

$$\max_{1 \leq t \leq n} \sup_{\mathcal{B}_0} \|D_t(\boldsymbol{\theta}, 0)\| = o_p(\sqrt{n}).$$

(ii) 注意到

$$\frac{1}{n} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0) - \frac{1}{n} \sum_{t=1}^n D_t(\boldsymbol{\theta}_0, 0) = \frac{1}{n} \sum_{t=1}^n (D'_{t,1}(\boldsymbol{\theta}) - D'_{t,1}(\boldsymbol{\theta}_0), D_{t,2}(\boldsymbol{\theta}, 0) - D_{t,2}(\boldsymbol{\theta}_0, 0))'.$$

要证明 (ii), 只要证明

$$\sup_{\mathcal{B}_0} \left\| \frac{1}{n} \sum_{t=1}^n (D_{t,1}(\boldsymbol{\theta}) - D_{t,1}(\boldsymbol{\theta}_0)) \right\| = O_p\left(\frac{1}{\sqrt{n}}\right),$$

和

$$\sup_{\mathcal{B}_0} \left| \frac{1}{n} \sum_{t=1}^n (D_{t,2}(\boldsymbol{\theta}, 0) - D_{t,2}(\boldsymbol{\theta}_0, 0)) \right| = O_p\left(\frac{1}{\sqrt{n}}\right).$$

根据泰勒展开和引理 4.1, 我们有

$$\begin{aligned} & \sup_{\mathcal{B}_0} \|D_{t,1}(\boldsymbol{\theta}) - D_{t,1}(\boldsymbol{\theta}_0)\| \\ &= \sup_{\mathcal{B}_0} \left\| w_{t-1}^{-2} \varepsilon_t(\boldsymbol{\theta}) \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - w_{t-1}^{-2} \varepsilon_t \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \right\| \\ &= \sup_{\mathcal{B}_0} w_{t-1}^{-2} \left\| (\varepsilon_t(\boldsymbol{\theta}) - \varepsilon_t) \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \varepsilon_t \left(\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \right) \right\| \\ &= \sup_{\mathcal{B}_0} w_{t-1}^{-2} \left\| (\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \varepsilon_t \frac{\partial^2 \varepsilon_t(\boldsymbol{\theta}^{**})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \right\| \\ &\leq w_{t-1}^{-2} \left\{ \left(\sup_{\mathcal{B}_0} \left\| \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\|^2 \right) \left(\sup_{\mathcal{B}_0} \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \right) \right. \\ &\quad \left. + \left(\sup_{\mathcal{B}_0} \left\| \frac{\partial^2 \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| \right) \left(\sup_{\mathcal{B}_0} |\varepsilon_t(\boldsymbol{\theta})| \right) \left(\sup_{\mathcal{B}_0} \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \right) \right\} \\ &\leq \frac{2C_0 C^2}{\sqrt{n}} w_{t-1}^{-2} \xi_{\rho,t-1}^2, \end{aligned}$$

其中 $\boldsymbol{\theta}^*$ 和 $\boldsymbol{\theta}^{**}$ 位于 $\boldsymbol{\theta}_0$ 和 $\boldsymbol{\theta}$ 中间。由假设 4.3 可得:

$$\frac{1}{n} \sum_{t=1}^n \sup_{\mathcal{B}_0} \|D_{t,1}(\boldsymbol{\theta}) - D_{t,1}(\boldsymbol{\theta}_0)\| = O_p\left(\frac{1}{\sqrt{n}}\right).$$

由引理 4.1, 我们有

$$\left| \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \right| \leq \sup_{\mathcal{B}_0} \left\| \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \frac{CC_0}{\sqrt{n}} \xi_{\rho,t-1}.$$

因此, 由泰勒展开和假设 4.3-4.4, 我们有

$$\begin{aligned} & E \sup_{\mathcal{B}_0} w_{t-1}^{-1} |\text{sgn}(\varepsilon_t(\boldsymbol{\theta})) - \text{sgn}(\varepsilon_t)| \\ &= E \sup_{\mathcal{B}_0} w_{t-1}^{-1} \left[|I(\varepsilon_t(\boldsymbol{\theta}) > 0) - I(\varepsilon_t > 0)| + |I(\varepsilon_t(\boldsymbol{\theta}) < 0) - I(\varepsilon_t < 0)| \right] \\ &= E \sup_{\mathcal{B}_0} w_{t-1}^{-1} \left[\left| I\left(\frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) < \varepsilon_t \leq 0 \right) - I\left(0 < \varepsilon_t \leq \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) \right) \right| \right. \\ &\quad \left. + \left| I\left(0 \leq \varepsilon_t < \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) \right) - I\left(\frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) \leq \varepsilon_t < 0 \right) \right| \right] \\ &\leq 2E \sup_{\mathcal{B}_0} w_{t-1}^{-1} I(|\varepsilon_t| \leq |(\partial \varepsilon_t(\boldsymbol{\theta}^*)/\partial \boldsymbol{\theta}')(\boldsymbol{\theta}_0 - \boldsymbol{\theta})|) \\ &= 2E \{ w_{t-1}^{-1} I(|\eta_t| \leq CC_0 n^{-\frac{1}{2}} \xi_{\rho,t-1}/\sigma_t) \} \\ &= \frac{2g(0)CC_0}{\sqrt{n}} E \left\{ w_{t-1}^{-1} \frac{\xi_{\rho,t-1}}{\sigma_t} \right\} + o\left(\frac{1}{\sqrt{n}}\right) = O\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \tag{4.6}$$

其中 θ^* 位於 θ_0 和 θ 之間。因此，由 (4.6)，我們有

$$\frac{1}{n} \sum_{t=1}^n \sup_{B_0} |D_{t,2}(\theta, 0) - D_{t,2}(\theta_0, 0)| = \frac{1}{n} \sum_{t=1}^n \sup_{B_0} w_{t-1}^{-1} |\text{sgn}(\varepsilon_t(\theta)) - \text{sgn}(\varepsilon_t)| = O_p\left(\frac{1}{\sqrt{n}}\right),$$

这样就证明了 (ii)。

(iii) 要证明 (iii)，只需要证明

$$\sup_{B_0} \left\| \frac{1}{n} \sum_{t=1}^n \{D_{t,1}(\theta) D_{t,2}(\theta, 0)\} - E\{D_{t,1}(\theta_0) D_{t,2}(\theta_0, 0)\} \right\| = o_p(1), \quad (4.7)$$

其余交叉乘积项的部分和的证明均类似。首先，

$$\begin{aligned} D_{t,1}(\theta) &= w_{t-1}^{-2} \varepsilon_t(\theta) \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \\ &= w_{t-1}^{-2} (\varepsilon_t(\theta) - \varepsilon_t) \frac{\partial \varepsilon_t(\theta)}{\partial \theta} + w_{t-1}^{-2} \varepsilon_t \left(\frac{\partial \varepsilon_t(\theta)}{\partial \theta} - \frac{\partial \varepsilon_t(\theta_0)}{\partial \theta} \right) + w_{t-1}^{-2} \varepsilon_t \frac{\partial \varepsilon_t(\theta_0)}{\partial \theta} \\ &=: L_1(t) + L_2(t) + L_3(t), \end{aligned}$$

和

$$\begin{aligned} D_{t,2}(\theta, 0) &= w_{t-1}^{-1} \text{sgn}(\varepsilon_t(\theta)) = w_{t-1}^{-1} \{ \text{sgn}(\varepsilon_t(\theta)) - \text{sgn}(\varepsilon_t) \} + w_{t-1}^{-1} \text{sgn}(\varepsilon_t) \\ &=: M_1(t) + M_2(t). \end{aligned}$$

由泰勒展开、(4.6) 和引理 4.1，我們有

$$\begin{aligned} \sup_{B_0} \|L_1(t) + L_2(t)\| &= \sup_{B_0} w_{t-1}^{-2} \left\| (\theta - \theta_0)' \frac{\partial \varepsilon_t(\theta^*)}{\partial \theta} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} + \varepsilon_t \frac{\partial^2 \varepsilon_t(\tilde{\theta})}{\partial \theta \partial \theta'} (\theta - \theta_0) \right\| \\ &\leq \frac{2C_0 C^2}{\sqrt{n}} w_{t-1}^{-2} \xi_{\rho, t-1}^2, \end{aligned} \quad (4.8)$$

其中 θ^* 和 $\tilde{\theta}$ 位於 θ_0 和 θ 之間。此外，

$$\|L_3(t)\| \leq \sup_{B_0} w_{t-1}^{-2} \left\| \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \right\| \|\varepsilon_t(\theta)\| \leq C^2 w_{t-1}^{-2} \xi_{\rho, t-1}^2,$$

和

$$\sup_{B_0} |M_1(t)| = \sup_{B_0} w_{t-1}^{-1} |\text{sgn}(\varepsilon_t(\theta)) - \text{sgn}(\varepsilon_t)| \leq 2w_{t-1}^{-1} I\left\{ |\eta_t| \leq CC_0 n^{-\frac{1}{2}} \frac{\xi_{\rho, t-1}}{\sigma_t} \right\}. \quad (4.9)$$

因此，由 (4.8)、(4.9)、(4.9) 和假设 4.3，我們可以证明：

$$\frac{1}{n} \sum_{t=1}^n \sup_{B_0} \left\{ |M_1(t)| (\|L_1(t) + L_2(t)\|) \right\} \xrightarrow{p} 0, \quad (4.10)$$

$$\frac{1}{n} \sum_{t=1}^n \sup_{B_0} \left\{ |M_1(t)| (\|L_3(t)\|) \right\} \xrightarrow{p} 0, \quad (4.11)$$

和

$$\frac{1}{n} \sum_{t=1}^n \sup_{\mathcal{B}_0} \left\{ |M_2(t)| (\|L_1(t) + L_2(t)\|) \right\} \xrightarrow{p} 0. \quad (4.12)$$

因此, 由 (4.10)、(4.11) 和 (4.12) 可得:

$$\begin{aligned} & \sup_{\mathcal{B}_0} \left\| \frac{1}{n} \sum_{t=1}^n \{D_{t,1}(\boldsymbol{\theta}) D_{t,2}(\boldsymbol{\theta}, 0)\} - E\{D_{t,1}(\boldsymbol{\theta}_0) D_{t,2}(\boldsymbol{\theta}_0, 0)\} \right\| \\ &= \sup_{\mathcal{B}_0} \left\| \frac{1}{n} \sum_{t=1}^n \{M_1(t)(L_1(t) + L_2(t))\} + \frac{1}{n} \sum_{t=1}^n \{M_2(t)(L_1(t) + L_2(t))\} \right. \\ & \quad \left. + \frac{1}{n} \sum_{t=1}^n \{M_1(t)L_3(t)\} + \frac{1}{n} \sum_{t=1}^n \{M_2(t)L_3(t)\} - E\{D_{t,1}(\boldsymbol{\theta}_0, 0) D_{t,2}(\boldsymbol{\theta}_0, 0)\} \right\| \\ &= \sup_{\mathcal{B}_0} \left\| \frac{1}{n} \sum_{t=1}^n \{M_2(t)L_3(t)\} - E\{D_{t,1}(\boldsymbol{\theta}_0, 0) D_{t,2}(\boldsymbol{\theta}_0, 0)\} \right\| + o_p(1) \\ &= o_p(1), \end{aligned}$$

即 (4.7) 成立。 $\square$

引理 4.3: 假设定理 4.1 的条件成立, 则当 $n \rightarrow \infty$ 时,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}_0, 0) \xrightarrow{d} N\left(\mathbf{0}, E\{D_1(\boldsymbol{\theta}_0, 0) D_1'(\boldsymbol{\theta}_0, 0)\}\right).$$

证明: 注意到 $\mathcal{F}_t$ 是由序列 $\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$ 生成的 σ 域, 则

$$E(D_{t,1}(\boldsymbol{\theta}_0) | \mathcal{F}_{t-1}) = E\left(w_{t-1}^{-2} \varepsilon_t \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \middle| \mathcal{F}_{t-1}\right) = w_{t-1}^{-2} \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} E(\varepsilon_t | \mathcal{F}_{t-1}) = 0,$$

和

$$E(D_{t,2}(\boldsymbol{\theta}_0, 0) | \mathcal{F}_{t-1}) = E\left(w_{t-1}^{-1} \text{sgn}(\varepsilon_t) | \mathcal{F}_{t-1}\right) = 0.$$

因此, $\{D_t(\boldsymbol{\theta}_0, 0)\}$ 是鞅差序列。

由引理 4.2 (i) 和 (iii), 我们有

$$\max_{1 \leq t \leq n} \left\| \frac{1}{\sqrt{n}} D_t(\boldsymbol{\theta}_0, 0) \right\| = o_p(1),$$

$$\frac{1}{n} \sum_{t=1}^n \{D_t(\boldsymbol{\theta}_0, 0) D_t'(\boldsymbol{\theta}_0, 0)\} = E\{D_1(\boldsymbol{\theta}_0, 0) D_1'(\boldsymbol{\theta}_0, 0)\} + o_p(1).$$

并且由控制收敛定理可得:

$$E \left\| \max_{1 \leq t \leq n} \frac{1}{n} D_t(\boldsymbol{\theta}_0, 0) D_t'(\boldsymbol{\theta}_0, 0) \right\| = o(1).$$

因此，鞅差序列的中心极限定理的条件满足。由鞅差序列的中心极限定理可知引理 4.3 成立（参考 Hall and Heyde (1980) 第 58 页）。□

引理 4.4: 在定理 4.1 的条件下，当 $n \rightarrow \infty$ 时，我们有

$$\sup_{B_0} \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n \{w_{t-1}^{-1} [\text{sgn}(\varepsilon_t(\boldsymbol{\theta})) - \text{sgn}(\varepsilon_t)]\} - 2g(0)E\left\{\frac{w_{t-1}^{-1}}{\sigma_t} \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right\} [\sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)] \right\| = o_p(1). \quad (4.13)$$

证明: 令

$$d_{1t}(\boldsymbol{\theta}) = w_{t-1}^{-1} \left\{ I\left(\frac{1}{\sigma_t} \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) < \eta_t\right) - I(\eta_t > 0) \right\},$$

和

$$d_{2t}(\boldsymbol{\theta}) = w_{t-1}^{-1} \left\{ I\left(\frac{1}{\sigma_t} \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) > \eta_t\right) - I(\eta_t < 0) \right\},$$

则我们有

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} [\text{sgn}(\varepsilon_t(\boldsymbol{\theta})) - \text{sgn}(\varepsilon_t)] = \frac{1}{\sqrt{n}} \sum_{t=1}^n [d_{1t}(\boldsymbol{\theta}) - d_{2t}(\boldsymbol{\theta})].$$

因此，我们只需要证明

$$\sup_{B_0} \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ [d_{1t}(\boldsymbol{\theta}) - d_{2t}(\boldsymbol{\theta})] - E[(d_{1t}(\boldsymbol{\theta}) - d_{2t}(\boldsymbol{\theta})) | \mathcal{F}_{t-1}] \right\} \right\| = o_p(1), \quad (4.14)$$

和

$$\sup_{B_0} \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n E[(d_{1t}(\boldsymbol{\theta}) - d_{2t}(\boldsymbol{\theta})) | \mathcal{F}_{t-1}] - 2g(0)E\left\{\frac{w_{t-1}^{-1}}{\sigma_t} \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right\} [\sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)] \right\| = o_p(1). \quad (4.15)$$

对于 (4.14), 我们只需要证明

$$\sup_{B_0} \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n \{d_{it}(\boldsymbol{\theta}) - E(d_{it}(\boldsymbol{\theta}) | \mathcal{F}_{t-1})\} \right\| =: \sup_{B_0} \|M_{n,i}(\boldsymbol{\theta})\| = o_p(1), \quad (4.16)$$

对 $i = 1, 2$ 成立。类似 Zhu and Ling (2011) 中引理 2.2 的证明，我们可以得出：对任意 $\eta > 0$,

$$\sup_{\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < \eta} \frac{\|M_{n,i}(\boldsymbol{\theta})\|}{1 + \sqrt{n}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|} = o_p(1), \quad (4.17)$$

对 $i = 1, 2$ 成立。当 $\boldsymbol{\theta} \in \mathcal{B}_0$ 时, 存在一个 $\eta_0 > 0$ 使得 $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq C_0/\sqrt{n} < \eta_0$ 。由 (4.17), 我们有

$$\sup_{\mathcal{B}_0} \|M_{n,i}(\boldsymbol{\theta})\| \leq (C_0 + 1) \sup_{\mathcal{B}_0} \frac{\|M_{n,i}(\boldsymbol{\theta})\|}{1 + \sqrt{n}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|} = o_p(1),$$

因此可得出 (4.16) 成立。

接下来, 我们证明 (4.15)。由泰勒展开和假设 4.5, 我们有

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{t=1}^n E[d_{1t}(\boldsymbol{\theta}) - d_{2t}(\boldsymbol{\theta}) | \mathcal{F}_{t-1}] \\ &= 2g(0) \left\{ \frac{1}{n} \sum_{t=1}^n \frac{w_{t-1}^{-1}}{\sigma_t} \frac{\partial \varepsilon_t(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} \right\} [\sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)] (1 + o_p(1)). \end{aligned} \quad (4.18)$$

由 (4.18), Ling and McAleer (2003) 中的定理 3.1 和控制收敛定理, 很容易可以得到 (4.15)。因此, 引理 4.4 成立。□

定理 4.1 的证明: 令 $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \frac{\mathbf{u}}{\sqrt{n}}$, 其中 $\mathbf{u}$ 是一个 $(r + s + 1)$ 维向量, 定理 4.2 中定义了 $\boldsymbol{\Lambda} = E\{\mathbf{D}_1(\boldsymbol{\theta}_0, 0)\mathbf{D}_1'(\boldsymbol{\theta}_0, 0)\}$ 。记

$$h(\boldsymbol{\theta}, 0, \boldsymbol{\lambda}) = \frac{1}{n} \sum_{t=1}^n \frac{\mathbf{D}_t(\boldsymbol{\theta}, 0)}{1 + \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, 0)},$$

和 $\delta_t(\boldsymbol{\theta}, 0) = \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, 0)$, 其中 $\boldsymbol{\lambda}$ 是等式 $h(\boldsymbol{\theta}, 0, \boldsymbol{\lambda}) = \mathbf{0}$ 的一个解。

首先, 我们证明 $\|\boldsymbol{\lambda}\| = O_p(1/\sqrt{n})$ 。记 $\boldsymbol{\lambda} = \varrho \mathbf{v}$, 其中 $\|\mathbf{v}\| = 1$, 并且定义 $Z_n(\boldsymbol{\theta}, 0) = \max_{1 \leq t \leq n} \|\mathbf{D}_t(\boldsymbol{\theta}, 0)\|$ 和 $\mathbf{T}_n(\boldsymbol{\theta}, 0) = \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0)\mathbf{D}_t'(\boldsymbol{\theta}, 0)$ 。因为

$$\begin{aligned} 0 &= \|h(\boldsymbol{\theta}, 0, \varrho \mathbf{v})\| \geq \|\mathbf{v}' h(\boldsymbol{\theta}, 0, \varrho \mathbf{v})\| = \left| \frac{1}{n} \sum_{t=1}^n \frac{\mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0)}{1 + \varrho \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0)} \right| \\ &= \left| \frac{1}{n} \sum_{t=1}^n \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0) - \frac{1}{n} \sum_{t=1}^n \frac{\varrho \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0) \mathbf{D}_t'(\boldsymbol{\theta}, 0) \mathbf{v}}{1 + \varrho \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0)} \right| \\ &\geq \frac{1}{n} \left| \sum_{t=1}^n \frac{\varrho \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0) \mathbf{D}_t'(\boldsymbol{\theta}, 0) \mathbf{v}}{1 + \varrho \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0)} \right| - \frac{1}{n} \left| \sum_{t=1}^n \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0) \right| \\ &\geq \frac{\varrho \mathbf{v}' \mathbf{T}_n(\boldsymbol{\theta}, 0) \mathbf{v}}{1 + \varrho Z_n(\boldsymbol{\theta}, 0)} - \frac{1}{n} \left| \sum_{t=1}^n \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0) \right|, \end{aligned}$$

由引理 4.2 (ii) 和引理 4.3 可得:

$$\begin{aligned} \frac{\varrho \mathbf{v}' \mathbf{T}_n(\boldsymbol{\theta}, 0) \mathbf{v}}{1 + \varrho Z_n(\boldsymbol{\theta}, 0)} &\leq \frac{1}{n} \left| \sum_{t=1}^n \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}, 0) \right| \\ &\leq \frac{1}{n} \left| \sum_{t=1}^n \mathbf{v}' \mathbf{D}_t(\boldsymbol{\theta}_0, 0) \right| + \frac{1}{n} \left| \sum_{t=1}^n \mathbf{v}' (\mathbf{D}_t(\boldsymbol{\theta}, 0) - \mathbf{D}_t(\boldsymbol{\theta}_0, 0)) \right| \\ &= O_p(1/\sqrt{n}), \end{aligned}$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。

令 a 为 $\boldsymbol{\Lambda}$ 最小的特征值。由引理 4.2 (iii), $\mathbf{v}' \mathbf{T}_n(\boldsymbol{\theta}, 0) \mathbf{v} \geq a + o_p(1)$ 对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。因此, 我们有

$$\varrho = \|\boldsymbol{\lambda}\| = O_p(1/\sqrt{n}), \quad (4.19)$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。

接下来, 我们将推导出 $\boldsymbol{\lambda}$ 的解析表达式, 即为 $\mathbf{D}_t(\boldsymbol{\theta}, 0)$ 的一个函数。由引理 4.2 (i) 和 (4.19) 可得:

$$\max_{1 \leq t \leq n} |\delta_t(\boldsymbol{\theta}, 0)| \leq \|\boldsymbol{\lambda}\| \max_{1 \leq t \leq n} \|\mathbf{D}_t(\boldsymbol{\theta}, 0)\| = O_p(1/\sqrt{n}) o_p(\sqrt{n}) = o_p(1), \quad (4.20)$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。由引理 4.2 (ii) 和引理 4.3, 我们有

$$\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) = \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}_0, 0) + O_p(1/\sqrt{n}) = O_p(1/\sqrt{n}) + O_p(1/\sqrt{n}) = O_p(1/\sqrt{n}),$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。由 (4.20) 可得

$$\begin{aligned} \mathbf{0} &= h(\boldsymbol{\theta}, 0, \boldsymbol{\lambda}) = \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \left(1 - \boldsymbol{\lambda}' \mathbf{D}_t(\boldsymbol{\theta}, 0) + \frac{\delta_t^2(\boldsymbol{\theta}, 0)}{1 + \delta_t(\boldsymbol{\theta}, 0)} \right) \\ &= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) - \mathbf{T}_n(\boldsymbol{\theta}, 0) \boldsymbol{\lambda} + \frac{1}{n} \sum_{t=1}^n \frac{\mathbf{D}_t(\boldsymbol{\theta}, 0) \delta_t^2(\boldsymbol{\theta}, 0)}{1 + \delta_t(\boldsymbol{\theta}, 0)} \\ &\leq \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) - \mathbf{T}_n(\boldsymbol{\theta}, 0) \boldsymbol{\lambda} + \frac{\max_{1 \leq t \leq n} |\delta_t(\boldsymbol{\theta}, 0)|^2}{1 - \max_{1 \leq t \leq n} |\delta_t(\boldsymbol{\theta}, 0)|} \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) \\ &= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) - \mathbf{T}_n(\boldsymbol{\theta}, 0) \boldsymbol{\lambda} + o_p(1) O_p(1/\sqrt{n}) \\ &= \frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) - \mathbf{T}_n(\boldsymbol{\theta}, 0) \boldsymbol{\lambda} + o_p(1/\sqrt{n}). \end{aligned}$$

由引理 4.2 (iii), $\mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) \geq C$ 以趋近于 1 的概率成立。因此,

$$\boldsymbol{\lambda} = \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) + \mathbf{R}_n,$$

其中 $\|\mathbf{R}_n\| = o_p(1/\sqrt{n})$ 。

其次, 我们推导出 $l(\boldsymbol{\theta}, 0)$ 的表达式, 即为 $\mathbf{D}_t(\boldsymbol{\theta}, 0)$ 的函数。由泰勒展开和 (4.20) 可知:

$$\begin{aligned}
l(\boldsymbol{\theta}, 0) &= -2 \ln L(\boldsymbol{\theta}, 0) \\
&= 2 \sum_{t=1}^n \ln(1 + \delta_t(\boldsymbol{\theta}, 0)) \\
&= 2\boldsymbol{\lambda}' \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) - n\boldsymbol{\lambda}' \mathbf{T}_n(\boldsymbol{\theta}, 0)\boldsymbol{\lambda} + 2 \sum_{t=1}^n U_t(\boldsymbol{\theta}, 0) \\
&= 2 \left\{ \left[\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right]' \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) + \mathbf{R}_n' \right\} \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) \\
&\quad - n \left\{ \left[\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right]' \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) + \mathbf{R}_n' \right\} \mathbf{T}_n(\boldsymbol{\theta}, 0) \times \\
&\quad \left\{ \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) \left[\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right] + \mathbf{R}_n \right\} + 2 \sum_{t=1}^n U_t(\boldsymbol{\theta}, 0) \\
&= 2n \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right)' \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) + 2\mathbf{R}_n' \left(\sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) \\
&\quad - n \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right)' \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) - n \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right)' \mathbf{R}_n \\
&\quad - n\mathbf{R}_n' \left(\frac{1}{n} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) - n\mathbf{R}_n' \mathbf{T}_n(\boldsymbol{\theta}, 0)\mathbf{R}_n + 2 \sum_{t=1}^n U_t(\boldsymbol{\theta}, 0) \\
&= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right)' \mathbf{T}_n^{-1}(\boldsymbol{\theta}, 0) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}, 0) \right) \\
&\quad - n\mathbf{R}_n' \mathbf{T}_n(\boldsymbol{\theta}, 0)\mathbf{R}_n + 2 \sum_{t=1}^n U_t(\boldsymbol{\theta}, 0),
\end{aligned}$$

其中 $U_t(\boldsymbol{\theta}, 0)$ 是残差项。因为 $\|\mathbf{R}_n\| = o_p(1/\sqrt{n})$, 我们有

$$n\mathbf{R}_n' \mathbf{T}_n(\boldsymbol{\theta}, 0)\mathbf{R}_n = o_p(1)。$$

对于残差项 $U_t(\boldsymbol{\theta}, 0)$, 我们将证明

$$2 \sum_{t=1}^n U_t(\boldsymbol{\theta}, 0) = o_p(1)。$$

由引理 4.2 (iii), 我们有

$$\begin{aligned}
\sum_{t=1}^n \|D_t(\boldsymbol{\theta}, 0)\|^3 &\leq \max_{1 \leq t \leq n} \|D_t(\boldsymbol{\theta}, 0)\| \left(\sum_{t=1}^n \|D_t(\boldsymbol{\theta}, 0)\|^2 \right) \\
&= n o_p(\sqrt{n}) \left(\frac{1}{n} \sum_{t=1}^n D_t'(\boldsymbol{\theta}, 0) D_t(\boldsymbol{\theta}, 0) \right) \\
&= n o_p(\sqrt{n}) \text{tr} \left(\frac{1}{n} \sum_{t=1}^n D_t'(\boldsymbol{\theta}, 0) D_t(\boldsymbol{\theta}, 0) \right) \\
&= n o_p(\sqrt{n}) \text{tr} \left(\frac{1}{n} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0) D_t'(\boldsymbol{\theta}, 0) \right) \\
&= n o_p(\sqrt{n}) O_p(1) = o_p(n^{\frac{3}{2}}),
\end{aligned}$$

其中 tr 为矩阵的迹。故由 $\|\boldsymbol{\lambda}\|^3 = O_p(n^{-\frac{3}{2}})$, 我们有

$$\begin{aligned}
|2 \sum_{t=1}^n U_t(\boldsymbol{\theta}, 0)| &\leq 2B \sum_{t=1}^n |\delta_t(\boldsymbol{\theta}, 0)|^3 \leq 2B \sum_{t=1}^n \|\boldsymbol{\lambda}\|^3 \|D_t(\boldsymbol{\theta}, 0)\|^3 \\
&= O_p(n^{-\frac{3}{2}}) o_p(n^{\frac{3}{2}}) = o_p(1).
\end{aligned}$$

因此, 由引理 4.2 (iii) 可得

$$\begin{aligned}
l(\boldsymbol{\theta}, 0) &= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0) \right)' T_n^{-1}(\boldsymbol{\theta}, 0) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0) \right) + o_p(1) \\
&= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0) \right)' \Lambda^{-1} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0) \right) + o_p(1),
\end{aligned}$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。特别地, 我们有

$$l(\boldsymbol{\theta}_0, 0) = \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}_0, 0) \right)' \Lambda^{-1} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}_0, 0) \right) + o_p(1).$$

有了 $l(\boldsymbol{\theta}, 0)$ 和 $l(\boldsymbol{\theta}_0, 0)$ 的表达式, 我们接下来就可以推导出使得 $l(\boldsymbol{\theta}, 0) - l(\boldsymbol{\theta}_0, 0)$ 达到最小值的点 $\hat{\boldsymbol{\theta}}$ 。

记 $\mathbf{S}_n(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n D_t(\boldsymbol{\theta}, 0)$, $l(\boldsymbol{\theta}, 0) - l(\boldsymbol{\theta}_0, 0)$ 可以被写为以下形式:

$$\begin{aligned}
&l(\boldsymbol{\theta}, 0) - l(\boldsymbol{\theta}_0, 0) \\
&= \mathbf{S}_n'(\boldsymbol{\theta}) \Lambda^{-1} \mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n'(\boldsymbol{\theta}_0) \Lambda^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0) + o_p(1) \\
&= (\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0))' \Lambda^{-1} \mathbf{S}_n(\boldsymbol{\theta}) + \mathbf{S}_n'(\boldsymbol{\theta}_0) \Lambda^{-1} \mathbf{S}_n(\boldsymbol{\theta}) \\
&\quad - \mathbf{S}_n'(\boldsymbol{\theta}_0) \Lambda^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0) + o_p(1) \\
&= (\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0))' \Lambda^{-1} (\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0)) + (\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0))' \Lambda^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0) \\
&\quad + \mathbf{S}_n'(\boldsymbol{\theta}_0) \Lambda^{-1} (\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0)) + o_p(1),
\end{aligned}$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立, 其中

$$\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left(\mathbf{D}'_{t,1}(\boldsymbol{\theta}) - \mathbf{D}'_{t,1}(\boldsymbol{\theta}_0), D_{t,2}(\boldsymbol{\theta}, 0) - D_{t,2}(\boldsymbol{\theta}_0, 0) \right)'.$$

下面我们推导 $\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0)$ 的表达式。

令 $\boldsymbol{\Sigma}_1 = E \left\{ \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right\} = E \left\{ w_{t-1}^{-2} \left(\frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} + \varepsilon_t \frac{\partial^2 \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right) \right\}$, 由泰勒展开和引理 4.2 (ii), 我们有

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{D}_{t,1}(\boldsymbol{\theta}) - \mathbf{D}_{t,1}(\boldsymbol{\theta}_0)) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) + o_p\left(\frac{1}{\sqrt{n}}\right) \right\} \\ &= \left(\frac{1}{n} \sum_{t=1}^n \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right) \mathbf{u} + o_p(1) \\ &= \boldsymbol{\Sigma}_1 \mathbf{u} + o_p(1), \end{aligned}$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。定义 $\boldsymbol{\Sigma}_2 = 2g(0)E \left\{ \frac{w_{t-1}^{-1}}{\sigma_t} \frac{\partial \varepsilon_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right\}$, 由引理 4.4, 我们有

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{t=1}^n (D_{t,2}(\boldsymbol{\theta}, 0) - D_{t,2}(\boldsymbol{\theta}_0, 0)) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} (\text{sgn}(\varepsilon_t(\boldsymbol{\theta})) - \text{sgn}(\varepsilon_t)) \\ &= \boldsymbol{\Sigma}_2 \mathbf{u} + o_p(1), \end{aligned}$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。若记 $\boldsymbol{\Gamma} = (\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$, 则:

$$\mathbf{S}_n(\boldsymbol{\theta}) - \mathbf{S}_n(\boldsymbol{\theta}_0) = (\boldsymbol{\Sigma}'_1, \boldsymbol{\Sigma}'_2)' \mathbf{u} + o_p(1) = \boldsymbol{\Gamma}' \mathbf{u} + o_p(1),$$

对 $\boldsymbol{\theta} \in \mathcal{B}_0$ 一致成立。

这样, $l(\boldsymbol{\theta}, 0) - l(\boldsymbol{\theta}_0, 0)$ 就可以写为:

$$l(\boldsymbol{\theta}, 0) - l(\boldsymbol{\theta}_0, 0) = \mathbf{u}' \boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \boldsymbol{\Gamma}' \mathbf{u} + \mathbf{u}' \boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0) + \mathbf{S}_n'(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}^{-1} \boldsymbol{\Gamma}' \mathbf{u} + o_p(1).$$

类似 Qin and Lawless (1994) 中引理 1 的证明, 我们知道使得上式达到最小值的点 $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \frac{\mathbf{u}}{\sqrt{n}}$ 位于 $\mathcal{B}_0$ 内, 且有以下的表达式:

$$\hat{\mathbf{u}} = -(\boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \boldsymbol{\Gamma}')^{-1} (\boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0)) + o_p(1).$$

相应地, 我们就可以写出 $l(\hat{\boldsymbol{\theta}}, 0)$:

$$\begin{aligned} l(\hat{\boldsymbol{\theta}}, 0) &= \mathbf{S}_n'(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0) - (\boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0))' (\boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \boldsymbol{\Gamma}')^{-1} (\boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \mathbf{S}_n(\boldsymbol{\theta}_0)) + o_p(1) \\ &= (\boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{S}_n(\boldsymbol{\theta}_0))' (\mathbf{I} - \boldsymbol{\Lambda}^{-\frac{1}{2}} \boldsymbol{\Gamma}' (\boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-1} \boldsymbol{\Gamma}')^{-1} \boldsymbol{\Gamma} \boldsymbol{\Lambda}^{-\frac{1}{2}}) (\boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{S}_n(\boldsymbol{\theta}_0)) + o_p(1). \end{aligned}$$

注意到引理 4.3 表明了 $\Lambda^{-\frac{1}{2}}\mathbf{S}_n(\boldsymbol{\theta}_0)$ 依分布收斂到多元標準正態分布；又有

$$\begin{aligned} \text{tr}(\mathbf{I} - \Lambda^{-\frac{1}{2}}\boldsymbol{\Gamma}'(\boldsymbol{\Gamma}\Lambda^{-1}\boldsymbol{\Gamma}')^{-1}\boldsymbol{\Gamma}\Lambda^{-\frac{1}{2}}) &= \text{tr}(\mathbf{I}) - \text{tr}(\Lambda^{-\frac{1}{2}}\boldsymbol{\Gamma}'(\boldsymbol{\Gamma}\Lambda^{-1}\boldsymbol{\Gamma}')^{-1}\boldsymbol{\Gamma}\Lambda^{-\frac{1}{2}}) \\ &= (r + s + 2) - (r + s + 1) = 1. \end{aligned} \quad (4.21)$$

因此，當 $n \rightarrow \infty$ 時， $l(\hat{\boldsymbol{\theta}}, 0) \xrightarrow{d} \chi_1^2$ 。 $\square$

定理 4.2 的证明： 类似定理 4.1 的证明，我們可以寫出 $l(\hat{\boldsymbol{\theta}}, 0)$ 的表达式：

$$l(\hat{\boldsymbol{\theta}}, 0) = (\Lambda^{-\frac{1}{2}}\mathbf{S}_n^*(\boldsymbol{\theta}_0))'(\mathbf{I} - \Lambda^{-\frac{1}{2}}\boldsymbol{\Gamma}'(\boldsymbol{\Gamma}\Lambda^{-1}\boldsymbol{\Gamma}')^{-1}\boldsymbol{\Gamma}\Lambda^{-\frac{1}{2}})(\Lambda^{-\frac{1}{2}}\mathbf{S}_n^*(\boldsymbol{\theta}_0)) + o_p(1), \quad (4.22)$$

其中， $\mathbf{S}_n^*(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t(\boldsymbol{\theta}_0, 0)$ 。若記 $D_{t,2}^* = w_{t-1}^{-1}\{\text{sgn}(\varepsilon_t) - E(\text{sgn}(\varepsilon_t))\}$ 和 $\mathbf{D}_t^* = (\mathbf{D}_{t,1}'(\boldsymbol{\theta}_0), D_{t,2}^*)'$ ，則 $\mathbf{S}_n^*(\boldsymbol{\theta}_0)$ 就可以寫為：

$$\begin{aligned} &\mathbf{S}_n^*(\boldsymbol{\theta}_0) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{D}_{t,1}'(\boldsymbol{\theta}_0), D_{t,2}(\boldsymbol{\theta}_0, 0))' \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{D}_{t,1}'(\boldsymbol{\theta}_0), w_{t-1}^{-1}\{\text{sgn}(\varepsilon_t) - E(\text{sgn}(\varepsilon_t)|\mathcal{F}_{t-1})\})' \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}E((\text{sgn}(\varepsilon_t) - \text{sgn}(\varepsilon_t - d_0))|\mathcal{F}_{t-1}))' \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}E(\text{sgn}(\varepsilon_t - d_0)|\mathcal{F}_{t-1}))' \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t^* + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}E\{(\text{sgn}(\eta_t) - \text{sgn}(\eta_t - d_0/\sigma_t))|\mathcal{F}_{t-1}\})' \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}\{E(\text{sgn}(\eta_t - d_0/\sigma_t)|\mathcal{F}_{t-1}) - E(\text{sgn}(\eta_t - d_0/\sigma_t))\})' \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t^* + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', 2w_{t-1}^{-1}\{F_\eta(d_0/\sigma_t) - F_\eta(0)\})' \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', 2w_{t-1}^{-1}\{F_\eta(d_0/\sigma_t) - E(F_\eta(d_0/\sigma_t))\})', \end{aligned} \quad (4.23)$$

其中， $\mathbf{0}$ 是一個 $(r + s + 1)$ 維向量， F_η 是 η_t 的分布。

由 $E\{w_{t-1}^{-1}[\text{sgn}(\varepsilon_t) - E(\text{sgn}(\varepsilon_t)|\mathcal{F}_{t-1})]|\mathcal{F}_{t-1}\} = 0$ 可知， $\{\mathbf{D}_t^*\}$ 是一個鞅差序列。类似引理 4.3，當 $d_0 \rightarrow 0$ 時，我們可以證明

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{D}_t^* \xrightarrow{d} N\left(\mathbf{0}, E\{\mathbf{D}_1(\boldsymbol{\theta}_0, 0)\mathbf{D}_1'(\boldsymbol{\theta}_0, 0)\}\right)。$$

此外, 因为 $d_0/\sigma_t \leq M/(n\omega_1)^{1/2}$ 对所有 $1 \leq t \leq n$ 一致成立, 则由泰勒展开可得:

$$\left| \frac{1}{\sqrt{n}} \sum_{t=1}^n \{w_{t-1}^{-1}(F_\eta(d_0/\sigma_t) - F_\eta(0))\} - g(0)M_1E\{w_{t-1}^{-1}\sigma_t^{-1}\} \right| = o_p(1),$$

和

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \{F_\eta(d_0/\sigma_t) - E(F_\eta(d_0/\sigma_t))\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} g(0) [(d_0/\sigma_t) - E(d_0/\sigma_t)] (1 + o_p(1)) \\ &= \frac{M_1 g(0)}{n} \sum_{t=1}^n w_{t-1}^{-1} (\sigma_t^{-1} - E(\sigma_t^{-1})) (1 + o_p(1)) \\ &= M_1 g(0) E\{w_{t-1}^{-1} (\sigma_t^{-1} - E(\sigma_t^{-1}))\} + o_p(1), \end{aligned}$$

其中最后一个等式可以由遍历性得出。故 $\mathbf{\Lambda}^{-\frac{1}{2}} \mathbf{S}_n^*(\boldsymbol{\theta}_0)$ 收敛到一个多维正态分布, 均值为 $\mathbf{\Lambda}^{-1/2}(\mathbf{0}', 2g(0)M_1E[w_{t-1}^{-1}E(\sigma_t^{-1})])'$, 协方差为 $\mathbf{I}$ 。最后, 利用 (4.22), 我们可以得到 $l(\hat{\boldsymbol{\theta}}, 0)$ 依分布收敛到自由度为 1 的非中心卡方分布, 非中心参数为

$$(\mathbf{0}', 2g(0)M_1E(w_{t-1}^{-1})E(\sigma_t^{-1}))\mathbf{\Lambda}^{-1}(\mathbf{0}', 2g(0)M_1E(w_{t-1}^{-1})E(\sigma_t^{-1}))'.$$

□

5 基于随机加权自助法的 ARMA-GARCH 模型误差零均值检验

5.1 引言

本章我们考虑经典的 ARMA(r, s)-GARCH(p, q) 模型:

$$\begin{cases} X_t = \mu + \sum_{i=1}^r \phi_i X_{t-i} + \sum_{j=1}^s \psi_j \varepsilon_{t-j} + \varepsilon_t, \\ \varepsilon_t = \sigma_t \eta_t, \sigma_t^2 = \omega + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \end{cases} \quad (5.1)$$

其中 $\mu \in \mathbb{R}$, $\phi_i \in \mathbb{R}$, $i = 1, \dots, r$, $\psi_j \in \mathbb{R}$, $j = 1, \dots, s$, $\omega > 0$, $a_i \geq 0$, $i = 1, 2, \dots, p$, $b_j \geq 0$, $j = 1, 2, \dots, q$, $\{\eta_t\}$ 是一个 i.i.d. 的随机变量, 均值为 0, 方差为 1。令 $\mathcal{F}_t$ 表示由 $\{\eta_s : s \leq t\}$ 生成的 σ 域。当 $\{w_t > 0\}$ 平稳、 w_t 是 $\mathcal{F}_t$ 可测且 $E(\sigma_t/w_{t-1}) \in (0, \infty)$ 时, 原假设 $H_0 : E(\eta_t) = 0$ 相当于 $H_0 : E(\varepsilon_t/w_{t-1}) = 0$, 因此我们可以考虑在不估计 GARCH 模型且无需 $E(\varepsilon_t^4) < \infty$ 的情况下, 检验 $H_0 : E(\varepsilon_t/w_{t-1}) = 0$ 。和第 3 章中 ARMA 和 GARCH 部分都需要估计的经验似然检验相比, 本章所提出的检验更加稳健。

本章首先考虑了 Zhu and Ling (2015) 中的基于中位数的统计推断方法来估计 ARMA 模型, 当 $E(|\varepsilon_t|^\delta) < \infty$ 对某个 $\delta > 0$ 成立时, 估计量有正态极限。接下来我们用上一步估计的 ε_t 的样本均值来检验 $H_0 : E(\varepsilon_t/w_{t-1}) = 0$ 。然而, 一些模拟结果表明第 3-4 章所用的 profile 经验似然检验在这种情况下的有限样本表现并不理想, 因为使用基于中位数的推断使得 profile 经验似然方法的计算更加复杂。因此, 我们提出使用 Zhu (2016) 提出的随机加权自助法来进行零均值检验。

本章的结构如下。第 5.2 节我们给出了估计方法和渐近理论。第 5.3 和 5.4 节分别给出了模拟研究和实证分析。第 5.5 节进行了总结。证明将在第 5.6 节给出。

5.2 方法和理论结果

考虑模型 (5.1) 中的 ARMA(r, s)-GARCH(p, q) 模型, 其中 η_t 的中位数为 0 而不是均值为 0。令 $\boldsymbol{\theta} = (\mu, \phi_1, \dots, \phi_r, \psi_1, \dots, \psi_s)'$, $\boldsymbol{\delta} = (\omega, a_1, \dots, a_p, b_1, \dots, b_q)'$ 和 $\boldsymbol{\gamma} = (\boldsymbol{\theta}', \boldsymbol{\delta}')$, 其中 $\boldsymbol{\gamma}_0$ 表示真值。为了方便, 我们记 $\eta_t(\boldsymbol{\gamma}_0) = \eta_t$, $\varepsilon_t(\boldsymbol{\theta}_0) = \varepsilon_t$, $\sigma_t(\boldsymbol{\gamma}_0) = \sigma_t$, 并且定义 $\phi(z) = 1 - \sum_{i=1}^r \phi_i z^i$, $\psi(z) = 1 + \sum_{j=1}^s \psi_j z^j$, $a(z) = \sum_{i=1}^p a_i z^i$ 和 $b(z) = 1 - \sum_{j=1}^q b_j z^j$ 。

给定观测值 $\{X_1, \dots, X_n\}$ 和初值 $\{X_0, X_{-1}, \dots\}$, 我们将 (5.1) 写为参数形式:

$$\begin{cases} \varepsilon_t(\boldsymbol{\theta}) = X_t - \mu - \sum_{i=1}^r \phi_i X_{t-i} - \sum_{j=1}^s \psi_j \varepsilon_{t-j}(\boldsymbol{\theta}), \\ \sigma_t^2(\boldsymbol{\gamma}) = \omega + \sum_{i=1}^p a_i \varepsilon_{t-i}^2(\boldsymbol{\theta}) + \sum_{j=1}^q b_j \sigma_{t-j}^2(\boldsymbol{\gamma}), \quad \eta_t(\boldsymbol{\gamma}) = \frac{\varepsilon_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\gamma})}. \end{cases}$$

本章的目标是检验

$$H_0 : E(\varepsilon_t/w_{t-1}) = 0 \quad v.s. \quad H_a : E(\varepsilon_t/w_{t-1}) \neq 0,$$

相当于检验

$$H_0 : E(\eta_t) = 0 \quad v.s. \quad H_a : E(\eta_t) \neq 0,$$

其中 $\{w_t = w(X_t, X_{t-1}, \dots) > 0\}$ 将在下文定义。

首先我们在零中位数的假设下估计 ε_t 。我们先用 Zhu and Ling (2015) 中的加权 LAD 方法来估计参数 θ ，即

$$\hat{\theta} = \arg \min_{\theta} \sum_{t=1}^n w_{t-1}^{-1} |\varepsilon_t(\theta)|.$$

估计量 $\hat{\theta}$ 当 $E(|\varepsilon_t|^\delta) < \infty$ 对某个 $\delta > 0$ 成立时渐近收敛于正态分布。接下来，我们用 $\varepsilon_t(\hat{\theta})$ 估计 ε_t ，其中 $t = 1, \dots, n$ 。

为了检验 H_0 ，我们用

$$\hat{\nu} = \frac{1}{n} \sum_{t=1}^n w_{t-1}^{-1} \varepsilon_t(\hat{\theta})$$

来估计 $\nu = E(\varepsilon_t/w_{t-1})$ 。为了避免估计复杂的渐近方差，我们采用了 Zhu (2016) 的随机加权自助抽样法：

- **步骤 1):** 从一个均值和方差均为 1 的分布（这里我们使用标准指数分布）中生成样本量为 n 的随机样本，记为 $\{\delta_t^b\}_{t=1}^n$ 。

- **步骤 2):** 计算

$$\hat{\theta}^b = \arg \min_{\theta} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} |\varepsilon_t(\theta)|,$$

和

$$\hat{\nu}^b = \frac{\sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \varepsilon_t(\hat{\theta}^b)}{\sum_{t=1}^n \delta_t^b}.$$

- **步骤 3):** 重复上面两个步骤 B 次，计算 $\{\hat{\nu}^b\}_{b=1}^B$ 。

因此，若

$$\hat{\nu}^2 / \left\{ \frac{1}{B} \sum_{b=1}^B (\hat{\nu}^b - \hat{\nu})^2 \right\} \geq \chi_{1,1-a}^2,$$

那么我们在显著性水平 a 上拒绝原假设 $H_0 : \nu = 0$ ，其中 $\chi_{1,1-a}^2$ 为自由度为 1 的卡方分布的 $(1-a)$ 分位数。

为了建立上述检验的理论性质，我们引入以下的常规假设。

假设 5.1: θ_0 为 Θ 的一个内点。对每个 $\theta \in \Theta$, 当 $|z| \leq 1$ 时, $\phi(z) \neq 0$, $\psi(z) \neq 0$, 并且 $\phi(z)$ 和 $\psi(z)$ 无共同根, $\phi_r \neq 0$ 或 $\psi_s \neq 0$ 。

假设 5.2: $\omega > 0$, 如第 4 章中所定义的 Lyapunov 指数 $\beta < 0$, 并且 $E(\ln(\max(|\eta_1|, 1))) < \infty$ 。

假设 5.3: 对任意 $\rho \in (0, 1)$, $E(w_{t-1}^{-4} \xi_{\rho,t-1}^4) < \infty$, 其中 $\xi_{\rho,t} = 1 + \sum_{i=0}^{\infty} \rho^i |X_{t-i}|$, $\{w_t = w(X_t, X_{t-1}, \dots)\}$ 是一个满足 $\inf_{t \geq 1} w_t > c_0 > 0$ 的平稳序列, 并且 w_t 为 $\mathcal{F}_t$ 可测。

假设 5.4: $\{\eta_t\}$ 为 i.i.d. 随机变量序列, 中位数为 0, $E(\eta_t^2) = 1$ 。

假设 5.5: $\{\eta_t\}$ 有连续密度函数 $g(x)$, 满足 $g(0) > 0$ 和 $\sup_{x \in \mathbb{R}} g(x) < \infty$ 。

假设 5.1 和 5.2 保证了模型 (5.1) 的第一个和第二个等式均存在唯一严平稳因果解。假设 5.3 的权重 w_t 用来减少 σ_t 的作用, 我们将在下文中给出其定义。假设 5.4 和 5.5 为对 ARMA 模型使用基于中位数的估计方法的常规假设。

定理 5.1: 在假设 5.1-5.5 和原假设 $H_0: \nu = 0$ 下, 当 $B \rightarrow \infty$ 和 $n \rightarrow \infty$ 时,

$$\sqrt{n}\hat{\nu} \xrightarrow{d} N(0, \sigma^2), \quad \frac{n}{B} \sum_{b=1}^B (\hat{\nu}^b - \hat{\nu})^2 / \sigma^2 \xrightarrow{p} 1, \quad (5.2)$$

其中

$$\begin{aligned} \sigma^2 &= (-\Gamma(2g(0)\Sigma)^{-1}, 1)E[\tilde{D}_1 \tilde{D}_1'](-\Gamma(2g(0)\Sigma)^{-1}, 1)', \\ \Gamma &= E\left\{w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta'} \middle|_{\theta=\theta_0}\right\}, \quad \Sigma = E\left[(w_{t-1}\sigma_t)^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \frac{\partial \varepsilon_t(\theta)}{\partial \theta'} \middle|_{\theta=\theta_0}\right], \\ \tilde{D}_t &= (w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta'} \middle|_{\theta=\theta_0} \text{sgn}(\varepsilon_t), w_{t-1}^{-1} \varepsilon_t)'. \end{aligned}$$

上述定理证明了我们提出的检验有精确的渐近 size。为了研究检验的 power, 我们考虑如下的局部备择假设:

$$H_a: \nu = \frac{M}{\sqrt{n}}, \quad M \text{ 为常数}. \quad (5.3)$$

下面的定理 5.2 表明, 当 $|M| \rightarrow \infty$ 时, 检验的 power 趋于 1。

定理 5.2: 若对模型 (5.1) 有假设 5.1-5.5 成立, 那么在备择假设 (5.3) 下, 当 $n \rightarrow \infty$ 和 $B \rightarrow \infty$ 时,

$$\frac{\sqrt{n}\hat{\nu}}{\sqrt{nB^{-1} \sum_{t=1}^B (\hat{\nu}^b - \hat{\nu})^2}} \xrightarrow{d} N(M/\sigma, 1),$$

其中 σ 如定理 5.1 中定義。

在此，我們建議採用和第 4 章中一樣的權重函數：

$$w_t(h) = \max(C, \sum_{i=0}^{t-1} e^{\log(h) \log^2(i+1)} |X_{t-i}|) \text{ 其中 } h \in (0, 1), t = 1, \dots, n, \quad (5.4)$$

其中 C 為 $|X_t|$ 的 90% 分位數，且 $w_0(h) = 1$ 。

5.3 模擬研究

本節我們研究了本章中的檢驗在有限樣本下的經驗 size 和 power。我們從 ARM-A(1,0)-GARCH(1,1) 模型和 ARMA(1,1)-GARCH(1,1) 模型中生成了 5000 個隨機樣本，樣本量 $n = 1000$ 和 2500。參數 $\mu = 0.1, \phi_1 = 0.5, \psi_1 = 0.2, \omega = 0.1, a_1 = 0.1, b_1 = 0.8, \eta_t = V/\sqrt{E(V^2)}$ ，其中

$$V = I(U < 1/2)V_1 - I(U \geq 1/2)V_2,$$

$U \sim \text{Uniform}(0, 1), V_1 \sim \text{Pareto}(1, \alpha_1)$ （即當 $x \geq 0$ 時 $P(V_1 \leq x) = 1 - (1+x)^{-\alpha_1}$ ）和 $V_2 \sim \text{Pareto}(1, \alpha_2)$ 相互獨立。顯然，

$$E(V) = \frac{1}{2} \left\{ \frac{1}{\alpha_1 - 1} - \frac{1}{\alpha_2 - 1} \right\}, E(V^2) = \frac{1}{(\alpha_1 - 1)(\alpha_1 - 2)} + \frac{1}{(\alpha_2 - 1)(\alpha_2 - 2)},$$

η_t 有零均值、右尾指數 α_1 、左尾指數 α_2 ；若 $\alpha_1 = \alpha_2$ 則有零均值，若 $\alpha_1 \neq \alpha_2$ 則有非零均值。

我們在 $\alpha_1 = \alpha_2 = 3$ 下計算檢驗的 size；在 $\alpha_1 = 3.2$ 或 3.5、 $\alpha_2 = 3$ 下計算檢驗的 power。在隨機加權自助法中，我們取 $B = 1000$ ，對於 (5.4) 中的權重函數 $w_t(h)$ ，取 $h = 0.2$ 和 0.4。表 5.1 和表 5.2 分別給出了檢驗在顯著性水平 10% 和 5% 下的經驗 size 和經驗 power，可以看出我們的檢驗有精確的 size 和良好的 power，使用 $h = 0.2$ 和 0.4 都有穩健的結果，當 n 變大或 α_1 更偏离原假設時，檢驗的 power 也隨之提升。

表 5.1 ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型檢驗的 size

a_1	α_1	n	Level	ARMA(1,0)-GARCH(1,1)		ARMA(1,1)-GARCH(1,1)	
				$h = 0.2$	$h = 0.4$	$h = 0.2$	$h = 0.4$
0.1	3	1000	10%	0.0980	0.0990	0.1002	0.1008
			5%	0.0442	0.0434	0.0468	0.0446
		2500	10%	0.0976	0.0990	0.0996	0.1010
			5%	0.0472	0.0476	0.0472	0.0498

表 5.2 ARMA(1,0)-GARCH(1,1) 和 ARMA(1,1)-GARCH(1,1) 模型检验的 power

a_1	α_1	n	Level	ARMA(1,0)-GARCH(1,1)		ARMA(1,1)-GARCH(1,1)	
				$h = 0.2$	$h = 0.4$	$h = 0.2$	$h = 0.4$
0.1	3.2	1000	10%	0.2244	0.2168	0.2182	0.2140
			5%	0.1386	0.1338	0.1292	0.1292
		2500	10%	0.3672	0.3648	0.3676	0.3668
			5%	0.2500	0.2526	0.2546	0.2570
			5%				
0.1	3.5	1000	10%	0.6568	0.6508	0.6490	0.6484
			5%	0.5188	0.5242	0.5200	0.5180
		2500	10%	0.9334	0.9322	0.9330	0.9324
			5%	0.8842	0.8862	0.8842	0.8842
			5%				

5.4 实证分析：汇率序列

本节研究了从 2011 年 5 月 3 日至 2021 年 5 月 2 日六种外汇的每日对数收益率 ($\times 100$)，这六种外汇分别是：HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 INR/USD。在图 5.1 中我们画出了这些外汇的时间序列图。

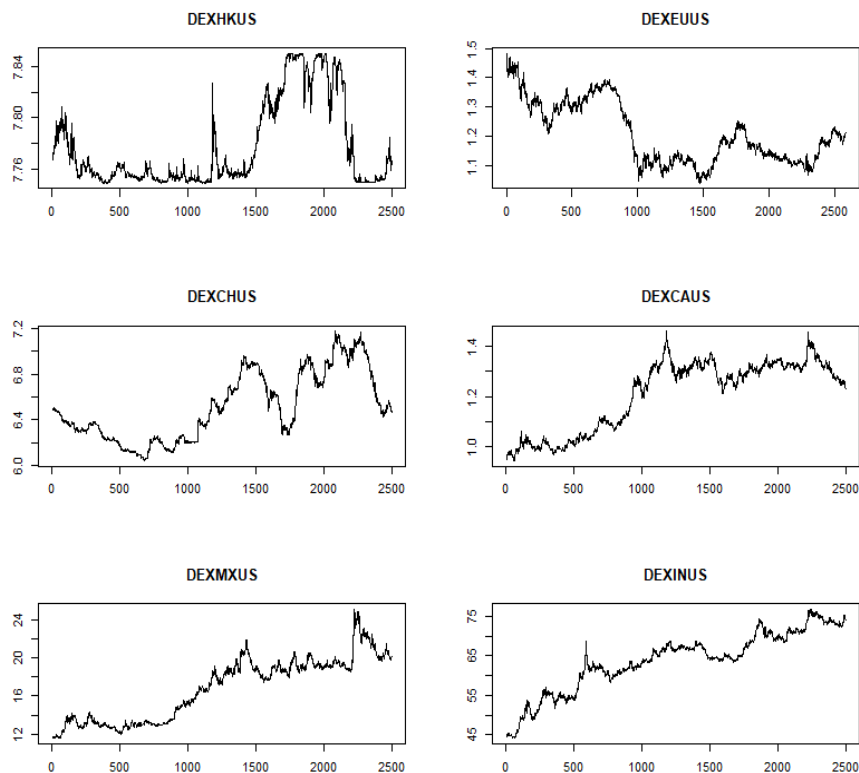


图 5.1 2011 年 5 月 3 日到 2021 年 5 月 2HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 INR/USD 外汇的时间序列图

我們首先用 Hill (1975) 提出的 Hill 估計量估計了 $\{|X_t|\}$ 的尾指數，定義為：

$$\hat{\alpha}(k) = \left[\frac{1}{k} \sum_{i=1}^k \log \left(\frac{X_{(n-i+1)}}{X_{(n-k)}} \right) \right]^{-1},$$

其中 $\{X_{(t)}\}$ 為 $\{|X_t|\}$ 的升序統計量。圖 5.2 中給出了 Hill 估計量和不同的 k 之間的關係圖。我們可以看出，除了 CAD/USD，其餘 5 個我們所考慮的外匯的對數收益率均有小於 4 的尾指數，即 $EX_t^4 = \infty$ 。因此，當我們用 QMLE 擬合 ARMA-GARCH 模型時，其估計量可能沒有正態極限。為了探索使用 SWQMELE 來擬合 ARMA-GARCH 模型的可能性，我們需檢驗在使用 LAD 之後 GARCH 模型是否有零均值。

為了應用我們提出的檢驗，我們首先使用 R 軟件包 “forecast” 中的 “auto.arima” 函數和 AIC 準則來選擇 ARMA 模型的階數，其結果呈現在表 5.3 中。接着，我們用 (5.4) 中的權重函數 $w_t(h)$ 來計算檢驗的 p 值，其中 $h = 0.2$ 和 0.4 。對於隨機加權自助抽樣法，我們取 $B = 5000$ 。表 5.3 表明，我們傾向於對 INR/USD 和 MXN/USD 外匯拒絕原假設，而對其他外匯傾向於接收原假設。因此，在使用 SWQMELE 對 INR/USD 和 MXN/USD 外匯的對數收益率擬合 ARMA-GARCH 模型時應該謹慎。

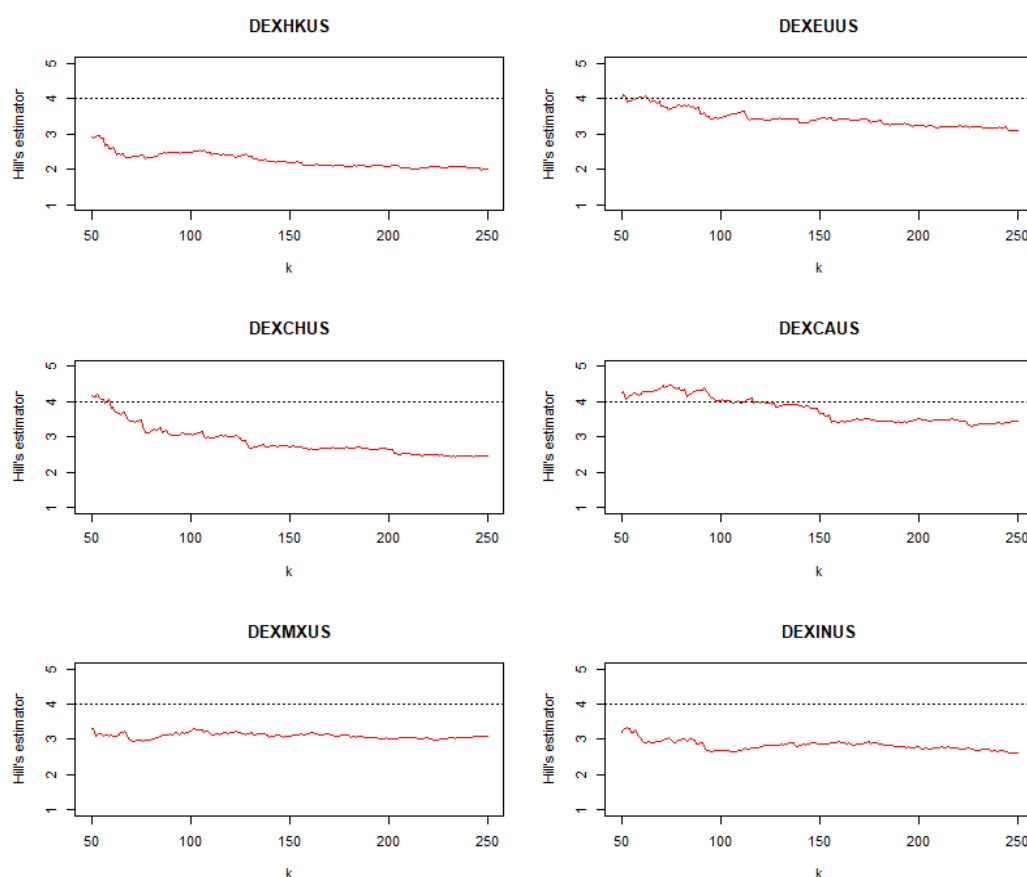


圖 5.2 2011 年 5 月 3 日到 2021 年 5 月 2 日 HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 INR/USD 外匯的 Hill 估計量

表 5.3 对 2011 年 5 月 3 日到 2021 年 5 月 2 日 HKD/USD、EUR/USD、CNY/USD、CAD/USD、MXN/USD 和 INR/USD 外汇的对数收益率拟合的 ARMA 模型和检验的 p 值

外汇	ARMA 模型	$h = 0.2$	$h = 0.4$
HKD/USD	ARMA(1,3)	0.3931	0.2159
EUR/USD	ARMA(0,1)	0.5625	0.4433
CNY/USD	ARMA(1,2)	0.9138	0.9024
CAD/USD	ARMA(2,0)	0.2306	0.1622
MXN/USD	ARMA(0,1)	0.0932	0.1044
INR/USD	ARMA(3,1)	0.0025	0.0009

5.5 结论

重尾现象使得基于中位数的统计推断在对金融收益率拟合 ARMA-GARCH 模型时很受欢迎。为了确保在使用基于中位数的推断后，所使用的 ARMA-GARCH 模型仍然刻画了序列的条件均值，我们提出使用随机加权自助抽样法对误差的零均值进行检验。我们所提出的检验对异方差和重尾有稳健的特性，因为它不需要估计模型的 GARCH 部分，并且只需要较宽松的矩条件。本章的模拟研究也证实了在有限样本下检验的 size 和 power 均有良好的表现。实证分析表明，在使用基于中位数的推断来拟合 INR/USA 和 MXN/USD 外汇的对数收益率时进行检验很有必要。

5.6 定理 5.1 和 5.2 的证明

定义

$$\tilde{D}_{t,1} = w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \text{sgn}(\varepsilon_t), \quad \tilde{D}_{t,2} = w_{t-1}^{-1} \varepsilon_t, \quad \tilde{\boldsymbol{D}}_t = (\tilde{D}_{t,1}', \tilde{D}_{t,2}')'.$$

首先，我们需要以下两个引理。

引理 5.1: 在定理 5.1 的条件下，存在常数 $\rho \in (0, 1)$ 、常数 $C > 0$ 和 $\boldsymbol{\theta}_0$ 的邻域 Θ_0 满足

$$\sup_{\Theta_0} |\varepsilon_t(\boldsymbol{\theta})| \leq C \xi_{\rho,t-1}, \quad \sup_{\Theta_0} \left\| \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| \leq C \xi_{\rho,t-1} \quad \text{和} \quad \sup_{\Theta_0} \left\| \frac{\partial^2 \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| \leq C \xi_{\rho,t-1},$$

其中 $\xi_{\rho,t}$ 如假设 5.3 中定义。

证明: 参考 Ling (2007) 的引理 A.1。 □

引理 5.2: 在定理 5.1 的条件下，当 $n \rightarrow \infty$ 时，

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \tilde{\boldsymbol{D}}_t \xrightarrow{d} N\left(0, E\{\tilde{\boldsymbol{D}}_1 \tilde{\boldsymbol{D}}_1'\}\right).$$

证明： 因为 $\mathcal{F}_t$ 是序列 $\{\eta_t, \eta_{t-1}, \dots\}$ 生成的 σ 域，显然可以验证：

$$E(\tilde{D}_{t,1}|\mathcal{F}_{t-1}) = E\left(w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \text{sgn}(\varepsilon_t) \Big| \mathcal{F}_{t-1}\right) = w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} E(\text{sgn}(\eta_t)) = 0,$$

和

$$E(\tilde{D}_{t,2}|\mathcal{F}_{t-1}) = E(w_{t-1}^{-1} \varepsilon_t | \mathcal{F}_{t-1}) = w_{t-1}^{-1} \sigma_t E(\eta_t) = 0,$$

即 $\{\tilde{D}_t\}$ 是一个鞅差序列。由假设 5.3、引理 5.1 和控制收敛定理可得：

$$\max_{1 \leq t \leq n} \left\| \frac{1}{\sqrt{n}} \tilde{D}_t \right\| = o_p(1), \quad \frac{1}{n} \sum_{t=1}^n \{\tilde{D}_t \tilde{D}_t'\} = E\{\tilde{D}_1 \tilde{D}_1'\} + o_p(1), \quad E \left\| \max_{1 \leq t \leq n} \frac{1}{n} \tilde{D}_t \tilde{D}_t' \right\| = o(1).$$

因此，鞅差序列的中心极限定理的条件满足（参考 Hall and Heyde (1980) 的定理 3.2），引理 5.2 成立。 $\square$

定理 5.1 的证明： 由 Zhu and Ling (2015) 的定理 5.2 可得：

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = -\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \text{sgn}(\varepsilon_t) + o_p(1), \quad (5.5)$$

其中 $\boldsymbol{\Sigma}$ 如定理 5.1 中定义。由泰勒展开、引理 5.1 和 (5.5)，我们有

$$\begin{aligned} \sqrt{n}\hat{\nu} &= \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \varepsilon_t(\hat{\boldsymbol{\theta}}) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} [\varepsilon_t(\hat{\boldsymbol{\theta}}) - \varepsilon_t] + \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \varepsilon_t \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \varepsilon_t + o_p(1), \\ &= -\boldsymbol{\Gamma} \frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \text{sgn}(\varepsilon_t) + \frac{1}{\sqrt{n}} \sum_{t=1}^n w_{t-1}^{-1} \varepsilon_t + o_p(1) \\ &\xrightarrow{d} N\left(0, (-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1) E[\tilde{D}_1 \tilde{D}_1'] (-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1)'\right), \end{aligned} \quad (5.6)$$

其中 $\boldsymbol{\Gamma} = E\left\{w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right\}$ 如定理 5.1 中定义。因此，(5.2) 的第一个等式成立。

类似 Zhu and Ling (2015) 定理 2 的证明，我们有

$$\sqrt{n}(\hat{\boldsymbol{\theta}}^b - \boldsymbol{\theta}_0) = -\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \text{sgn}(\varepsilon_t) + o_p(1). \quad (5.7)$$

类似 (5.6) 的证明, 并且由 (5.7), 我们可以得到:

$$\begin{aligned}
\sqrt{n}\hat{\nu}^b &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \varepsilon_t(\hat{\theta}^b) + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta'} \Big|_{\theta=\theta_0} (\hat{\theta}^b - \theta_0) + \frac{1}{\sqrt{n}} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \varepsilon_t + o_p(1), \\
&= -\Gamma \frac{(2g(0)\Sigma)^{-1}}{\sqrt{n}} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} \text{sgn}(\varepsilon_t) + \frac{1}{\sqrt{n}} \sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \varepsilon_t + o_p(1).
\end{aligned} \tag{5.8}$$

结合 (5.6) 和 (5.8), 我们可以得到 $\sqrt{n}(\hat{\nu}^b - \hat{\nu})$ 的表达式:

$$\begin{aligned}
&\sqrt{n}(\hat{\nu}^b - \hat{\nu}) \\
&= -\Gamma \frac{(2g(0)\Sigma)^{-1}}{\sqrt{n}} \sum_{t=1}^n (\delta_t^b - 1) w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} \text{sgn}(\varepsilon_t) + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\delta_t^b - 1) w_{t-1}^{-1} \varepsilon_t + o_p(1).
\end{aligned} \tag{5.9}$$

记

$$\begin{aligned}
\mathbf{Z}_{t,1}^b &= (\delta_t^b - 1) w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} \text{sgn}(\varepsilon_t), \quad \mathbf{Z}_{t,2}^b = (\delta_t^b - 1) w_{t-1}^{-1} \varepsilon_t, \\
\mathbf{Z}_{t,1} &= w_{t-1}^{-1} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} \text{sgn}(\varepsilon_t), \quad \mathbf{Z}_{t,2} = w_{t-1}^{-1} \varepsilon_t.
\end{aligned}$$

基于表达式 (5.9), 当 $B \rightarrow \infty$ 和 $n \rightarrow \infty$ 时,

$$\begin{aligned}
&\frac{n}{B} \sum_{b=1}^B (\hat{\nu}^b - \hat{\nu})^2 \\
&= \frac{1}{B} \sum_{b=1}^B \left\{ \Gamma(2g(0)\Sigma)^{-1} \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_{t,1}^b (\mathbf{Z}_{t,1}^b)' (\Gamma(2g(0)\Sigma)^{-1})' \right. \\
&\quad \left. + \frac{1}{n} \sum_{t=1}^n (\mathbf{Z}_{t,2}^b)^2 - 2\Gamma(2g(0)\Sigma)^{-1} \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_{t,1}^b \mathbf{Z}_{t,2}^b \right\} + o_p(1) \\
&= \Gamma(2g(0)\Sigma)^{-1} \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_{t,1} (\mathbf{Z}_{t,1})' (\Gamma(2g(0)\Sigma)^{-1})' \\
&\quad + \frac{1}{n} \sum_{t=1}^n (\mathbf{Z}_{t,2})^2 - 2\Gamma(2g(0)\Sigma)^{-1} \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_{t,1} \mathbf{Z}_{t,2} + o_p(1) \\
&= (-\Gamma(2g(0)\Sigma)^{-1}, 1) E[\tilde{\mathbf{D}}_1 \tilde{\mathbf{D}}_1'] (-\Gamma(2g(0)\Sigma)^{-1}, 1)' + o_p(1),
\end{aligned}$$

这样就可以得到 (5.2) 的第二个等式成立。 $\square$

定理 5.2 的证明： 令 $\tilde{D}_{t,2}^* = w_{t-1}^{-1}[\varepsilon_t - E(\varepsilon_t|\mathcal{F}_{t-1})]$ 和 $\tilde{D}_t^* = (\tilde{D}_{t,1}', \tilde{D}_{t,2}^*)'$ 。我们有

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=1}^n \tilde{D}_t &= \frac{1}{\sqrt{n}} \sum_{t=1}^n (\tilde{D}_{t,1}', w_{t-1}^{-1}[\varepsilon_t - E(\varepsilon_t|\mathcal{F}_{t-1})])' + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}E(\varepsilon_t|\mathcal{F}_{t-1}))' \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^n \tilde{D}_t^* + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}E(\varepsilon_t|\mathcal{F}_{t-1}))' \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^n \tilde{D}_t^* + \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}\sigma_t E(\eta_t))' \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^n \tilde{D}_t^* + \frac{M}{n} \sum_{t=1}^n (\mathbf{0}', w_{t-1}^{-1}\sigma_t / E(w_{t-1}^{-1}\sigma_t))', \tag{5.10}
\end{aligned}$$

其中, $\mathbf{0}$ 是 $(r+s+1)$ 维向量, 最后一个等式是根据 $E(w_{t-1}^{-1}\varepsilon_t) = E(\eta_t) E(w_{t-1}^{-1}\sigma_t) = M/\sqrt{n}$ 得到的。

因为 $E\{w_{t-1}^{-1}[\varepsilon_t - E(\varepsilon_t|\mathcal{F}_{t-1})]|\mathcal{F}_{t-1}\} = 0$, 所以 $\{\tilde{D}_t^*\}$ 是一个鞅差序列。类似引理 5.2 的证明, 当 $n \rightarrow \infty$ 时, 我们有

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \tilde{D}_t^* \xrightarrow{d} N(\mathbf{0}, E\{\tilde{D}_1 \tilde{D}_1'\}).$$

此外, 由平稳序列的弱大数定律, 当 $n \rightarrow \infty$ 时, 我们有

$$\frac{M}{n} \sum_{t=1}^n w_{t-1}^{-1}\sigma_t / E(w_{t-1}^{-1}\sigma_t) \xrightarrow{p} M.$$

因此, 类似定理 5.1 的证明, 当 $B \rightarrow \infty$ 且 $n \rightarrow \infty$ 时,

$$\sqrt{n}\hat{\nu} \xrightarrow{d} N\left(M, \{-\Gamma(2g(0)\Sigma)^{-1}, 1\}E[\tilde{D}_1 \tilde{D}_1']\{-\Gamma(2g(0)\Sigma)^{-1}, 1\}'\right),$$

$$\frac{n}{B} \sum_{b=1}^B (\hat{\nu}^b - \hat{\nu})^2 = \sigma^2 + o_p(1),$$

即定理 5.2 成立。 □

6 重尾噪声下的 Portmanteau 型单位根检验

6.1 引言

第 3-5 章考虑了在 ARMA-GARCH 模型的误差四阶矩不存在的情况下, 利用 SWQMELE 拟合 ARMA-GARCH 模型时的误差检验问题, 我们所提出的三种检验均要求误差的二阶矩存在。然而实际中, 金融时间序列经常出现无限方差的情形, 并且此现象通常伴随着非平稳性出现。因此在本章, 我们考虑了无限方差的噪声下单位根过程的检验问题。

本章将 Zhang and Chan (2018) 所提出的 Portmanteau 型统计量扩展到了重尾情况。更确切地说, 考虑一个 AR(1) 模型:

$$Y_t = \beta Y_{t-1} + \varepsilon_t, t = 1, \dots, n, \quad (6.1)$$

其中 ε_t 是一个重尾线性过程, 即 $\varepsilon_t = \sum_{j=0}^{\infty} d_j \eta_{t-j}$, $\{\eta_t\}$ 是 i.i.d. 序列, 属于 α -stable 分布的吸引域, 这里 $\alpha \in (0, 2)$ 。在此模型下, 我们的目标是检验 $\beta = 1$ 。我们证明了该 Portmanteau 型统计量的极限分布是一个 stable 过程的泛函。我们的方法不仅保留了 Zhang and Chan (2018) 中检验的易于实施的特性, 而且相应的模拟研究也表明, 与传统的检验统计量 PP、DF 和 ADF 检验相比, 我们的检验在 size 和 power 上有更好的数值效果。对每日世界原油价格和 3 个月 AA 金融商业票据利率的检验证实了我们的方法表现良好。

本章的其余部分组织如下。第 6.2 节介绍了所提出的检验方法和其渐近性质。在第 6.3 节中, 我们阐述了 wild 自助抽样算法如何应用于我们所提出的检验方法。第 6.4 节和第 6.5 节分别给出了模拟结果和实证分析。第 6.6 节给出了结论。证明均在第 6.7 节给出。

6.2 方法和理论结果

考虑 AR(1) 模型 (6.1), 其中 ε_t 是一个平稳、可逆且方差可能为无限的线性过程。更确切地说, 存在实系数 d_j 和 i.i.d. 序列 u_t , 使得 ε_t 有以下的形式:

$$\varepsilon_t = \sum_{j=0}^{\infty} d_j u_{t-j}, \quad (6.2)$$

其中 $d_0 = 1$ 且 $\sum_{j=0}^{\infty} |d_j| < \infty$ 。不失一般性地, 我们假设 $Y_0 = 0$ 。最经典的线性过程为 ARMA 过程, 参考 Brockwell and Davis (1991)。

本章主要目的是检验原假设 $H_0: \beta = 1$ 和备择假设 $H_1: |\beta| < 1$ 。类似 Zhang

and Chan (2018) 所采取的方法，我们提出使用以下统计量：

$$T_n = \frac{n}{M(M+1)} \sum_{i=1}^M (\hat{\rho}(i) - 1),$$

其中 $\hat{\rho}(i) = \hat{\gamma}(i)/\hat{\gamma}(0)$ 为样本自相关函数， $\hat{\gamma}(i) = \frac{1}{n} \sum_{t=1}^{n-i} Y_t Y_{t+i}$ 为样本自协方差函数。为了推导出 T_n 的渐近性质，我们需要以下假设。

假设 6.1: $\{u_t\}$ 是一个 i.i.d. 的序列，属于 α -stable law 的吸引域，其中 $0 < \alpha < 2$ 。更确切地说， u_t 分布的尾部有指数下降形式：

$$P(|u_t| > x) = x^{-\alpha} L(x), \text{ 对 } x > 0, \quad (6.3)$$

其中 $L(x)$ 是一个在无穷点缓变的函数，且当 $x \rightarrow \infty$ 时，

$$\frac{P(u_t > x)}{P(|u_t| > x)} \rightarrow p, \quad \frac{P(u_t < -x)}{P(|u_t| > x)} \rightarrow q,$$

其中 $0 \leq p \leq 1$, $q = 1 - p$ 。当 $\alpha > 1$ 时，我们假设 $Eu_t = 0$ 。

假设 6.2: 系数 d_j 满足 $\sum_{j=1}^{\infty} j|d_j|^\delta < \infty$, 其中 $0 < \delta < \alpha \wedge 1$ 。

注 6.1: 假设 6.1 中的参数 α 控制了 u_t 分布尾部的厚度，因此经常被称为尾指数，更多细节参考 Feller (1971) 的第 XVII 章。假设 6.2 保证了序列 $\sum_{j=0}^{\infty} d_j u_{t-j}$ 几乎处处收敛。

令 $\{a_n\}$ 表示满足以下条件的正常数序列：

$$a_n = \inf\{x : P(|u_1| > x) \leq \frac{1}{n}\}, \quad (6.4)$$

这样由假设 6.1 可以推出

$$\frac{1}{a_n} \sum_{t=1}^n u_t \xrightarrow{d} U_\alpha(1), \text{ 和 } \left(\frac{1}{a_n} \sum_{t=1}^{[nr]} u_t, \frac{1}{a_n^2} \sum_{t=1}^{[nr]} u_t^2 \right) \xrightarrow{J_1} \left(U_\alpha(r), U_{\frac{\alpha}{2}}(r) \right), \quad (6.5)$$

在 $D[0, 1]$ 上成立，其中 $U_\alpha(r)$ 和 $U_{\frac{\alpha}{2}}(r)$ 分别为 Lévy α -stable 和 $\alpha/2$ -stable 过程。这里 $\xrightarrow{J_1}$ 表示在 Skorohod J_1 拓扑下的弱收敛。(6.5) 的证明可以参考 Ibragimov and Linnik (1971) 和 Davis and Resnick (1986)。

在假设 6.1 和 6.2 下，我们对 T_n 有以下定理成立。

定理 6.1: 在假设 6.1 和 6.2 下，对于模型 (6.1)，当 $\beta = 1$ 成立时，则

(i) 对任意给定的 $M \geq 1$,

$$T_n \xrightarrow{d} -\frac{\omega^2 U_\alpha^2(1) + D_M U_{\frac{\alpha}{2}}(1)}{4\omega^2 \int_0^1 U_\alpha^2(t) dt},$$

其中 $D_M = f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right]$, $f_j(1) = \sum_{k=0}^{\infty} d_k d_{k+j}$, $j = 0, 1, \dots, M-1$, $\omega = \sum_{k=0}^{\infty} d_k$.

(ii) 当 $M := M(n) \rightarrow \infty$, 且 $M = o(a_n \wedge n)$ 时, 若 a_n 如 (6.4) 中定义, 那么

$$T_n \xrightarrow{d} -\frac{\omega^2 U_\alpha^2(1) + [f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1)] U_{\frac{\alpha}{2}}(1)}{4\omega^2 \int_0^1 U_\alpha^2(t) dt}.$$

6.3 Wild 自助抽样算法

实际中, 若我们已知尾指数 α 和系数 d_j , 我们可以在原假设下利用参数真值, 通过大量重复生成 T_n 来计算经验分位数, 将其作为检验所需的临界值。然而, 若 α 和 d_j 未知, 我们可以用 wild 自助抽样算法来进行检验。

通常我们所使用的自助重抽样方法 (例如 i.i.d. 重抽样法) 在噪声方差无限时无法逼近检验统计量的渐近分布。因此, Romano and Wolf (1999) 提出的子抽样法 (subsampling technique) 和 Arcones and Giné (1989) 提出的 m -out-of- n ($m \leq n$) 自助抽样法在重尾情况下经常被采用。Horvath and Kokoszka (2003) 曾在 $\alpha \in (1, 2)$ 情况下, 提出了一个 m -out-of- n 残差自助抽样法来逼近检验统计量 $n(\hat{\beta} - 1)$ 的分布, 其中 $\hat{\beta}$ 为 β 的 OLS 估计。Jach and Kokoszka (2004) 也使用了子抽样方法来逼近当 $\alpha \in (1, 2)$ 时 $n(\hat{\beta} - 1)$ 在原假设下的分布。然而, 上述两个方法被证实在有限样本下检验结果的可靠性很低, 参考 Cornea and Davidson (2015), 并且他们的方法只对 $\alpha \in (1, 2)$ 有效。我们本章所采用的 wild 自助抽样算法使用了和原样本一致的样本量, 并且对任意 $\alpha \in (0, 2)$ 在有限样本下都有很好的表现。

本章的 wild 自助抽样算法的思路主要源于 Cavaliere et al. (2018)。他们在线性过程具有无限方差的背景下, 提出了针对 ADF 检验统计量的 sieve wild 自助抽样算法。在此情况下, 当 wild 自助抽样算法的噪声由 sieve 残差乘上一个 i.i.d. 对称两点 (Rademacher) 分布随机变量序列所生成时, sieve wild 自助抽样算法下的 ADF 检验渐近有效。唯一需要的条件是噪声的分布是对称的, 而本章中假设 6.1 满足此条件。因此, 基于 Cavaliere et al. (2018), 我们提出了针对 T_n 的 wild 自助抽样算法:

- **步骤 1:** 基于 ADF 型回归计算 OLS 残差: $\hat{\varepsilon}_t := \Delta Y_t - \hat{\phi}_k Y_{t-1} - \sum_{i=1}^k \hat{\pi}_i \Delta Y_{t-i}$, $t = k+1, \dots, n$ 。这里, $\hat{\pi}_i$ 和 $\hat{\phi}_k$ 为 ADF 型回归系数的估计量。

- **步骤 2:** 生成自助残差 $\varepsilon_t^b := \hat{\varepsilon}_t w_t, t = k+1, \dots, n$, 这里 w_t 是 i.i.d. Rademacher 分布随机变量, 即 $w_t \in \{-1, +1\}$, 且每个结果都有概率 $\frac{1}{2}$, 并和原数据独立。
- **步骤 3:** 构造自助抽样样本:

$$Y_t^b = \sum_{i=1}^t \varepsilon_i^b, t = 1, \dots, n,$$

这里初值取为 $(\varepsilon_1^b, \dots, \varepsilon_k^b) = (0, \dots, 0)$ 。

- **步骤 4:** 基于自助抽样样本 Y_t^b 计算 T_n , 即 $T_n^b = \frac{n}{M(M+1)} \sum_{i=1}^M (\hat{\rho}^b(i) - 1)$, 其中 $\hat{\rho}^b(i)$ 为基于自助抽样样本的样本自相关函数。
- **步骤 5:** 重复步骤 2-4 B 次, 计算 p 值: $\hat{p} = \frac{1}{B} \sum_{b=1}^B \mathbb{I}(T_n^b \leq T_n)$, 其中 $\mathbb{I}(\cdot)$ 为示性函数。若 $\hat{p}$ 小于显著性水平, 则拒绝原假设。

类似 Cavaliere et al. (2018) 的定理 5, 我们可以得出 wild 自助抽样统计量逼近了原数据下的极限分布。从第 6.4 节的模拟结果我们也可以看出, wild 自助抽样算法有更稳定的结果。

6.4 模拟研究

本节我们探究了我们提出的单位根检验在有限样本下的表现, 并且和 PP、DF 和 ADF 检验进行了比较。本节所用的数据由以下模型生成:

$$Y_t = \beta Y_{t-1} + \varepsilon_t,$$

其中 ε_t 由 ARMA(1,1) 模型生成:

$$\varepsilon_t = \phi \varepsilon_{t-1} + \varphi u_{t-1} + u_t.$$

这里, u_t 由 $\beta_s = 0$ 和 $\alpha = 0.5, 1, 1.5$ 的 stable 分布生成。 $\alpha \in (0, 2)$ 代表 stable 分布的尾指数, $\beta_s \in [-1, 1]$ 代表 stable 分布的偏斜参数, 当 $\beta_s = 0$ 时, stable 分布的特征函数为对称的拉长的指数函数。对于噪声 ε_t , 我们取 $(\phi, \varphi) = (0, 0)$ 对应独立的情况, 取 $(\phi, \varphi) = (0, -0.8), (0, -0.5), (0, 0.5)$ 和 $(0, 0.8)$ 对应 MA 模型的情况, 取 $(\phi, \varphi) = (0.3, 0)$ 对应 AR 模型的情况, 取 $(\phi, \varphi) = (0.3, -0.8), (0.3, -0.5), (0.3, 0.5)$ 和 $(0.3, 0.8)$ 对应 ARMA 模型的情况。

为了得到定理 6.1 中 T_n 的极限分布的经验临界值, 我们用不同的 α 和 (ϕ, φ) 生成 Y_t , 并且在 H_0 下计算检验统计量 T_n 。对每种情况, 我们均取样本量 $n = 10000$,

$M = [n^{1/3}]$ 、 $[2n^{1/3}]$ 和 $[3n^{1/3}]$ ，重复次数 10000 次。表 6.1-6.3 给出了每种情况下重复生成 10000 次后所得到的 T_n 的经验分位数。

接下来，我们在显著性水平 5% 下来探究我们提出的单位根检验的 size 和 power。我们取 $\beta = 1 - \frac{c}{n}$ ，其中 c 为一个常数，样本量 $n = 200$ 、500 和 1000。当 $c = 0$ 时，我们计算检验的 size；当 $c = 40$ 时，我们计算检验的 power（即对 $n = 200$ 、500 和 1000， β 分别为 0.8、0.92 和 0.96）。在每种设定下，我们重复模拟 1000 次，并且取表 6.1-6.3 中 $M = [n^{1/3}]$ 下的经验分位数作为我们检验的临界值。

表 6.1 $\alpha = 1.5$ 时，不同 M 下 T_n 的经验分位数

(ϕ, φ)	M	0.01	0.05	0.1	0.9	0.95	0.99
(0,-0.8)	$n^{1/3}$	-48.9647	-31.7676	-25.0198	-4.2963	-3.5162	-2.5832
	$2n^{1/3}$	-30.0277	-19.3105	-15.1698	-2.9480	-2.4978	-1.8664
	$3n^{1/3}$	-24.6529	-15.9597	-12.6407	-2.5191	-2.1342	-1.5978
(0,-0.5)	$n^{1/3}$	-17.1711	-10.8595	-8.5734	-1.8190	-1.5661	-1.2157
	$2n^{1/3}$	-15.4659	-9.4602	-7.5175	-1.6659	-1.4381	-1.1313
	$3n^{1/3}$	-14.6574	-9.0293	-7.1388	-1.6104	-1.3875	-1.0748
(0,0)	$n^{1/3}$	-13.1634	-8.2485	-6.6255	-1.5115	-1.2981	-1.0060
	$2n^{1/3}$	-13.6470	-8.3771	-6.7150	-1.5131	-1.3064	-0.9950
	$3n^{1/3}$	-13.1224	-8.3443	-6.6272	-1.5253	-1.3218	-1.0183
(0,0.5)	$n^{1/3}$	-13.1128	-8.0993	-6.3951	-1.4792	-1.2724	-0.9842
	$2n^{1/3}$	-13.0610	-8.1679	-6.3983	-1.4930	-1.2832	-0.9970
	$3n^{1/3}$	-13.5744	-8.1827	-6.5219	-1.4929	-1.3032	-1.0121
(0,0.8)	$n^{1/3}$	-12.5828	-7.9164	-6.4214	-1.4709	-1.2721	-0.9847
	$2n^{1/3}$	-13.2019	-8.1040	-6.5123	-1.4761	-1.2679	-0.9656
	$3n^{1/3}$	-12.9688	-8.1529	-6.5223	-1.4797	-1.2880	-0.9919
(0.3,-0.8)	$n^{1/3}$	-30.1177	-19.4141	-15.6532	-3.0040	-2.5383	-1.9266
	$2n^{1/3}$	-22.7123	-14.4461	-11.4287	-2.3012	-1.9273	-1.4750
	$3n^{1/3}$	-18.9509	-12.3020	-9.9492	-2.0458	-1.7162	-1.3146
(0.3,-0.5)	$n^{1/3}$	-15.4310	-9.3955	-7.3043	-1.6197	-1.3978	-1.0750
	$2n^{1/3}$	-14.5267	-8.8990	-6.9891	-1.5661	-1.3546	-1.0755
	$3n^{1/3}$	-14.1327	-8.6220	-6.8318	-1.5486	-1.3362	-1.0639
(0.3,0)	$n^{1/3}$	-13.5556	-8.1236	-6.3925	-1.4637	-1.2662	-0.9684
	$2n^{1/3}$	-13.3686	-8.2448	-6.4476	-1.4908	-1.2868	-0.9980
	$3n^{1/3}$	-12.7757	-8.1206	-6.5004	-1.4861	-1.2783	-1.0220
(0.3,0.5)	$n^{1/3}$	-13.6847	-8.0434	-6.4096	-1.4452	-1.2435	-0.9650
	$2n^{1/3}$	-12.9987	-7.9864	-6.3503	-1.4690	-1.2745	-0.9961
	$3n^{1/3}$	-13.6080	-8.1021	-6.3452	-1.4780	-1.2833	-0.9806
(0.3,0.8)	$n^{1/3}$	-13.4281	-7.9730	-6.2401	-1.4212	-1.2286	-0.9571
	$2n^{1/3}$	-13.6698	-8.1072	-6.3745	-1.4739	-1.2796	-0.9987
	$3n^{1/3}$	-13.4488	-8.2534	-6.4590	-1.4718	-1.2736	-1.0056

我们也比较了每种设定下 PP、DF 和 ADF 检验的表现。对于 PP 检验，我们基于带有常数趋势项的回归，计算 Phillips-Perron 的 t 检验统计量（参考 Hamilton (1994) 中 (17.6.12) 的 Case 2）。也就是说，在随机游走模型 $Y_t = Y_{t-1} + \varepsilon_t$ 下，基于回归

表 6.2 $\alpha = 1$ 时, 不同 M 下 T_n 的经验分位数

(ϕ, φ)	M	0.01	0.05	0.1	0.9	0.95	0.99
(0,-0.8)	$n^{1/3}$	-63.4898	-32.5832	-23.9419	-4.4967	-3.7629	-2.8000
	$2n^{1/3}$	-45.5288	-21.1783	-15.7773	-3.0047	-2.5576	-1.9578
	$3n^{1/3}$	-32.3396	-16.9800	-12.5754	-2.5030	-2.1731	-1.6930
(0,-0.5)	$n^{1/3}$	-25.0250	-11.8528	-8.7670	-1.8288	-1.6045	-1.2665
	$2n^{1/3}$	-23.2887	-10.8226	-7.8591	-1.6828	-1.4717	-1.1732
	$3n^{1/3}$	-23.0045	-10.5451	-7.5117	-1.6268	-1.4358	-1.1614
(0,0)	$n^{1/3}$	-23.2540	-9.7449	-7.1644	-1.4878	-1.3214	-1.0781
	$2n^{1/3}$	-21.3980	-9.9214	-7.0846	-1.5039	-1.3412	-1.0632
	$3n^{1/3}$	-20.6695	-9.5055	-7.0417	-1.4903	-1.3260	-1.0488
(0,0.5)	$n^{1/3}$	-22.3850	-9.8752	-7.0113	-1.4699	-1.3085	-1.0266
	$2n^{1/3}$	-22.1959	-9.5291	-6.8818	-1.4906	-1.3197	-1.0739
	$3n^{1/3}$	-21.2040	-9.3074	-6.7265	-1.4921	-1.3361	-1.0561
(0,0.8)	$n^{1/3}$	-22.8270	-9.6592	-6.8835	-1.4581	-1.2993	-1.0567
	$2n^{1/3}$	-23.8670	-9.7084	-6.9790	-1.4898	-1.3278	-1.0628
	$3n^{1/3}$	-20.4179	-9.1932	-6.7464	-1.4838	-1.3230	-1.0802
(0.3,-0.8)	$n^{1/3}$	-45.8994	-22.1912	-16.5621	-3.0585	-2.6381	-1.9862
	$2n^{1/3}$	-33.0213	-15.5285	-11.5184	-2.2777	-1.9701	-1.5544
	$3n^{1/3}$	-28.1760	-13.1594	-9.6626	-2.0472	-1.7853	-1.4151
(0.3,-0.5)	$n^{1/3}$	-24.9993	-10.6365	-7.6177	-1.6220	-1.4391	-1.1515
	$2n^{1/3}$	-23.4015	-10.5522	-7.5818	-1.5610	-1.3923	-1.1210
	$3n^{1/3}$	-22.3311	-9.7750	-7.1390	-1.54016	-1.3624	-1.1013
(0.3,0)	$n^{1/3}$	-20.0154	-8.9447	-6.6184	-1.4445	-1.2836	-1.0267
	$2n^{1/3}$	-22.2945	-9.3954	-6.8422	-1.4809	-1.3128	-1.0608
	$3n^{1/3}$	-21.6116	-9.5220	-6.9252	-1.4819	-1.3237	-1.0582
(0.3,0.5)	$n^{1/3}$	-20.1311	-9.0442	-6.5051	-1.4187	-1.2564	-1.0183
	$2n^{1/3}$	-27.2554	-9.4947	-6.9245	-1.4568	-1.2881	-1.0477
	$3n^{1/3}$	-21.3161	-9.4514	-6.8369	-1.4840	-1.3242	-1.0540
(0.3,0.8)	$n^{1/3}$	-21.9844	-9.1346	-6.5580	-1.4325	-1.2785	-1.0267
	$2n^{1/3}$	-18.7975	-8.9660	-6.6282	-1.4558	-1.3035	-1.0458
	$3n^{1/3}$	-20.2876	-9.3011	-6.74412	-1.4713	-1.3036	-1.0548

表 6.3 $\alpha = 0.5$ 时, 不同 M 下 T_n 的经验分位数

(ϕ, φ)	M	0.01	0.05	0.1	0.9	0.95	0.99
(0,-0.8)	$n^{1/3}$	-98.8651	-36.7751	-24.5864	-4.8812	-4.1956	-3.2146
	$2n^{1/3}$	-66.8225	-24.6136	-16.2070	-3.1864	-2.7663	-2.1301
	$3n^{1/3}$	-57.3926	-21.4173	-13.6921	-2.6836	-2.3484	-1.8680
(0,-0.5)	$n^{1/3}$	-47.0132	-15.2823	-9.8346	-1.8647	-1.6868	-1.3490
	$2n^{1/3}$	-42.1031	-13.5819	-8.8144	-1.6885	-1.5503	-1.2663
	$3n^{1/3}$	-40.6267	-13.5953	-8.5271	-1.6259	-1.4913	-1.2259
(0,0)	$n^{1/3}$	-39.0005	-13.1585	-8.0611	-1.5093	-1.3997	-1.1565
	$2n^{1/3}$	-44.9054	-12.5844	-7.6736	-1.5123	-1.4125	-1.1658
	$3n^{1/3}$	-38.5103	-12.5659	-7.7533	-1.5161	-1.4275	-1.1712
(0,0.5)	$n^{1/3}$	-40.7118	-12.4666	-7.8754	-1.4764	-1.3767	-1.1521
	$2n^{1/3}$	-38.2451	-11.9609	-7.6090	-1.4941	-1.3844	-1.1441
	$3n^{1/3}$	-37.6569	-12.6619	-7.8060	-1.5002	-1.4028	-1.1416
(0,0.8)	$n^{1/3}$	-39.5048	-11.9822	-7.5551	-1.4688	-1.3601	-1.1076
	$2n^{1/3}$	-39.5129	-12.4299	-7.7578	-1.4856	-1.3759	-1.1397
	$3n^{1/3}$	-41.9460	-12.6892	-7.8764	-1.4981	-1.3890	-1.1610
(0.3,-0.8)	$n^{1/3}$	-70.2221	-25.5682	-16.3834	-3.2679	-2.8546	-2.1957
	$2n^{1/3}$	-57.0094	-19.9725	-12.7122	-2.4108	-2.1417	-1.7159
	$3n^{1/3}$	-52.0281	-16.9192	-10.9152	-2.1297	-1.9152	-1.5506
(0.3,-0.5)	$n^{1/3}$	-45.2799	-13.7290	-8.7316	-1.6445	-1.5108	-1.2386
	$2n^{1/3}$	-42.6986	-13.2707	-8.3175	-1.5812	-1.4747	-1.1971
	$3n^{1/3}$	-44.6812	-13.1987	-8.2425	-1.5565	-1.4454	-1.1924
(0.3,0)	$n^{1/3}$	-47.0298	-12.4729	-7.5799	-1.4544	-1.3506	-1.1108
	$2n^{1/3}$	-42.7606	-12.2616	-7.8842	-1.4820	-1.3784	-1.1482
	$3n^{1/3}$	-35.2937	-12.0681	-7.6500	-1.4897	-1.3876	-1.1442
(0.3,0.5)	$n^{1/3}$	-39.1284	-12.4871	-7.5440	-1.4323	-1.3358	-1.1005
	$2n^{1/3}$	-37.3248	-12.4646	-7.7619	-1.4766	-1.3842	-1.1477
	$3n^{1/3}$	-38.5823	-12.1599	-7.6325	-1.4823	-1.3666	-1.1344
(0.3,0.8)	$n^{1/3}$	-36.8563	-12.4337	-7.7127	-1.4301	-1.3326	-1.0939
	$2n^{1/3}$	-36.9934	-12.0701	-7.5395	-1.4699	-1.3699	-1.1304
	$3n^{1/3}$	-39.9906	-12.6922	-7.6797	-1.4799	-1.3701	-1.1378

$Y_t = \omega + \beta Y_{t-1} + \varepsilon_t$, 我们计算 t 检验统计量, 其中方差 σ^2 是由 Newey-West 估计量在截断阶数取为 $[4(n/100)^{1/4}]$ 下估计, 此估计方法和 R 中 “PP.test” 中的估计方法一样。DF 统计量定义为

$$n(\hat{\beta} - 1) = n\left(\frac{\sum_{t=1}^n Y_{t-1}Y_t}{\sum_{t=1}^n Y_{t-1}^2} - 1\right),$$

其中 $\hat{\beta}$ 为 OLS 估计量。对于 ADF 检验, 我们使用 R 软件包 “tseries” 的函数 “adf.test”, 滞后长度 p 取为 $[(n-1)^{1/3}]$ 。这三个检验的临界值是随样本量大小 n 变化的。对于 PP 和 ADF 检验, 在显著性水平 5% 下, 临界值在 $n = 200$ 下取 -2.88, 在 $n = 500$ 下取 -2.87, 在 $n = 1000$ 下取 -2.86。对于 DF 检验, 临界值在 $n = 200$ 和 500 时选为 -14.0, 在 $n = 1000$ 时选为 -14.1, 参考 Hamilton (1994) 的表 B.5-B.6。

当 $\alpha = 0.5$ 、1 和 1.5 时, 检验的 power 和 size 分别在表 6.4-6.6 和表 6.7-6.9 中给出。

表 6.4 $\alpha = 1.5$ 和 $c = 40$ 时 Portmanteau 型检验、PP、DF 和 ADF 检验的 power

(ϕ, φ)	n=200				n=500				n=1000			
	New	PP	DF	ADF	New	PP	DF	ADF	New	PP	DF	ADF
(0,-0.8)	0.961	1.000	1.000	0.995	1.000	1.000	1.000	0.998	1.000	1.000	1.000	0.999
(0,-0.5)	1.000	0.999	1.000	0.992	1.000	1.000	1.000	0.991	1.000	1.000	1.000	0.995
(0,0)	1.000	0.997	1.000	0.972	1.000	0.998	0.999	0.990	1.000	0.997	1.000	0.995
(0,0.5)	1.000	0.969	0.989	0.973	1.000	0.985	0.988	0.992	1.000	0.995	0.986	0.996
(0,0.8)	0.998	0.956	0.970	0.983	1.000	0.987	0.966	0.994	1.000	0.996	0.954	0.996
(0.3,-0.8)	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	0.999	1.000	0.997
(0.3,-0.5)	1.000	1.000	1.000	0.980	1.000	0.999	1.000	0.993	1.000	1.000	1.000	0.998
(0.3,0)	0.997	0.970	0.988	0.962	0.999	0.992	0.982	0.995	1.000	0.996	0.976	0.996
(0.3,0.5)	0.992	0.938	0.622	0.951	1.000	0.985	0.625	0.986	0.999	0.990	0.610	0.994
(0.3,0.8)	0.992	0.922	0.509	0.980	0.998	0.980	0.485	0.989	1.000	0.990	0.474	0.995

从表 6.4-6.6 可以看出, 总的来说, 我们提出的检验比 PP、DF 和 ADF 检验有更大的 power, 且 power 随着样本量的增加而变大。当 ε_t 服从 ARMA 和 AR 模型时 (这种情形 ε_t 可以表示为 MA(∞) 的形式), power 随样本量的增加而变大的特性更加明显, 特别是在 $\alpha = 0.5$ 的设定下。此外, 当 α 变小时, 四个检验的 power 均在变小。其中, DF 检验表现最差, 在所有情况下都给出了最小的 power。当 $\alpha = 1.5$ 和 1 时, 我们提出的检验比 ADF 和 PP 的 power 更大。当 $\alpha = 0.5$ 时, 我们的检验和 PP、ADF 检验表现相似。然而, ADF 和 PP 检验都有较差的 size, 这可以从表 6.9 看出。另外, 这种设定下我们的检验的 power 可以通过利用第 6.3 节中提到的 wild 自助抽样算法来得到改善, 参考下文的表 6.10。

表 6.7-6.9 表明, 我们提出的检验有精确的 size。当样本量增加时, size 更接近 0.05, 甚至是对系数为负数的 MA 模型 (即 $(\phi, \varphi) = (0, -0.8)$ 和 $(0.3, -0.8)$) 也是如

表 6.5 $\alpha = 1$ 和 $c = 40$ 时 Portmanteau 型检验、PP、DF 和 ADF 检验的 power

(ϕ, φ)	n=200				n=500				n=1000			
	New	PP	DF	ADF	New	PP	DF	ADF	New	PP	DF	ADF
(0,-0.8)	0.911	0.999	1.000	0.991	1.000	1.000	1.000	0.998	1.000	1.000	1.000	0.998
(0,-0.5)	1.000	1.000	0.999	0.986	1.000	1.000	1.000	0.990	1.000	1.000	1.000	0.991
(0,0)	1.000	0.995	1.000	0.969	1.000	0.995	1.000	0.981	1.000	1.000	1.000	0.996
(0,0.5)	0.990	0.961	0.985	0.957	0.999	0.983	0.987	0.984	1.000	0.991	0.985	0.967
(0,0.8)	0.990	0.961	0.967	0.977	0.999	0.983	0.972	0.986	0.999	0.990	0.973	0.992
(0.3,-0.8)	1.000	0.999	0.999	0.990	1.000	1.000	1.000	0.995	1.000	1.000	1.000	0.994
(0.3,-0.5)	1.000	0.999	1.000	0.985	1.000	0.998	1.000	0.982	1.000	1.000	1.000	0.988
(0.3,0)	0.990	0.977	0.988	0.966	0.999	0.978	0.984	0.978	0.999	0.993	0.981	0.993
(0.3,0.5)	0.976	0.938	0.734	0.958	0.993	0.989	0.724	0.989	0.999	0.987	0.692	0.991
(0.3,0.8)	0.976	0.935	0.498	0.965	0.997	0.980	0.532	0.982	1.000	0.991	0.466	0.993

表 6.6 $\alpha = 0.5$ 和 $c = 40$ 时 Portmanteau 型检验、PP、DF 和 ADF 检验的 power

(ϕ, φ)	n=200				n=500				n=1000			
	New	PP	DF	ADF	New	PP	DF	ADF	New	PP	DF	ADF
(0,-0.8)	0	0.997	1.000	0.980	1.000	1.000	1.000	0.988	1.000	0.999	1.000	0.993
(0,-0.5)	1.000	0.997	1.000	0.969	1.000	0.999	1.000	0.988	1.000	1.000	1.000	0.983
(0,0)	0.975	0.996	1.000	0.965	0.993	0.992	1.000	0.983	0.993	0.992	1.000	0.985
(0,0.5)	0.950	0.965	0.985	0.964	0.988	0.979	0.986	0.980	0.991	0.988	0.988	0.989
(0,0.8)	0.942	0.970	0.972	0.973	0.983	0.984	0.973	0.986	0.995	0.981	0.979	0.984
(0.3,-0.8)	1.000	0.997	1.000	0.981	1.000	0.999	1.000	0.988	1.000	1.000	1.000	0.994
(0.3,-0.5)	1.000	0.998	1.000	0.971	1.000	0.997	1.000	0.984	1.000	0.999	1.000	0.992
(0.3,0)	0.897	0.964	0.981	0.953	0.966	0.985	0.977	0.984	0.989	0.982	0.985	0.986
(0.3,0.5)	0.363	0.962	0.836	0.967	0.954	0.976	0.849	0.975	0.978	0.981	0.826	0.985
(0.3,0.8)	0.307	0.944	0.611	0.956	0.960	0.977	0.632	0.981	0.986	0.983	0.382	0.989

表 6.7 $\alpha = 1.5$ 时 Portmanteau 型检验的 size

(ϕ, φ)	n=200				n=500				n=1000			
	New	PP	DF	ADF	New	PP	DF	ADF	New	PP	DF	ADF
(0,-0.8)	0.004	0.986	0.915	0.573	0.098	0.979	0.928	0.426	0.099	0.955	0.915	0.336
(0,-0.5)	0.098	0.386	0.259	0.159	0.097	0.317	0.279	0.152	0.078	0.240	0.265	0.152
(0,0)	0.033	0.055	0.007	0.148	0.043	0.061	0.009	0.147	0.051	0.066	0.005	0.145
(0,0.5)	0.026	0.057	0	0.176	0.048	0.063	0	0.142	0.046	0.055	0	0.158
(0,0.8)	0.026	0.046	0	0.190	0.042	0.056	0	0.159	0.041	0.055	0.001	0.172
(0.3,-0.8)	0.084	0.821	0.601	0.450	0.164	0.741	0.647	0.392	0.143	0.652	0.643	0.291
(0.3,-0.5)	0.060	0.128	0.070	0.157	0.059	0.099	0.079	0.148	0.044	0.099	0.057	0.149
(0.3,0)	0.026	0.047	0	0.141	0.030	0.057	0.001	0.143	0.041	0.067	0.001	0.144
(0.3,0.5)	0.019	0.047	0	0.145	0.025	0.050	0	0.134	0.046	0.050	0	0.148
(0.3,0.8)	0.020	0.053	0	0.194	0.033	0.049	0	0.165	0.037	0.054	0	0.177

表 6.8 $\alpha = 1$ 时 Portmanteau 型检验的 size

(ϕ, φ)	n=200				n=500				n=1000			
	New	PP	DF	ADF	New	PP	DF	ADF	New	PP	DF	ADF
(0,-0.8)	0.008	0.993	0.917	0.554	0.087	0.987	0.943	0.368	0.111	0.972	0.948	0.239
(0,-0.5)	0.073	0.356	0.213	0.149	0.073	0.253	0.229	0.117	0.063	0.187	0.263	0.120
(0,0)	0.043	0.080	0.006	0.129	0.039	0.085	0.003	0.132	0.051	0.070	0.012	0.136
(0,0.5)	0.033	0.063	0.001	0.132	0.039	0.056	0	0.122	0.047	0.073	0.002	0.128
(0,0.8)	0.032	0.058	0	0.145	0.038	0.063	0.002	0.137	0.042	0.061	0	0.144
(0.3,-0.8)	0.050	0.828	0.597	0.418	0.101	0.668	0.598	0.322	0.097	0.613	0.606	0.222
(0.3,-0.5)	0.055	0.127	0.042	0.136	0.058	0.123	0.042	0.143	0.067	0.115	0.065	0.140
(0.3,0)	0.037	0.075	0	0.163	0.050	0.071	0.002	0.120	0.050	0.074	0	0.130
(0.3,0.5)	0.036	0.069	0	0.144	0.041	0.072	0	0.128	0.043	0.077	0	0.143
(0.3,0.8)	0.037	0.077	0	0.154	0.035	0.076	0	0.146	0.042	0.073	0	0.149

表 6.9 $\alpha = 0.5$ 时 Portmanteau 型检验的 size

(ϕ, φ)	n=200				n=500				n=1000			
	New	PP	DF	ADF	New	PP	DF	ADF	New	PP	DF	ADF
(0,-0.8)	0	0.996	0.957	0.462	0.051	0.995	0.959	0.265	0.055	0.990	0.970	0.159
(0,-0.5)	0.066	0.311	0.180	0.117	0.058	0.219	0.183	0.101	0.053	0.203	0.211	0.116
(0,0)	0.045	0.098	0.004	0.112	0.062	0.091	0.005	0.107	0.044	0.104	0.002	0.103
(0,0.5)	0.048	0.075	0.001	0.099	0.049	0.095	0.005	0.109	0.046	0.094	0.002	0.095
(0,0.8)	0.051	0.080	0.003	0.116	0.050	0.076	0.002	0.112	0.043	0.107	0.002	0.109
(0.3,-0.8)	0.056	0.866	0.536	0.314	0.084	0.668	0.578	0.234	0.090	0.561	0.566	0.167
(0.3,-0.5)	0.045	0.805	0.036	0.103	0.052	0.121	0.033	0.121	0.045	0.108	0.044	0.104
(0.3,0)	0.035	0.084	0.001	0.092	0.037	0.082	0.001	0.083	0.052	0.097	0.001	0.113
(0.3,0.5)	0.031	0.089	0	0.128	0.047	0.088	0.002	0.105	0.049	0.088	0.001	0.108
(0.3,0.8)	0.041	0.085	0	0.129	0.036	0.089	0.001	0.118	0.040	0.095	0	0.126

表 6.10 Wild 自助法下, $\alpha = 0.5$ 和 $c = 40$ 时 Portmanteau 型检验的 power

(ϕ, φ)	n=200	n=500	n=1000
	Wild Bootstrap		
(0,-0.8)	1.000	1.000	1.000
(0,-0.5)	1.000	0.998	0.995
(0,0)	0.987	0.985	0.987
(0,0.5)	0.976	0.977	0.976
(0,0.8)	0.975	0.973	0.976
(0.3,-0.8)	1.000	1.000	1.000
(0.3,-0.5)	1.000	0.988	0.992
(0.3,0)	0.977	0.974	0.976
(0.3,0.5)	0.959	0.977	0.979
(0.3,0.8)	0.959	0.969	0.971

此, 而此情况下在大多数检验中 size 都会过大。这证实了我们的检验比其他三种方法有更稳健的 size。另外, 对于较小的 α , size 明显接近 0.05, 这也表明我们提出的检验在尾指数 α 较小时表现更好。

注 6.2: 我们用 wild 自助抽样法在不同的模型设定下也做了更多的模拟研究。因为结果大体上相似, 所以我们只给出 $\alpha = 0.5$ 时的 power, 见表 6.10。Wild 自助抽样法和使用经验分位数作为临界值的方法相比更加稳健。比如, 当 ε_t 服从一个系数为 $(\phi, \varphi) = (0.3, 0.5)$ 和 $(0.3, 0.8)$ 的 ARMA(1,1) 模型时, 使用 wild 自助抽样法的 power 比使用经验分布作为临界值的方法更大。

综上所述, 我们提出的检验表现很好, 并且比 PP, DF 和 ADF 检验更稳健, 尤其对尾指数 α 较小的情况。

6.5 实证分析

在本节中, 我们研究了 2018 年 7 月 2 日至 2021 年 6 月 29 日的每日世界原油价格 (样本量为 745) 和 2018 年 4 月 2 日至 2020 年 6 月 26 日的每日 3 个月 AA 金融商业票据利率 (DCPF3M) (样本量为 508), 其中每日世界原油价格从 <https://finance.yahoo.com> 下载, 每日 3 个月 AA 金融商业票据利率从 <https://fred.stlouisfed.org> 下载。我们在图 6.1 中画出了这两个序列。

令 $Y_{1,t}$ 和 $Y_{2,t}$ 分别表示每日世界原油价格和 DCPF3M。基于 R 软件包 “tseries” 里的函数 “arma”, 我们首先对序列拟合下列 AR(1) 模型:

$$Y_{1,t} = 0.0444 + 0.9887Y_{1,t-1} + \varepsilon_{1,t}, \quad Y_{2,t} = 0.0223 + 0.9585Y_{2,t-1} + \varepsilon_{2,t}. \quad (6.6)$$

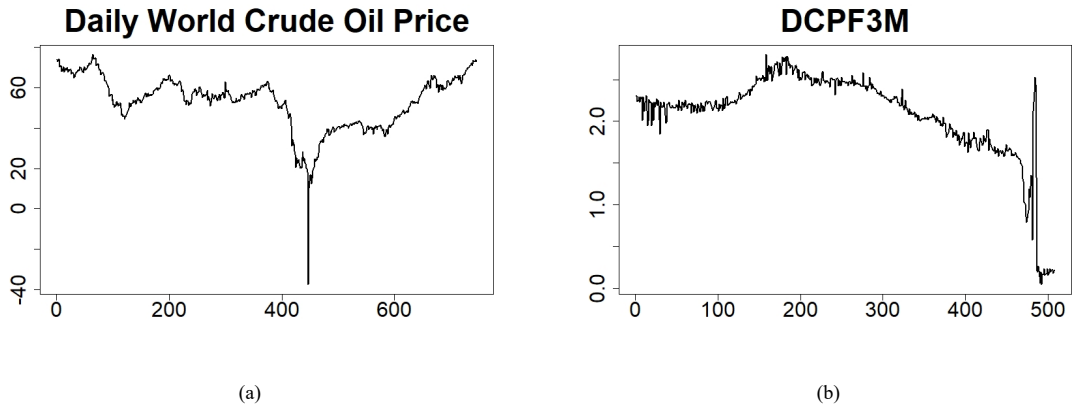


图 6.1 2018 年 7 月 2 日至 2021 年 6 月 29 日的每日世界原油价格（左）和 2018 年 4 月 2 日至 2020 年 6 月 26 日的每日 3 个月 AA 金融商业票据利率（DCPF3M）（右）的时间序列图

接下来我们用 Hill 估计量 (参考 Hill (1975)) 估计 (6.6) 里 $\varepsilon_{i,t}$ ($i = 1, 2$) 的尾指数:

$$\hat{\alpha}_i(k) = \left[\frac{1}{k} \sum_{j=1}^k \log \frac{\varepsilon_{i,(n-j)}}{\varepsilon_{i,(n-k)}} \right]^{-1}, i = 1, 2,$$

其中 $\{\varepsilon_{i,(t)}\}_{t=1}^n$ 为 $\{\varepsilon_{i,t}\}_{t=1}^n$ 的升序统计量。我们在图 6.2 里画出了 $\{\hat{\alpha}_1(k)\}_{k=1}^{200}$ 和 $\{\hat{\alpha}_2(k)\}_{k=1}^{150}$ 。由 Hill 图可知, 当 $k \in [30, 200]$, 两个序列的尾指数的 Hill 估计量均稳定在 1 到 2 之间。因此, 我们取 $\hat{\alpha}_1 = 1.7$ 和 $\hat{\alpha}_2 = 1.4$ 作为尾指数的估计。接下来, R 软件包 “forecast” 里的 “auto.arima” 对 (6.6) 里的 $\varepsilon_{1,t}$ 给出了 ARMA(4,1) 模型:

$$\varepsilon_{1,t} = u_{1,t} + 0.4467\varepsilon_{1,t-1} - 0.0582\varepsilon_{1,t-2} - 0.0314\varepsilon_{1,t-3} + 0.2267\varepsilon_{1,t-4} - 0.5222u_{1,t-1},$$

(0.1243)
(0.0398)
(0.0412)
(0.0357)
(0.1261)

对 (6.6) 里的 $\varepsilon_{2,t}$ 给出了 AR(1) 模型:

$$\varepsilon_{2,t} = u_{2,t} - 0.3448\varepsilon_{2,t-1},$$

(0.0416)

括号里的数字为标准差。有了估计的系数和尾指数, 我们用和模拟中一样的步骤得出显著性水平 5% 上的临界值: 对原油价格为 -6.6882, 对 DCPF3M 为 -8.8841。

我们将本章中提出的检验与 PP、DF 和 ADF 检验一起应用于这两个序列。表 6.11 给出了每个统计量的值。对于原油价格, 我们提出的检验统计量给出的值为 -4.2069, 大于 -6.6882, 这表明我们不能拒绝 H_0 , 每日原油价格序列中确实存在单位根。PP 和 ADF 的值都大于 -2.86, DF 值大于 -14.1, 这与我们提出的检验结果一致。对于 3 个月的 AA 金融商业票据利率, 我们的检验结果为 -11.2020, 小于 -8.8841, 所以我们应该拒绝原假设。然而, ADF、PP 和 DF 检验都表明, H_0 不能被拒绝。这些结果也可以从拟合的 AR(1) 系数中看出。总之, 应用结果表明, 我们提出的检验方法表现良好, 当存在重尾现象时能给出更准确的结果。

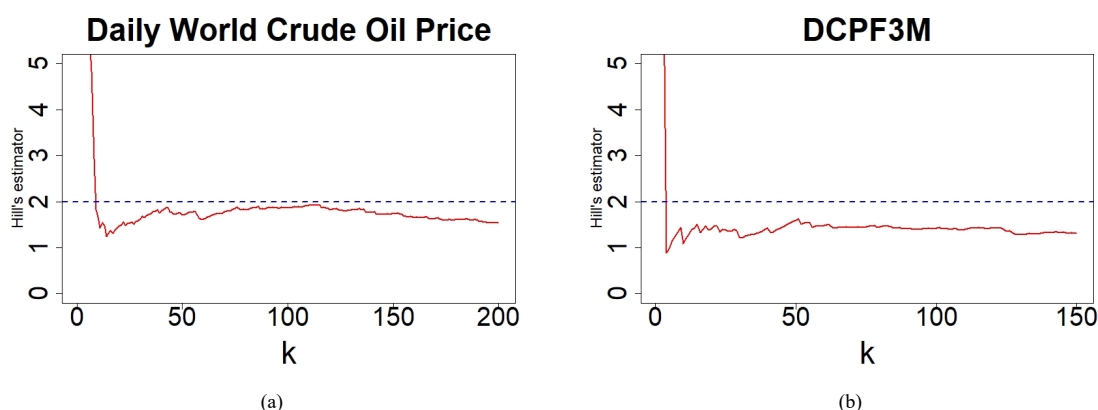


图 6.2 每日原油价格（左）和每日 3 个月 AA 金融商业票据利率（DCPF3M）（右）的 Hill 估计量

表 6.11 原油价格和 3 个月 AA 金融商业票据利率（DCPF3M）在 4 种检验统计量下的值

序列	New	PP	DF	ADF
原油价格	-4.2069	-2.3698	-0.0501	-2.4507
DCPF3M	-11.2020	-1.4452	-10.9513	2.4489

6.6 结论

本章考虑了重尾情况下，应用 Portmanteau 型统计量来进行单位根检验。我们假设了噪声是由重尾线性过程驱动，该过程属于 α -stable 分布的吸引域，其中 $\alpha \in (0, 2)$ 。我们研究了检验统计量的渐近行为，并证明了其极限分布为标准 stable 分布的一个泛函。通过模拟研究和对真实数据的应用，我们证实了该检验在有限样本下具有良好的数值效果。该方法仍有一些地方可以改进。例如，本章只考虑了短记忆过程下的单位根检验，该检验还可以被扩展到长记忆过程的情形。此外，我们还可以考虑当 ε_t 为重尾 GARCH 模型时的情况。这些问题，我们将在未来工作中加以考虑。

6.7 定理 6.1 的证明

我们将 T_n 写为

$$\begin{aligned}
 T_n &= \frac{n}{M(M+1)} \sum_{k=1}^M (\hat{\rho}(k) - 1) \\
 &= \frac{n}{M(M+1)} \sum_{k=1}^M \left(\frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} - 1 \right) \\
 &= \frac{n}{M(M+1)} \sum_{k=1}^M \left(\frac{\hat{\gamma}(k) - \hat{\gamma}(0)}{\hat{\gamma}(0)} \right), \tag{6.7}
 \end{aligned}$$

其中 $\hat{\gamma}(k)$ 如第 6.2 节中定义。经过基本的计算, $\hat{\gamma}(k) - \hat{\gamma}(0)$ 可以写为如下的形式:

$$\hat{\gamma}(k) - \hat{\gamma}(0) = -\frac{1}{2n} \left[\sum_{t=n-k+1}^n Y_t^2 + \sum_{t=1}^k Y_t^2 + \sum_{t=1}^{n-k} \left(\sum_{j=1}^k \varepsilon_{t+j} \right)^2 \right].$$

因此,

$$\begin{aligned} \sum_{k=1}^M \left(\hat{\gamma}(k) - \hat{\gamma}(0) \right) &= -\frac{1}{2n} \left[\sum_{k=1}^M \sum_{t=n-k+1}^n Y_t^2 + \sum_{k=1}^M \sum_{t=1}^k Y_t^2 + \sum_{k=1}^M \sum_{t=1}^{n-k} \left(\sum_{j=1}^k \varepsilon_{t+j} \right)^2 \right] \\ &:= -\frac{1}{2n} [S_{n1} + S_{n2} + S_{n3}], \end{aligned} \quad (6.8)$$

其中 $S_{n1} = \sum_{k=1}^M \sum_{t=n-k+1}^n Y_t^2$, $S_{n2} = \sum_{k=1}^M \sum_{t=1}^k Y_t^2$, $S_{n3} = \sum_{k=1}^M \sum_{t=1}^{n-k} \left(\sum_{j=1}^k \varepsilon_{t+j} \right)^2$ 。结合 (6.8) 和 (6.7), 统计量 T_n 可以被写为如下形式:

$$T_n = -\frac{S_{n1} + S_{n2} + S_{n3}}{2M(M+1)\hat{\gamma}(0)}.$$

为了对 Y_t 应用 B-N 分解, 记 $d(L) = \sum_{j=0}^{\infty} d_j L^j$, 其中 L 为滞后算子, d_j 为 (6.2) 中定义的系数且满足假设 6.2, 那么在原假设下, 我们有 $Y_t - Y_{t-1} = \sum_{j=0}^{\infty} d_j u_{t-j} = d(L)u_t$ 。记 $\tilde{d}(L) = \sum_{j=0}^{\infty} \tilde{d}_j L^j$, 其中 $\tilde{d}_j = -\sum_{k=j+1}^{\infty} d_k$ 。利用 B-N 分解, 经过简单的计算, 我们可以得到:

$$\begin{aligned} Y_t &= d(L) \sum_{j=1}^t u_j \\ &= (d(1) + (1-L)\tilde{d}(L)) \left(\sum_{j=1}^t u_j \right) \\ &= \left(\sum_{k=0}^{\infty} d_k \right) \left(\sum_{j=1}^t u_j \right) + e_t - e_0 \\ &= \sum_{j=1}^t \omega u_j + e_t - e_0, \end{aligned} \quad (6.9)$$

其中 $\omega = d(1) = \sum_{k=0}^{\infty} d_k$, $e_t = \tilde{d}(L)u_t = \sum_{j=0}^{\infty} \tilde{d}_j u_{t-j}$ 。注意 $\sum_{j=0}^{\infty} |\tilde{d}_j|^\delta < \sum_{j=0}^{\infty} j|d_j|^\delta < \infty$, 其中 $0 < \delta < \alpha \wedge 1$, 因此 e_t 几乎处处收敛, 且属于 $\mathcal{D}(\alpha)$ 。类似地, 因为 $e_t - e_0 = \sum_{j=0}^{t-1} \tilde{d}_j u_{t-j} + \sum_{j=t}^{\infty} (\tilde{d}_{j+t} - \tilde{d}_j) u_{t-j}$, 我们有 $e_t - e_0$ 属于 $\mathcal{D}(\alpha)$ 。

6.7.1 S_{n1} 的收敛性

首先我们证明 $S_{n1} = \sum_{k=1}^M \sum_{t=n-k+1}^n Y_t^2$ 的收敛性。注意到 $\frac{\sum_{t=1}^{[nr]} \varepsilon_t}{a_n} \xrightarrow{M_1} wU_\alpha(r)$, 其中 $\xrightarrow{M_1}$ 表示 M_1 拓扑下的弱收敛, 关于 M_1 拓扑的定义和性质, 可以参见 Avram and

Taqqu (1992)。因此,

$$\sup_{1 \leq t \leq n} \frac{|Y_t|}{a_n} = O_p(1). \quad (6.10)$$

下一步, 我们证明: $\sum_{t=n-M+1}^n \frac{|u_t|}{a_n} = o_p(1)$ 。若 $\alpha > 1$, 那么 $E|u_t|$ 存在, 则对任意 $\delta > 0$, 当 $M = o(a_n \wedge n)$ 时,

$$P\left(\sum_{t=n-M+1}^n \frac{|u_t|}{a_n} > \delta\right) \leq \frac{E(\sum_{t=n-M+1}^n |u_t|)}{\delta a_n} = \sum_{t=n-M+1}^n \frac{E(|u_t|)}{\delta a_n} = \frac{ME(|u_t|)}{\delta a_n} \rightarrow 0.$$

若 $0 < \alpha \leq 1$,

$$\begin{aligned} \sum_{t=n-M+1}^n \frac{|u_t|}{a_n} &= \frac{1}{a_n} \sum_{t=n-M+1}^n \left(|u_t|I(|u_t| \leq \delta a_n) + |u_t|I(|u_t| > \delta a_n) \right) \\ &= \frac{1}{a_n} \sum_{t=n-M+1}^n |u_t|I(|u_t| \leq \delta a_n) + \frac{1}{a_n} \sum_{t=n-M+1}^n |u_t|I(|u_t| > \delta a_n) \\ &=: Z_{1n} + Z_{2n}. \end{aligned}$$

当 $n \rightarrow \infty$ 时, 由 Karamata 定理和 $M = o(a_n \wedge n)$,

$$\begin{aligned} P(Z_{1n} > \delta) &= P\left(\sum_{t=n-M+1}^n |u_t|I(|u_t| \leq \delta a_n) > \delta a_n\right) \\ &\leq \frac{1}{\delta a_n} \sum_{t=n-M+1}^n E|u_t|I(|u_t| \leq \delta a_n) \\ &= \frac{M}{\delta a_n} \int_0^{\delta a_n} P(|u_t| > x) dx \\ &= \frac{M}{\delta} P(|u_t| > a_n) \int_0^{\delta} \frac{P(|u_t| > a_n y)}{P(|u_t| > a_n)} dy \\ &= \frac{M}{n\delta^\alpha} \frac{1}{1-\alpha} \rightarrow 0, \end{aligned}$$

因此, $Z_{1n} = o_p(1)$ 。当 $n \rightarrow \infty$ 时, 由 $M = o(a_n \wedge n)$,

$$\begin{aligned} P(Z_{2n} > \delta) &= P\left(\sum_{t=n-M+1}^n |u_t|I(|u_t| > \delta a_n) > \delta a_n\right) \\ &\leq \sum_{t=n-M+1}^n P(|u_t| > \delta a_n) \\ &= MP(|u_t| > a_n) \frac{P(|u_t| > \delta a_n)}{P(|u_t| > a_n)} \\ &= \frac{M}{n\delta^\alpha} \rightarrow 0, \end{aligned}$$

因此 $Z_{2n} = o_p(1)$ 。综上，我们有 $\sum_{t=n-M+1}^n \frac{|u_t|}{a_n} = o_p(1)$ 。

因为 e_t 严平稳且属于 $\mathcal{D}(\alpha)$ ，我们有 $\frac{|e_n|}{a_n} = o_p(1)$ 。此外，利用 $M = o(a_n \wedge n)$ ，

$$P\left(\sup_{1 \leq k \leq M} \frac{|e_{n-k}|}{a_n} > \delta\right) \leq \sum_{k=1}^M P(|e_{n-k}| > a_n \delta) = O\left(\frac{M}{n}\right) = o(1),$$

可以得到 $\sup_{1 \leq k \leq M} \frac{|e_{n-k}|}{a_n} = o_p(1)$ 。因此，

$$\begin{aligned} \frac{1}{a_n^2} \sup_{1 \leq k \leq M} |Y_{n-k}^2 - Y_n^2| &= \frac{1}{a_n^2} \sup_{1 \leq k \leq M} |(Y_{n-k} - Y_n)(Y_{n-k} + Y_n)| \\ &= \frac{1}{a_n^2} \sup_{1 \leq k \leq M} \left| \left(\sum_{t=n-k+1}^n \omega u_t \right) + e_n - e_{n-k} \right| (Y_{n-k} + Y_n) \\ &\leq 2 \left(|\omega| \sum_{t=n-M+1}^n \frac{|u_t|}{a_n} + \frac{|e_n|}{a_n} + \sup_{1 \leq k \leq M} \frac{|e_{n-k}|}{a_n} \right) \left(\sup_{1 \leq t \leq n} \frac{|Y_t|}{a_n} \right) \\ &= o_p(1). \end{aligned} \quad (6.11)$$

此外，将 $\frac{Y_n^2}{a_n^2}$ 写为

$$\begin{aligned} \frac{Y_n^2}{a_n^2} &= \frac{\omega^2 (\sum_{t=1}^n u_t)^2}{a_n^2} + \frac{2\omega (\sum_{t=1}^n u_t)(e_n - e_0)}{a_n^2} + \frac{(e_n - e_0)^2}{a_n^2} \\ &:= F_1 + F_2 + F_3. \end{aligned}$$

利用假设 6.1 可以得到 $F_1 \xrightarrow{d} \omega^2 U_\alpha^2(1)$ 。又因为 $e_n - e_0$ 属于 $\mathcal{D}(\alpha)$ ，我们有 $\frac{e_n - e_0}{a_n} = o_p(1)$ 。另一方面，利用 (6.5) 可推出 $\frac{\sum_{t=1}^n u_t}{a_n} = O_p(1)$ 。因此， $F_2 = o_p(1)$ 和 $F_3 = o_p(1)$ 成立，并且

$$\frac{Y_n^2}{a_n^2} \xrightarrow{d} \omega^2 U_\alpha^2(1). \quad (6.12)$$

对于 S_{n1} ，我们有

$$\begin{aligned} \frac{|S_{n1}|}{a_n^2} &= \frac{1}{a_n^2} |MY_n^2 + (M-1)Y_{n-1}^2 + \dots + Y_{n-M+1}^2| \\ &\leq M \frac{Y_n^2}{a_n^2} + (M-1) \frac{|Y_{n-1}^2 - Y_n^2|}{a_n^2} + (M-1) \frac{Y_n^2}{a_n^2} \\ &\quad + (M-2) \frac{|Y_{n-2}^2 - Y_n^2|}{a_n^2} + (M-2) \frac{Y_n^2}{a_n^2} \\ &\quad + \dots + \frac{|Y_{n-M+1}^2 - Y_n^2|}{a_n^2} + \frac{Y_n^2}{a_n^2} \\ &\leq \frac{M(M+1)}{2} \frac{Y_n^2}{a_n^2} + \frac{M(M-1)}{2} \sup_{1 \leq k \leq M} \frac{|Y_{n-k} + Y_n|}{a_n} \frac{|Y_n - Y_{n-k}|}{a_n}, \end{aligned}$$

那么基于 (6.11) 和 (6.12), 可以得到:

$$\frac{2S_{n1}}{a_n^2 M(M+1)} = \frac{Y_n^2}{a_n^2} + o_p(1) \xrightarrow{d} \omega^2 U_\alpha^2(1). \quad (6.13)$$

6.7.2 S_{n2} 的收敛性

其次, 我们证明 S_{n2} 的收敛性。(6.10) 表明 $\sup_{1 \leq t \leq M} Y_t^2 = O_p(a_M^2)$, 因此, 由 $M = o(a_n \wedge n)$ 可得:

$$\begin{aligned} \frac{2S_{n2}}{a_n^2 M(M+1)} &= \frac{2 \sum_{k=1}^M \sum_{t=1}^k Y_t^2}{a_n^2 M(M+1)} \leq \frac{2M^2}{M(M+1)} \frac{a_M^2}{a_n^2} \left(\max_{1 \leq t \leq M} \frac{Y_t^2}{a_M^2} \right) \\ &= O_p\left(\frac{a_M^2}{a_n^2}\right) = o_p(1). \end{aligned} \quad (6.14)$$

6.7.3 S_{n3} 的收敛性

我们首先考虑 $\varepsilon_t \varepsilon_{t+h}$ 的二阶 B-N 分解形式, 其中 $h = 0, 1, \dots, n$ 。对 $j < 0$, 我们取 $d_j = 0$ 。记 $f_j(L) = \sum_{k=0}^{\infty} d_k d_{k+j} L^k = \sum_{k=0}^{\infty} f_{jk} L^k$, $\tilde{f}_j(L) = \sum_{k=0}^{\infty} \tilde{f}_{jk} L^k = \sum_{k=0}^{\infty} (\sum_{s=k+1}^{\infty} f_{js}) L^k = \sum_{k=0}^{\infty} (\sum_{s=k+1}^{\infty} d_s d_{s+j}) L^k$, 类似 Phillips and Solo (1992) 第 999 页上的证明, $\varepsilon_t \varepsilon_{t+h}$ 有以下分解形式:

$$\begin{aligned} \varepsilon_t \varepsilon_{t+h} &= f_h(1) u_t^2 + u_{t+h} u_t^h - (1-L) \tilde{f}_h(L) u_t^2 \\ &\quad - (1-L) \sum_{r=1}^{\infty} [\tilde{f}_{h+r}(L) u_{t+h-r} u_{t+h} + \tilde{f}_{h-r}(L) u_{t+h+r} u_{t+h}] \\ &:= E_{1,t,h} + E_{2,t,h} - E_{3,t,h} - E_{4,t,h} - E_{5,t,h}, \end{aligned} \quad (6.15)$$

其中, $u_t^h = \sum_{r=1}^{\infty} [f_{h+r}(1) u_{t+h-r} + f_{h-r}(1) u_{t+h+r}]$ 。特别地, 对于 ε_t^2 有:

$$\begin{aligned} \varepsilon_t^2 &= f_0(1) u_t^2 + u_t u_t^0 - (1-L) \tilde{f}_0(L) u_t^2 \\ &\quad - (1-L) \sum_{r=1}^{\infty} [\tilde{f}_r(L) u_{t-r} u_t + \tilde{f}_{-r}(L) u_{t+r} u_t] \\ &:= D_{1,t} + D_{2,t} - D_{3,t} - D_{4,t} - D_{5,t}, \end{aligned} \quad (6.16)$$

其中, $u_t^0 = \sum_{r=1}^{\infty} [f_r(1) u_{t-r} + f_{-r}(1) u_{t+r}]$ 。

接下来, 我们对 S_{n3} 证明其收敛性。注意到 $S_{n3} = \sum_{k=1}^M \sum_{t=1}^{n-k} (\sum_{j=1}^k \varepsilon_{t+j})^2$ 可以

写成以下形式：

$$\begin{aligned}
S_{n3} &= \sum_{k=1}^M \sum_{t=1}^{n-k} \left(\sum_{j=1}^k \varepsilon_{t+j} \right)^2 \\
&= \sum_{k=1}^M \sum_{t=1}^{n-k} \left(\sum_{j=1}^k \varepsilon_{t+j}^2 + \sum_{j_1=1}^k \sum_{j_2=1, j_2 \neq j_1}^k \varepsilon_{t+j_1} \varepsilon_{t+j_2} \right) \\
&= \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k \varepsilon_{t+j}^2 + \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{j_2=1, j_2 \neq j_1}^k \varepsilon_{t+j_1} \varepsilon_{t+j_2} \\
&=: \Delta_{n1} + \Delta_{n2}.
\end{aligned}$$

我们只需分别证明 Δ_{n1} 和 Δ_{n2} 的收敛性即可。

首先我们证明 Δ_{n1} 的收敛性。根据 (6.16), Δ_{n1} 可以写为：

$$\Delta_{n1} = \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k \varepsilon_{t+j}^2 = \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k (D_{1,t+j} + D_{2,t+j} - D_{3,t+j} - D_{4,t+j} - D_{5,t+j}).$$

对于 $D_{2,t+j} = u_{t+j} u_{t+j}^0$, 由 Phillips and Solo (1992) 的引理 5.19 (b) 和假设 6.2, 我们可以得出 $\sum_{r=-\infty, \neq 0}^{\infty} |f_r(1)|^\delta < \infty$, 故 u_{t+j}^0 属于 $\mathcal{D}(\alpha)$ 。因为 u_{t+j} 和 u_{t+j}^0 独立, 根据 Breiman (1965) 的命题 3 可得 $D_{2,t+j}$ 属于 $\mathcal{D}(\alpha)$, 因此 $\frac{\sum_{t=1}^n D_{2,t+j}}{a_n^2} = o_p(1)$ 。所以,

$$\frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{2,t+j}}{a_n^2 M(M+1)} = o_p(1). \quad (6.17)$$

对于 $D_{3,t+j} = (1-L)\tilde{f}_0(L)u_{t+j}^2$, 由 Phillips and Solo (1992) 的引理 5.19 (a), 我们有 $\sum_{j=0}^{\infty} |\tilde{f}_{kj}|^{\delta/2} < \infty$ 和 $\sum_{r=1}^{\infty} |\sum_{k=0}^{\infty} \tilde{f}_{rk}|^\delta < \infty$ 成立。因此, $\tilde{f}_0(L)u_t^2 = \sum_{j=1}^{\infty} \tilde{f}_{0j}u_{t-j}^2$ 几乎处处收敛且属于 $\mathcal{D}(\alpha/2)$ 。故对任意 $\delta > 0$, $P(\frac{\tilde{f}_0(L)u_t^2}{a_n^2} > \delta) = O(\frac{1}{n})$ 和 $\frac{\tilde{f}_0(L)u_t^2}{a_n^2} = o_p(1)$ 成立。进一步, 我们有

$$\frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{3,t+j}}{a_n^2 M(M+1)} = o_p(1). \quad (6.18)$$

对于 $D_{4,t+j} = (1-L)\sum_{r=1}^{\infty} \tilde{f}_r(L)u_{t+j-r}u_{t+j}$, 利用 u_{t+j-r} 和 u_{t+j} 独立, 其中 $1 \leq j \leq M$, 可以得到 $u_{t+j-r}u_{t+j}$ 属于 $\mathcal{D}(\alpha)$ 。此外, 根据假设 6.2 可得出:

$$\begin{aligned}
\sum_{r=1}^{\infty} \sum_{k=0}^{\infty} |\tilde{f}_{rk}|^\delta &\leq \sum_{r=1}^{\infty} \sum_{k=0}^{\infty} \sum_{s=k+1}^{\infty} |d_s d_{s+r}|^\delta = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} s |d_s|^\delta |d_{s+r}|^\delta \\
&< \left(\sum_{s=1}^{\infty} s |d_s|^\delta \right) \left(\sum_{r=0}^{\infty} |d_s|^\delta \right) < \infty,
\end{aligned}$$

因此, $\sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t+j-r} u_{t+j} = \sum_{r=1}^{\infty} \sum_{k=0}^{\infty} \tilde{f}_{rk} u_{t+j-r-k} u_{t+j-k}$ 几乎处处收敛, 即 $\sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t+j-r} u_{t+j}$ 属于 $\mathcal{D}(\alpha)$ 。故对任意 $\delta > 0$,

$$\begin{aligned} P\left(\sup_{1 \leq t \leq M} \left| \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t+j-r} u_{t+j} \right| > \delta a_n\right) &\leq \sum_{t=1}^M P\left(\left| \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t+j-r} u_{t+j} \right| > \delta a_n\right) \\ &\leq O\left(\frac{M}{n}\right) = o(1), \end{aligned}$$

且 $\sup_{1 \leq t \leq M} \left| \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t+j-r} u_{t+j} \right| = o_p(a_n)$ 。再次利用 $M = o(a_n \wedge n)$, 我们可以得到:

$$\begin{aligned} &\frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{4,t+j}}{a_n^2 M(M+1)} \\ &= \frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k \sum_{r=1}^{\infty} (1-L) \tilde{f}_r(L) u_{t+j-r} u_{t+j}}{a_n^2 M(M+1)} \\ &= \frac{2}{a_n^2 M(M+1)} \left[\sum_{t=2}^{M+1} (t-1) \left(M+1 - \frac{1}{2}t\right) \left((1-L) \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t-r} u_t\right) \right. \\ &\quad \left. + \frac{M(M+1)}{2} \sum_{t=M+2}^{n-M} \left((1-L) \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t-r} u_t\right) \right. \\ &\quad \left. + \sum_{t=n-M+1}^n (n-t+1) \left(M - \frac{1}{2}n + \frac{1}{2}t\right) \left((1-L) \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t-r} u_t\right) \right] \\ &= \frac{o_p(a_n)}{a_n^2 M(M+1)} \sum_{t=2}^{M+1} (t-1) \left(M+1 - \frac{1}{2}t\right) + \frac{\sum_{t=M+2}^{n-M} \left((1-L) \sum_{r=1}^{\infty} \tilde{f}_r(L) u_{t-r} u_t\right)}{a_n^2} \\ &= \frac{o_p(a_n)}{a_n^2 M(M+1)} O_p(M^3) + o_p(1) \\ &= o_p(1). \end{aligned} \tag{6.19}$$

类似地, 对 $D_{5,t+j} = (1-L) \sum_{r=1}^{\infty} \tilde{f}_{-r}(L) u_{t+j+r} u_{t+j}$,

$$\frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{5,t+j}}{a_n^2 M(M+1)} = o_p(1). \tag{6.20}$$

最后, 对 $D_{1,t+j}$, 我们有:

$$\begin{aligned}
& \frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{1,t+j}}{a_n^2 M(M+1)} \\
&= \frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k f_0(1) u_{t+j}^2}{a_n^2 M(M+1)} \\
&= \frac{2f_0(1)}{a_n^2 M(M+1)} \left[\sum_{t=2}^{M+1} (t-1)(M+1 - \frac{1}{2}t) u_t^2 \right. \\
&\quad \left. + \frac{M(M+1)}{2} \sum_{t=M+2}^{n-M} u_t^2 + \sum_{t=n-M+1}^n (n-t+1)(M - \frac{1}{2}n + \frac{1}{2}t) u_t^2 \right].
\end{aligned}$$

由 (6.5) 可推出 $\sum_{t=1}^M u_t^2 = O_p(a_M^2)$ 。因为 $\frac{a_M^2}{a_n^2} = o_p(1)$, 我们可以得到:

$$\begin{aligned}
\frac{2f_0(1)}{a_n^2 M(M+1)} \sum_{t=2}^{M+1} (t-1)(M+1 - \frac{1}{2}t) u_t^2 &\leq \frac{2f_0(1)M^2}{a_n^2 M(M+1)} \sum_{t=2}^{M+1} u_t^2 \\
&= \frac{2f_0(1)M^2 a_M^2}{a_n^2 M(M+1)} \sum_{t=2}^{M+1} \frac{u_t^2}{a_M^2} = o_p(1).
\end{aligned}$$

类似地,

$$\frac{2f_0(1)}{a_n^2 M(M+1)} \sum_{t=n-M+1}^n (n-t+1)(M - \frac{1}{2}n + \frac{1}{2}t) u_t^2 = o_p(1).$$

所以,

$$\frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{1,t+j}}{a_n^2 M(M+1)} = f_0(1) \frac{\sum_{t=M+2}^{n-M} u_t^2}{a_n^2} + o_p(1) \xrightarrow{d} f_0(1) U_{\frac{\alpha}{2}}(1). \quad (6.21)$$

对 (6.17)、(6.18)、(6.19)、(6.20) 和 (6.21) 求和, 我们可以得到 Δ_{n1} 的收敛性:

$$\begin{aligned}
\frac{2\Delta_{n1}}{a_n^2 M(M+1)} &= \frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{1,t+j}}{a_n^2 M(M+1)} + o_p(1) \\
&\xrightarrow{d} f_0(1) U_{\frac{\alpha}{2}}(1). \quad (6.22)
\end{aligned}$$

接下来, 我们证明 Δ_{n2} 的收敛性。由分解式 (6.15), Δ_{n2} 可以写为:

$$\begin{aligned}
\Delta_{n2} &= \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{j_2=1, j_2 \neq j_1}^k \varepsilon_{t+j_1} \varepsilon_{t+j_2} \\
&= 2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{j_2=j_1+1}^k \varepsilon_{t+j_1} \varepsilon_{t+j_2} \\
&= 2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} \varepsilon_{t+j_1} \varepsilon_{t+j_1+h} \\
&= 2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} (E_{1,t+j_1,h} + E_{2,t+j_1,h} - E_{3,t+j_1,h} - E_{4,t+j_1,h} - E_{5,t+j_1,h}).
\end{aligned}$$

对于 $E_{2,t+j_1,h} = u_{t+j_1+h} u_{t+j_1}^h$, 由假设 6.2 知 $\sum_{r=-\infty, \neq 0}^{\infty} |f_{h+r}(1)|^\delta < \infty$, 因为 $u_{t+j_1}^h$ 属于 $\mathcal{D}(\alpha)$, 且 u_{t+j_1+h} 和 $u_{t+j_1}^h$ 独立, 所以 $E_{2,t+j_1,h}$ 属于 $\mathcal{D}(\alpha)$ 。因此,

$$\frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} E_{2,t+j_1,h}}{a_n^2 M(M+1)} = o_p(1). \quad (6.23)$$

对于 $E_{3,t+j_1,h} = (1-L)\tilde{f}_h(L)u_{t+j_1}^2$, 由 Phillips and Solo (1992) 的引理 5.19 (a), 我们有 $\sum_{j=0}^{\infty} |\tilde{f}_{kj}|^{\delta/2} < \infty$ 和 $\sum_{r=1}^{\infty} |\sum_{k=0}^{\infty} \tilde{f}_{h+r,k}|^\delta < \infty$ 成立, 故 $\tilde{f}_h(L)u_{t+j_1}^2 = \sum_{j=1}^{\infty} \tilde{f}_{hj}u_{t+j_1-j}^2$ 几乎处处收敛并且属于 $\mathcal{D}(\alpha/2)$ 。因此, $\frac{\tilde{f}_h(L)u_{t+j_1}^2}{a_n^2} = o_p(1)$, 并且

$$\frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} E_{3,t+j_1,h}}{a_n^2 M(M+1)} = o_p(1). \quad (6.24)$$

对于 $E_{4,t+j_1,h} = (1-L)\sum_{r=1}^{\infty} \tilde{f}_{h+r}(L)u_{t+j_1+h-r}u_{t+j_1+h}$, 由假设 6.2 可得:

$$\sum_{r=1}^{\infty} \sum_{k=0}^{\infty} |\tilde{f}_{h+r,k}|^\delta \leq \sum_{r=1}^{\infty} \sum_{k=0}^{\infty} \sum_{s=k+1}^{\infty} |d_s d_{s+h+r}|^\delta = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} s |d_s|^\delta |d_{s+h+r}|^\delta < \infty,$$

故 $\sum_{r=1}^{\infty} \tilde{f}_{h+r}(L)u_{t+j_1+h-r}u_{t+j_1+h}$ 几乎处处收敛并且属于 $\mathcal{D}(\alpha)$ 。因此,

$$\sup_{1 \leq t \leq M} \left| \sum_{h=1}^{M-1} \sum_{r=1}^{\infty} \tilde{f}_{h+r}(L)u_{t+j_1+h-r}u_{t+j_1+h} \right| = O_p(a_M).$$

再次利用 $M = o(a_n \wedge n)$, 我们有

$$\begin{aligned}
& \frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} E_{4,t+j_1,h}}{a_n^2 M(M+1)} \\
&= \frac{4}{a_n^2 M(M+1)} \left[\sum_{t=2}^{M+1} (t-1)(M+1 - \frac{1}{2}t) ((1-L) \sum_{h=1}^{M-1} \sum_{r=1}^{\infty} \tilde{f}_{h+r}(L) u_{t+h-r} u_{t+h}) \right. \\
&\quad + \frac{M(M+1)}{2} \sum_{t=M+2}^{n-M} ((1-L) \sum_{h=1}^{M-1} \sum_{r=1}^{\infty} \tilde{f}_{h+r}(L) u_{t+h-r} u_{t+h}) \\
&\quad \left. + \sum_{t=n-M+1}^n (n-t+1)(M - \frac{1}{2}n + \frac{1}{2}t) ((1-L) \sum_{h=1}^{M-1} \sum_{r=1}^{\infty} \tilde{f}_{h+r}(L) u_{t+h-r} u_{t+h}) \right] \\
&= \frac{4}{a_n^2 M(M+1)} O_p(M^3 a_M) + \frac{2}{a_n^2} \sum_{h=1}^{M-1} \sum_{t=M+2}^{n-M} ((1-L) \sum_{r=1}^{\infty} \tilde{f}_{h+r}(L) u_{t+h-r} u_{t+h}) \\
&= \frac{4}{a_n^2 M(M+1)} O_p(M^3 a_M) + O_p\left(\frac{M}{a_n}\right) \\
&= o_p(1). \tag{6.25}
\end{aligned}$$

类似地,

$$\frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} E_{5,t+j_1,h}}{a_n^2 M(M+1)} = o_p(1). \tag{6.26}$$

对于 $E_{1,t+j_1,h}$, 在给定 $M \geq 1$ 时, 我们有

$$\begin{aligned}
& \frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} E_{1,t+j_1,h}}{a_n^2 M(M+1)} \\
&= \frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} f_h(1) u_{t+j_1}^2}{a_n^2 M(M+1)} \\
&\xrightarrow{d} 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] U_{\frac{\alpha}{2}}(1). \tag{6.27}
\end{aligned}$$

结合 (6.23)、(6.24)、(6.25)、(6.26) 和 (6.27), 我们可以得到 Δ_{n2} 的收敛性:

$$\begin{aligned}
\frac{2\Delta_{n2}}{a_n^2 M(M+1)} &= \frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j_1=1}^k \sum_{h=1}^{k-1} E_{1,t+j_1,h}}{a_n^2 M(M+1)} + o_p(1) \\
&\xrightarrow{d} 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] U_{\frac{\alpha}{2}}(1). \tag{6.28}
\end{aligned}$$

最终, 结合 (6.22) 和 (6.28), 对给定 $M \geq 1$, 我们可以得到 S_{n3} 的收敛性:

$$\frac{2S_{n3}}{a_n^2 M(M+1)} \xrightarrow{d} \left(f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] \right) U_{\frac{\alpha}{2}}(1). \tag{6.29}$$

对 $D_M = f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right]$, 因为 $\sum_{i=1}^{\infty} |f_i(1)| < \infty$, 则对任意 $\nu > 0$, 存在 N 使得 $\sum_{i=N+1}^{\infty} |f_i(1)| < \nu$ 成立。当 $M > \frac{2N}{\nu}$ 时, $1 - \nu < \frac{(1+M-i)(M-i)}{M(M+1)} < 1$ 对 $1 \leq i \leq N$ 成立, 我们有

$$\begin{aligned} D_M &= f_0(1) + 2 \sum_{i=1}^N \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) + 2 \sum_{i=N+1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \\ &\geq f_0(1) + 2 \sum_{i=1}^N f_i(1) - 2\nu \sum_{i=1}^N |f_i(1)| - 2\nu \\ &\geq f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1) - \nu(4 + 2 \sum_{i=1}^N |f_i(1)|), \end{aligned}$$

和 $D_M \leq f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1) + \nu(4 + 2 \sum_{i=1}^N |f_i(1)|)$ 。此外, Phillips and Solo (1992) 的引理 5.19 证明了 $f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1) < \infty$ 。由 ν 的任意性, 我们有

$$\lim_{M \rightarrow \infty} D_M = f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1)。$$

6.7.4 T_n 的收敛性

对 $\hat{\gamma}(0)$, 我们有

$$\frac{\hat{\gamma}(0)}{a_n^2} = \frac{1}{na_n^2} \sum_{t=1}^n Y_t^2 \xrightarrow{d} \omega^2 \int_0^1 U_{\alpha}^2(t) dt, \quad (6.30)$$

更多细节参见 Phillips (1990) 的定理 2.1。

对给定的 $M \geq 1$, 记

$$S_n = \left(\frac{2S_{n1}}{a_n^2 M(M+1)}, \frac{2S_{n2}}{a_n^2 M(M+1)}, \frac{2S_{n3}}{a_n^2 M(M+1)}, \frac{\hat{\gamma}(0)}{a_n^2} \right)$$

和

$$Z = \left(\omega^2 U_{\alpha}^2(1), 0, \left(f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] \right) U_{\frac{\alpha}{2}}(1), \omega^2 \int_0^1 U_{\alpha}^2(t) dt \right),$$

我们需要证明

$$S_n \xrightarrow{d} Z。$$

令 $c = (c_1, c_2, c_3, c_4)' \in \mathbf{R}^4$, 由 Chan and Zhang (2009) 的引理 3.3, 我们有

$$\begin{aligned}
& c_1 \frac{2S_{n1}}{a_n^2 M(M+1)} + c_2 \frac{2S_{n2}}{a_n^2 M(M+1)} + c_3 \frac{2S_{n3}}{a_n^2 M(M+1)} + c_4 \frac{\hat{\gamma}(0)}{a_n^2} \\
= & c_1 \frac{Y_n^2}{a_n^2} + c_3 \frac{2 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k D_{1,t+j}}{a_n^2 M(M+1)} + c_3 \frac{4 \sum_{k=1}^M \sum_{t=1}^{n-k} \sum_{j=1}^k \sum_{h=1}^{k-1} E_{1,t+j_1,h}}{a_n^2 M(M+1)} \\
& + c_4 \frac{1}{na_n^2} \sum_{t=1}^n Y_t^2 + o_p(1) \\
\stackrel{d}{\rightarrow} & c_1 \omega^2 U_\alpha^2(1) + c_3 \left(f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] \right) U_{\frac{\alpha}{2}}(1) \\
& + c_4 \omega^2 \int_0^1 U_\alpha^2(t) dt.
\end{aligned}$$

由 Cramér-Wold 定理, 我们有 $S_n \xrightarrow{d} Z$ 成立。

因此, 由 (6.13)、(6.14) 和 (6.29), 对给定 $M \geq 1$, 我们有

$$\begin{aligned}
& \frac{2(S_{n1} + S_{n2} + S_{n3})}{a_n^2 M(M+1)} \\
\stackrel{d}{\rightarrow} & \omega^2 U_\alpha^2(1) + \left(f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] \right) U_{\frac{\alpha}{2}}(1). \quad (6.31)
\end{aligned}$$

并且由等式 (6.30) 和 (6.31),

$$\begin{aligned}
\frac{2(S_{n1} + S_{n2} + S_{n3})}{M(M+1)\hat{\gamma}(0)} & \stackrel{d}{\rightarrow} \frac{\omega^2 U_\alpha^2(1) + \left(f_0(1) + 2 \left[\sum_{i=1}^{M-1} \frac{(1+M-i)(M-i)}{M(M+1)} f_i(1) \right] \right) U_{\frac{\alpha}{2}}(1)}{\omega^2 \int_0^1 U_\alpha^2(t) dt} \\
& = \frac{\omega^2 U_\alpha^2(1) + D_M U_{\frac{\alpha}{2}}(1)}{\omega^2 \int_0^1 U_\alpha^2(t) dt}.
\end{aligned}$$

类似地, 当 $M = M(n) \rightarrow \infty$ 且 $M = o(a_n \wedge n)$ 时, 记 $Z^* = \left(\omega^2 U_\alpha^2(1), 0, [f_0(1) + 2 \sum_{i=1}^\infty f_i(1)] U_{\frac{\alpha}{2}}(1), \omega^2 \int_0^1 U_\alpha^2(t) dt \right)$, 对 $c = (c_1, c_2, c_3, c_4)' \in \mathbf{R}^4$, 我们有

$$\begin{aligned}
& c_1 \frac{2S_{n1}}{a_n^2 M(M+1)} + c_2 \frac{2S_{n2}}{a_n^2 M(M+1)} + c_3 \frac{2S_{n3}}{a_n^2 M(M+1)} + c_4 \frac{\hat{\gamma}(0)}{a_n^2} \\
\stackrel{d}{\rightarrow} & c_1 \omega^2 U_\alpha^2(1) + c_3 [f_0(1) + 2 \sum_{i=1}^\infty f_i(1)] U_{\frac{\alpha}{2}}(1) + c_4 \omega^2 \int_0^1 U_\alpha^2(t) dt.
\end{aligned}$$

由 Cramér-Wold 定理, 当 $M = M(n) \rightarrow \infty$ 和 $M = o(a_n \wedge n)$ 时, 我们有 $S_n \xrightarrow{d} Z^*$ 成立。因此, 当 $M = M(n) \rightarrow \infty$ 和 $M = o(a_n \wedge n)$ 时,

$$\frac{2(S_{n1} + S_{n2} + S_{n3})}{a_n^2 M(M+1)} \stackrel{d}{\rightarrow} \omega^2 U_\alpha^2(1) + [f_0(1) + 2 \sum_{i=1}^\infty f_i(1)] U_{\frac{\alpha}{2}}(1). \quad (6.32)$$

因此，由等式 (6.30) 和 (6.32) 可得：

$$\frac{2(S_{n1} + S_{n2} + S_{n3})}{M(M+1)\hat{\gamma}(0)} \xrightarrow{d} \frac{\omega^2 U_\alpha^2(1) + [f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1)] U_{\frac{\alpha}{2}}(1)}{\omega^2 \int_0^1 U_\alpha^2(t) dt}.$$

结合 (6.7) 和 (6.8)，对任意 $M \geq 1$ ，我们有

$$T_n = -\frac{S_{n1} + S_{n2} + S_{n3}}{2M(M+1)\hat{\gamma}(0)} \xrightarrow{d} -\frac{\omega^2 U_\alpha^2(1) + D_M U_{\frac{\alpha}{2}}(1)}{4\omega^2 \int_0^1 U_\alpha^2(t) dt}, \quad (6.33)$$

即定理 6.1 的第一个结论成立；当 $M = M(n) \rightarrow \infty$ 和 $M = (a_n \wedge n)$ 时，

$$T_n = -\frac{S_{n1} + S_{n2} + S_{n3}}{2M(M+1)\hat{\gamma}(0)} \xrightarrow{d} -\frac{\omega^2 U_\alpha^2(1) + [f_0(1) + 2 \sum_{i=1}^{\infty} f_i(1)] U_{\frac{\alpha}{2}}(1)}{4\omega^2 \int_0^1 U_\alpha^2(t) dt}, \quad (6.34)$$

即定理 6.1 的第二个结论成立。 □

7 总结与展望

重尾现象经常出现在许多科学领域，例如金融和经济学，相应的重尾理论也在实际问题中起着越来越重要的作用。本文主要研究了基于重尾时间序列的非参数假设检验。

本文的前半部分主要研究了在重尾 ARMA-GARCH 模型中的假设检验。在对 ARMA-GARCH 模型进行统计推断时，由于一般常用的估计方法 QMLE 不能保证在重尾情况下有渐近正态性成立，所以我们经常采用 SWQMELE 方法。无论误差的四阶矩是否存在，此方法都可以保证渐近正态性成立。然而，由于 SWQMELE 要求误差的中位数为 0，而在经典的 ARMA-GARCH 模型及 QMLE 中，我们一般假设均值为 0。为了保证零中位数的假设成立，需要采取的模型变换会导致 ARMA-GARCH 结构不再成立。因此，我们需要在应用 SWQMELE 前检验其假设是否成立。在第 3 章，我们考虑了应用 profile 经验似然方法在误差中位数为 0、绝对矩为 1 的假设下检验误差均值是否为 0，以此来检验 SWQMELE 的条件是否满足。在第 4 章，为了提高检验的稳健性，我们利用 GARCH 结构而无需估计 GARCH 模型，在误差均值为 0、方差为 1 的假设下，对误差的中位数是否为 0 进行检验。为了进一步提高零均值检验的稳健性，在第 5 章中，我们结合 LAD 和随机加权自助抽样法，在误差中位数为 0、二阶矩为 1 的假设下，进行零均值检验。针对以上提出的三种检验方法，我们都考虑了原假设和局部备择假设下的渐近理论。相应的模拟研究和实证分析都表明，我们的检验表现良好，且在应用 SWQMELE 模型前进行假设检验是非常有必要的。

然而，我们的研究也存在一些局限性。在现实中，我们不仅仅会对单个序列感兴趣，往往会对一组序列进行研究，如何利用序列之间的共同波动性来做假设检验，即如何将我们对单个 ARMA-GARCH 模型的检验拓展到多维的情况也值得进一步的研究。

本文前半部分的研究主要针对 ARMA-GARCH 模型中误差四阶矩不存在但方差存在的情形。为了进一步研究方差无限时的情形，我们考虑了检验带有无限方差噪声的自回归模型中单位根的存在性，即当噪声项具有 $\varepsilon_t = \sum_{j=0}^{\infty} d_j \eta_{t-j}$ 的形式，其中 $\{\eta_t\}$ 独立同分布，且属于 α -stable 分布吸引域，其中 $\alpha \in (0, 2)$ 。我们扩展了 Zhang and Chan (2018) 中的 Portmanteau 型检验，得到了该检验统计量的渐近分布。模拟研究和实证分析都表明，我们的检验优于传统的 PP、DF 和 ADF 检验。然而，我们的方法还可以拓展到很多有趣的情形。首先，我们只对 AR(1) 单位根模型进行了

检验，然而在实际中，经常存在 $AR(p)$ 的情况。如何将我们的检验拓展到高阶自回归的设定是一个值得考虑的研究方向。此外，我们只在短记忆过程的情形下研究了 Portmanteau 型单位根检验，如何在长记忆的重尾噪声的情形下进行假设检验也是目前的研究热点之一，这些我们将在未来的工作中考虑。

参考文献

- [1] Ahn, S. K., Fotopoulos, S. B. and He, L. (2001). Unit root tests with infinite variance errors[J]. *Econometric Reviews* **20**: 461–483.
- [2] Arcones, M. A. and Giné, E. (1989). The bootstrap of the mean with arbitrary bootstrap sample size[J]. *Annales de l'I.H.P. Probabilités et Statistiques, Section B* **25**: 457–481.
- [3] Arvanitis, S. (2017). A note on the limit theory of a Dickey–Fuller unit root test with heavy tailed innovations[J]. *Statistics & Probability Letters* **126**: 198–204.
- [4] Avram, F. and Taqqu, M. S. (1992). Weak convergence of sums of moving averages in the α -stable domain of attraction[J]. *The Annals of Probability* **20**: 483–503.
- [5] Basrak B., Davis R.A. and Mikosch T. (2002). Regular variation of GARCH processes[J]. *Stochastic Processes and their Applications* **99**: 95–116.
- [6] Berkes, I. and Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes[J]. *The Annals of Statistics* **32(2)**: 633–655.
- [7] Berkes, I., Horváth, L. and Kokoszka, P. (2003). GARCH processes: structure and estimation[J]. *Bernoulli* **9**: 201–227.
- [8] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity[J]. *Journal of Econometrics* **31**: 307–327.
- [9] Box, G. E. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models[J]. *Journal of the American Statistical Association* **65(332)**: 1509–1526.
- [10] Breiman, L. (1965). On some limit theorems similar to the arc-sin law[J]. *Theory of Probability & Its Applications*, **10**: 323–331.
- [11] Brockwell, P.J. and Davis, R.A. (1991). *Time Series Theory and Methods*, 2nd Edition[M]. Springer, New York.
- [12] Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection[J]. *Sociological Methods & Research* **33(2)**: 261–304.

- [13] Cai, Y. and Shintani, M. (2006). On the alternative long-run variance ratio test for a unit root[J]. *Econometric Theory* **22**: 347–372.
- [14] Carbon, M. and Francq, C. (2011). Portmanteau goodness-of-fit test for asymmetric power GARCH models[J]. *Austrian Journal of Statistics* **40(1&2)**: 55–64.
- [15] Cavaliere, G., Georgiev, I. and Taylor, A. (2018). Unit root inference for non-stationary linear processes driven by infinite variance innovations[J]. *Econometric Theory* **34**: 302–348.
- [16] Chan, N. H. and Ling, S. (2006). Empirical likelihood for models[J]. *Econometric Theory* **22**: 403–428.
- [17] Chan N. H., Peng L. and Zhang R. (2012). Interval estimation of the tail index of a GARCH(1,1) model[J]. *Test* **21**: 546–565.
- [18] Chan, N. H. and Tran, L. T. (1989). On the first-order autoregressive process with infinite variance[J]. *Econometric Theory* **5**: 354–362.
- [19] Chan, N. H. and Zhang, R. M. (2009). Inference for nearly nonstationary processes under strong dependence with infinite variance[J]. *Statistica Sinica* **19**: 925–947.
- [20] Chan, N. H. and Zhang, R. M. (2010). Inference for unit-root models with infinite variance GARCH errors[J]. *Statistica Sinica* **20**: 1363–1393.
- [21] Chen, H. Q., Chong, T. T. L and Bai, J. S. (2012). Theory and applications of TAR model with two threshold variables[J]. *Econometric Reviews* **31**: 142–170.
- [22] Chen, K., Guo, S., Lin, Y. and Ying, Z. (2010). Least absolute relative error estimation[J]. *Journal of the American Statistical Association* **105(491)**: 1104–1112.
- [23] Chen, K., Ying, Z., Zhang, H. and Zhao, L. (2008). Analysis of least absolute deviation[J]. *Biometrika* **95(1)**: 107–122.
- [24] Chuang, C. S. and Chan, N. H. (2002). Empirical likelihood for autoregressive models with applications to unstable time series[J]. *Statistica Sinica* **12**: 387–407.
- [25] Ciuperca, G. and Salloum, Z. (2015). Empirical likelihood test in a posteriori change-point nonlinear model[J]. *Metrika: International Journal for Theoretical and Applied Statistics* **78**: 919–952.
- [26] Cornea-Madeira, A. and Davidson, R. (2015). A parametric bootstrap for heavy-tailed distributions[J]. *Econometric Theory* **31**: 449–470.

- [27] Davis, R. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages[J]. *The Annals of Statistics* **14**: 533–558.
- [28] De Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*[M]. Springer, Berlin.
- [29] Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root[J]. *Journal of the American Statistical Association* **74**: 427–431.
- [30] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation[J]. *Econometrica* **50**: 987–1007.
- [31] Escanciano, J. C. (2008). Joint and marginal specification tests for conditional mean and variance models[J]. *Journal of Econometrics* **143(1)**: 74–87.
- [32] Fan, J. and Huang, T. (2005). Profile likelihood inferences on semiparametric varying-coefficient partially linear models[J]. *Bernoulli* **11**: 1031–1057.
- [33] Fan, J., Qi, L. and Xiu, D. (2014). Quasi maximum likelihood estimation of GARCH models with heavy-tailed likelihoods[J]. *Journal of Business and Economics Statistics* **32**: 178–205.
- [34] Fan, Y. and Gençay, R. (2010). Unit root tests with wavelets[J]. *Econometric Theory* **26**: 1305–1331.
- [35] Feller, W. (1971) *An Introduction to Probability Theory and its Applications, Vol. 2*[M]. John Wiley & Sons.
- [36] Francq, C. and Zakoian, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes[J]. *Bernoulli* **10**: 605–637.
- [37] Georgiev, I., Rodrigues, P. M. and Robert Taylor, A. M. (2017). Unit root tests and heavy-tailed innovations[J]. *Journal of Time Series Analysis* **38**: 733–768.
- [38] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application*[M]. New York: Academic Press.
- [39] Hall, P. and Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors[J]. *Econometrica* **71**: 285–317.
- [40] Hamilton, J. D. (1994). *Time Series Analysis*[M]. Princeton University Press, New York.

- [41] He, Y., Hou, Y., Peng, L. and Shen, H. (2020). Inference for conditional value-at-risk of a predictive regression[J]. *The Annals of Statistics* **8(6)**: 3442–3464.
- [42] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution[J]. *Annals of Statistics* **3**: 1163–1174.
- [43] Hong, Y., and Lee, T. H. (2003). Diagnostic checking for the adequacy of nonlinear time series models[J]. *Econometric Theory* **19(6)**: 1065–1121.
- [44] Horváth, L. and Kokoszka, P. (2003). A bootstrap approximation to a unit root test statistic for heavy-tailed observations[J]. *Statistics & Probability Letters* **62**: 163–173.
- [45] Horváth, L. and Liese, F. (2004). L_p -estimators in ARCH models[J]. *Journal of Statistical Planning and Inference* **119(2)**: 277–309.
- [46] Huang, H., Leng, X., Liu, X. and Peng, L. (2020). Unified inference for an AR process with possible infinite variance GARCH errors[J]. *Journal of Financial Econometrics* **18**: 425–470.
- [47] Huang, H., Peng, L. and Yao, V.W. (2019). Comovements and asymmetric tail dependence in state housing prices in the USA: A nonparametric approach[J]. *Journal of Applied Econometrics* **34**: 843–849.
- [48] Ibragimov, I.A. and Linnik, J.V.(1971). *Independent and Stationary Sequences of Random Variables*[M]. Groningen, Wolters-Noordhoff.
- [49] Jach, A. and Kokoszka, P. (2004). Subsampling unit root tests for heavy-tailed observations[J]. *Methodology and Computing in Applied Probability* **6**: 73–97.
- [50] Jansson, M. (2008). Semiparametric power envelopes for tests of the unit root hypothesis[J]. *Econometrica* **76**, 1103–1142.
- [51] Jansson, M. and Moreira, M. J. (2006). Optimal inference in regression models with nearly integrated regressors[J]. *Econometrica* **74**: 681–714.
- [52] Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models[J]. *Econometric Theory* **14**: 70–86.
- [53] Jin, Z., Ying, Z. and Wei, L. J. (2001). A simple resampling method by perturbing the minimand[J]. *Biometrika* **88(2)**: 381–390.
- [54] Knight, K. (1989). Limit theory for autoregressive-parameter estimates in an infinite-variance random walk[J]. *The Canadian Journal of Statistics* **17**: 261–278.

- [55] Knight, K. (1991). Limit theory for M-estimates in an integrated infinite variance[J]. *Econometric Theory* **7**: 200–212.
- [56] Kolaczyk, E. D. (1994). Empirical Likelihood for generalized linear models[J]. *Statistica Sinica* **4**: 2007–2020.
- [57] Li, G., Leng, C. and Tsai, C. L. (2014). A hybrid bootstrap approach to unit root tests[J]. *Journal of Time Series Analysis* **35(4)**: 299–321.
- [58] Li, G. and Li, W. K. (2005). Diagnostic checking for time series models with conditional heteroscedasticity estimated by the least absolute deviation approach[J]. *Biometrika* **92(3)**: 691–701.
- [59] Li, G. and Li, W. K. (2008). Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity[J]. *Biometrika* **95(2)**: 399–414.
- [60] Li, G., Li, Y. and Tsai, C. L. (2015). Quantile correlations and quantile autoregressive modeling[J]. *Journal of the American Statistical Association* **110(509)**: 246–261.
- [61] Li, J. Y., Liang, W. and He, S. Y. (2012). Empirical likelihood for AR-ARCH models based on LAD estimation[J]. *Acta Mathematicae Applicatae Sinica, English Series* **28(2)**: 371–382.
- [62] Li, W. K. and Mak, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity[J]. *Journal of Time Series Analysis* **15(6)**: 627–636.
- [63] Ling, S. (1999). On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model[J]. *Journal of Applied Probability* **36**: 688–705.
- [64] Ling, S. (2007). Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models[J]. *Journal of Econometrics* **140**: 849–873.
- [65] Ling, S. and Li, W.K. (1997a). On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity[J]. *Journal of the American Statistical Association* **92**: 1184–1194.
- [66] Ling, S. and Li, W. K. (1997b). Diagnostic checking of nonlinear multivariate time series with multivariate ARCH errors[J]. *Journal of Time Series Analysis* **18(5)**: 447–464.

- [67] Ling, S. and McAleer, M. (2003). Asymptotic theory for a vector ARMA-GARCH model[J]. *Econometric Theory* **19**: 280–310.
- [68] Ling, S. and Tong, H. (2011). Score based goodness-of-fit tests for time series[J]. *Statistica Sinica* **21**:1807–1829.
- [69] Liu, Y., Zou, L. and Zhang, R. (2008). Empirical likelihood ratio test for a change-point in linear regression model[J]. *Communications in Statistics - Theory and Methods* **37**: 2551–2563.
- [70] Ljung, G. M. and Box, G. E. (1978). On a measure of lack of fit in time series models[J]. *Biometrika* **65(2)**: 297–303.
- [71] Mikosch T. and Starica C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) sequence[J]. *Annals of Statistics* **28**: 1427–1451.
- [72] Mykland, P. A. (1995). Dual likelihood[J]. *Annals of Statistics* **23(2)**: 396–421.
- [73] Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional[J]. *Biometrika* **75**: 237–249.
- [74] Owen, A. B. (1990). Empirical likelihood ratio confidence regions[J]. *Annals of Statistics* **18**: 90–120.
- [75] Owen, A. B. (1991). Empirical likelihood for linear models[J]. *Annals of Statistics* **19**: 1725–1747.
- [76] Owen, A.B. (2001). *Empirical Likelihood*[M]. Chapman Hall.
- [77] Pan, J., Wang, H. and Yao, Q. (2007). Weighted least absolute deviations estimation for ARMA models with infinite variance[J]. *Econometric Theory* **23**: 852–879.
- [78] Peng, L. and Yao, Q. (2003). Least absolute deviations estimation for ARCH and GARCH models[J]. *Biometrika* **90(4)**: 967–975.
- [79] Phillips, P. C. (1990). Time series regression with a unit root and infinite-variance errors[J]. *Econometric Theory* **6**: 44–62.
- [80] Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression[J]. *Biometrika* **75**: 335–346.
- [81] Phillips, P. C. and Solo, V. (1992). Asymptotics for linear processes[J]. *The Annals of Statistics* **20**: 971–1001.

- [82] Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations[J]. *Annals of Statistics* **22**: 300–325.
- [83] Resnick, S. and Starica, C. (1995). Consistency of Hill’s estimator for dependent data[J]. *Journal of Applied probability* **32(1)**: 139–167.
- [84] Resnick, S. and Starica, C. (1997). Asymptotic behavior of Hill’s estimator for autoregressive data[J]. *Communications in Statistics - Stochastic Models* **13(4)**: 703–721.
- [85] Romano, J. P. and Wolf, M. (1999). Subsampling inference for the mean in the heavy-tailed case[J]. *Metrika* **50**: 55–69.
- [86] Rootzen, H., Leadbetter, M. R. and De Haan, L. (1990). Tail and quantile estimation for strongly mixing stationary sequences[J]. *North Carolina Univ at Chapel Hill Center for Stochastic Processes*.
- [87] Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order[J]. *Biometrika* **71**: 599–607.
- [88] Samarakoon, D. M. and Knight, K. (2009). A note on unit root tests with infinite variance noise[J]. *Econometric Reviews* **28**: 314–334.
- [89] Sargan, J. D. and Bhargava, A. (1983). Testing residuals from least squares regression for being generated by the Gaussian random walk[J]. *Econometrica* **51**: 153–174.
- [90] Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis[J]. *Annals of Statistics* **14(4)**: 1261–1295.
- [91] Zhang, R. and Chan, N. H. (2018). Portmanteau-type tests for unit-root and cointegration[J]. *Journal of Econometrics* **207**: 307–324.
- [92] Zhang, R. and Chan, N. H. (2020). Nonstationary linear processes with infinite variance GARCH errors[J]. *Econometric Theory*: 1–34.
- [93] Zhang R., Li C. and Peng L. (2019). Inference for the tail index of a GARCH(1,1) model and an AR(1) model with ARCH(1) errors[J]. *Econometric Reviews* **38**: 151–169.
- [94] Zhang, R. and Ling, S. (2015). Asymptotic inference for AR models with heavy-tailed G-GARCH noises[J]. *Econometric Theory* **31(4)**: 880–890.

- [95] Zhu, K. (2016). Bootstrapping the Portmanteau tests in weak auto-regressive moving average models[J]. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **78(2)**: 463–485.
- [96] Zhu, K. and Ling, S. (2011). Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA–GARCH/IGARCH models[J]. *Annals of Statistics* **39**: 2131–2163.
- [97] Zhu, K. and Ling, S. (2015). LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises[J]. *Journal of the American Statistical Association* **110(510)**: 784–794.
- [98] Zimmer, D. M. (2012). The role of copulas in the housing crisis[J]. *Review of Economics and Statistics* **94**: 607-620.
- [99] Zimmer, D. M. (2015). Asymmetric dependence in house prices: evidence from USA and international data[J]. *Empirical Economics* **49**: 161–183.

附录

第 3 章证明

引理 3.2 的证明:

(a) 为了证明 $E \sup_{\Theta_0} \|D_t(\boldsymbol{\theta}, \mu_0)\|^2 < \infty$, 注意到

$$E \sup_{\Theta_0} \|D_t(\boldsymbol{\theta}, \mu_0)\|^2 = E \sup_{\Theta_0} \left(\|D_{t,1}(\boldsymbol{\theta})\|^2 + |D_{t,2}(\boldsymbol{\theta}, \mu_0)|^2 \right).$$

因此, 只需要证明 $E \sup_{\Theta_0} \|D_{t,1}(\boldsymbol{\theta})\|^2 < \infty$ 和 $E \sup_{\Theta_0} |D_{t,2}(\boldsymbol{\theta}, \mu_0)|^2 < \infty$ 即可。

首先, 我们有

$$h_t(\boldsymbol{\theta}) = \alpha_0 \beta^{-1}(1) + \sum_{i=1}^{\infty} a_{\delta}(i) \varepsilon_{t-i}^2(\boldsymbol{\gamma}),$$

其中, 序列 $\{a_{\delta}(i)\}$ 对某个 $\rho \in (0, 1)$ 满足 $\alpha(z)\beta^{-1}(z) = \sum_{i=1}^{\infty} a_{\delta}(i)z^i$ 和 $\sup_{\Theta_{\delta}} a_{\delta}(i) = O(\rho^i)$ 。因此,

$$h_t(\boldsymbol{\theta}) \geq \alpha_0 \beta^{-1}(1) \quad \text{且} \quad \beta^{-1}(1) = \frac{1}{1 - \sum_{i=1}^s \beta_i} > 0,$$

表明 $\frac{h_t(\boldsymbol{\theta})}{\alpha_0 \beta^{-1}(1)} \geq 1$ 和 $\alpha_0 > 0$ 。

由引理 3.1 (e)、(f) 和 (i), 我们有

$$\begin{aligned} \sup_{\Theta_0} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right\|^2 &\leq \sup_{\Theta_0} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}'} \right\|^2 + \sup_{\Theta_0} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}'} \right\|^2 \\ &\leq \frac{C^2}{\alpha_0} \xi_{\rho t-1}^2 + C^2 \xi_{\rho t-1}^{2\iota_1}. \end{aligned} \quad (\text{A.1})$$

由引理 3.1 (参考 Ling (2007) 的 (A.13)), 我们有 $\sup_{\Theta_0} |\varepsilon_t(\boldsymbol{\gamma})| \leq C_1 \xi_{\rho t-1} (1 + |\eta_t|)$ 。

再由引理 3.1(e) 可得, 存在常数 C 使得

$$\sup_{\Theta_0} |\eta_t(\boldsymbol{\theta})| = \sup_{\Theta_0} \frac{|\varepsilon_t(\boldsymbol{\gamma})|}{\sqrt{h_t(\boldsymbol{\theta})}} \leq C \xi_{\rho t-1} (1 + |\eta_t|). \quad (\text{A.2})$$

对 $\mathbf{D}_{t,1}(\boldsymbol{\theta})$ 来说, 由引理 3.1 (b)-(e) 和 (A.1), 我们有

$$\begin{aligned}
& E \sup_{\Theta_0} \|\mathbf{D}_{t,1}(\boldsymbol{\theta})\|^2 \\
&= E \sup_{\Theta_0} \left\| w_t \frac{1}{2h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (1 - |\eta_t(\boldsymbol{\theta})|) + w_t \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \boldsymbol{\theta}} \text{sgn}(\eta_t(\boldsymbol{\theta})) \right\|^2 \\
&\leq E \sup_{\Theta_0} \left\| w_t \frac{1}{2h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (1 - |\eta_t(\boldsymbol{\theta})|) \right\|^2 + E \sup_{\Theta_0} \left\| w_t \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \boldsymbol{\theta}} \text{sgn}(\eta_t(\boldsymbol{\theta})) \right\|^2 \\
&\quad + E \sup_{\Theta_0} \left\| w_t^2 (1 - |\eta_t(\boldsymbol{\theta})|) \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \boldsymbol{\theta}} \text{sgn}(\eta_t(\boldsymbol{\theta})) \right\|^2 \\
&\leq C^2 E \left(\sup_{\Theta_0} \frac{w_t^2}{4} (1 - |\eta_t(\boldsymbol{\theta})|)^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) + \frac{C^2}{\underline{\alpha}_0} E(w_t^2 \xi_{\rho t-1}^2) \\
&\quad + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E \left(\sup_{\Theta_0} w_t^2 (1 + |\eta_t(\boldsymbol{\theta})|) \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) \\
&\leq C^2 E \left(\sup_{\Theta_0} \frac{w_t^2}{4} (1 + |\eta_t(\boldsymbol{\theta})|^2) \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) \\
&\quad + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E \left(\sup_{\Theta_0} w_t^2 |\eta_t(\boldsymbol{\theta})| \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) \\
&\quad + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E \left(w_t^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) + \frac{C^2}{\underline{\alpha}_0} E(w_t^2 \xi_{\rho t-1}^2) \\
&\leq \frac{C^2}{4} E \left(\sup_{\Theta_0} w_t^2 |\eta_t(\boldsymbol{\theta})|^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E \left(\sup_{\Theta_0} w_t^2 |\eta_t(\boldsymbol{\theta})| \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) \\
&\quad + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E \left(w_t^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) + \frac{5C^2}{4\underline{\alpha}_0} E(w_t^2 \xi_{\rho t-1}^2) + \frac{C^2}{4} E(w_t^2 \xi_{\rho t-1}^{2\iota_1}). \tag{A.3}
\end{aligned}$$

因为 $\{\eta_t\}$ 与 w_t 和 $\xi_{\rho t-1}$ 独立, 根据 (A.2)、Cauchy-Schwarz 不等式和假设 3.5 里的 $E|\eta_t| = 1$, 我们有

$$\begin{aligned}
& \frac{C^2}{4} E \left(\sup_{\Theta_0} w_t^2 |\eta_t(\boldsymbol{\theta})|^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) \\
&\leq \frac{C^2}{4} E \left(w_t^2 C^2 \xi_{\rho t-1}^2 (1 + |\eta_t|)^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) \\
&\leq \frac{C^4}{4} E(1 + |\eta_t|)^2 E \left(w_t^2 \xi_{\rho t-1}^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) \\
&\leq \frac{C^4}{2} (1 + E|\eta_t|^2) E \left(w_t^2 \xi_{\rho t-1}^2 \left(\frac{1}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right) \right) \\
&\leq \frac{C^4}{2\underline{\alpha}_0} (1 + E|\eta_t|^2) E(w_t^2 \xi_{\rho t-1}^4) + \frac{C^4}{2} (1 + E|\eta_t|^2) E(w_t^2 \xi_{\rho t-1}^{2+2\iota_1}),
\end{aligned}$$

和

$$\begin{aligned}
& \frac{C^2}{\sqrt{\alpha_0}} E \left(\sup_{\Theta_0} w_t^2 |\eta_t(\boldsymbol{\theta})| \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) \\
& \leq \frac{C^3}{\sqrt{\alpha_0}} E \left(w_t^2 \xi_{\rho t-1} (1 + |\eta_t|) \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) \\
& \leq \frac{C^3}{\sqrt{\alpha_0}} E(1 + |\eta_t|) E \left(w_t^2 \xi_{\rho t-1}^2 \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right) \\
& = \frac{2C^3}{\sqrt{\alpha_0}} E \left(w_t^2 \xi_{\rho t-1}^2 \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right).
\end{aligned}$$

因此, 结合 (A.3), 再根据假设 3.4, 我们可以得到:

$$\begin{aligned}
& E \sup_{\Theta_0} \|\mathbf{D}_{t,1}(\boldsymbol{\theta})\|^2 \\
& \leq \frac{C^4}{2\alpha_0} (1 + E|\eta_t|^2) E(w_t^2 \xi_{\rho t-1}^4) + \frac{C^4}{2} (1 + E|\eta_t|^2) E(w_t^2 \xi_{\rho t-1}^{2+2\iota_1}) \\
& \quad + \frac{2C^3}{\sqrt{\alpha_0}} E \left(w_t^2 \xi_{\rho t-1}^2 \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right) + \frac{C^2}{\sqrt{\alpha_0}} E \left(w_t^2 \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \xi_{\rho t-1} \right) \\
& \quad + \frac{5C^2}{4\alpha_0} E(w_t^2 \xi_{\rho t-1}^2) + \frac{C^2}{4} E(w_t^2 \xi_{\rho t-1}^{2\iota_1}) \\
& < \infty.
\end{aligned}$$

对 $D_{t,2}(\boldsymbol{\theta}, \mu_0)$ 来说, 由 (A.2)、Cauchy-Schwarz 不等式和假设 3.4 可得:

$$\begin{aligned}
E \sup_{\Theta_0} |D_{t,2}(\boldsymbol{\theta}, \mu_0)|^2 &= E \sup_{\Theta_0} |w_t \eta_t(\boldsymbol{\theta})|^2 \leq C^2 E(w_t^2 \xi_{\rho t-1}^2 (1 + |\eta_t|)^2) \\
&\leq 2C^2 (1 + E|\eta_t|^2) E(w_t^2 \xi_{\rho t-1}^2) < \infty.
\end{aligned}$$

综上, $E \sup_{\Theta_0} \|\mathbf{D}_t(\boldsymbol{\theta}, \mu_0)\|^2 < \infty$.

(b) 注意到 $\mathbf{P}_t(\boldsymbol{\theta}, \mu_0)$ 是 $(p+q+r+s+3) \times (p+q+r+s+3)$ 的矩阵, 且有以下形式:

$$\mathbf{P}_t(\boldsymbol{\theta}, \mu_0) = \begin{pmatrix} \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} & \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \mu} \\ \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}'} & \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu)}{\partial \mu} \Big|_{\mu=\mu_0} \end{pmatrix}, \quad (\text{A.4})$$

其中 $\frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$ 是一个 $(p+q+r+s+2) \times (p+q+r+s+2)$ 的矩阵, $\frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \mu}$ 是一个 $(p+q+r+s+2)$ 维的向量, $\frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}'}$ 是 $(p+q+r+s+2)$ 维的横向量, $\frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu)}{\partial \mu} \Big|_{\mu=\mu_0}$ 是一个标量。要证明 $E \sup_{\Theta_0} \|\mathbf{P}_t(\boldsymbol{\theta}, \mu_0)\| < \infty$, 只需要证明 $E \sup_{\Theta_0} \left\| \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right\| < \infty$, $E \sup_{\Theta_0} \left\| \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \mu} \right\| < \infty$, $E \sup_{\Theta_0} \left\| \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}'} \right\| < \infty$ 和 $E \sup_{\Theta_0} \left\| \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu)}{\partial \mu} \right\|_{\mu=\mu_0} < \infty$ 。

对于 $\frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$, 注意到 $\frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$, 其中

$$\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{pmatrix} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \gamma \partial \gamma'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \gamma \partial \delta'} \\ \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \delta \partial \gamma'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \delta \partial \delta'} \end{pmatrix}.$$

因为 $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$ 存在且连续, 要证明 $E \sup_{\Theta_0} \left\| \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right\| < \infty$, 我们只需要证明

$$E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \gamma \partial \gamma'} \right\| < \infty, \quad E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \delta \partial \delta'} \right\| < \infty \quad \text{和} \quad E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \delta \partial \gamma'} \right\| < \infty.$$

首先,

$$\begin{aligned} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \gamma \partial \gamma'} \\ = & -\frac{1}{2h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma'} (1 - |\eta_t(\boldsymbol{\theta})|) \\ & + \frac{1}{2h_t(\boldsymbol{\theta})} \left[\frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \gamma \partial \gamma'} (1 - |\eta_t(\boldsymbol{\theta})|) - \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma} \text{sgn}(\eta_t(\boldsymbol{\theta})) \left(\frac{\partial \varepsilon_t(\gamma)}{\partial \gamma'} \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \right. \right. \\ & \left. \left. - \frac{1}{2} \eta_t(\boldsymbol{\theta}) \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma'} \right) \right] - \frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma'} \text{sgn}(\eta_t(\boldsymbol{\theta})) \\ & + \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma \partial \gamma'} \text{sgn}(\eta_t(\boldsymbol{\theta})) \\ = & \frac{1}{4h_t^2(\boldsymbol{\theta})} |\eta_t(\boldsymbol{\theta})| \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma'} + \frac{1}{2h_t(\boldsymbol{\theta})} \left[\frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \gamma \partial \gamma'} - \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma'} \right] (1 - |\eta_t(\boldsymbol{\theta})|) \\ & + \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma \partial \gamma'} \text{sgn}(\eta_t(\boldsymbol{\theta})) - \frac{\text{sgn}(\eta_t(\boldsymbol{\theta}))}{2h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \left[\frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma'} + \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma'} \right]. \end{aligned}$$

则根据引理 3.1 (b)、(c)、(e)、(i) 和 (j) 可得：

$$\begin{aligned}
& E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}'} \right\| \\
& \leq E \left(\sup_{\Theta_0} \frac{w_t |\eta_t(\boldsymbol{\theta})|}{4} \frac{1}{\underline{\alpha}_0} \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} \right\|^2 \right) \\
& \quad + E \left(\sup_{\Theta_0} w_t \left| \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2} \right| \left(\frac{1}{\sqrt{\underline{\alpha}_0}} \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}'} \right\| + \frac{1}{\underline{\alpha}_0} \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} \right\|^2 \right) \right) \\
& \quad + E \left(\sup_{\Theta_0} \frac{w_t}{\sqrt{\underline{\alpha}_0}} \left\| \frac{\partial^2 \varepsilon_t(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}'} \right\| \right) + E \left(\sup_{\Theta_0} \frac{w_t}{\underline{\alpha}_0} \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} \right\| \left\| \frac{\partial \varepsilon_t(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} \right\| \right) \\
& \leq \frac{C^2}{4\underline{\alpha}_0} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^2) + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(\sup_{\Theta_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) \xi_{\rho t-1}) \\
& \quad + \frac{C^2}{2\underline{\alpha}_0} E(\sup_{\Theta_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) \xi_{\rho t-1}^2) + \frac{C}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) + \frac{C^2}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) \\
& = \frac{3C^2}{4\underline{\alpha}_0} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^2) + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) \\
& \quad + \frac{3C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) + \frac{3C^2}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2). \tag{A.5}
\end{aligned}$$

因为 $\{\eta_t\}$ 与 w_t 和 $\xi_{\rho t-1}$ 独立，结合 (A.2) 和假设 3.5，我们有

$$E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^2) \leq CE(w_t (1 + |\eta_t|) \xi_{\rho t-1}^3) = 2CE(w_t \xi_{\rho t-1}^3),$$

和

$$E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) \leq CE(w_t (1 + |\eta_t|) \xi_{\rho t-1}^2) = 2CE(w_t \xi_{\rho t-1}^2).$$

因此，利用 (A.5) 和假设 3.4 可得：

$$\begin{aligned}
& E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}'} \right\| \\
& \leq \frac{3C^3}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^3) + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^2) + \frac{3C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) + \frac{3C^2}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) \\
& = \frac{3C^3}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^3) + \frac{(2\underline{\alpha}_0 + 3\sqrt{\underline{\alpha}_0})C^2}{2\underline{\alpha}_0\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^2) + \frac{3C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) \\
& < \infty.
\end{aligned}$$

同理，

$$\begin{aligned}
& \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \\
& = \frac{1}{4h_t^2(\boldsymbol{\theta})} |\eta_t(\boldsymbol{\theta})| \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}'} + \frac{1}{2h_t(\boldsymbol{\theta})} \left[\frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} - \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}'} \right] (1 - |\eta_t(\boldsymbol{\theta})|),
\end{aligned}$$

由引理 3.1 (f)-(g) 可得：

$$\begin{aligned}
& E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \right\| \\
& \leq E \left(\sup_{\Theta_0} \frac{w_t |\eta_t(\boldsymbol{\theta})|}{4} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \right\|^2 \right) \\
& \quad + E \left[\sup_{\Theta_0} w_t \left| \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2} \right| \left(\left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \right\| + \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \right\|^2 \right) \right] \\
& \leq \frac{C^2}{4} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E \left(\sup_{\Theta_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) (\xi_{\rho t-1}^{\iota_1} + C \xi_{\rho t-1}^{2\iota_1}) \right) \\
& = \frac{3C^2}{4} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1}) \\
& \quad + \frac{C}{2} E(w_t \xi_{\rho t-1}^{\iota_1}) + \frac{C^2}{2} E(w_t \xi_{\rho t-1}^{2\iota_1}). \tag{A.6}
\end{aligned}$$

根据 (A.2) 和假设 3.5，很容易可以看出：

$$E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{2\iota_1}) \leq C E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{2\iota_1+1}) = 2C E(w_t \xi_{\rho t-1}^{2\iota_1+1}),$$

和

$$E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1}) \leq C E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{\iota_1+1}) = 2C E(w_t \xi_{\rho t-1}^{\iota_1+1}).$$

因此，由 (A.6) 和假设 3.4，我们有

$$\begin{aligned}
E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \right\| & \leq \frac{3C^3}{2} E(w_t \xi_{\rho t-1}^{2\iota_1+1}) + C^2 E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2} E(w_t \xi_{\rho t-1}^{\iota_1}) + \frac{C^2}{2} E(w_t \xi_{\rho t-1}^{2\iota_1}) \\
& < \infty.
\end{aligned}$$

最后，

$$\begin{aligned}
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\gamma}'} & = \frac{1}{4h_t^2(\boldsymbol{\theta})} |\eta_t(\boldsymbol{\theta})| \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}'} + \frac{1}{2h_t(\boldsymbol{\theta})} \left[\frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\gamma}'} - \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}'} \right] \\
& \quad \times (1 - |\eta_t(\boldsymbol{\theta})|) - \frac{\text{sgn}(\eta_t(\boldsymbol{\theta}))}{2h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \left[\frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \frac{\partial \varepsilon_t(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}'} \right],
\end{aligned}$$

根据引理 3.1 (b)、(e)、(f)、(i) 和 (l)，我们有

$$\begin{aligned}
& E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\gamma}'} \right\| \\
& \leq E \left[\sup_{\Theta_0} \frac{w_t |\eta_t(\boldsymbol{\theta})|}{4\sqrt{\alpha_0}} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \right\| \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} \right\| \right] \\
& \quad + E \left[\sup_{\Theta_0} w_t \left| \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2\sqrt{\alpha_0}} \right| \left(\left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\gamma}'} \right\| + \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \right\| \left\| \frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} \right\| \right) \right] \\
& \quad + E \left(\sup_{\Theta_0} \frac{w_t}{2\sqrt{\alpha_0}} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \right\| \left\| \frac{\partial \varepsilon_t(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} \right\| \right) \\
& \leq \frac{C^2}{4\sqrt{\alpha_0}} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+1}) + \frac{1}{2\sqrt{\alpha_0}} E \left(\sup_{\Theta_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) (C \xi_{\rho t-1} + C^2 \xi_{\rho t-1}^{\iota_1+1}) \right) \\
& \quad + \frac{C^2}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) \\
& \leq \frac{3C^2}{4\sqrt{\alpha_0}} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\alpha_0}} E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) \\
& \quad + \frac{C}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) + \frac{C^2}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}). \tag{A.7}
\end{aligned}$$

利用 (A.2) 和假设 3.5，我们可以推出

$$E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+1}) \leq CE(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{\iota_1+2}) = 2CE(w_t \xi_{\rho t-1}^{\iota_1+2}),$$

和

$$E(\sup_{\Theta_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) \leq CE(w_t (1 + |\eta_t|) \xi_{\rho t-1}^2) = 2CE(w_t \xi_{\rho t-1}^2).$$

因此, 由 (A.7) 和假设 3.4，我们可得：

$$\begin{aligned}
E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\gamma}'} \right\| & \leq \frac{3C^3}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \frac{C^2}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^2) \\
& \quad + \frac{C}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) + \frac{C^2}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) \\
& < \infty.
\end{aligned}$$

综上所述，我们可以得到：

$$E \sup_{\Theta_0} \left\| \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right\| = E \sup_{\Theta_0} \left\| w_t \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| < \infty.$$

显然， $\frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \mu} = 0$ 可推出 $E \sup_{\Theta_0} \left\| \frac{\partial \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \mu} \right\| < \infty$ 。

对于 $D_{t,2}(\boldsymbol{\theta}, \mu_0)$, 由引理 3.1 (b)-(e), 我们有

$$\begin{aligned}
& E \sup_{\boldsymbol{\theta}_0} \left\| \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}'} \right\| \\
&= E \sup_{\boldsymbol{\theta}_0} \left\| w_t \frac{\partial \eta_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right\| \\
&= E \sup_{\boldsymbol{\theta}_0} \left\| w_t \left(\frac{1}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\boldsymbol{\gamma})}{\partial \boldsymbol{\theta}'} - \frac{\eta_t(\boldsymbol{\theta})}{2h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right) \right\| \\
&\leq E \left(\sup_{\boldsymbol{\theta}_0} \frac{w_t}{\sqrt{\alpha_0}} \left\| \frac{\partial \varepsilon_t(\boldsymbol{\gamma})}{\partial \boldsymbol{\theta}'} \right\| \right) + E \left(\sup_{\boldsymbol{\theta}_0} w_t \frac{|\eta_t(\boldsymbol{\theta})|}{2} \left\| \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} \right\| \right) \\
&\leq \frac{C}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) + \frac{C}{2} E \left(\sup_{\boldsymbol{\theta}_0} w_t |\eta_t(\boldsymbol{\theta})| \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right). \quad (\text{A.8})
\end{aligned}$$

由 (A.2) 和假设 3.5,

$$\begin{aligned}
E \left(\sup_{\boldsymbol{\theta}_0} w_t |\eta_t(\boldsymbol{\theta})| \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right) &\leq CE \left(w_t \xi_{\rho t-1} (1 + |\eta_t|) \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right) \\
&= 2CE \left(w_t \xi_{\rho t-1} \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right),
\end{aligned}$$

再结合假设 3.4, 可以推出:

$$\begin{aligned}
& E \sup_{\boldsymbol{\theta}_0} \left\| \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}'} \right\| \\
&\leq \frac{C}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) + C^2 E \left(w_t \xi_{\rho t-1} \left(\frac{1}{\alpha_0} \xi_{\rho t-1}^2 + \xi_{\rho t-1}^{2\iota_1} \right)^{\frac{1}{2}} \right) \\
&< \infty.
\end{aligned}$$

此外,

$$E \sup_{\boldsymbol{\theta}_0} \left| \frac{\partial D_{t,2}(\boldsymbol{\theta}, \mu)}{\partial \mu} \right|_{\mu=\mu_0} = E \sup_{\boldsymbol{\theta}_0} \left| w_t \frac{\partial(\eta_t(\boldsymbol{\theta}) - \mu)}{\partial \mu} \right|_{\mu=\mu_0} = E w_t < \infty.$$

因此, 由 (A.4), 我们有

$$E \sup_{\boldsymbol{\theta}} \| \boldsymbol{P}_t(\boldsymbol{\theta}, \mu_0) \| < \infty,$$

即引理 3.2 (b) 成立。 □

引理 3.3 的证明:

(a) 由切贝雪夫不等式可知, 当 $n \rightarrow \infty$ 时, 对于任意 $\varepsilon > 0$,

$$\begin{aligned} & P\left(\max_{1 \leq t \leq n} \sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| \geq n\varepsilon^2\right) \\ & \leq \frac{1}{n\varepsilon^2} \sum_{t=1}^n E\left(\sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| I\left(\sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| > n\varepsilon^2\right)\right) \\ & \leq \frac{1}{\varepsilon^2} \max_{1 \leq t \leq n} E\left(\sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| I\left(\sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| > n\varepsilon^2\right)\right) \rightarrow 0, \end{aligned}$$

其中, 最后一个不等式利用 $E(\sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\|) < \infty$ 可得。因此, $\max_{1 \leq t \leq n} \sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| = o_p(n)$ 。

(b) 对任意 $\varepsilon > 0$, 当 $n \rightarrow \infty$ 时,

$$\begin{aligned} & P\left(\max_{1 \leq t \leq n} \|D_t(\boldsymbol{\theta}_0, \mu_0)\| \geq \sqrt{n}\varepsilon\right) \\ & \leq \frac{1}{n\varepsilon^2} \sum_{t=1}^n E\left(\|D_t(\boldsymbol{\theta}_0, \mu_0)\|^2 I(\|D_t(\boldsymbol{\theta}_0, \mu_0)\| > \sqrt{n}\varepsilon)\right) \\ & \leq \frac{1}{\varepsilon^2} \max_{1 \leq t \leq n} E\left(\|D_t(\boldsymbol{\theta}_0, \mu_0)\|^2 I(\|D_t(\boldsymbol{\theta}_0, \mu_0)\| > \sqrt{n}\varepsilon)\right) \rightarrow 0, \end{aligned}$$

其中, 最后一个不等式可以根据 $E\|D_t(\boldsymbol{\theta}_0, \mu_0)\|^2 < \infty$ 得出。因此,

$$\max_{1 \leq t \leq n} \|D_t(\boldsymbol{\theta}_0, \mu_0)\| = o_p(\sqrt{n}).$$

再次利用泰勒展开可得:

$$\frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \sup_{V_0} \|D_t(\boldsymbol{\theta}, \mu_0)\| \leq \frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \|D_t(\boldsymbol{\theta}_0, \mu_0)\| + \frac{1}{n} \max_{1 \leq t \leq n} \sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0)\| = o_p(1),$$

因此 (b) 成立。

(c) 首先, 由引理 3.2 和遍历定理, 我们有 $\frac{1}{n} \sum_{t=1}^n P_t(\boldsymbol{\theta}_0, \mu_0) = \boldsymbol{\Psi} + o_p(1)$ 。

另一方面, 我们先证明 $E \sup_{V_0} \left\| \frac{\partial P_t(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}} \right\| < \infty$ 。

对 $i, j, k = 1, 2, \dots, r + s + p + q + 2$, 由引理 3.1、(A.2) 和假设 3.4-3.5, 我们有

$$\begin{aligned}
& E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} \right| \\
= & E \left[\sup_{V_0} w_t \left| \frac{3 \operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{4 h_t^2(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} - \frac{|\eta_t(\boldsymbol{\theta})|}{4 h_t^3(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \right. \right. \\
& + \frac{3 |\eta_t(\boldsymbol{\theta})| - 2}{4 h_t^2(\boldsymbol{\theta})} \left(\frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_i} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \right) \\
& + \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2 h_t(\boldsymbol{\theta})} \frac{\partial^3 h_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} + \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{\sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial^3 \varepsilon_t(\gamma)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} \\
& - \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{2 h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \left(\frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_k} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j} + \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma_i \partial \gamma_j} \right) \\
& + \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{2 h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \left(\frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_k} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_j} + \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_k} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_i} + \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma_j \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_i} + \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma_i \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} \right) \\
& \left. - \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{4 h_t^2(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \left(\frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_i} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_j} + \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_i} \right) \right| \Big] \\
\leq & \frac{3C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^3) + \frac{C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^3) + \frac{3C^2}{4\underline{\alpha}_0} E(\sup_{V_0} w_t (3|\eta_t(\boldsymbol{\theta})| + 2) \xi_{\rho t-1}^2) \\
& + \frac{C}{2\sqrt{\alpha_0}} E(\sup_{V_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) \xi_{\rho t-1}) + \frac{C}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) + \frac{C^2}{\alpha_0} E(w_t \xi_{\rho t-1}^2) \\
& + \frac{2C^2}{\alpha_0} E(w_t \xi_{\rho t-1}^2) + \frac{C^3}{2\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^3) \\
= & \frac{C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^3) + \frac{9C^2}{4\underline{\alpha}_0} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^2) + \frac{C}{2\sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) \\
& + \frac{5C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^3) + \frac{9C^2}{\alpha_0} E(w_t \xi_{\rho t-1}^2) + \frac{3C^2}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
\leq & \frac{C^4}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^4) + \frac{9C^3}{4\underline{\alpha}_0} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^3) + \frac{C^2}{2\sqrt{\alpha_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^2) \\
& + \frac{5C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^3) + \frac{9C^2}{\alpha_0} E(w_t \xi_{\rho t-1}^2) + \frac{3C^2}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
= & \frac{C^4}{2\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^4) + \frac{9C^3}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^3) + \frac{C^2}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^2) \\
& + \frac{5C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^3) + \frac{9C^2}{\alpha_0} E(w_t \xi_{\rho t-1}^2) + \frac{3C^2}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
= & \frac{C^4}{2\underline{\alpha}_0 \sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^4) + \left(\frac{9C^3}{2\underline{\alpha}_0} + \frac{5C^3}{4\underline{\alpha}_0 \sqrt{\alpha_0}} \right) E(w_t \xi_{\rho t-1}^3) \\
& + \left(\frac{9C^2}{\alpha_0} + \frac{C^2}{\sqrt{\alpha_0}} \right) E(w_t \xi_{\rho t-1}^2) + \frac{3C^2}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
< & \infty,
\end{aligned}$$

$$\begin{aligned}
& E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_j \partial \gamma_k} \right| \\
= & E \left[\sup_{V_0} w_t \left| \frac{3 \operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{4 h_t^2(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} - \frac{|\eta_t(\boldsymbol{\theta})|}{4 h_t^3(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} \frac{h_t(\boldsymbol{\theta})}{\partial \gamma_k} \right. \right. \\
& + \frac{3|\eta_t(\boldsymbol{\theta})| - 2}{4 h_t^2(\boldsymbol{\theta})} \left(\frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_j} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_j} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \right) \\
& + \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2 h_t(\boldsymbol{\theta})} \frac{\partial^3 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_j \partial \gamma_k} - \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{2 h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_k} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_j} \\
& + \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{2 h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \left(\frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_k} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_j} + \frac{\partial^2 \varepsilon_t(\gamma)}{\partial \gamma_j \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \right) \\
& \left. \left. - \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{4 h^2(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_j} \right| \right] \\
\leq & \frac{3C^3}{4\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \frac{C^3}{4\underline{\alpha}_0} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+2}) \\
& + \frac{C^2}{4} E \left(\sup_{V_0} w_t (3|\eta_t(\boldsymbol{\theta})| + 2) \left(\frac{2}{\underline{\alpha}_0} \xi_{\rho t-1}^2 + \frac{1}{\sqrt{\underline{\alpha}_0}} \xi_{\rho t-1}^{\iota_1+1} \right) \right) \\
& + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(\sup_{V_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) \xi_{\rho t-1}) + \frac{C^2}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) + \frac{C^2}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) \\
& + \frac{C^2}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C^3}{4\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+2}) \\
= & \frac{C^3}{4\underline{\alpha}_0} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+2}) + \frac{3C^2}{2\underline{\alpha}_0} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^2) + \frac{3C^2}{4\sqrt{\underline{\alpha}_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+1}) \\
& + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) + \frac{C^3}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \frac{2C^2}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) \\
& + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) \\
\leq & \frac{C^4}{4\underline{\alpha}_0} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{\iota_1+3}) + \frac{3C^3}{2\underline{\alpha}_0} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^3) + \frac{3C^3}{4\sqrt{\underline{\alpha}_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{\iota_1+2}) \\
& + \frac{C^2}{2\sqrt{\underline{\alpha}_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^2) + \frac{C^3}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \frac{2C^2}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) \\
& + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) \\
= & \frac{C^4}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+3}) + \frac{3C^3}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^3) + \frac{3C^3}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^2) + \frac{C^3}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+2}) \\
& + \frac{2C^2}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^2) + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) \\
= & \frac{C^4}{2\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^{\iota_1+3}) + \frac{3C^3}{\underline{\alpha}_0} E(w_t \xi_{\rho t-1}^3) + \left(\frac{3C^3}{2\sqrt{\underline{\alpha}_0}} + \frac{C^3}{\underline{\alpha}_0} \right) E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \left(\frac{C^2}{\sqrt{\underline{\alpha}_0}} + \frac{2C^2}{\underline{\alpha}_0} \right) E(w_t \xi_{\rho t-1}^2) \\
& + \frac{C^2}{\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\underline{\alpha}_0}} E(w_t \xi_{\rho t-1}) < \infty,
\end{aligned}$$

$$\begin{aligned}
& E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j \partial \gamma_k} \right| \\
= & E \left[\sup_{V_0} w_t \left| \frac{3 \operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{4 h_t^2(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_j} - \frac{|\eta_t(\boldsymbol{\theta})|}{4 h_t^3(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_j} \frac{h_t(\boldsymbol{\theta})}{\partial \gamma_k} \right. \right. \\
& + \frac{3|\eta_t(\boldsymbol{\theta})| - 2}{4 h_t^2(\boldsymbol{\theta})} \left(\frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_j} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_j \partial \gamma_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \gamma_k} \right) \\
& \left. + \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2 h_t(\boldsymbol{\theta})} \frac{\partial^3 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j \partial \gamma_k} - \frac{\operatorname{sgn}(\eta_t(\boldsymbol{\theta}))}{2 h_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta})}} \frac{\partial \varepsilon_t(\gamma)}{\partial \gamma_k} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j} \right| \Big] \\
\leq & \frac{3C^3}{4\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{2\iota_1+1}) + \frac{C^3}{4\sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{2\iota_1+1}) + \frac{3C^2}{4\sqrt{\alpha_0}} E(\sup_{V_0} w_t (3|\eta_t(\boldsymbol{\theta})| + 2) \xi_{\rho t-1}^{\iota_1+1}) \\
& + \frac{C}{2\sqrt{\alpha_0}} E(\sup_{V_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) \xi_{\rho t-1}) + \frac{C^2}{2\alpha_0} E(w_t \xi_{\rho t-1}^{\iota_1+1}) \\
= & \frac{C^3}{4\sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{2\iota_1+1}) + \frac{9C^2}{4\sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\alpha_0}} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}) \\
& + \frac{3C^3}{4\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{2\iota_1+1}) + \left(\frac{3C^2}{2\sqrt{\alpha_0}} + \frac{C^2}{2\alpha_0} \right) E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
\leq & \frac{C^4}{4\sqrt{\alpha_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{2\iota_1+2}) + \frac{9C^3}{4\sqrt{\alpha_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{\iota_1+2}) + \frac{C^2}{2\sqrt{\alpha_0}} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^2) \\
& + \frac{3C^3}{4\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{2\iota_1+1}) + \left(\frac{3C^2}{2\sqrt{\alpha_0}} + \frac{C^2}{2\alpha_0} \right) E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
= & \frac{C^4}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{2\iota_1+2}) + \frac{9C^3}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{\iota_1+2}) + \frac{C^2}{\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^2) \\
& + \frac{3C^3}{4\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}^{2\iota_1+1}) + \left(\frac{3C^2}{2\sqrt{\alpha_0}} + \frac{C^2}{2\alpha_0} \right) E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{C}{2\sqrt{\alpha_0}} E(w_t \xi_{\rho t-1}) \\
< & \infty,
\end{aligned}$$

和

$$\begin{aligned}
& E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j \partial \delta_k} \right| \\
&= E \left[\sup_{V_0} w_t \left| -\frac{|\eta_t(\boldsymbol{\theta})|}{4h_t^3(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_j} \frac{h_t(\boldsymbol{\theta})}{\partial \delta_k} \right. \right. \\
&\quad + \frac{3|\eta_t(\boldsymbol{\theta})| - 2}{4h_t^2(\boldsymbol{\theta})} \left(\frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_j} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_j \partial \delta_k} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_i} + \frac{\partial h_t^2(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \delta_k} \right) \\
&\quad \left. \left. + \frac{1 - |\eta_t(\boldsymbol{\theta})|}{2h_t(\boldsymbol{\theta})} \frac{\partial^3 h_t(\boldsymbol{\theta})}{\partial \delta_i \partial \delta_j \partial \delta_k} \right| \right] \\
&\leq \frac{C^3}{4} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{3\iota_1}) + \frac{3C^2}{4} E(\sup_{V_0} w_t (3|\eta_t(\boldsymbol{\theta})| + 2) \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E(\sup_{V_0} w_t (1 + |\eta_t(\boldsymbol{\theta})|) \xi_{\rho t-1}^{\iota_1}) \\
&= \frac{C^3}{4} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{3\iota_1}) + \frac{9C^2}{4} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E(\sup_{V_0} w_t |\eta_t(\boldsymbol{\theta})| \xi_{\rho t-1}^{\iota_1}) \\
&\quad + \frac{3C^2}{2} E(w_t \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E(w_t \xi_{\rho t-1}^{\iota_1}) \\
&\leq \frac{C^4}{4} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{3\iota_1+1}) + \frac{9C^e}{4} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{2\iota_1+1}) + \frac{C^2}{2} E(w_t (1 + |\eta_t|) \xi_{\rho t-1}^{\iota_1+1}) \\
&\quad + \frac{3C^2}{2} E(w_t \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E(w_t \xi_{\rho t-1}^{\iota_1}) \\
&= \frac{C^4}{2} E(w_t \xi_{\rho t-1}^{3\iota_1+1}) + \frac{9C^e}{2} E(w_t \xi_{\rho t-1}^{2\iota_1+1}) + C^2 E(w_t \xi_{\rho t-1}^{\iota_1+1}) + \frac{3C^2}{2} E(w_t \xi_{\rho t-1}^{2\iota_1}) + \frac{C}{2} E(w_t \xi_{\rho t-1}^{\iota_1}) \\
&< \infty.
\end{aligned}$$

$$\begin{aligned}
& \text{同理, } E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j \partial \delta_k} \right| < \infty, E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \delta_j \partial \gamma_k} \right| < \infty, E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \gamma_i \partial \delta_j \partial \delta_k} \right| < \\
& \infty \text{ 和 } E \sup_{V_0} \left| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \delta_i \partial \gamma_j \partial \delta_k} \right| < \infty. \text{ 因此,}
\end{aligned}$$

$$E \sup_{V_0} \left\| \frac{\partial \mathbf{P}_t(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}} \right\| = E \sup_{V_0} \left\| \frac{\partial^2 \mathbf{D}_{t,1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| = E \sup_{\Theta_0} \left\| w_t \frac{\partial^3 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}' \partial \boldsymbol{\theta}} \right\| < \infty. \quad (\text{A.9})$$

最后, 当 $n \rightarrow \infty$ 时, 对任意 $\varepsilon > 0$,

$$\begin{aligned}
& P\left(\frac{1}{n} \sum_{t=1}^n \sup_{V_{\theta_0}} \|\mathbf{P}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{P}_t(\boldsymbol{\theta}_0, \mu_0)\| > \varepsilon\right) \\
&\leq \frac{1}{\varepsilon} E \left[\sup_{V_0} \|\mathbf{P}_t(\boldsymbol{\theta}, \mu_0) - \mathbf{P}_t(\boldsymbol{\theta}_0, \mu_0)\| \right] \\
&\leq \frac{1}{\varepsilon} E \left[\sup_{V_0} \left\| \frac{\partial \mathbf{P}_t(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \right\| \right] \\
&\leq \frac{M}{\varepsilon \sqrt{n}} E \sup_{V_0} \left\| \frac{\partial \mathbf{P}_t(\boldsymbol{\theta}, \mu_0)}{\partial \boldsymbol{\theta}} \right\| \rightarrow 0,
\end{aligned}$$

即 (c) 成立。

(d) 再次应用引理 3.2 和遍历定理, 我们有

$$\frac{1}{n} \sum_{t=1}^n [D_t(\boldsymbol{\theta}_0, \mu_0) D_t'(\boldsymbol{\theta}_0, \mu_0)] = \boldsymbol{\Omega} + o_p(1). \quad (\text{A.10})$$

注意到

$$\frac{1}{n} \sum_{t=1}^n \|P_t(\boldsymbol{\theta}_0, \mu_0)\| = E\|P_t(\boldsymbol{\theta}_0, \mu_0)\| + o_p(1) = O_p(1),$$

因此由 (c) 可得 $\sup_{V_0} \|P_t(\boldsymbol{\theta}, \mu_0) - P_t(\boldsymbol{\theta}_0, \mu_0)\| = o_p(1)$ 。所以,

$$\sup_{V_0} \frac{1}{n} \sum_{t=1}^n \|P_t(\boldsymbol{\theta}, \mu_0)\| = O_p(1). \quad (\text{A.11})$$

对任意 $\boldsymbol{\theta} \in V_0$, 由泰勒展开可得:

$$D_t(\boldsymbol{\theta}, \mu_0) = D_t(\boldsymbol{\theta}_0, \mu_0) + P_t(\boldsymbol{\theta}^*, \mu_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0) + o_p\left(\frac{1}{\sqrt{n}}\right),$$

其中 $P_t(\boldsymbol{\theta}^*, \mu_0)$ 是引理 3.2 中定义的 $(r+s+p+q+3) \times (r+s+p+q+3)$ 矩阵, $\boldsymbol{\theta}^*$ 位于 $\boldsymbol{\theta}_0$ 和 $\boldsymbol{\theta}$ 中间, $\boldsymbol{\omega}_0$ 是 $\boldsymbol{\omega}$ 的真值。因此,

$$\begin{aligned} & \sup_{V_0} \frac{1}{n} \left\| \sum_{t=1}^n \left[D_t(\boldsymbol{\theta}, \mu_0) D_t'(\boldsymbol{\theta}, \mu_0) - D_t(\boldsymbol{\theta}_0, \mu_0) D_t'(\boldsymbol{\theta}_0, \mu_0) \right] \right\| \\ &= \sup_{V_0} \frac{1}{n} \left\| \sum_{t=1}^n \left[P_t(\boldsymbol{\theta}^*, \mu_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0) D_t'(\boldsymbol{\theta}_0, \mu_0) + D_t(\boldsymbol{\theta}_0, \mu_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0)' P_t'(\boldsymbol{\theta}^*, \mu_0) \right. \right. \\ & \quad \left. \left. + P_t(\boldsymbol{\theta}^*, \mu_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0)' P_t'(\boldsymbol{\theta}^*, \mu_0) \right] \right\| \\ &\leq \sup_{V_0} \frac{2}{n} \left\| \sum_{t=1}^n \left[P_t(\boldsymbol{\theta}, \mu_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0) D_t'(\boldsymbol{\theta}_0, \mu_0) \right] \right\| \\ & \quad + \sup_{V_0} \frac{1}{n} \left\| \sum_{t=1}^n \left[P_t(\boldsymbol{\theta}, \mu_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0)(\boldsymbol{\omega} - \boldsymbol{\omega}_0)' P_t'(\boldsymbol{\theta}, \mu_0) \right] \right\| \\ &\leq 2M \max_{1 \leq t \leq n} \left\| \frac{D_t(\boldsymbol{\theta}_0, \mu_0)}{\sqrt{n}} \right\| \sup_{V_0} \left(\frac{1}{n} \sum_{t=1}^n \|P_t(\boldsymbol{\theta}, \mu_0)\| \right) \\ & \quad + M^2 \max_{1 \leq t \leq n} \sup_{V_0} \left\| \frac{P_t(\boldsymbol{\theta}, \mu_0)}{n} \right\| \sup_{V_0} \left(\frac{1}{n} \sum_{t=1}^n \|P_t(\boldsymbol{\theta}, \mu_0)\| \right) \\ &= o_p(1) O_p(1) + o_p(1) O_p(1) \\ &= o_p(1), \end{aligned}$$

其中最后一个等式可由此引理的 (a)-(c) 和 (A.11) 得到。因此 (d) 成立。 $\square$

作者简历及在学期间所取得的科研成果

周沫，女，数学科学学院在读博士研究生。

1. 教育经历

- 2012-2016 年，浙江大学数学科学学院，本科，专业：统计学
- 2016-2021 年，浙江大学数学科学学院，直接攻博，专业：概率论与数理统计

2. 研究兴趣

- 非参数方法，经验似然，重尾理论

3. 攻读学位期间发表的论文

- [1] Zhou, M., Peng, L. and Zhang, R. (2021). Empirical likelihood test for the application of SWQMELE in fitting an ARMA-GARCH model. *Journal of Time Series Analysis* **42(2)**: 222-239.
- [2] Ma, Y., Zhou, M., Peng, L. and Zhang, R. (2021). Test for zero median of errors in an ARMA-GARCH model. *Econometric Theory*, in press.
- [3] Zhou, M., Ma, Y. and Zhang, R. (2021). Portmanteau-type test for unit root with heavy-tailed noise. *Journal of Statistical Planning and Inference*, in press.
- [4] Ma, Y., Zhou, M., Peng, L. and Zhang, R. (2021). Test for zero mean of errors after using median inference. Revised&resubmitted.
- [5] Zhou, M. and Zhang, R. (2021). Estimation of precision matrix in time series with latent factors. Manuscript.