

学校代码: 10200

研究生学号: 2016200217

分类号: O212

密 级: 无



东北师范大学

博士学位论文

高维协方差矩阵的单样本和双样本检验方法

One-sample and two-sample testing methods for high-dimensional covariance matrices

作者: 邹婷婷

指导教师: 郑术蓉 教授

一级学科: 统计学

研究方向: 高维统计推断

东北师范大学学位评定委员会

2021 年 5 月



东北师范大学

博士学位论文

学位论文评阅专家及答辩委员会人员信息

论 文 题 目	高维协方差矩阵的单样本和双样本检验方法			
作 者	邹婷婷			
指 导 教 师	郑术蓉 教授			
论 文 评 阅 人	评阅人 1	匿名评阅		
	评阅人 2	匿名评阅		
	评阅人 3	匿名评阅		
学 位 论 文 答 辩 委 员 会	主 席	朱复康	吉林大学	教授
	委 员	高巍	东北师范大学	教授
		郭建华	东北师范大学	教授
		史宁中	东北师范大学	教授
		徐平峰	长春工业大学	教授

独创性声明

本人声明所提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。据我所知,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,也不包含为获得东北师范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: 邹婷婷 日期: 2021.05.28

学位论文授权使用授权书

本学位论文作者完全了解东北师范大学有关保留、使用学位论文的规定,即:东北师范大学有权保留并向国家有关部门或机构送交学位论文的复印件和磁盘,允许论文被查阅和借阅。本人授权东北师范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,可以采用影印、缩印或其它复制手段保存、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: 邹婷婷 指导教师签名: 郑术蓉
日 期: 2021.05.28 日 期: 2021.05.28

学位论文作者毕业后去向:

工作单位: _____

电话: _____

通讯地址: _____

邮编: _____

摘要

随着大数据时代的到来,许多数据的维数不仅相对于样本量较大,数据的结构也变得复杂。对于这样的高维数据,很多经典的统计方法可能表现较差或失效。因此,寻找有效的统计方法分析这些高维数据有着重要的意义。协方差矩阵是分析高维数据的重要工具之一。在经济学、生物学等领域,许多处理数据的方法都需要对协方差矩阵进行统计推断。基于此,本文将研究与总体协方差矩阵相关的高维假设检验问题。

首先,本文研究了两个高维总体协方差矩阵相等的假设检验问题。本文基于两个样本协方差矩阵的差和比,提出了三个适用于稠密备择假设和稀疏备择假设的检验方法:一个检验方法适用于稠密备择假设,另外两个检验方法适用于一般情况,包括稠密备择假设、稀疏备择假设,或两者的混合。基于大维随机矩阵理论,本文在数据维数和样本量成比例增长条件和原假设下导出了这三个检验方法的渐近性质。此外,本文还在具有代表性的备择假设下,研究了这三个检验方法的渐近功效函数。大量的数值模拟表明,与现有的检验方法相比,本文提出的检验方法在控制犯第一类错误的概率和经验功效方面都具有很好的表现。

其次,本文给出了检验一组高维数据是否来自平稳向量自回归过程的检验方法,该检验方法也严格依赖样本协方差矩阵。本文考虑的向量自回归过程是一种特殊的向量自回归过程,该过程中的系数矩阵是对角线元素一致的对角矩阵。因此,在原假设下,可以将这组数据看作是来自平稳自回归过程的独立同分布的样本。本文首先给出了自回归过程中未知系数的最小二乘估计方法,并且导出了该估计的渐近正态性。此外,本文基于自回归过程的协方差矩阵和未知系数的估计,提出了一个方法来检验数据是否来自上述的平稳向量自回归过程。该检验方法不仅适用于高维数据还适用于低维数据。特别地,该检验方法由于没有分布的假定,因而具有普适性。数值模拟表明,该检验方法在不同模型、不同参数设置下都有着很好的表现。

关键词: 高维数据; 协方差矩阵; 假设检验; 向量自回归过程; 大维随机矩阵理论

Abstract

With the arrival of the era of big data, the dimension of many data is not only large relative to the sample size, but the structure of the data also becomes complicated. For such high-dimensional data, many classical statistical methods may perform poorly or fail. Therefore, it is of great significance to find effective methods to analyze these high-dimensional data. Covariance matrix is one of the important tools for analyzing high-dimensional data. In the fields of economics, biology and etc., many methods of processing data require to make statistical inference on covariance matrix. For this reason, this paper will study the high-dimensional hypothesis testing problems related to population covariance matrix.

First of all, this paper studies the hypothesis testing problem of the equality of two high-dimensional population covariance matrices. Based on the difference and ratio of two sample covariance matrices, this paper proposes three testing methods for dense alternative hypothesis and sparse alternative hypothesis: one testing method is shown to be powerful against dense alternative hypothesis, and the other two testing methods are suitable for general cases, including dense alternative hypothesis, sparse alternative hypothesis, or a mixture of the two. Based on large-dimensional random matrix theory, the asymptotic properties of these three testing methods are derived under the condition that the data dimension and sample sizes tend to infinity proportionally and the null hypothesis. In addition, the asymptotic power function of these three testing methods are also studied under the representative alternative hypotheses. Extensive simulation studies demonstrate that, compared with the existing testing methods, the proposed testing methods in this paper have good performance in terms of the type I error rate and the empirical power.

Secondly, this paper proposes a testing method to test whether high-dimensional data comes from the stationary vector autoregressive process, and this testing method also strictly relies on the sample covariance matrix. The vector autoregressive process considered in this paper is a special kind of vector autoregressive process, in which the

coefficient matrices are diagonal with the same diagonal elements. Therefore, under the null hypothesis, the data can be regarded as an independent and identically distributed sample from a stationary autoregressive process. In this paper, the least square estimation method of unknown coefficients in the autoregressive process is given, and the asymptotic normality of the estimation is derived. In addition, based on the covariance matrix of the autoregressive process and the estimation of unknown coefficients, a new testing method is proposed to test whether a set of data comes from the above-mentioned stationary vector autoregressive process. This testing method is not only suitable for high-dimensional data but also for low-dimensional data. In particular, the testing method is universal in the sense that it makes no assumption on distribution. Simulation studies show that the testing method performs well under different models and different parameter settings.

Key words: high-dimensional data; covariance matrix; hypothesis testing; vector autoregressive process; large-dimensional random matrix theory

目 录

中文摘要	I
英文摘要	II
第一章 绪论	1
1.1 概述	1
1.2 检验两个高维总体协方差矩阵相等的研究简介	1
1.3 检验数据是否来自平稳向量自回归过程的研究简介	6
1.4 本文主要工作和结构安排	7
1.5 符号说明	9
第二章 检验两个高维总体协方差矩阵的相等性	10
2.1 研究动机	10
2.2 检验方法	11
2.2.1 T_d 和 T_r 在原假设下的联合渐近分布	12
2.2.2 三种检验方法	15
2.3 功效分析	17
2.3.1 稠密备择集合 Π_1 下的功效分析	18
2.3.2 稠密备择集合 Π_2 下的功效分析	20
2.3.3 备择集合 Π_3 下的功效分析	22
2.4 数值模拟	22
2.5 理论证明	35
2.5.1 证明定理2.2.1, 定理2.3.1和定理2.3.2的准备工作	35
2.5.2 定理和性质的证明	51
2.5.3 定理和性质证明过程中需要的引理	65
2.6 附录	86
2.6.1 假定的条件	86
2.6.2 验证数值模拟中考虑模型设置	87

第三章 检验数据是否来自平稳向量自回归过程	91
3.1 检验问题	91
3.2 主要结果	92
3.2.1 系数的最小二乘估计	92
3.2.2 检验方法	94
3.3 数值模拟	97
3.4 理论证明	108
3.4.1 定理3.2.1的证明	108
3.4.2 定理3.2.2的证明	114
3.4.3 性质3.2.1的证明	121
第四章 总结和未来研究工作	126
4.1 总结	126
4.2 未来研究工作	127
参考文献	128
致谢	133
在学期间公开发表论文及著作情况	135

第一章 绪论

§1.1 概述

在多元统计分析中,协方差矩阵的假设检验在统计推断中具有很广泛的应用。例如,可以通过检验一组数据的总体协方差矩阵是否正比于单位阵,来检验这组数据的分量是否独立(如果数据服从正态分布)并且具有相同的方差^[1, 2, 3, 4];在两总体的判别分析中,很多方法都依赖于两个总体协方差矩阵相等的基本假设^[5, 6, 7, 8, 9, 10]。因此,在使用这些方法之前,应该检验这两个总体协方差矩阵是否相等。在传统的低维设定下,即样本的维数是固定的而样本量是趋于无穷的,协方差矩阵的假设检验问题已经被很好地研究。

随着计算机技术的飞速发展和广泛应用,海量数据的采集和存储变得越来越普遍,高维数据出现在各个领域。在金融领域,每天来自世界各地市场的网络数据在千兆网基站上累积。在遗传实验中,从单一组织中就可能记录成千上万的基因表达。经研究发现,当数据维数比较大时,很多传统的检验协方差矩阵的方法将会失效,即检验方法犯第一类错误的概率随着数据维数增加而趋向于1。例如, Bai et al. (2009)^[11]给出了两个传统似然比方法在检验高维协方差矩阵时失效的例子,并利用样本协方差矩阵和 F 随机矩阵的线性谱统计量的中心极限定理对似然比方法做出修正,以使其适用于高维数据。除了维数较高,在医疗、金融、工业等领域,许多数据的结构也比较复杂。在这种大数据背景下,如何对高维数据的协方差矩阵进行检验?如何通过协方差矩阵的假设检验来处理具有复杂结构的高维数据?这些问题值得我们深入思考。

本文考虑两个涉及总体协方差矩阵的高维假设检验问题。第一个是检验两组高维数据的总体协方差矩阵是否相等;第二个是基于一组高维数据的总体协方差矩阵结构,检验这组数据是否来自平稳向量自回归过程。

§1.2 检验两个高维总体协方差矩阵相等的研究简介

为了检验两组高维数据的总体协方差矩阵是否相等,一些检验方法被相继提出。接下来,主要介绍这些检验方法的基本思想。令 $\mathbf{x}_{k1}, \mathbf{x}_{k2}, \dots, \mathbf{x}_{kN_k}$ 为来自

第 k 个 p 维总体的独立同分布(i.i.d.)的样本, 总体均值向量为 $\boldsymbol{\mu}_k$, 总体协方差矩阵为 $\boldsymbol{\Sigma}_k$, 样本量为 N_k , $k = 1, 2$ 。令 $n_k = N_k - 1$ 和 $y_{n_k} = p/n_k$ 。第 k 个总体的样本协方差矩阵为

$$\mathbf{S}_k = n_k^{-1} \sum_{i=1}^{N_k} (\mathbf{x}_{ki} - \bar{\mathbf{x}}_k)(\mathbf{x}_{ki} - \bar{\mathbf{x}}_k)^T,$$

其中 $\bar{\mathbf{x}}_k = N_k^{-1} \sum_{i=1}^{N_k} \mathbf{x}_{ki}$ 是第 k 个总体的样本均值, $k = 1, 2$ 。本文感兴趣的假设检验问题为

$$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} \quad \text{versus} \quad H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2, \quad (1.1)$$

其中矩阵 $\boldsymbol{\Sigma}$ 是未知的。

对于上述问题的检验方法, 大致可以分为以下三种类型。第一种类型的检验方法是修正传统的检验方法。这类方法基于传统检验方法中的统计量或者给出该统计量的修正, 在数据维数和样本量成比例增长的假设条件下得到这些统计量的渐近分布, 进而得到新的检验方法。例如, 利用 F 矩阵的线性谱统计量的中心极限定理^[12], 在 $y_{n_k} \rightarrow y_k \in (0, 1)$ 的条件下, Bai et al. (2009)^[11]给出了似然比统计量 $-2 \log L$ 的中心极限定理, 其中

$$L = \frac{|\mathbf{A}_1|^{N_1/2} \cdot |\mathbf{A}_2|^{N_2/2}}{|c_{N_1} \mathbf{A}_1 + c_{N_2} \mathbf{A}_2|^{N/2}},$$

$\mathbf{A}_k = N_k^{-1} \sum_{i=1}^{N_k} (\mathbf{x}_{ki} - \boldsymbol{\mu}_k)(\mathbf{x}_{ki} - \boldsymbol{\mu}_k)^T$, $N = N_1 + N_2$, $c_{N_k} = N_k/N$ 。在 $y_k \in (0, 1]$ 和原假设条件下, Jiang和Yang (2013)^[13]得到了统计量 $\log L^*$ 的渐近分布, 其中

$$L^* = \frac{|\mathbf{S}_1|^{n_1/2} \cdot |\mathbf{S}_2|^{n_2/2}}{|c_{n_1} \mathbf{S}_1 + c_{n_2} \mathbf{S}_2|^{n/2}},$$

$n = n_1 + n_2$ 和 $c_{n_k} = n_k/n$ 。虽然上述两篇文献给出了似然比统计量在高维框架下的渐近分布, 但是他们的方法仍然要求数据的维数 p 小于 n_1 和 n_2 。为了解决数据维数大于样本量的情况, 在维数 p 小于 n 的条件下, Zhang et al. (2019)^[14]基于Beta矩阵 $\mathbf{B}(1, 2) = n_1 \mathbf{S}_1 (n_1 \mathbf{S}_1 + n_2 \mathbf{S}_2)^{-1}$ 的特征值 $\{\lambda_j^{\mathbf{B}(1,2)}, 1 \leq j \leq p\}$ ($\lambda_j^{\mathbf{B}(1,2)}$ 为矩阵 $\mathbf{B}(1, 2)$ 的第 j 大特征值)提出了以下两个检验统计量

$$\begin{aligned} T_{ZHB}^1 &= \sum_{\lambda_j^{\mathbf{B}(1,2)} \in (0,1)} [c_{n_1} \log \lambda_j^{\mathbf{B}(1,2)} + c_{n_2} \log(1 - \lambda_j^{\mathbf{B}(1,2)})], \\ T_{ZHB}^2 &= \sum_{\lambda_j^{\mathbf{B}(1,2)} \in (0,1)} \log \lambda_j^{\mathbf{B}(1,2)}. \end{aligned}$$

这两个统计量源于似然比统计量

$$2 \log L^*/n = c_{n_1} \log(c_{n_1}^{-1} |\mathbf{B}(1, 2)|) + c_{n_2} \log(c_{n_2}^{-1} |\mathbf{I}_p - \mathbf{B}(1, 2)|)$$

和Wilks (1932)^[15]给出的统计量 $\tilde{L} = \log |\mathbf{B}(1, 2)|$, 当 $p > n_1$ 或 $p > n_2$ 时, 为了保证对数函数有定义, 通过去除矩阵 $\mathbf{B}(1, 2)$ 的非0和非1特征值, 他们得到了统计量 T_{ZHB}^1 和 T_{ZHB}^2 。类似地, 基于Pillai迹统计量^[16] $\text{tr} \mathbf{B}(1, 2)$ 和

$$c_{n_1} \text{tr}(c_{n_1}^{-1} \mathbf{B}(1, 2) - \mathbf{I}_p)^2 + c_{n_2} \text{tr}(c_{n_2}^{-1} \mathbf{B}(2, 1) - \mathbf{I}_p)^2,$$

Zhang et al. (2020)^[17]又提出了两个检验统计量

$$\begin{aligned} T_{ZHB}^3 &= \sum_{\lambda_j^{\mathbf{B}(1,2)} \in (0,1)} \lambda_j^{\mathbf{B}(1,2)}, \\ T_{ZHB}^4 &= \sum_{\lambda_j^{\mathbf{B}(1,2)} \in (0,1)} \sum_{\lambda_{j'}^{\mathbf{B}(2,1)} \in (0,1)} [c_{n_1} (c_{n_1}^{-1} \lambda_j^{\mathbf{B}(1,2)} - 1)^2 + c_{n_2} (c_{n_2}^{-1} \lambda_{j'}^{\mathbf{B}(2,1)} - 1)^2], \end{aligned}$$

其中 $\mathbf{B}(2, 1) = n_2 \mathbf{S}_2 (n_1 \mathbf{S}_1 + n_2 \mathbf{S}_2)^{-1}$, $\{\lambda_{j'}^{\mathbf{B}(2,1)}, 1 \leq j' \leq p\}$ 为 $\mathbf{B}(2, 1)$ 的特征值。利用Beta矩阵的线性谱统计量的中心极限定理^[18], Zhang et al. (2019)^[14]和Zhang et al. (2020)^[17]在原假设 H_0 下给出了统计量 T_{ZHB}^1 , T_{ZHB}^2 , T_{ZHB}^3 和 T_{ZHB}^4 的渐近分布。

第二种类型的检验方法是通过构建目标函数 $h(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$, 使其满足: 在原假设下, $h(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) = 0$; 在备择假设下, $h(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) > 0$ 。然后基于这个目标函数给出一个合理的检验统计量 T_h , 当 $T_h > h_0$ 时, 就拒绝原假设, 其中 h_0 由 $P(T_h > h_0 | H_0) = \alpha$ 确定, α 为给定的显著水平。例如, 在高斯总体假设下, 通过修正统计量 $T_d = \text{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$ 的期望值和目标函数 $\text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$ 的偏差, Schott (2007)^[19]提出了统计量

$$t_{np} = \text{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2 - \sum_{k=1}^2 (n_k \eta_k)^{-1} [n_k (n_k - 2) \text{tr}(\mathbf{S}_k^2) + n_k^2 \text{tr}^2(\mathbf{S}_k)],$$

其中 $\eta_k = (n_k - 1)(n_k + 2)$, $k = 1, 2$ 。 t_{np} 是目标函数 $\text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$ 的无偏估计, 即

$$E(t_{np}) = \text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2.$$

当条件 $y_{n_k} \rightarrow y_k \in [0, \infty)$ 和 $\lim_{p \rightarrow +\infty} \text{tr}(\boldsymbol{\Sigma}^i)/p = \gamma_{i0} \in (0, \infty)$, $i = 1, \dots, 8$ 成立时, 在原假设下, Schott (2007)^[19]得到了统计量 t_{np} 的渐近分布。由于目标函数 $\text{tr}(\boldsymbol{\Sigma}_1 -$

Σ_2)²仅度量了 $\text{tr}(\Sigma_1^2)$ 和 $\text{tr}(\Sigma_2^2)$ 之间的差异，考虑到 $\text{tr}(\Sigma_1)$ 和 $\text{tr}(\Sigma_2)$ 之间的差异也可能揭示 Σ_1 和 Σ_2 之间的差异，Srivastava和Yanagihara (2010)^[20]基于目标函数

$$\text{tr}(\Sigma_1^2)/\text{tr}^2(\Sigma_1) - \text{tr}(\Sigma_2^2)/\text{tr}^2(\Sigma_2)$$

构造检验统计量。当样本来自高斯总体时，在 $n_k = O(p^\delta)$, $\delta > 1/2$ 和 $\lim_{p \rightarrow +\infty} \text{tr}(\Sigma_k^i)/p = \gamma_{ik} \in (0, \infty)$, $i = 1, \dots, 4$ 的条件下，他们给出了 $a_{1k} = \text{tr}(\Sigma_k)/p$ 和 $a_{2k} = \text{tr}(\Sigma_k^2)/p$ 的无偏相合估计

$$\hat{a}_{1k} = \frac{1}{p} \text{tr}(\mathbf{S}_k), \quad \hat{a}_{2k} = \frac{n_k}{p(n_k - 1)(n_k + 2)} [n_k \text{tr}(\mathbf{S}_k^2) - \text{tr}^2(\mathbf{S}_k)], \quad k = 1, 2,$$

然后提出了检验统计量

$$Q_2 = \hat{\gamma}_1 - \hat{\gamma}_2 = \frac{\hat{a}_{21}}{\hat{a}_{11}^2} - \frac{\hat{a}_{22}}{\hat{a}_{12}^2}$$

并在原假设下给出了 Q_2 的渐近分布。注意到上面提到的两个检验方法都依赖于高斯总体假定。

对于一般总体，Li和Chen (2012)^[21]利用 U 统计量导出了目标函数 $\text{tr}(\Sigma_1 - \Sigma_2)^2$ 的无偏估计

$$T_{N_1, N_2} = A_{N_1} + A_{N_2} - 2C_{N_1, N_2},$$

并把 T_{N_1, N_2} 做为检验统计量，其中

$$\begin{aligned} A_{N_k} &= \frac{1}{N_k(N_k - 1)} \sum_{i \neq j} (\mathbf{x}_{ki}^T \mathbf{x}_{kj})^2 - \frac{2}{N_k(N_k - 1)(N_k - 2)} \sum_{i, j, \ell}^* \mathbf{x}_{ki}^T \mathbf{x}_{kj} \mathbf{x}_{kj}^T \mathbf{x}_{k\ell} \\ &\quad + \frac{1}{N_k(N_k - 1)(N_k - 2)(N_k - 3)} \sum_{i, j, \ell, g}^* \mathbf{x}_{ki}^T \mathbf{x}_{kj} \mathbf{x}_{k\ell}^T \mathbf{x}_{kg}, \quad k = 1, 2, \\ C_{N_1, N_2} &= \frac{1}{N_1 N_2} \sum_i \sum_j (\mathbf{x}_{1i}^T \mathbf{x}_{2j})^2 - \frac{1}{N_1 N_2 (N_1 - 1)} \sum_{i, \ell}^* \sum_j \mathbf{x}_{1i}^T \mathbf{x}_{2j} \mathbf{x}_{2j}^T \mathbf{x}_{1\ell} \\ &\quad - \frac{1}{N_1 N_2 (N_2 - 1)} \sum_{i, \ell}^* \sum_j \mathbf{x}_{2i}^T \mathbf{x}_{1j} \mathbf{x}_{1j}^T \mathbf{x}_{2\ell} \\ &\quad + \frac{1}{N_1 N_2 (N_1 - 1)(N_2 - 1)} \sum_{i, \ell}^* \sum_{j, g}^* \mathbf{x}_{1i}^T \mathbf{x}_{2j} \mathbf{x}_{1\ell}^T \mathbf{x}_{2g} \end{aligned}$$

分别是 $\text{tr}(\Sigma_k^2)$ 和 $\text{tr}(\Sigma_1 \Sigma_2)$ 的无偏估计， \sum^* 代表在相互不同的指标上求和。值得一提的是，在获得统计量 T_{N_1, N_2} 的渐近分布时，Li和Chen (2012)^[21]并没有假定数据的

维数 p 和样本量之间的关系，因此他们的方法不仅无需假定总体的分布，而且适用于超高维数据。在许多应用中，例如基因组学中的基因选择，两个群体的协方差矩阵可以相等或者非常相似，因为它们可能只有少量的项存在差异。在这种稀疏的备择假设下，Cai et al. (2013)^[22]指出上面提到的检验方法可能会失效，并基于目标函数

$$\max_{1 \leq \ell_1 \leq \ell_2 \leq p} |\sigma_{1\ell_1\ell_2} - \sigma_{2\ell_1\ell_2}|$$

提出了一个新的检验方法，其中 $\sigma_{k\ell_1\ell_2}$ 是第 k 个总体协方差矩阵 Σ_k 的第 (ℓ_1, ℓ_2) 个元素， $k = 1, 2$ 。在他们的方法中，检验统计量为

$$T_x = \max_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{(s_{1\ell_1\ell_2} - s_{2\ell_1\ell_2})^2}{\hat{\theta}_{1\ell_1\ell_2}/n_1 + \hat{\theta}_{2\ell_1\ell_2}/n_2}, \quad (1.2)$$

其中 $s_{k\ell_1\ell_2}$ 是第 k 个样本协方差矩阵 S_k 的第 (ℓ_1, ℓ_2) 个元素，

$$\hat{\theta}_{k\ell_1\ell_2} = n_k^{-1} \sum_{i=1}^{N_k} \{(x_{k\ell_1 i} - \bar{x}_{k\ell_1})(x_{k\ell_2 i} - \bar{x}_{k\ell_2}) - s_{k\ell_1\ell_2}\}^2, \quad k = 1, 2,$$

$x_{k\ell i}$ 和 $\bar{x}_{k\ell}$ 分别是向量 \mathbf{x}_{ki} 和 $\bar{\mathbf{x}}_k$ 的第 ℓ 个元素， $\ell = 1, \dots, p$ 。在原假设和其他一些条件下， $T_x - 4 \log p + \log \log p$ 的渐近分布是第I类极值分布。

不同的目标函数是在不同的角度刻画 Σ_1 和 Σ_2 之间的差异，所以在构造检验方法时如果只考虑一个目标函数，这会导致该方法使用的局限性。比如，Li和Chen (2012)^[21]中的方法在 $\Sigma_1 - \Sigma_2$ 是稠密的情况下效果很好，但是当 $\Sigma_1 - \Sigma_2$ 是稀疏的，他们的方法可能会失效。反之，Cai et al. (2013)^[22]中的方法在 $\Sigma_1 - \Sigma_2$ 是稀疏的情况下效果很好；当 $\Sigma_1 - \Sigma_2$ 是稠密的，他们的方法或许表现很差。在实际应用中，一般不知道两个总体协方差矩阵的差到底是稠密的还是稀疏的，因此产生了第三种类型的检验方法：将基于不同目标函数生成的统计量以某种方式合并成新的统计量，然后通过这个新的统计量构造检验方法。例如，Yang和Pan (2017)^[23]和Zheng et al. (2020)^[24]给出的统计量都包含两项，一项用来捕捉 Σ_1 和 Σ_2 之间的稠密差异，另一项用来提高稀疏备择假设下的检验功效。因此，这两篇文献中的检验方法既适用于 $\Sigma_1 - \Sigma_2$ 是稠密的，又适用于 $\Sigma_1 - \Sigma_2$ 是稀疏的，或者两者的混合。在Zheng et al. (2020)^[24]的方法中，他们的检验统计量为

$$T_2 = T_d + p^2 I(T_x > s(N_1, N_2, p)),$$

其中 $I(\cdot)$ 表示示性函数, $s(N_1, N_2, p)$ 为事先指定的阈值, 其值依赖于样本维数 p 和样本量 N_1, N_2 。具体来说, 通过精心选择阈值 $s(N_1, N_2, p)$, T_2 中的第二项在原假设下依概率收敛到0, 并且只要 T_x 超过阈值 $s(N_1, N_2, p)$, 它的值就会变得很大。因此, 在稠密备择假设下, T_2 中的第一项 T_d 起主导作用; 在稀疏备择假设下, T_2 中的第二项会提升检验功效。综上, 统计量 T_2 能够自适应地结合来自 T_d 和 T_x 的信息。

本文的第二章根据目标函数 $\text{tr}(\mathbf{\Sigma}_1 \mathbf{\Sigma}_2^{-1} - \mathbf{I}_p)^2$ 提出了统计量 $T_r = \text{tr}(\mathbf{S}_1 \mathbf{S}_2^{-1} - \mathbf{I}_p)^2$, 并利用 T_d , T_r 和 T_x 提出了三个检验方法。在数据维数和样本量成比例增长的假设条件下, 研究了这三个方法在原假设和三个特殊的备择假设下的渐近性质。

§1.3 检验数据是否来自平稳向量自回归过程的研究简介

随着科技的发展, 维数随着样本量增加而增加或者维数相对于样本量而言较大的时间序列数据越来越常见。例如, 随时间变化的空间数据、功能性磁共振成像(fMRI)数据集和光谱成像数据集等。在上述高维数据的分析中, 一些文献假设模型中的误差向量是平稳向量过程, 例如因子模型^[25]和动态因子模型^[26, 27, 28]。

在平稳向量过程中, 一类非常重要的过程是向量自回归滑动平均(VARMA)过程。令 $\mathbf{X} = (x_{it})_{N \times T} = (\mathbf{x}_1, \dots, \mathbf{x}_T)$ 是观测到的时间序列构成的矩阵, $\mathbf{E} = (e_{it})_{N \times T} = (\mathbf{e}_1, \dots, \mathbf{e}_T)$ 是未观测到的白噪声序列构成的矩阵。如果 \mathbf{x}_t 是一个VARMA(p, q)过程, 那么

$$\phi(B)\mathbf{x}_t = \theta(B)\mathbf{e}_t, \quad (1.3)$$

其中 $\phi(B) = \mathbf{I}_N - \phi_1 B - \dots - \phi_p B^p$ 和 $\theta(B) = \mathbf{I}_N + \theta_1 B + \dots + \theta_q B^q$ 是 B 的实矩阵多项式函数, ϕ_1, \dots, ϕ_p 和 $\theta_1, \dots, \theta_q$ 是 $N \times N$ 的系数矩阵, B 是滞后算子满足 $B^j \mathbf{x}_t = \mathbf{x}_{t-j}$, $j = 0, 1, \dots$ 。特别地, 当 $\phi_j = \phi_j \mathbf{I}_N$ 和 $\theta_j = \theta_j \mathbf{I}_N$ 时, 记 $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ 和 $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$, 则(1.3)变成下列简单的VARMA(p, q)模型:

$$\phi(B)\mathbf{x}_t = \theta(B)\mathbf{e}_t. \quad (1.4)$$

当 $\phi(B) = 1$ 时, 模型(1.4)为简单的向量滑动平均(VMA)模型; 当 $\theta(B) = 1$ 时, 模型(1.4)为简单的向量自回归(VAR)模型。若 \mathbf{x}_t 是平稳的, 则 \mathbf{x}_t 的总体自协方差矩

阵 Γ_u 和样本自协方差矩阵 $\hat{\Gamma}_u$ 分别为

$$\Gamma_u = E(\mathbf{x}_t \mathbf{x}_{t+u}^T), \quad u = 0, 1, \dots, \quad \hat{\Gamma}_u = \frac{1}{T} \sum_{t=1}^{T-u} \mathbf{x}_t \mathbf{x}_{t+u}^T, \quad 0 \leq u \leq T-1.$$

由于统计学界对高维VARMA模型的研究刚刚起步,即使 \mathbf{X} 满足模型(1.4),所得到的结果大部分都是关于样本协方差矩阵 $\mathbf{B}_N = N^{-1} \mathbf{X}^T \mathbf{X}$ 的极限谱分布^[29, 30, 31]和 $\Pi_{\text{sym}}(\hat{\Gamma}_u, \hat{\Gamma}_u^T : u \geq 0)$ 的极限谱分布^[32, 33, 34, 35, 36],其中 $\Pi_{\text{sym}}(\cdot)$ 是一个对称多项式。关于VARMA模型检验问题的文献很少。当 N 和 T 成比例趋于无穷时, Bhattacharjee和Bose (2016)^[36]用 $\hat{\Gamma}_0$ 的迹 $\text{tr}(\hat{\Gamma}_0)$ 作为统计量来检验一组数据是否来自简单的VMA模型,但在他们的方法中,原假设下VMA模型的系数 $\theta_1, \dots, \theta_q$ 是已知的。令

$$\begin{aligned} R_{\Pi_{\text{sym}}} &= \sqrt{TN^{-1}} (\Pi_{\text{sym}}(\hat{\Gamma}_u, \hat{\Gamma}_u^T : u \geq 0) - \Pi_{\text{sym}}(\Gamma_u, \Gamma_u^T : u \geq 0)), \\ R_{\Pi} &= \sqrt{TN^{-1}} (\Pi(\hat{\Gamma}_u, \hat{\Gamma}_u^T : u \geq 0) - \Pi(\Gamma_u, \Gamma_u^T : u \geq 0)), \end{aligned}$$

其中 $\Pi(\cdot)$ 为一个多项式。当 N 和 T 满足 $N/T \rightarrow 0$ 时,根据 $R_{\Pi_{\text{sym}}}$ 的极限谱分布和 $\text{tr}(R_{\Pi})$ 的渐近正态性, Bhattacharjee和Bose (2019)^[37]给出了一系列检验VARMA模型中系数矩阵的方法。基于最大特征值 $\lambda_{\max}(\hat{\Gamma}_0)$ 的Tracy-Widom分布,在 N 和 T 成比例趋于无穷的条件下,张阳春(2020)^[38]提出了方法LE来检验一组数据是否来自简单的VARMA模型(1.4),这个方法要求系数 ϕ_1, \dots, ϕ_p 和 $\theta_1, \dots, \theta_q$ 是已知的。为了处理系数未知的情况,他们又提出了一个检验方法LN。这两个方法不需要假定 e_{it} 的分布,但它们要求 e_{it} 的前三阶矩和标准正态分布的前三阶矩相等。

通过VARMA模型的检验问题和检验方法可知,在 N 和 T 都比较大时,关于这种时间序列模型的统计推断仍有待发展。本文的第三章给出了一个方法来检验一组数据是否来自系数未知的VAR模型。这个检验方法无需假定 e_{it} 的分布,同时可以适用于高维时间序列数据。

§1.4 本文主要工作和结构安排

本文共分为四章。

第一章,绪论部分。首先介绍在大数据时代,很多传统的协方差矩阵检验方法已不适用于高维数据。其次引入两个高维假设检验问题,并简单地介绍这两个问题的研究现状。最后介绍本文的结构安排和符号说明。

第二章, 主要考虑两个高维总体协方差矩阵相等的假设检验问题。首先, 介绍研究这个假设检验问题的动机。其次, 提出了三个新的检验方法。一个检验方法适用于稠密备择假设, 另外两个检验方法适用于一般备择假设。在原假设下, 研究了这三个方法的渐近性质。在三个具有代表性的备择假设下, 分析了这三个方法的渐近功效函数。最后通过数值模拟验证了这三个检验方法的优越性。

第三章, 主要考虑检验一组数据是否来自系数未知的平稳向量自回归过程。在原假设下, 可以将这组数据看作是来自平稳自回归过程的独立同分布的样本。首先给出了自回归过程中未知系数的最小二乘估计以及这些估计的渐近正态性。然后根据该过程的协方差矩阵结构和系数的估计值, 提出了一个新的检验统计量, 并且通过这个检验统计量在原假设下的渐近分布给出了一个检验方法。最后通过数值模拟验证了这个检验方法的正确性和有效性。

第四章, 概括本文的研究内容和所得结论, 指出目前工作的不足之处, 并对未来的工作做了一些设想和展望。

§1.5 符号说明

本文中常用的缩略语和符号。

i.i.d.	独立同分布
\mathbf{A}^T	矩阵 \mathbf{A} 的转置
$\text{tr}\mathbf{A}$	矩阵 \mathbf{A} 的迹
$\Im z$	复数 z 的虚部
a.s.	几乎必然
$X_n \xrightarrow{p} X$	随机变量序列 $\{X_n\}$ 依概率收敛到随机变量 X
$X_n \xrightarrow{d} X$	随机变量序列 $\{X_n\}$ 依分布收敛到随机变量 X
$X_n = o_p(1)$	随机变量序列 $\{X_n\}$ 依概率收敛到0
$a_n = o(1)$	数列 $\{a_n\}$ 收敛到0
$I(\cdot)$	示性函数
\mathbb{R}	实数域
\mathbb{C}^+	复平面的上半平面, 不含实轴
δ_a	点 a 的狄拉克测度
ESD	经验谱分布
LSD	极限谱分布
\mathbf{I}_p	$p \times p$ 单位矩阵
$\lambda_{\max}(\mathbf{A})$	矩阵 \mathbf{A} 的最大特征值
$\lambda_{\min}(\mathbf{A})$	矩阵 \mathbf{A} 的最小特征值
$\ \mathbf{A}\ $	矩阵 \mathbf{A} 的谱模
$\text{card}(\Omega)$	集合 Ω 的基数

第二章 检验两个高维总体协方差矩阵的相等性

本章考虑两个高维总体协方差矩阵相等的假设检验问题。基于新的目标函数提出了检验统计量，并利用该统计量和原有的统计量给出了三个检验方法。在原假设和三个具有代表性的备择假设下，研究了这三个检验方法的理论性质。大量的数值模拟表明，新的检验方法的功效会等于或优于现有的检验方法。

§2.1 研究动机

令 $\mathbf{x}_{k1}, \mathbf{x}_{k2}, \dots, \mathbf{x}_{kN_k}$ 为来自第 k 个 p 维总体的 i.i.d. 样本，总体均值向量为 $\boldsymbol{\mu}_k$ ，总体协方差矩阵为 $\boldsymbol{\Sigma}_k$ ，样本量为 N_k ， $k = 1, 2$ 。令 $n_k = N_k - 1$ 和 $y_{n_k} = p/n_k$ 。第 k 个总体的样本协方差矩阵为

$$\mathbf{S}_k = n_k^{-1} \sum_{i=1}^{N_k} (\mathbf{x}_{ki} - \bar{\mathbf{x}}_k)(\mathbf{x}_{ki} - \bar{\mathbf{x}}_k)^T,$$

其中 $\bar{\mathbf{x}}_k = N_k^{-1} \sum_{i=1}^{N_k} \mathbf{x}_{ki}$ 是第 k 个总体的样本均值， $k = 1, 2$ 。本章考虑的是假设检验问题(1.1)，即

$$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} \quad \text{versus} \quad H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2.$$

在传统的低维假定下，即样本的维数 p 是固定的而样本量是趋于无穷的，上述的假设检验问题已经得到了很好的研究。比如，Sugiura和Nagao (1968)^[39]，Gupta和Giri (1973)^[40]，Perlman (1980)^[41]，Gupta和Tang (1984)^[42]，O'Brien (1992)^[43]和Anderson (2003)^[44]。在高维假定下，即样本的维数 p 相对于样本量 N_1 和 N_2 较大，为了解决传统检验方法失效的问题，几种不同的检验方法被相继提出。比如，Bai et al. (2009)^[11]，Schott (2007)^[19]，Srivastava和Yanagihara (2010)^[20]，Li和Chen (2012)^[21]，Jiang和Yang (2013)^[13]，Cai et al. (2013)^[22]，Yang和Pan (2017)^[23]，Chang et al. (2017)^[45]，Zhu et al. (2017)^[46]，Han et al. (2018)^[47]，Zheng et al. (2020)^[24]等。在这些检验方法中，许多检验方法都是基于 $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$ 的Frobenius范数的平方 $\text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$ 得到的。例如，在高斯总体假设下，通过修正统计量 $T_d = \text{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$ 的期望值和 $\text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$ 的偏差，Schott (2007)^[19] 给出了 $\text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$ 的一个无偏估计并基于该估计提出了一个检验方法。Li和Chen (2012)^[21] 使用 U 统计量构造了 $\text{tr}(\boldsymbol{\Sigma}_1 -$

Σ_2)²的另一个无偏估计,并将此估计作为检验统计量。他们提出的方法在许多情况下都表现不错,尤其是当两个总体协方差矩阵有稠密差异时。在一般总体假设下,Zheng et al. (2020)^[24]研究了统计量 T_d 的渐近分布并给出了新的检验方法。

另一方面,作为与差度量 $\text{tr}(\Sigma_1 - \Sigma_2)^2$ 对应的度量,比度量 $\text{tr}(\Sigma_1 \Sigma_2^{-1} - \mathbf{I}_p)^2$ 也可以用来刻画 Σ_1 和 Σ_2 之间的差异。当 \mathbf{S}_2 可逆时,很自然地可以用 $T_r = \text{tr}(\mathbf{S}_1 \mathbf{S}_2^{-1} - \mathbf{I}_p)^2$ 估计 $\text{tr}(\Sigma_1 \Sigma_2^{-1} - \mathbf{I}_p)^2$,但是很少有文献研究统计量 T_r ^[48, 49]。这也许是因为统计量 T_r 涉及 \mathbf{S}_2 的逆,这就要求样本的维数 p 不能超过样本量 N_2 ,并且当 p 接近 N_2 时, T_r 的表现比较复杂。实际上,基于统计量 T_d 和统计量 T_r 的检验方法的表现因情况而异,后面的模拟研究和理论功效分析显示:在所有情况下,就检验功效而言, T_d 和 T_r 中的任何一个都不能一直优于另外一个。因此,本章的目标是给出一个检验方法,这个方法的表现能够与 T_d 和 T_r 中表现较好的那一个相同。

§2.2 检验方法

在给出新的检验方法之前,首先介绍一些基本符号。若 \mathbf{A} 是 $p \times p$ 非负定矩阵,则其经验谱分布(ESD)为

$$F^{\mathbf{A}} = p^{-1} \sum_{j=1}^p \delta_{\lambda_j^{\mathbf{A}}},$$

其中 $\{\lambda_j^{\mathbf{A}}, 1 \leq j \leq p\}$ 为 \mathbf{A} 的特征值, δ_a 表示在点 a 的狄拉克测度。对两个 $m \times n$ 矩阵 $\mathbf{A} = (a_{ij})_{i=1, \dots, m, j=1, \dots, n}$ 和 $\mathbf{B} = (b_{ij})_{i=1, \dots, m, j=1, \dots, n}$,令

$$\mathbf{A} \circ \mathbf{B} = (a_{ij} b_{ij})_{i=1, \dots, m, j=1, \dots, n}.$$

两个总体的样本协方差阵由下式给出

$$\mathbf{S}_1 = n_1^{-1} \sum_{i=1}^{N_1} (\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)^{\text{T}}, \quad \mathbf{S}_2 = n_2^{-1} \sum_{i=1}^{N_2} (\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)^{\text{T}},$$

其中 $\bar{\mathbf{x}}_1 = N_1^{-1} \sum_{i=1}^{N_1} \mathbf{x}_{1i}$ 和 $\bar{\mathbf{x}}_2 = N_2^{-1} \sum_{i=1}^{N_2} \mathbf{x}_{2i}$ 为样本均值。上一节中提到的两个统计量为

$$T_d = \text{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2, \quad T_r = \text{tr}(\mathbf{S}_1 \mathbf{S}_2^{-1} - \mathbf{I}_p)^2.$$

接下来,为了研究本章所考虑的统计量的极限行为,给出在高维随机矩阵理论中常用的两个假设条件。

假设2.2.1. 随机向量 \mathbf{x}_{ki} 具有独立成分结构 $\mathbf{x}_{ki} = \boldsymbol{\mu}_k + \boldsymbol{\Sigma}_k^{1/2} \mathbf{w}_{ki}$, 其中

$$\mathbf{w}_{ki} = (w_{k1i}, \dots, w_{kpi})^T,$$

元素 $\{w_{k\ell i}, k = 1, 2; \ell = 1, \dots, p; i = 1, \dots, N_k\}$ 是独立同分布的, 满足

$$Ew_{k\ell i} = 0, \quad Ew_{k\ell i}^2 = 1, \quad \beta_k = Ew_{k\ell i}^4 - 3.$$

此外, $\boldsymbol{\Sigma}_k$ 是正定的并且是谱模有界的。 $\boldsymbol{\Sigma}_k$ 的ESD收敛到极限谱分布(LSD) L_k 。

假设2.2.2. 维数 p 和 n_1, n_2 成比例趋于无穷, 即

$$y_{n_1} = p/n_1 \rightarrow y_1 \in (0, +\infty), \quad y_{n_2} = p/n_2 \rightarrow y_2 \in (0, 1),$$

其中 $n_1 = N_1 - 1$ 和 $n_2 = N_2 - 1$ 。

假设2.2.2要求数据维数 p 小于 n_2 , 这样可以保证第二个总体的样本协方差阵 \mathbf{S}_2 是可逆的。

§2.2.1 T_d 和 T_r 在原假设下的联合渐近分布

在(1.1)中的原假设 H_0 下, 令 $L(x)$ 为 $\boldsymbol{\Sigma}$ 的LSD。下面的定理建立了统计量 T_d 和统计量 T_r 在原假设下的联合渐近分布。

定理2.2.1. 在原假设 H_0 下, 若假设2.2.1-假设2.2.2成立, 则

$$\begin{pmatrix} T_d - \mu_0 - \mu_{10} \\ T_r - \mu_{20} \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}_2, \begin{pmatrix} \sigma_{110} & \sigma_{120} \\ \sigma_{210} & \sigma_{220} \end{pmatrix}\right),$$

其中 $\mathbf{0}_2 = (0, 0)^T$,

$$\begin{aligned} \mu_0 &= n_1^{-1} \text{tr}^2(\mathbf{S}_1) + n_2^{-1} \text{tr}^2(\mathbf{S}_2), \quad m_{10} = \int x dL(x), \quad m_{20} = \int x^2 dL(x), \\ \mu_{10} &= (y_{n_1} + y_{n_2}) p^{-1} \text{tr}(\boldsymbol{\Sigma}^2) + (\beta_1 y_{n_1} + \beta_2 y_{n_2}) p^{-1} \text{tr}(\boldsymbol{\Sigma} \circ \boldsymbol{\Sigma}), \\ \mu_{20} &= p \left[-\frac{2}{1 - y_{n_2}} + \frac{y_{n_1}}{(1 - y_{n_2})^2} + \frac{1}{(1 - y_{n_2})^3} \right] + p \\ &\quad - \frac{2y_{n_2}}{(1 - y_{n_2})^2} + \frac{2y_{n_1}y_{n_2} + y_{n_1}}{(1 - y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1 - y_{n_2})^4} \\ &\quad - \frac{2\beta_2 y_{n_2}}{1 - y_{n_2}} + \frac{2\beta_2 y_{n_1}y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1 - y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1 - y_{n_2})^3}, \end{aligned}$$

$$\begin{aligned}
 \sigma_{110} &= 4(y_1 + y_2)^2 m_{20}^2, \\
 \sigma_{120} &= \sigma_{210} = \left[\frac{8y_2(y_1 + y_2)^2 + 4y_1^2 + 4y_2^2}{(1 - y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1 - y_2)^3} \right] m_{10}^2, \\
 \sigma_{220} &= \frac{8y_1^3 + 16y_1^2y_2}{(1 - y_2)^5} + \frac{4y_1^2 + 40y_1^2y_2 + 64y_1y_2^2}{(1 - y_2)^6} \\
 &\quad + \frac{8y_1y_2^4 + 56y_1y_2^2 + 48y_2^3 + 8y_1y_2}{(1 - y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1 - y_2)^8} \\
 &\quad + 4(\beta_1y_1 + \beta_2y_2) \left[\frac{(y_1 + y_2)^2}{(1 - y_2)^4} + \frac{2y_2(y_1 + y_2)}{(1 - y_2)^5} + \frac{y_2^2}{(1 - y_2)^6} \right].
 \end{aligned}$$

注释2.2.1. 根据Hölder不等式, $m_{20} \geq m_{10}^2$ 且 $\beta_k \geq -2$, $k = 1, 2$ 。因此

$$\begin{aligned}
 &\sigma_{110}\sigma_{220} - \sigma_{120}\sigma_{210} \\
 \geq &16 \left[\frac{(4y_2^3 + 2y_1y_2)(y_1 + y_2)^4 + 4y_1^2y_2^2(y_1 + y_2)^2 + y_1^4y_2}{(1 - y_2)^5} \right. \\
 &\quad + \frac{(2y_2^3 + 4y_2^2)(y_1 + y_2)^4 + 4y_1y_2^2(y_1 + y_2)^3 + 2y_1^2y_2^4}{(1 - y_2)^6} \\
 &\quad + \frac{(4y_1y_2^3 + 4y_1y_2^2)(y_1 + y_2)^2 + y_1^3y_2(y_1 + y_2) + 3y_1^3y_2^2}{(1 - y_2)^6} \\
 &\quad + \frac{14y_2^3(y_1 + y_2)^3 + (8y_2^5 + 4y_1y_2^2)(y_1 + y_2)^2}{(1 - y_2)^7} \\
 &\quad + \frac{(y_2^5 + 2y_1y_2^3)(y_1 + y_2) + y_2^7 + 3y_1y_2^5 + y_2^5 + 2y_1y_2^3}{(1 - y_2)^7} \\
 &\quad \left. + \frac{(4y_2^7 + 4y_2^3)(y_1 + y_2)^2 + y_2^4(y_1 + y_2) + 2y_1y_2^4}{(1 - y_2)^8} \right] m_{10}^4 > 0.
 \end{aligned}$$

综上, 协方差矩阵 $\begin{pmatrix} \sigma_{110} & \sigma_{120} \\ \sigma_{210} & \sigma_{220} \end{pmatrix}$ 是正定的。

定理2.2.1表明, 在原假设 H_0 下, T_d 和 T_r 的联合渐近分布是一个二元正态分布。 T_d 的边缘极限分布与Zheng et al. (2020)^[24]中定理1给出的结果是一致的。由于 $\sigma_{120} > 0$, 因此 T_d 和 T_r 是渐近正相关的。

因为 μ_{10} , σ_{110} , σ_{120} 和 σ_{210} 的表达式中含有未知总体协方差矩阵 Σ 的泛函, 在实际应用中, 需要给出这些量的估计。令

$$\begin{aligned}
 \mathbf{S} &= (n_1 + n_2)^{-1}(n_1\mathbf{S}_1 + n_2\mathbf{S}_2), \quad \hat{h}_{10} = p^{-1}\text{tr}(\mathbf{S} \circ \mathbf{S}), \\
 \hat{m}_{10} &= p^{-1}\text{tr}(\mathbf{S}), \quad \hat{m}_{20} = p^{-1}[\text{tr}(\mathbf{S}^2) - (n_1 + n_2)^{-1}\text{tr}^2(\mathbf{S})].
 \end{aligned}$$

定理2.2.2. 在定理2.2.1的条件下, μ_{10} , σ_{110} , σ_{120} , σ_{210} 和 σ_{220} 的弱相合估计为

$$\begin{aligned}
 \hat{\mu}_{10} &= (y_{n_1} + y_{n_2})\hat{m}_{20} + (\beta_1 y_{n_1} + \beta_2 y_{n_2})\hat{h}_{10}, \\
 \hat{\sigma}_{110} &= 4(y_{n_1} + y_{n_2})^2 \hat{m}_{20}^2, \\
 \hat{\sigma}_{120} &= \hat{\sigma}_{210} = \left[\frac{8y_{n_2}(y_{n_1} + y_{n_2})^2 + 4y_{n_1}^2 + 4y_{n_2}^2}{(1 - y_{n_2})^2} + \frac{8y_{n_1}y_{n_2} + 8y_{n_2}^3}{(1 - y_{n_2})^3} \right] \hat{m}_{10}^2, \\
 \hat{\sigma}_{220} &= \frac{8y_{n_1}^3 + 16y_{n_1}^2 y_{n_2}}{(1 - y_{n_2})^5} + \frac{4y_{n_1}^2 + 40y_{n_1}^2 y_{n_2} + 64y_{n_1} y_{n_2}^2}{(1 - y_{n_2})^6} \\
 &\quad + \frac{8y_{n_1} y_{n_2}^4 + 56y_{n_1} y_{n_2}^2 + 48y_{n_2}^3 + 8y_{n_1} y_{n_2}}{(1 - y_{n_2})^7} + \frac{8y_{n_2}^5 + 24y_{n_2}^3 + 4y_{n_2}^2}{(1 - y_{n_2})^8} \\
 &\quad + 4(\beta_1 y_{n_1} + \beta_2 y_{n_2}) \left[\frac{(y_{n_1} + y_{n_2})^2}{(1 - y_{n_2})^4} + \frac{2y_{n_2}(y_{n_1} + y_{n_2})}{(1 - y_{n_2})^5} + \frac{y_{n_2}^2}{(1 - y_{n_2})^6} \right] \\
 &\quad + 8y_{n_1} \left[\frac{2(y_{n_1} + y_{n_2} - y_{n_1} y_{n_2})(y_{n_1} + 3y_{n_2} - y_{n_1} y_{n_2} - y_{n_2}^2)}{(1 - y_{n_2})^6} \right. \\
 &\quad \quad \left. + \frac{(\beta_1 y_{n_1} + \beta_2 y_{n_2})(y_{n_1} + 2y_{n_2} - y_{n_1} y_{n_2} - y_{n_2}^2)}{(1 - y_{n_2})^4} \right] v_{10p} \\
 &\quad + 8y_{n_1} \left[\frac{(y_{n_1} + y_{n_2} - 1)^2}{(1 - y_{n_2})^2} + \frac{y_{n_1}}{(1 - y_{n_2})^3} \right] v_{20p} \\
 &\quad + \frac{16y_{n_1}(y_{n_1} + y_{n_2} - 1)}{(1 - y_{n_2})} v_{30p} + 8y_{n_1} v_{40p} \\
 &\quad + 4y_{n_1}^2 \left[\frac{2(y_{n_1} + y_{n_2} - y_{n_1} y_{n_2})}{(1 - y_{n_2})^4} + \frac{\beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1 - y_{n_2})^2} \right] v_{10p}^2 \\
 &\quad + \frac{16y_{n_1}^2 (y_{n_1} + y_{n_2} - 1)}{(1 - y_{n_2})} v_{10p} v_{20p} + 16y_{n_1}^2 v_{10p} v_{30p} \\
 &\quad + 4y_{n_1}^2 v_{20p}^2 + 8y_{n_1}^3 v_{10p} v_{20p},
 \end{aligned}$$

其中

$$\begin{aligned}
 v_{10p} &= \frac{1}{p} \left[\frac{y_{n_2}}{(1 - y_{n_2})^2} + \frac{\beta_2 y_{n_2}}{(1 - y_{n_2})} \right], \\
 v_{20p} &= \frac{1}{p} \left[\frac{y_{n_2}^2 + 3y_{n_2}}{(1 - y_{n_2})^4} - \frac{\beta_2 (y_{n_2}^2 - 3y_{n_2})}{(1 - y_{n_2})^3} \right], \\
 v_{30p} &= \frac{1}{p} \left[\frac{y_{n_2}^3 + 9y_{n_2}^2 + 6y_{n_2}}{(1 - y_{n_2})^6} + \frac{6\beta_2 y_{n_2}}{(1 - y_{n_2})^5} \right], \\
 v_{40p} &= \frac{1}{p} \left[\frac{y_{n_2}^4 + 18y_{n_2}^3 + 35y_{n_2}^2 + 10y_{n_2}}{(1 - y_{n_2})^8} + \frac{10\beta_2 (y_{n_2}^2 + y_{n_2})}{(1 - y_{n_2})^7} \right].
 \end{aligned}$$

注释2.2.2. $\hat{\sigma}_{220}$ 的表达式不是简单地用 y_{n_1} 和 y_{n_2} 分别替换 σ_{220} 中的 y_1 和 y_2 , 而是添加了含有 v_{i0p} 的附加项, $i = 1, 2, 3, 4$ 。这是因为即使 v_{i0p} 的阶是 $o(1)$, 当 p 很小时, v_{i0p} 的值也远大于 0 。

根据Slutsky定理, 可以进一步得到随机向量

$$\left(\frac{T_d - \mu_0 - \hat{\mu}_{10}}{\sqrt{\hat{\sigma}_{110}}}, \frac{T_r - \mu_{20}}{\sqrt{\hat{\sigma}_{220}}} \right)^T$$

依分布收敛到 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$, 其中 $\rho = \sigma_{120} / \sqrt{\sigma_{110}\sigma_{220}}$ 。

§2.2.2 三种检验方法

为了检验假设(1.1), 本节给出了以下三种检验方法:

检验1: 令第一个检验统计量为

$$T_{dr} = \max\left\{|T_d - \mu_0 - \hat{\mu}_{10}| / \sqrt{\hat{\sigma}_{110}}, |T_r - \mu_{20}| / \sqrt{\hat{\sigma}_{220}}\right\}. \quad (2.1)$$

换句话说, T_{dr} 由标准化的统计量 T_d 和 T_r 的最大绝对值构成。对于给定的显著水平 α , 第一个检验的拒绝域为

$$\{\mathbf{x}_{11}, \dots, \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : T_{dr} > t_\alpha\}, \quad (2.2)$$

拒绝域(2.2)中的临界值 t_α 满足

$$\alpha = 1 - \int_{-t_\alpha}^{t_\alpha} \int_{-t_\alpha}^{t_\alpha} f(x_d, x_r) dx_d dx_r, \quad (2.3)$$

其中 $f(x_d, x_r)$ 为二元正态分布 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \hat{\rho} \\ \hat{\rho} & 1 \end{pmatrix}\right)$ 的密度函数, $\hat{\rho} = \hat{\sigma}_{120} / \sqrt{\hat{\sigma}_{110}\hat{\sigma}_{220}}$ 。

检验2: T_d 和 T_r 能够很好地度量 Σ_1 和 Σ_2 之间的稠密差异, T_{dr} 也是如此。为了提高检验(2.2)在稀疏备择假设下的功效, 可以使用Cai et al. (2013)^[22]中的理论结果。若Cai et al. (2013)^[22]中的条件(C1), (C2)(或(C2*))和(C3)(见本章附录§2.6.1)成立, 在原假设 H_0 下, 对任意 $t \in \mathbb{R}$, 当 $N_1, N_2, p \rightarrow \infty$ 时,

$$P(T_x - 4 \log p + \log \log p \leq t) \rightarrow \exp\left(-\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{t}{2}\right)\right), \quad (2.4)$$

其中

$$T_x = \max_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{(s_{1\ell_1\ell_2} - s_{2\ell_1\ell_2})^2}{\hat{\theta}_{1\ell_1\ell_2}/n_1 + \hat{\theta}_{2\ell_1\ell_2}/n_2},$$

$$\hat{\theta}_{k\ell_1\ell_2} = n_k^{-1} \sum_{i=1}^{N_k} \{(x_{k\ell_1 i} - \bar{x}_{k\ell_1})(x_{k\ell_2 i} - \bar{x}_{k\ell_2}) - s_{k\ell_1\ell_2}\}^2,$$

$\mathbf{S}_1 = (s_{1\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$ 和 $\mathbf{S}_2 = (s_{2\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$ 。类似 Zheng et al. (2020)^[24] 中统计量 T_2 的构造(见§1.2节), 令第二个检验统计量为

$$T_{\text{drx}}^w = T_{\text{dr}} + p^2 I(T_x > s(N_1, N_2, p)),$$

其中 $s(N_1, N_2, p)$ 为事先指定的阈值, 其值依赖于样本维数 p 和样本量 N_1, N_2 。第二个检验的拒绝域为

$$\{\mathbf{x}_{11}, \dots, \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : T_{\text{drx}}^w > t_\alpha\}, \quad (2.5)$$

其中 t_α 满足等式(2.3)。

检验3: 为了整合 T_{dr} 和 T_x 的信息, 令第三个检验统计量为

$$T_{\text{drx}}^m = \max\{T_{\text{dr}}, c_\alpha T_x\},$$

其中 $c_\alpha = t_{\alpha/2}/q_{\alpha/2}$, $t_{\alpha/2}$ 满足

$$\alpha/2 = 1 - \int_{-t_{\alpha/2}}^{t_{\alpha/2}} \int_{-t_{\alpha/2}}^{t_{\alpha/2}} f(x_d, x_r) dx_d dx_r, \quad (2.6)$$

$$q_{\alpha/2} = -\log(8\pi) - 2 \log \log(1 - \alpha/2)^{-1} + 4 \log p - \log \log p.$$

统计量 T_{drx}^m 的形式与统计量 T_{dr} 的形式类似, 但它包含了来自统计量 T_x 的附加信息。第三个检验的拒绝域为

$$\{\mathbf{x}_{11}, \dots, \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : T_{\text{drx}}^m > t_{\alpha/2}\}. \quad (2.7)$$

在下文中, 将基于统计量 T_{dr} , 统计量 T_{drx}^w 和统计量 T_{drx}^m 的检验方法简称为检验 T_{dr} , 检验 T_{drx}^w 和检验 T_{drx}^m 。

定理2.2.3. 对于上述三个检验方法, 有以下结果:

- 在定理2.2.1的条件下, 检验 T_{dr} 的渐近水平为 α ;
- 在定理2.2.1的条件和Cai et al. (2013)^[22]中的条件(C1), (C2)(或(C2*)), (C3)下, 若阈值 $s(N_1, N_2, p)$ 满足 $s(N_1, N_2, p) - 4 \log p \geq 0$, 检验 T_{drx}^w 的渐近水平为 α ;
- 在定理2.2.1的条件和Cai et al. (2013)^[22]中的条件(C1), (C2)(或(C2*)), (C3)下, 检验 T_{drx}^m 的渐近水平等于或小于 α 。

根据定理2.2.3, 在本章中, 令

$$s(N_1, N_2, p) = \left\{ [\log \log(N_1/2 + N_2/2) - 1]^2 + 4 \right\} (\log p - \log \log p/4) + q, \quad (2.8)$$

其中 q 满足

$$\exp\{-(8\pi)^{-1/2} \exp(-q/2)\} = 0.99.$$

很容易验证阈值(2.8)满足 $s(N_1, N_2, p) - 4 \log p \geq 0$ 。

注释2.2.3. 最后, 给出一个比较检验 T_{drx}^w 和检验 T_{drx}^m 的注释。为了提高在稀疏备择假设下的检验功效, 这两个检验方法都从极值统计量 T_x 中获取信息。但 T_x 以不同的形式发挥作用: T_{drx}^w 直接将与 T_x 有关的项加到主项 T_{dr} 上, 而不改变临界值; T_{drx}^m 使用与Tippett组合检验^[50, 51]类似的想法整合来自 T_{dr} 和 T_x 的信息, 它本质上是一个需要多重调整的多重检验方法。由于 $t_{\alpha/2} > t_\alpha$, 当信号主要来自 T_{dr} 时, T_{drx}^w 的功效高于 T_{drx}^m 的功效。另一方面, 为了保证 T_{drx}^w 的渐近无偏性, T_x 对 T_{drx}^w 的贡献受到了严重的惩罚。因此, 在稀疏备择假设下, T_{drx}^m 的表现会优于 T_{drx}^w 。这些结论都将在数值模拟中得到证实。

§2.3 功效分析

根据定理2.2.1和定理2.2.2, 基于统计量 T_d 和统计量 T_r 的检验的拒绝域为

$$\{\mathbf{x}_{11}, \dots, \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : |T_d - \mu_0 - \hat{\mu}_{10}| / \sqrt{\hat{\sigma}_{110}} > z_{1-\alpha/2}\}, \quad (2.9)$$

$$\{\mathbf{x}_{11}, \dots, \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : |T_r - \mu_{20}| / \sqrt{\hat{\sigma}_{220}} > z_{1-\alpha/2}\}, \quad (2.10)$$

其中 $z_{1-\alpha/2}$ 是标准正态分布 $N(0, 1)$ 的 $1 - \alpha/2$ 分位点。

本节在以下三个备择集合中研究检验 T_d ，检验 T_r ，检验 T_{dr} ，检验 T_{drx}^w 和检验 T_{drx}^m 的渐近功效函数：

$$\begin{aligned}\Pi_1 &= \{(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) : \mathbf{\Sigma}_2 = \tau_p \mathbf{\Sigma}_1, \tau_p = p/(p + a_1), a_1 > 0 \text{ 是一个常数}\}, \\ \Pi_2 &= \{(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) : \mathbf{\Sigma}_2 = \mathbf{\Sigma}_1 + \frac{a_2}{p} \mathbf{J}_p, \mathbf{\Sigma}_1 = \mathbf{I}_p, a_2 > 0 \text{ 是一个常数}\}, \\ \Pi_3 &= \{(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) : \max_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{(\sigma_{1\ell_1\ell_2} - \sigma_{2\ell_1\ell_2})^2}{\theta_{1\ell_1\ell_2}/n_1 + \theta_{2\ell_1\ell_2}/n_2} \geq 16 \log p\},\end{aligned}$$

其中 \mathbf{J}_p 是元素都为1的 $p \times p$ 矩阵， $\sigma_{k\ell_1\ell_2}$ 是第 k 个总体协方差矩阵 $\mathbf{\Sigma}_k$ 的第 (ℓ_1, ℓ_2) 个元素， $\theta_{k\ell_1\ell_2} = \text{Var}((x_{k\ell_1} - \mu_{k\ell_1})(x_{k\ell_2} - \mu_{k\ell_2}))$ ， $k = 1, 2$ 。前两个集合 Π_1 和 Π_2 为局部稠密备择，最后一个集合 Π_3 包含稀疏备择。

§2.3.1 稠密备择集合 Π_1 下的功效分析

下面的定理给出了当 $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_1$ 时 T_d 和 T_r 的联合渐近分布。

定理2.3.1. 当 $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_1$ 时，若假设2.2.1-假设2.2.2成立，则

$$\begin{pmatrix} T_d - \mu_0 - \mu_{11} \\ T_r - \mu_{21} \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}_2, \begin{pmatrix} \sigma_{111} & \sigma_{121} \\ \sigma_{211} & \sigma_{220} \end{pmatrix}\right),$$

其中

$$\begin{aligned}m_{11} &= \int x dL_1, \quad m_{21} = \int x^2 dL_1, \\ \mu_{11} &= (y_{n_1} + y_{n_2} \tau_p^2) p^{-1} \text{tr}(\mathbf{\Sigma}_1^2) + (\beta_1 y_{n_1} + \beta_2 y_{n_2} \tau_p^2) p^{-1} \text{tr}(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_1), \\ \mu_{21} &= p \left[-\frac{2}{(1 - y_{n_2}) \tau_p} + \frac{y_{n_1}}{(1 - y_{n_2})^2 \tau_p^2} + \frac{1}{(1 - y_{n_2})^3 \tau_p^2} \right] + p \\ &\quad - \frac{2y_{n_2}}{(1 - y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1 - y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1 - y_{n_2})^4} \\ &\quad - \frac{2\beta_2 y_{n_2}}{(1 - y_{n_2})} + \frac{2\beta_2 y_{n_1} y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1 - y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1 - y_{n_2})^3}, \\ \sigma_{111} &= 4(y_1 + y_2)^2 m_{21}^2, \\ \sigma_{121} &= \sigma_{211} = \left[\frac{8y_2(y_1 + y_2)^2 + 4y_1^2 + 4y_2^2}{(1 - y_2)^2} + \frac{8y_1 y_2 + 8y_2^3}{(1 - y_2)^3} \right] m_{11}^2.\end{aligned}$$

性质2.3.1. 在定理2.3.1的条件下，有以下结果：

(I) 对于检验 T_d ,

$$P(|T_d - \mu_0 - \hat{\mu}_{10}| / \sqrt{\hat{\sigma}_{110}} > z_{1-\alpha/2}) \rightarrow \alpha;$$

(II) 对于检验 T_r ,

$$P(|T_r - \mu_{20}| / \sqrt{\hat{\sigma}_{220}} > z_{1-\alpha/2}) \rightarrow 1 - [\Phi(z_{1-\alpha/2} - \Delta_1) - \Phi(-z_{1-\alpha/2} - \Delta_1)] > \alpha,$$

其中 $\Phi(\cdot)$ 是标准正态分布 $N(0, 1)$ 的分布函数,

$$\Delta_1 = \frac{2a_1}{\sqrt{\sigma_{220}}} \left[\frac{y_1 + y_2}{(1 - y_2)^2} + \frac{y_2}{(1 - y_2)^3} \right];$$

(III) 对于检验 T_{dr} ,

$$P(T_{dr} > t_\alpha) \rightarrow 1 - \int_{-t'_\alpha - \Delta_1}^{t'_\alpha - \Delta_1} \int_{-t'_\alpha}^{t'_\alpha} f'(x_d, x_r) dx_d dx_r > \alpha,$$

其中 $f'(x_d, x_r)$ 是二元正态分布 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix}\right)$ 的密度函数,

$$\rho_1 = \sigma_{121} / \sqrt{\sigma_{111}\sigma_{220}}, \quad \alpha = 1 - \int_{-t'_\alpha}^{t'_\alpha} \int_{-t'_\alpha}^{t'_\alpha} f'(x_d, x_r) dx_d dx_r; \quad (2.11)$$

(IV) 对于检验 T_{drx}^w ,

$$\lim_{p \rightarrow \infty} P(T_{drx}^w > t_\alpha) \geq \lim_{p \rightarrow \infty} P(T_{dr} > t_\alpha) > \alpha;$$

(V) 对于检验 T_{drx}^m ,

$$\lim_{p \rightarrow \infty} P(T_{drx}^m > t_{\alpha/2}) \geq 1 - \int_{-t'_{\alpha/2} - \Delta_1}^{t'_{\alpha/2} - \Delta_1} \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} f'(x_d, x_r) dx_d dx_r > \alpha/2,$$

其中 $t'_{\alpha/2}$ 满足(2.11)中用 $\alpha/2$ 替换 α 后的等式。

性质2.3.1表明, 只要 $(\Sigma_1, \Sigma_2) \in \Pi_1$, 检验 T_d 将会失效。由于检验 T_r , 检验 T_{dr} , 检验 T_{drx}^w 和检验 T_{drx}^m 的渐近功效函数是 Δ_1 的增函数, 因此当 Δ_1 较大时, 这些检验具有较高的功效。

§2.3.2 稠密备择集合 Π_2 下的功效分析

下面的定理给出了当 $(\Sigma_1, \Sigma_2) \in \Pi_2$ 时 T_d 和 T_r 的联合渐近分布。

定理2.3.2. 当 $(\Sigma_1, \Sigma_2) \in \Pi_2$ 以及 $\beta_2 = 0$ 时, 若假设2.2.1-假设2.2.2成立, 则

$$\begin{pmatrix} T_d - \mu_0 - \mu_{12} \\ T_r - \mu_{22} \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}_2, \begin{pmatrix} \sigma_{112} & \sigma_{122} \\ \sigma_{212} & \sigma_{222} \end{pmatrix}\right),$$

其中

$$\begin{aligned} \mu_{12} &= a_2^2 + y_{n_1} + y_{n_2} + \beta_1 y_{n_1}, \\ \mu_{22} &= p\left[-\frac{2}{(1-y_{n_2})} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3}\right] + p \\ &\quad + \frac{a_2^2 - 2a_2(1+a_2)(y_{n_1} + y_{n_2})}{(1-y_{n_2})^2(1+a_2)^2} - \frac{2a_2 y_{n_2}}{(1-y_{n_2})^3(1+a_2)} \\ &\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} + \frac{\beta_1 y_{n_1}}{(1-y_{n_2})^2}, \\ \sigma_{112} &= 4(y_1 + y_2)^2, \\ \sigma_{122} &= \sigma_{212} = \frac{8y_2(y_1 + y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3}, \\ \sigma_{222} &= \frac{8y_1^3 + 16y_1^2y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2y_2 + 64y_1y_2^2}{(1-y_2)^6} \\ &\quad + \frac{8y_1y_2^4 + 56y_1y_2^2 + 48y_2^3 + 8y_1y_2}{(1-y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1-y_2)^8} \\ &\quad + 4\beta_1 y_1 \left[\frac{(y_1 + y_2)^2}{(1-y_2)^4} + \frac{2y_2(y_1 + y_2)}{(1-y_2)^5} + \frac{y_2^2}{(1-y_2)^6} \right]. \end{aligned}$$

性质2.3.2. 在定理2.3.2的条件下, 有以下结果:

(I) 对于检验 T_d ,

$$\begin{aligned} &P(|T_d - \mu_0 - \hat{\mu}_{10}| / \sqrt{\hat{\sigma}_{110}} > z_{1-\alpha/2}) \\ &\rightarrow 1 - \left[\Phi(z_{1-\alpha/2} - \Delta_2) - \Phi(-z_{1-\alpha/2} - \Delta_2) \right] > \alpha, \end{aligned}$$

其中 $\Delta_2 = a_2^2/[2(y_1 + y_2)]$;

(II) 对于检验 T_r ,

$$\begin{aligned} & P(|T_r - \mu_{20}| / \sqrt{\hat{\sigma}_{220}} > z_{1-\alpha/2}) \\ \rightarrow & 1 - \left[\Phi(z_{1-\alpha/2} - \Delta_3) - \Phi(-z_{1-\alpha/2} - \Delta_3) \right] \geq \alpha, \end{aligned}$$

其中

$$\Delta_3 = \frac{1}{\sqrt{\sigma_{222}}} \left[\frac{a_2^2 - 2a_2(1+a_2)(y_1+y_2)}{(1-y_2)^2(1+a_2)^2} - \frac{2a_2y_2}{(1-y_2)^3(1+a_2)} \right];$$

(III) 对于检验 T_{dr} ,

$$P(T_{dr} > t_\alpha) \rightarrow 1 - \int_{-t_\alpha^* - \Delta_3}^{t_\alpha^* - \Delta_3} \int_{-t_\alpha^* - \Delta_2}^{t_\alpha^* - \Delta_2} f^*(x_d, x_r) dx_d dx_r > \alpha,$$

其中 $f^*(x_d, x_r)$ 是二元正态分布 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{pmatrix}\right)$ 的密度函数,

$$\rho_2 = \sigma_{122} / \sqrt{\sigma_{112}\sigma_{222}}, \quad \alpha = 1 - \int_{-t_\alpha^*}^{t_\alpha^*} \int_{-t_\alpha^*}^{t_\alpha^*} f^*(x_d, x_r) dx_d dx_r; \quad (2.12)$$

(IV) 对于检验 T_{drx}^w ,

$$\lim_{p \rightarrow \infty} P(T_{drx}^w > t_\alpha) \geq \lim_{p \rightarrow \infty} P(T_{dr} > t_\alpha) > \alpha;$$

(V) 对于检验 T_{drx}^m ,

$$\lim_{p \rightarrow \infty} P(T_{drx}^m > t_{\alpha/2}) \geq 1 - \int_{-t_{\alpha/2}^* - \Delta_3}^{t_{\alpha/2}^* - \Delta_3} \int_{-t_{\alpha/2}^* - \Delta_2}^{t_{\alpha/2}^* - \Delta_2} f^*(x_d, x_r) dx_d dx_r > \alpha/2,$$

其中 $t_{\alpha/2}^*$ 满足(2.12)中用 $\alpha/2$ 替换 α 后的等式。

性质2.3.2表明, 若 $(\Sigma_1, \Sigma_2) \in \Pi_2$ 和 $\beta_2 = 0$, 则检验 T_d , 检验 T_r , 检验 T_{dr} 和检验 T_{drx}^w 是渐近无偏的。由于 Δ_2 是 a_2 的增函数, 当 a_2 较大时, 检验 T_d , 检验 T_{dr} , 检验 T_{drx}^w 和检验 T_{drx}^m 具有较高的功效。随着 a_2 的增加, Δ_3 收敛到

$$\frac{1}{\sqrt{\sigma_{222}}} \left[\frac{1 - 2(y_1 + y_2)}{(1 - y_2)^2} - \frac{2y_2}{(1 - y_2)^3} \right].$$

因此, 在 Δ_3 较小的情况下, 检验 T_r 将会失效。例如, 当 $y_1 = 0.5 - (2y_2 - y_2^2)/(1 - y_2)$ 和 $a_2 \rightarrow \infty$ 时, 检验 T_r 的功效函数就会收敛到 α 。

§2.3.3 备择集合 Π_3 下的功效分析

下面的性质表明, 在 $(\Sigma_1, \Sigma_2) \in \Pi_3$ 和其他一些条件下, 检验 T_{drx}^w 和检验 T_{drx}^m 是相合的。令

$$\Sigma_d = \Sigma_1 - \Sigma_2, \quad \Sigma_w = \frac{n_1}{n_1 + n_2} \Sigma_1 + \frac{n_2}{n_1 + n_2} \Sigma_2.$$

性质2.3.3. 当 $(\Sigma_1, \Sigma_2) \in \Pi_3$ 时, 有以下结果:

(I) 若假设2.2.1-假设2.2.2成立, 在Cai et al. (2013)^[22]中的条件(C2)(或(C2*))下, 如果 $p^{-1}\text{tr}\Sigma_d$, $p^{-1}\text{tr}\Sigma_w$ 和 $p^{-1}\text{tr}(\Sigma_w^2)$ 收敛且

$$\max_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{(\sigma_{1\ell_1\ell_2} - \sigma_{2\ell_1\ell_2})^2}{\theta_{1\ell_1\ell_2}/n_1 + \theta_{2\ell_1\ell_2}/n_2} \geq 2s(N_1, N_2, p) + 8 \log p,$$

那么

$$P(T_{\text{drx}}^w > t_\alpha) \rightarrow 1;$$

(II) 在Cai et al. (2013)^[22]中的条件(C2)(或(C2*))下,

$$P(T_{\text{drx}}^m > t_{\alpha/2}) \rightarrow 1.$$

性质2.3.3表明, 如果 Σ_d 中的某些元素足够大, 那么检验 T_{drx}^w 和检验 T_{drx}^m 的功效函数就会趋于1。

§2.4 数值模拟

本节通过大量的数值模拟研究新的检验方法在有限样本情形下的表现。样本 $\mathbf{x}_{ki} = \Sigma_k^{1/2} \mathbf{w}_{ki}$, 其中 $\{w_{k\ell i}, k = 1, 2; \ell = 1, \dots, p; i = 1, \dots, N_k\}$ 是来自高斯分布 $N(0, 1)$ 或伽马分布Gamma(4,2)-2的i.i.d.样本。 Σ_1 和 Σ_2 的设置(Scenario)如下:

- **设置1.** 总体协方差矩阵 $\Sigma_k = \Gamma_k \Gamma_k^T$, 其中 $\Gamma_k = \mathbf{I}_p + \theta_k(u_{\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$, $\{u_{\ell_1\ell_2}, \ell_1, \ell_2 = 1, \dots, p\} \stackrel{i.i.d.}{\sim} \text{Unif}(-(2p)^{-2/3}, (2p)^{-2/3})$, $k = 1, 2$ 。当 $\theta_1 = \theta_2 = 0$ 时, $\Sigma_1 = \Sigma_2$; 当 $(\theta_1, \theta_2) = (0, 1)$ 时, $\Sigma_1 - \Sigma_2$ 是极其稠密的。用 $\theta_1 = \theta_2 = 0$ 评估经验水平, 用 $(\theta_1, \theta_2) = (0, 1)$ 评估经验功效。Zheng et al. (2020)^[24]使用了这种设置。

- **设置2.** 令 $\Sigma_k = (\rho_k^{|\ell_1 - \ell_2|})_{\ell_1, \ell_2=1}^p$, 其中 $|\rho_k| < 1$, $k = 1, 2$ 。当 $\rho_1 = \rho_2$ 时, $\Sigma_1 = \Sigma_2$; 当 $\rho_1 \neq \rho_2$ 时, $\Sigma_1 - \Sigma_2$ 是相对稠密的。用 $\rho_1 = \rho_2 = 0.5$ 评估经验水平, 用 $(\rho_1, \rho_2) = (0.5, 0.55)$ 评估经验功效。
- **设置3.** 记 $\mathbf{C} = ((0.1^{|\ell_1 - \ell_2|} + 0.2^{|\ell_1 - \ell_2|})/2)_{\ell_1, \ell_2=1}^p$ 。令 \mathbf{U} 为一个 $p \times p$ 对称矩阵, 在其上三角仅有4个来自 $\text{Unif}(0.4, 0.6)$ 的非零项。这4个非零项的位置是从 \mathbf{U} 的上三角中随机选择的。对于 $k = 1, 2$, 令总体协方差矩阵为

$$\Sigma_k = \mathbf{C} + \delta \mathbf{I}_p + \theta_k \mathbf{U},$$

其中 $\delta = |\min\{\lambda_{\min}(\mathbf{C} + \mathbf{U}), \lambda_{\min}(\mathbf{C})\}| + 0.05$ 。当 $\theta_1 = \theta_2 = 0$ 时, $\Sigma_1 = \Sigma_2$; 当 $(\theta_1, \theta_2) = (0, 1)$ 时, $\Sigma_1 - \Sigma_2$ 是极其稀疏的。用 $\theta_1 = \theta_2 = 0$ 评估经验水平, 用 $(\theta_1, \theta_2) = (0, 1)$ 评估经验功效。Cai et al. (2013)^[22]使用了这种设置。

- **设置4.** 记 $\mathbf{C} = ((0.5 - |\ell_1 - \ell_2|/10)I(|\ell_1 - \ell_2| \leq 4))_{\ell_1, \ell_2=1}^p + 0.5\mathbf{I}_p$ 。令 \mathbf{U} 为一个 $p \times p$ 对称矩阵, 在其上三角有 $p^2/4$ 个来自 $\text{Unif}(0, 0.04)$ 的非零项。令 \mathbf{E} 也为一个 $p \times p$ 对称矩阵, 在其上三角仅有1个非零项 $\log p/10$ 。 \mathbf{U} 和 \mathbf{E} 中非零项的位置是从它们的上三角中随机选择的。对于 $k = 1, 2$, 令总体协方差矩阵为

$$\Sigma_k = \mathbf{C} + \delta \mathbf{I}_p + \theta_k(\mathbf{U} + \mathbf{E}),$$

其中 $\delta = |\min\{\lambda_{\min}(\mathbf{C} + \mathbf{U} + \mathbf{E}), \lambda_{\min}(\mathbf{C})\}| + 0.05$ 。当 $\theta_1 = \theta_2 = 0$ 时, $\Sigma_1 = \Sigma_2$; 当 $(\theta_1, \theta_2) = (0, 1)$ 时, $\Sigma_1 - \Sigma_2$ 混合了稠密和稀疏的信号。用 $\theta_1 = \theta_2 = 0$ 评估经验水平, 用 $(\theta_1, \theta_2) = (0, 1)$ 评估经验功效。

以上四种设置均满足Cai et al. (2013)^[22]中的条件(C1), (C2*)和(C3)(见本章附录§2.6.2)。本节将新提出的三种检验方法与现有的两种检验方法LC(Li和Chen (2012)^[21])和CLX(Cai et al. (2013)^[22])进行比较。显著性水平 α 设置为0.05。根据每种设置下的10,000次独立重复试验, 表2.1-表2.4报告了在高斯分布和伽马分布下的经验水平, 表2.5-表2.8报告了在高斯分布和伽马分布下的经验功效。为了直观地展示这几种检验方法的经验功效, 图2.1和图2.2分别给出了在高斯分布和伽马分布下以表2.5-表2.8中的12种情况(Case)为横坐标, 每种方法的经验功效为纵坐标的折线图。

由表2.1-表2.4可知，所有检验方法的经验水平都接近给定的显著水平0.05。对于检验功效，根据表2.5-表2.8和图2.1-图2.2可以看出这七个检验方法的相对表现取决于不同的设置。

- 在设置1-设置2下，检验CLX的经验功效很低，这说明它无法检测到密集但很小的扰动。在设置1下，检验 T_d 的表现优于检验 T_r ；而在设置2下，检验 T_r 的经验功效远高于检验 T_d 。由于检验 T_{dr} ，检验 T_{drx}^w 和检验 T_{drx}^m 都结合了 T_d 和 T_r 的信息，因此这三个方法可以在设置1-设置2下具有较高的检验功效。除此之外，在这两种设置下，检验 T_{drx}^w 比检验 T_{drx}^m 具有更高的检验功效，这与注释2.2.3是一致的。
- 设置3考查在稀疏备择假设下所考虑的检验方法的表现，正如预期的那样，检验 T_d ，检验 T_r 和检验 T_{dr} 无法检测到稀疏信号并具有较低的经验功效。相比之下，检验 T_{drx}^w 和检验 T_{drx}^m 是适应的，具有较高的经验功效。设置4是一个混合了稠密和稀疏信号的备择假设，在这种情况下，检验 T_{drx}^w 和检验 T_{drx}^m 在经验功效上比其他方法有更好的表现。因为在后两种设置中， Σ_1 和 Σ_2 存在稀疏差异，所以检验 T_{drx}^m 的表现优于检验 T_{drx}^w 。

综上所述，数值模拟结果表明，新提出的三种检验方法能够很好地控制犯第一类错误的概率，并且可以有效地利用每个单独的统计量的信息。

表 2.1 当样本来自高斯分布和伽马分布时, 在设置1下, 新的检验方法与检验LC和检验CLX的经验水平(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	5.42	6.06	6.11	5.84	5.03	6.54	5.45
2	40	(120,120)	5.09	6.09	6.23	5.53	4.85	6.69	6.01
3	40	(120,160)	5.33	5.52	5.51	5.49	4.88	5.99	5.24
4	80	(120,160)	5.13	6.35	6.05	5.32	4.81	6.55	5.74
5	80	(160,240)	4.94	5.42	5.29	5.42	5.21	5.63	5.2
6	80	(240,240)	5.01	5.52	5.59	5.31	4.69	5.77	5.21
7	160	(160,240)	5.26	6.07	5.98	5.15	4.08	6.26	5.29
8	160	(240,240)	5.28	6.48	6.27	5	4.65	6.62	5.51
9	160	(240,320)	5.11	5.38	5.3	4.94	4.48	5.48	5.11
10	320	(240,480)	5.12	5.43	5.46	5.28	4.38	5.6	4.9
11	320	(320,480)	4.61	5.37	5.33	4.76	4.71	5.48	4.97
12	320	(480,480)	4.79	5.59	5.5	4.72	4.62	5.59	4.92
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	6.64	6.65	7.76	6.26	3.98	8.1	6.37
2	40	(120,120)	6.43	7.01	7.71	6.41	3.42	7.93	6.1
3	40	(120,160)	6.61	6.85	7.46	6.33	3.53	7.69	5.78
4	80	(120,160)	5.64	6.51	6.63	5.72	3.36	6.9	5.44
5	80	(160,240)	5.8	6.15	6.45	5.49	3.39	6.64	5.24
6	80	(240,240)	5.76	6.35	6.31	5.41	3.62	6.46	5
7	160	(160,240)	5.56	7.09	6.86	5.6	3.2	7.04	5.28
8	160	(240,240)	4.83	6.84	6.2	4.77	3.35	6.4	4.95
9	160	(240,320)	5.5	5.59	5.63	5.08	3.34	5.76	4.43
10	320	(240,480)	5.22	5.9	5.83	5.1	3.42	5.95	4.69
11	320	(320,480)	5.31	5.75	5.84	5.38	3.17	5.88	4.57
12	320	(480,480)	5.01	5.54	5.6	4.95	3.45	5.67	4.57

表 2.2 当样本来自高斯分布和伽马分布时,在设置2下,新的检验方法与检验LC和检验CLX的经验水平(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	5.75	6.11	6.53	6.59	5	7.08	6.23
2	40	(120,120)	6.12	6.23	6.74	6.48	4.45	7.2	5.92
3	40	(120,160)	5.85	6.13	6.39	6.5	4.63	6.8	5.87
4	80	(120,160)	5.31	6.01	6.1	6	4.83	6.55	5.77
5	80	(160,240)	5.37	5.53	5.57	5.65	4.7	5.93	5.1
6	80	(240,240)	5.68	5.46	5.85	5.98	4.71	6.14	5.26
7	160	(160,240)	5.36	6.25	5.99	5.65	4.66	6.26	5.52
8	160	(240,240)	5.17	5.56	5.44	5.13	4.64	5.73	5.29
9	160	(240,320)	4.84	5.35	5.3	4.77	4.46	5.56	4.65
10	320	(240,480)	4.85	5.69	5.48	5.08	4.48	5.57	5.13
11	320	(320,480)	4.99	5.86	5.45	5.13	4.76	5.61	4.89
12	320	(480,480)	5.09	5.72	5.41	5.32	4.75	5.55	5.33
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	6.65	6.89	7.79	6.91	4.09	8.19	6.49
2	40	(120,120)	6.25	7.2	7.8	6.67	3.27	8.12	6.12
3	40	(120,160)	6.73	6.83	7.86	7.14	4.3	8.05	6.56
4	80	(120,160)	6.18	7.01	7.14	6.51	3.62	7.43	5.83
5	80	(160,240)	5.91	5.75	6.09	5.97	3.93	6.36	5.02
6	80	(240,240)	5.92	5.85	6.31	6.23	3.51	6.54	5.1
7	160	(160,240)	5.42	6.42	6.63	5.46	3.2	6.79	5.24
8	160	(240,240)	5.64	6.97	6.57	5.63	3.68	6.71	5.27
9	160	(240,320)	5.84	5.79	6.06	5.62	3.15	6.26	4.76
10	320	(240,480)	5.11	5.39	5.08	5.52	3.63	5.14	4.06
11	320	(320,480)	5.23	5.53	5.72	5.35	3.7	5.8	4.64
12	320	(480,480)	5.43	5.79	5.98	5.51	3.74	6.06	4.98

表 2.3 当样本来自高斯分布和伽马分布时, 在设置3下, 新的检验方法与检验LC和检验CLX的经验水平(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	4.96	5.86	5.95	5.32	5.39	6.4	5.89
2	40	(120,120)	5.44	6.18	6.18	5.46	4.83	6.57	5.6
3	40	(120,160)	5.52	6.03	6.17	5.8	4.42	6.56	5.58
4	80	(120,160)	5.29	6.23	6.09	5.3	4.86	6.44	5.77
5	80	(160,240)	5.3	5.76	5.55	4.96	4.79	5.84	4.95
6	80	(240,240)	5.31	5.51	5.5	5.29	5.24	5.89	5.27
7	160	(160,240)	4.71	6.17	5.94	5.12	4.9	6.18	5.74
8	160	(240,240)	4.87	6.02	6.17	4.9	4.41	6.5	5.43
9	160	(240,320)	5.11	5.67	5.65	4.84	4.53	5.89	4.89
10	320	(240,480)	5.03	5.73	5.57	5.26	4.48	5.74	4.9
11	320	(320,480)	4.99	5.45	5.62	5.49	4.36	5.73	5.01
12	320	(480,480)	5.05	5.58	5.48	4.93	4.84	5.59	5.31
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	6.49	6.47	7.54	6.21	3.65	7.75	5.92
2	40	(120,120)	6.67	7.68	7.97	6.33	3.47	8.22	6.13
3	40	(120,160)	6.3	6.98	7.22	6.32	3.9	7.57	5.88
4	80	(120,160)	5.45	6.7	6.77	5.73	3.25	7.01	5.42
5	80	(160,240)	5.5	5.89	5.76	5.25	3.58	5.98	4.92
6	80	(240,240)	5.63	5.78	6.01	5.32	3.61	6.16	5.19
7	160	(160,240)	5.48	6.71	6.48	5.4	3.54	6.67	5.26
8	160	(240,240)	5.25	6.7	6.76	5.37	3.02	6.87	5.3
9	160	(240,320)	5.33	5.65	5.51	5.43	3.49	5.67	4.51
10	320	(240,480)	4.92	5.37	5.28	5.51	3.67	5.35	4.31
11	320	(320,480)	5.36	5.61	5.72	5.26	3.26	5.82	4.37
12	320	(480,480)	5.65	5.96	6.05	5.57	3.73	6.15	4.84

表 2.4 当样本来自高斯分布和伽马分布时, 在设置4下, 新的检验方法与检验LC和检验CLX的经验水平(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	6.17	5.73	6.7	6.53	5.29	7.2	6.19
2	40	(120,120)	5.65	6.26	6.46	6.19	4.92	6.98	6.03
3	40	(120,160)	5.48	5.77	5.95	6.2	4.78	6.4	5.62
4	80	(120,160)	5.76	6.18	6.23	5.54	4.85	6.62	5.67
5	80	(160,240)	5.97	5.77	6.13	6	4.52	6.46	5.34
6	80	(240,240)	5.36	5.88	5.93	5.56	4.54	6.29	5.29
7	160	(160,240)	5.14	6.09	5.83	5.35	4.83	6.16	5.34
8	160	(240,240)	5.16	6.39	6.33	5.37	4.31	6.66	5.54
9	160	(240,320)	5.12	5.47	5.48	5.22	4.14	5.7	4.84
10	320	(240,480)	4.88	5.05	4.94	5.05	4.3	5.14	4.77
11	320	(320,480)	5.24	5.46	5.53	5.48	4.6	5.7	4.82
12	320	(480,480)	5.44	5.82	5.86	5.28	4.47	6.03	5.05
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	7.27	6.82	7.91	7.03	4.24	8.29	6.5
2	40	(120,120)	6.83	7.22	7.84	6.74	3.54	8.05	6.62
3	40	(120,160)	7.36	6.82	7.95	7.46	3.99	8.27	6.18
4	80	(120,160)	6.02	6.46	6.57	6.33	3.66	6.8	5.54
5	80	(160,240)	5.94	5.94	6.53	6.02	3.66	6.69	5.15
6	80	(240,240)	5.96	5.96	6.43	5.99	3.61	6.6	5.36
7	160	(160,240)	5.29	6.36	6.2	5.74	3.54	6.42	5.06
8	160	(240,240)	5.7	6.94	6.67	5.62	3.51	6.75	5.33
9	160	(240,320)	5.49	5.76	5.85	5.41	3.7	5.98	4.81
10	320	(240,480)	5.65	5.85	5.97	5.49	3.68	6.12	4.95
11	320	(320,480)	5.08	5.62	5.66	5.01	3.82	5.74	4.62
12	320	(480,480)	5.13	6.03	5.69	5.45	3.55	5.78	4.77

表 2.5 当样本来自高斯分布和伽马分布时,在设置1下,新的检验方法与检验LC和检验CLX的经验功效(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	46.49	20.29	42.94	58.77	10.25	43.45	36.36
2	40	(120,120)	65.63	23.24	61.25	75.68	10.82	61.61	53.9
3	40	(120,160)	74.78	34.25	71.4	83.46	12.03	71.71	63.92
4	80	(120,160)	54.51	21.82	50.26	66.3	7.43	50.64	42.19
5	80	(160,240)	82.41	46.57	80.12	89.3	9.09	80.24	73.08
6	80	(240,240)	94.85	55.07	93.14	97.24	9.61	93.19	89.26
7	160	(160,240)	64.41	15.3	57.14	75.14	6.43	57.31	48.36
8	160	(240,240)	82.63	15.97	76.66	89.31	6.42	76.69	68.85
9	160	(240,320)	90.96	33.43	87.73	94.77	7.23	87.73	81.55
10	320	(240,480)	85.59	20.12	80.2	91.77	5.65	80.26	72.7
11	320	(320,480)	95.26	20.68	92.64	97.75	5.93	92.66	88.41
12	320	(480,480)	99.45	22.55	98.96	99.81	6.41	98.96	98.13
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	44.14	19.89	41.78	58.19	8.01	42.18	35.46
2	40	(120,120)	63.87	22.94	59.84	75.09	7.66	60.01	52.1
3	40	(120,160)	73.06	32.68	70.14	82.33	9.34	70.33	62.54
4	80	(120,160)	52.32	21.11	48.4	65.04	5.63	48.56	39.91
5	80	(160,240)	80.15	43.33	77.84	88.17	6.75	77.89	70.61
6	80	(240,240)	93.9	50.18	92.39	96.87	7.48	92.41	88.14
7	160	(160,240)	63.43	16.03	56.4	75.11	4.5	56.54	46.82
8	160	(240,240)	81.81	16.62	76.07	88.91	4.64	76.13	67.77
9	160	(240,320)	90.12	32.14	86.5	94.67	5.11	86.52	80.43
10	320	(240,480)	84.51	20.03	78.71	91.6	4.52	78.73	70.39
11	320	(320,480)	94.97	21.75	91.77	97.49	4.63	91.78	87.43
12	320	(480,480)	99.51	22.43	99.03	99.79	4.96	99.03	98.18

表 2.6 当样本来自高斯分布和伽马分布时, 在**设置2**下, 新的检验方法与检验LC和检验CLX的经验功效(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	6.59	28.64	25.07	9.24	6.2	25.64	21.32
2	40	(120,120)	8.32	30.77	27.52	11.01	5.11	27.87	22.87
3	40	(120,160)	7.83	39.17	34.6	11.14	6.16	34.87	28.93
4	80	(120,160)	7.32	37.44	32.54	10.67	5.63	32.91	27.12
5	80	(160,240)	8.08	68.49	61.18	12.4	6.08	61.3	53.34
6	80	(240,240)	10.79	74.19	68.1	15.87	5.75	68.16	60.83
7	160	(160,240)	7.66	46.33	40.3	12.15	5.84	40.6	33.73
8	160	(240,240)	9.63	48.79	43.37	15.26	5.37	43.6	36.19
9	160	(240,320)	10.01	87.06	81.78	15.99	5.56	81.88	75.25
10	320	(240,480)	10.48	93.17	90.38	17.8	6.72	90.39	85.86
11	320	(320,480)	14.85	94.25	91.58	23.66	6.31	91.6	87.74
12	320	(480,480)	21.36	95.88	93.67	31.03	6.02	93.69	89.92
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	7.89	27.61	24.98	9.91	5.44	25.44	20.69
2	40	(120,120)	8.54	29.84	27.22	11.44	4.11	27.45	22.34
3	40	(120,160)	9.06	37.6	34.22	12.15	5.65	34.46	28.61
4	80	(120,160)	7.6	37.4	32.49	10.76	4.62	32.75	27.27
5	80	(160,240)	8.18	63.73	56.82	12.94	5.64	57.07	49.81
6	80	(240,240)	11.46	69.29	62.34	16.65	4.28	62.44	53.8
7	160	(160,240)	7.42	44.57	39.29	12.42	4.32	39.52	32.82
8	160	(240,240)	10.21	47.85	43.17	15.24	4.31	43.27	35.86
9	160	(240,320)	10.59	84.11	78.44	16.54	4.53	78.49	71.62
10	320	(240,480)	11.23	92.9	89.25	18.93	6.01	89.27	84.58
11	320	(320,480)	14.54	93.86	90.87	22.6	5.16	90.88	86.33
12	320	(480,480)	21.46	95.25	92.78	31.33	4.73	92.8	89.1

表 2.7 当样本来自高斯分布和伽马分布时,在设置3下,新的检验方法与检验LC和检验CLX的经验功效(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	10.33	16.73	16.5	16.26	29.81	24.2	29.21
2	40	(120,120)	13.86	19	20.31	20.98	39.05	30.9	37.74
3	40	(120,160)	16.48	24.98	25.04	24.66	52.75	40.68	50.93
4	80	(120,160)	7.55	9.57	9.93	11.78	32.01	19.29	27.8
5	80	(160,240)	9.75	13.74	13.59	15.43	62.21	38.98	55.7
6	80	(240,240)	12.94	15.89	16.66	19.85	79.95	54.76	73.89
7	160	(160,240)	5.9	7.73	7.69	9.03	46.12	24.15	39.45
8	160	(240,240)	6.7	7.6	7.97	10.38	65.57	37.22	58.22
9	160	(240,320)	7	8.38	8.83	11.43	80.26	52.32	73.94
10	320	(240,480)	5.99	6.83	6.73	8.4	84.89	56.62	79.6
11	320	(320,480)	6.58	6.64	6.62	9.36	95.2	74.93	92.28
12	320	(480,480)	6.82	6.43	7.2	10.73	99.45	92.89	98.96
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	11.49	17.04	17.86	18.13	28.68	24.88	29.29
2	40	(120,120)	14.48	18.61	20.72	22.12	36.89	30.22	37.12
3	40	(120,160)	17.07	24.65	25.64	25.83	52.63	41.4	51.19
4	80	(120,160)	8	10.4	11	12.82	30.57	19.77	27.77
5	80	(160,240)	9.87	13.19	13.65	15.44	59.92	38.61	54.54
6	80	(240,240)	12.96	14.35	15.86	19.69	77.09	52.16	71.18
7	160	(160,240)	6.36	8.13	8.21	9.54	45.46	24.12	39.39
8	160	(240,240)	7.15	8.14	8.95	10.44	61.49	34.69	55.13
9	160	(240,320)	7.52	8.32	9.18	12	78.66	51.86	72.31
10	320	(240,480)	6.04	6.28	6.4	8.75	83.94	56.99	78.14
11	320	(320,480)	6.27	6.57	7.01	9.13	93.99	73.52	91.04
12	320	(480,480)	7.36	6.77	7.79	11.22	99.27	90.62	98.51

表 2.8 当样本来自高斯分布和伽马分布时,在设置4下,新的检验方法与检验LC和检验CLX的经验功效(百分比)。

Case	p	(N_1, N_2)	T_d	T_r	T_{dr}	LC	CLX	T_{drx}^w	T_{drx}^m
$N(0, 1)$									
1	40	(80,120)	6.37	9.95	9.74	8.32	7.79	10.89	10.48
2	40	(120,120)	6.41	10.48	10.74	8.8	8.54	12.07	11.5
3	40	(120,160)	6.65	11.82	11.33	9.31	10.83	13.04	13.21
4	80	(120,160)	6.8	13.82	12.91	9.94	15.04	16.27	17.58
5	80	(160,240)	8.45	23.84	21.64	12.7	28.51	29.9	33.08
6	80	(240,240)	10.02	27.07	24.07	15.02	42.55	38	45.05
7	160	(160,240)	12.94	23.77	23.41	20.24	33.38	34.02	38.55
8	160	(240,240)	17.53	24.68	27.29	26.22	48.96	45.36	53.45
9	160	(240,320)	20.77	53.66	50.06	30.21	63.57	68.04	73.47
10	320	(240,480)	76.75	99.98	99.95	85.39	94.43	100	99.99
11	320	(320,480)	89.17	100	99.99	94.12	98.54	100	100
12	320	(480,480)	97.78	100	100	98.87	99.91	100	100
$\text{Gamma}(4,2)-2$									
1	40	(80,120)	7.54	10.86	11.24	9.31	6.48	12.1	10.54
2	40	(120,120)	7.71	11.69	12.31	9.6	7.08	13.28	11.6
3	40	(120,160)	8.35	12.62	13.17	10.4	9.65	14.81	13.56
4	80	(120,160)	7.54	13.82	13.43	10.86	12.96	16.15	17.01
5	80	(160,240)	8.72	22.18	19.94	13.18	26.94	27.86	31.66
6	80	(240,240)	10.6	24.77	22.8	15.52	38.99	35.5	42.24
7	160	(160,240)	12.85	23.98	24.16	19.87	30.69	33.45	37.37
8	160	(240,240)	17.2	25.3	27.62	26.39	45.11	42.28	49.51
9	160	(240,320)	21.31	51.85	48.88	30.98	59.54	65.66	70.76
10	320	(240,480)	76.13	99.96	99.92	84.79	90.68	99.98	99.96
11	320	(320,480)	89.09	99.97	100	93.75	96.66	100	100
12	320	(480,480)	97.18	100	100	98.93	99.55	100	100

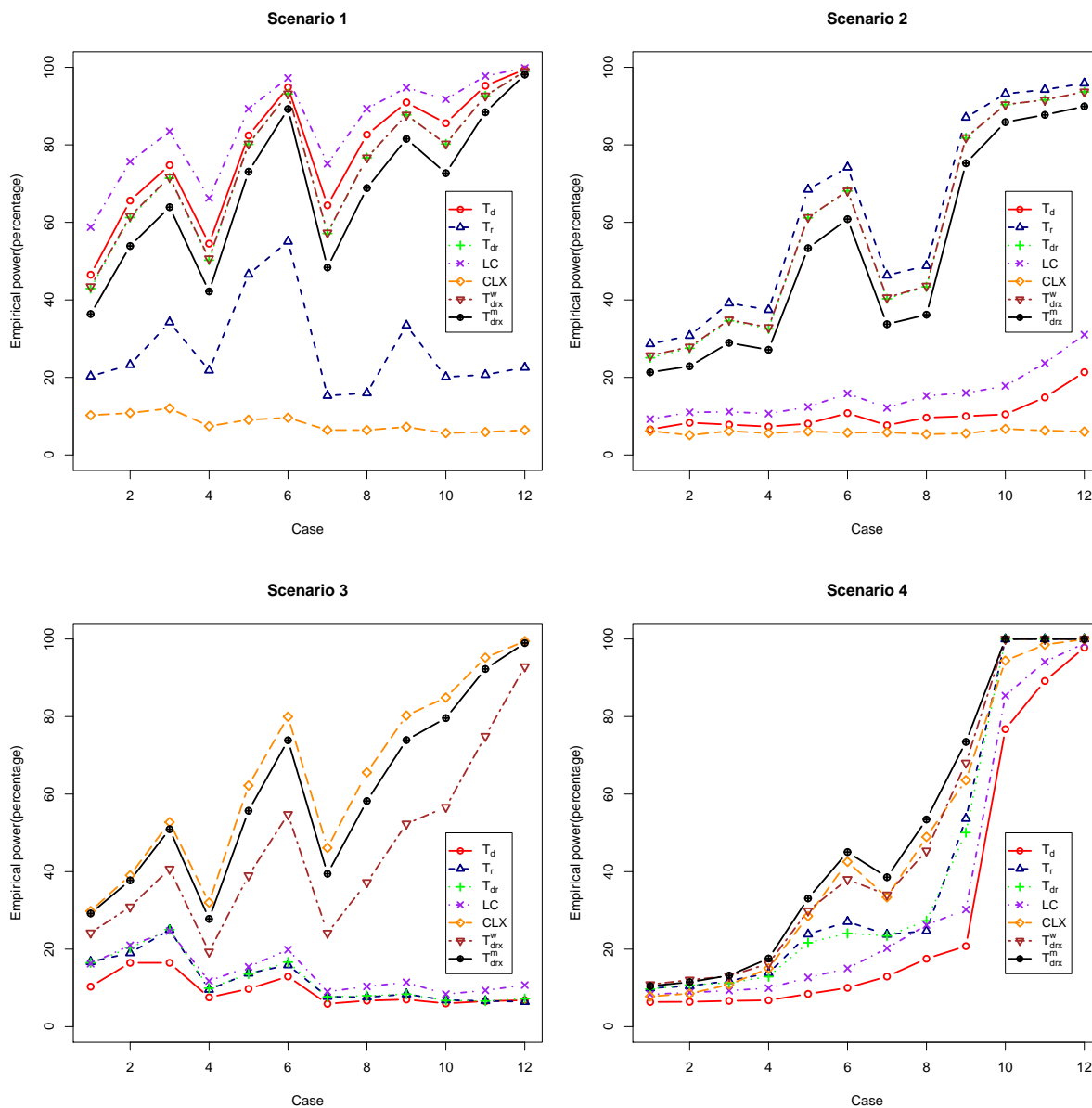


图 2.1 当样本来自高斯分布时，在4种设置下，新的检验方法与检验LC和检验CLX的经验功效(百分比)。Case的取值为1到12，对应于表2.5-表2.8中的第一列。

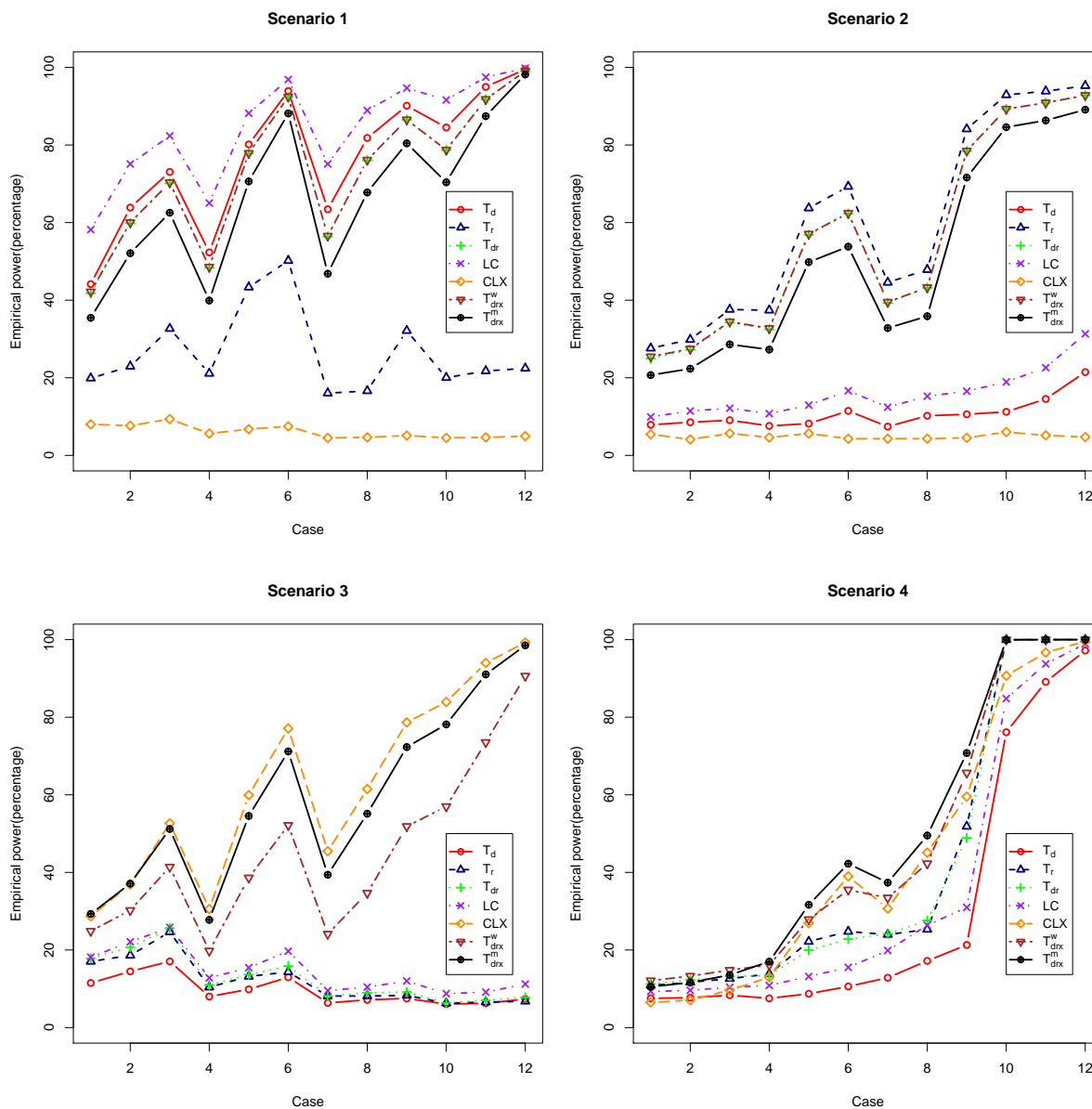


图 2.2 当样本来自伽马分布时，在4种设置下，新的检验方法与检验LC和检验CLX的经验功效(百分比)。Case的取值为1到12，对应于表2.5-表2.8中的第一列。

§2.5 理论证明

本节分为三小节。§2.5.1节给出了一些准备工作。§2.5.2节包含了本章所有定理和性质的证明。§2.5.3节给出了§2.5.1节中使用的所有引理。

§2.5.1 证明定理2.2.1, 定理2.3.1和定理2.3.2的准备工作

准备工作包括两个步骤, 这两个步骤的框架如下:

步骤1的框架. 定义 $T_d - \mu_0$ 和 T_r 的线性组合 $T = \omega_1(T_d - \mu_0) + \omega_2 T_r$, 其中 ω_1 和 ω_2 是两个任意常数。在这一步中, 将得到

$$T = T_A + T_B + \omega_1 \mu_1 + \omega_2 \mu_2 + o_p(1),$$

其中 T_A 和 T_B 在(2.24)和(2.25)中给出, 而 μ_1 和 μ_2 在(2.22)和(2.23)中给出。

步骤2的框架. 令 $E_0(\cdot)$ 代表期望, $E_j(\cdot)$ 代表关于 $\mathbf{w}_{21}, \dots, \mathbf{w}_{2j}$ 生成的 σ -域的条件期望, $E_{2,i}(\cdot)$ 代表关于 $\mathbf{w}_{21}, \dots, \mathbf{w}_{2N_2}, \mathbf{w}_{11}, \dots, \mathbf{w}_{1i}$ 生成的 σ -域的条件期望, 则有

$$T_A + T_B = \sum_{j=1}^{N_2} (E_j - E_{j-1})T_A + \sum_{i=1}^{N_1} (E_{2,i} - E_{2,i-1})T_B,$$

换句话说, $T_A + T_B$ 是鞅差序列

$$\{(E_j - E_{j-1})T_A, (E_{2,i} - E_{2,i-1})T_B, j = 1, \dots, N_2, i = 1, \dots, N_1\}$$

的和。因此, 根据鞅差序列的中心极限定理(Billingsley (1995)^[52], p.476, 定理35.12), 只需考虑

$$\sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1})T_A)^2] + \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1})T_B)^2].$$

经过冗长且繁琐的推导可得

$$\begin{aligned} & \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1})T_A)^2] + \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1})T_B)^2] \\ &= \omega_1^2 \sigma_{11p} + 2\omega_1 \omega_2 (\sigma_{12p}^1 + \sigma_{12p}^2) + \omega_2^2 (\sigma_{22p}^1 + \sigma_{22p}^2) + o_p(1), \end{aligned}$$

其中 σ_{11p} , σ_{12p}^1 , σ_{12p}^2 , σ_{22p}^1 和 σ_{22p}^2 分别在(2.34), (2.35), (2.36), (2.37)和(2.38)中给出。

当定理2.2.1, 定理2.3.1和定理2.3.2中的条件成立时, 基于 σ_{11p} , σ_{12p}^1 , σ_{12p}^2 , σ_{22p}^1 和 σ_{22p}^2 的表达式, §2.5.2节将给出这三个定理的证明。

接下来, 分别给出这两个步骤的详细过程。

步骤1: 根据 T_d , T_r 和 μ_0 的定义, 可得

$$\begin{aligned}
 T &= \omega_1(T_d - \mu_0) + \omega_2 T_r \\
 &= \omega_1[\text{tr}\mathbf{S}_1^2 - 2\text{tr}(\mathbf{S}_1\mathbf{S}_2) + \text{tr}\mathbf{S}_2^2 - n_1^{-1}\text{tr}^2\mathbf{S}_1 - n_2^{-1}\text{tr}^2\mathbf{S}_2] \\
 &\quad + \omega_2[\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - 2\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) + p] \\
 &= \omega_1[\text{tr}\mathbf{S}_1^2 - 2\text{tr}(\mathbf{S}_1\mathbf{S}_2) + \text{tr}\mathbf{S}_2^2 - n_1^{-1}(\text{tr}^2\mathbf{S}_1 - \text{tr}^2\boldsymbol{\Sigma}_1) - n_1^{-1}\text{tr}^2\boldsymbol{\Sigma}_1 \\
 &\quad + n_2^{-1}(\text{tr}^2\mathbf{S}_2 - \text{tr}^2\boldsymbol{\Sigma}_2) - n_2^{-1}\text{tr}^2\boldsymbol{\Sigma}_2] \\
 &\quad + \omega_2[\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - 2\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) + p].
 \end{aligned}$$

根据Zheng et al. (2020)^[24]的补充材料中引理1的证明和Zheng et al. (2015)^[53]中定理3.1的证明, 可以将随机变量 $w_{k\ell i}$ 截断成 $w_{k\ell i}I(|w_{k\ell i}| \leq \eta_{N_k} \sqrt{N_k})$, 其中 $\eta_{N_k} = o(1)$, 而不影响 T 的渐近分布。为了符号简便, 仍用 $w_{k\ell i}$ 表示中心化和标准化的截断变量 $w_{k\ell i}I(|w_{k\ell i}| \leq \eta_{N_k} \sqrt{N_k})$ 。将非中心化的样本协方差矩阵表示为

$$\mathbf{B}_k = N_k^{-1} \sum_{i=1}^{N_k} \boldsymbol{\Sigma}_k^{1/2} \mathbf{w}_{ki} \mathbf{w}_{ki}^T \boldsymbol{\Sigma}_k^{1/2}, \quad k = 1, 2.$$

令 $\boldsymbol{\gamma}_{ki} = (1/\sqrt{N_k})\mathbf{w}_{ki}$, 则有

$$\mathbf{S}_k = \mathbf{B}_k - n_k^{-1} \sum_{i \neq j} \boldsymbol{\Sigma}_k^{1/2} \boldsymbol{\gamma}_{ki} \boldsymbol{\gamma}_{kj}^T \boldsymbol{\Sigma}_k^{1/2}, \quad k = 1, 2. \quad (2.13)$$

基于(2.13)可得 $\text{tr}\mathbf{S}_1 = \text{tr}\mathbf{B}_1 - n_1^{-1} \sum_{i \neq j} \boldsymbol{\gamma}_{1j}^T \boldsymbol{\Sigma}_1 \boldsymbol{\gamma}_{1i}$ 。在假设2.2.1-假设2.2.2下, 根据等式(2.86), 可得 $n_1^{-1} \sum_{i \neq j} \boldsymbol{\gamma}_{1j}^T \boldsymbol{\Sigma}_1 \boldsymbol{\gamma}_{1i} = o_p(1)$ 和

$$E(\text{tr}\mathbf{B}_1 - \text{tr}\boldsymbol{\Sigma}_1)^2 = N_1^{-1} [2\text{tr}\boldsymbol{\Sigma}_1^2 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1)], \quad (2.14)$$

由此可得 $n_1^{-1}(\text{tr}\mathbf{S}_1 - \text{tr}\boldsymbol{\Sigma}_1)^2 = o_p(1)$ 。因此,

$$n_1^{-1}(\text{tr}^2\mathbf{S}_1 - \text{tr}^2\boldsymbol{\Sigma}_1) = 2n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1(\text{tr}\mathbf{S}_1 - \text{tr}\boldsymbol{\Sigma}_1) + o_p(1).$$

同理可得

$$n_2^{-1}(\text{tr}^2\mathbf{S}_2 - \text{tr}^2\boldsymbol{\Sigma}_2) = 2n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2(\text{tr}\mathbf{S}_2 - \text{tr}\boldsymbol{\Sigma}_2) + o_p(1).$$

根据上面的讨论, 令

$$\begin{aligned} T^* = & \omega_1[\text{tr}\mathbf{S}_1^2 - 2\text{tr}(\mathbf{S}_1\mathbf{S}_2) + \text{tr}\mathbf{S}_2^2 - 2n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1(\text{tr}\mathbf{S}_1 - \text{tr}\boldsymbol{\Sigma}_1) \\ & - n_1^{-1}\text{tr}^2\boldsymbol{\Sigma}_1 - 2n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2(\text{tr}\mathbf{S}_2 - \text{tr}\boldsymbol{\Sigma}_2) - n_2^{-1}\text{tr}^2\boldsymbol{\Sigma}_2] \\ & + \omega_2[\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - 2\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) + p]. \end{aligned}$$

将 T^* 改写成

$$T^* = T^* - \text{E}(T^*|\mathbf{S}_2) + \text{E}(T^*|\mathbf{S}_2),$$

其中

$$\begin{aligned} \text{E}(T^*|\mathbf{S}_2) = & \omega_1[\text{Etr}\mathbf{S}_1^2 - 2\text{tr}(\mathbf{S}_2\boldsymbol{\Sigma}_1) + \text{tr}\mathbf{S}_2^2 - n_1^{-1}\text{tr}^2\boldsymbol{\Sigma}_1 \\ & - 2n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2(\text{tr}\mathbf{S}_2 - \text{tr}\boldsymbol{\Sigma}_2) - n_2^{-1}\text{tr}^2\boldsymbol{\Sigma}_2] \\ & + \omega_2[\text{E}(\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2|\mathbf{S}_2) - 2\text{E}(\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})|\mathbf{S}_2) + p], \end{aligned}$$

$$\begin{aligned} T^* - \text{E}(T^*|\mathbf{S}_2) = & \omega_1[\text{tr}\mathbf{S}_1^2 - \text{Etr}\mathbf{S}_1^2 - 2(\text{tr}(\mathbf{S}_1\mathbf{S}_2) - \text{tr}(\mathbf{S}_2\boldsymbol{\Sigma}_1)) \\ & - 2n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1(\text{tr}\mathbf{S}_1 - \text{tr}\boldsymbol{\Sigma}_1)] \\ & + \omega_2[\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - \text{E}(\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2|\mathbf{S}_2) \\ & - 2(\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) - \text{E}(\text{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})|\mathbf{S}_2))]. \end{aligned}$$

由(2.13)和(2.88)可得

$$\text{Etr}\mathbf{S}_1^2 = n_1^{-1}\text{tr}^2\boldsymbol{\Sigma}_1 + \text{tr}\boldsymbol{\Sigma}_1^2 + n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1^2 + \beta_1 n_1^{-1}\text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1) + o(1), \quad (2.15)$$

$$\text{Etr}\mathbf{S}_2^2 = n_2^{-1}\text{tr}^2\boldsymbol{\Sigma}_2 + \text{tr}\boldsymbol{\Sigma}_2^2 + n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2^2 + \beta_2 n_2^{-1}\text{tr}(\boldsymbol{\Sigma}_2 \circ \boldsymbol{\Sigma}_2) + o(1). \quad (2.16)$$

当 $\boldsymbol{\Sigma}_1$ 可逆时, 令 $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}_1^{-1/2}\boldsymbol{\Sigma}_2^{1/2}$, $\tilde{\boldsymbol{\Sigma}}_2 = \boldsymbol{\Sigma}_1^{-1/2}\mathbf{S}_2\boldsymbol{\Sigma}_1^{-1/2}$ 和 $\mathbf{T}_p = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T$. 在假设2.2.1-假设2.2.2下, 若 \mathbf{T}_p 的经验谱分布 H_p 满足§2.5.3节中的假设c-d-f, 则 $\|\tilde{\boldsymbol{\Sigma}}_2^{-1}\|$ 是几乎处处有界的. 因此, 由(2.15)和(2.16)可得

$$\begin{aligned} \text{E}(T^*|\mathbf{S}_2) = & \omega_1[\text{tr}\mathbf{S}_2^2 - \text{Etr}\mathbf{S}_2^2 - 2(\text{tr}(\mathbf{S}_2\boldsymbol{\Sigma}_1) - \text{tr}(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1)) - 2n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2(\text{tr}\mathbf{S}_2 - \text{tr}\boldsymbol{\Sigma}_2) \\ & + \text{Etr}\mathbf{S}_1^2 + \text{Etr}\mathbf{S}_2^2 - 2\text{tr}(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1) - n_1^{-1}\text{tr}^2\boldsymbol{\Sigma}_1 - n_2^{-1}\text{tr}^2\boldsymbol{\Sigma}_2] \\ & + \omega_2[n_1^{-1}\text{tr}^2(\mathbf{S}_2^{-1}\boldsymbol{\Sigma}_1) + \text{tr}(\mathbf{S}_2^{-1}\boldsymbol{\Sigma}_1)^2 - 2\text{tr}(\mathbf{S}_2^{-1}\boldsymbol{\Sigma}_1) + n_1^{-1}\text{tr}(\mathbf{S}_2^{-1}\boldsymbol{\Sigma}_1)^2] \end{aligned}$$

$$\begin{aligned}
 & +\beta_1 n_1^{-1} \text{tr}(\mathbf{\Sigma}_1^{1/2} \mathbf{S}_2^{-1} \mathbf{\Sigma}_1^{1/2} \circ \mathbf{\Sigma}_1^{1/2} \mathbf{S}_2^{-1} \mathbf{\Sigma}_1^{1/2}) + p] + o_p(1) \\
 = & \omega_1 [\text{tr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1)^2 - \text{Etr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1^2) - \text{tr}(\mathbf{\Sigma}_2 \mathbf{\Sigma}_1)) \\
 & - 2n_2^{-1} \text{tr} \mathbf{\Sigma}_2 (\text{tr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1) - \text{tr} \mathbf{\Sigma}_2) + \text{tr}(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2 + n_1^{-1} \text{tr} \mathbf{\Sigma}_1^2 \\
 & + n_2^{-1} \text{tr} \mathbf{\Sigma}_2^2 + \beta_1 n_1^{-1} \text{tr}(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_1) + \beta_2 n_2^{-1} \text{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2)] \\
 + \omega_2 & [n_1^{-1} \text{tr}^2 \widetilde{\mathbf{S}}_2^{-1} - n_1^{-1} (\text{Etr} \widetilde{\mathbf{S}}_2^{-1})^2 + \text{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 - \text{Etr}(\widetilde{\mathbf{S}}_2^{-1})^2 \\
 & - 2(\text{tr} \widetilde{\mathbf{S}}_2^{-1} - \text{Etr} \widetilde{\mathbf{S}}_2^{-1}) + n_1^{-1} (\text{Etr} \widetilde{\mathbf{S}}_2^{-1})^2 + \text{Etr}(\widetilde{\mathbf{S}}_2^{-1})^2 \\
 & - 2\text{Etr} \widetilde{\mathbf{S}}_2^{-1} + n_1^{-1} \text{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 + \beta_1 n_1^{-1} \text{tr}(\widetilde{\mathbf{S}}_2^{-1} \circ \widetilde{\mathbf{S}}_2^{-1}) + p] + o_p(1).
 \end{aligned}$$

根据Zheng et al. (2015)^[53]中定理2.1的证明可知, $\text{tr} \widetilde{\mathbf{S}}_2^{-1} - \text{Etr} \widetilde{\mathbf{S}}_2^{-1}$ 的极限分布是正态的并且

$$\text{Etr}(\widetilde{\mathbf{S}}_2^{-1})^i = pc_i + \xi_i + o(1), \quad i = 1, 2, \quad (2.17)$$

其中

$$c_i = \int f_i(x) dF^{y_{n_2}, H_p}(x), \quad (2.18)$$

$$\begin{aligned}
 \xi_i = & -\frac{1}{2\pi i} \oint_{\mathcal{C}} f_i(z) \frac{y_2 \int \underline{m}_{y_2}^3(z) t^2 (1 + t \underline{m}_{y_2}(z))^{-3} dH(t)}{[1 - y_2 \int \underline{m}_{y_2}^2(z) t^2 (1 + t \underline{m}_{y_2}(z))^{-2} dH(t)]^2} dz \\
 & - \frac{\beta_2}{2\pi i} \oint_{\mathcal{C}} f_i(z) \frac{y_2 \int \underline{m}_{y_2}^3(z) t^2 (1 + t \underline{m}_{y_2}(z))^{-3} dH(t)}{1 - y_2 \int \underline{m}_{y_2}^2(z) t^2 (1 + t \underline{m}_{y_2}(z))^{-2} dH(t)} dz, \quad (2.19)
 \end{aligned}$$

这里 $f_1(x) = 1/x$, $f_2(x) = 1/x^2$, $F^{y_{n_2}, H_p}$ 是以 (y_{n_2}, H_p) 为指标的 Marčenko-Pastur 分布, $\underline{m}_{y_2}(z)$ 是伴随极限谱分布 $F^{y_2, H}$ 的 Stieltjes 变换 (见 §2.5.3 节), H 是 H_p 的 LSD, \mathcal{C} 是复平面中包含 $F^{y_2, H}$ 的支撑集的闭围道。由 (2.17) 可得

$$n_1^{-1} \text{tr}^2 \widetilde{\mathbf{S}}_2^{-1} - n_1^{-1} (\text{Etr} \widetilde{\mathbf{S}}_2^{-1})^2 = 2y_{n_1} c_1 (\text{tr} \widetilde{\mathbf{S}}_2^{-1} - \text{Etr} \widetilde{\mathbf{S}}_2^{-1}) + o_p(1).$$

由此可得

$$\begin{aligned}
 \text{E}(T^* | \mathbf{S}_2) = & \omega_1 [\text{tr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1)^2 - \text{Etr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1^2) - \text{tr}(\mathbf{\Sigma}_2 \mathbf{\Sigma}_1)) \\
 & - 2n_2^{-1} \text{tr} \mathbf{\Sigma}_2 (\text{tr}(\widetilde{\mathbf{S}}_2 \mathbf{\Sigma}_1) - \text{tr} \mathbf{\Sigma}_2) + \text{tr}(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2 + n_1^{-1} \text{tr} \mathbf{\Sigma}_1^2 \\
 & + n_2^{-1} \text{tr} \mathbf{\Sigma}_2^2 + \beta_1 n_1^{-1} \text{tr}(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_1) + \beta_2 n_2^{-1} \text{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2)] \\
 + \omega_2 & [2(y_{n_1} c_1 - 1) (\text{tr} \widetilde{\mathbf{S}}_2^{-1} - \text{Etr} \widetilde{\mathbf{S}}_2^{-1}) + \text{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 - \text{Etr}(\widetilde{\mathbf{S}}_2^{-1})^2]
 \end{aligned}$$

$$\begin{aligned}
 &+p(y_{n_1}c_1^2 - 2c_1 + c_2) + 2y_{n_1}c_1\xi_1 - 2\xi_1 + \xi_2 \\
 &+n_1^{-1}\text{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 + \beta_1 n_1^{-1}\text{tr}(\widetilde{\mathbf{S}}_2^{-1} \circ \widetilde{\mathbf{S}}_2^{-1}) + p] + o_p(1).
 \end{aligned}$$

令 $\widetilde{\mathbf{S}}_1 = N_1^{-1} \sum_{i=1}^{N_1} (\mathbf{w}_{1i} - \bar{\mathbf{w}}_1)(\mathbf{w}_{1i} - \bar{\mathbf{w}}_1)^T$, 则有

$$\begin{aligned}
 T^* - E(T^*|\mathbf{S}_2) &= \omega_1[\text{tr}(\widetilde{\mathbf{S}}_1 \boldsymbol{\Sigma}_1)^2 - E\text{tr}(\widetilde{\mathbf{S}}_1 \boldsymbol{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{S}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{S}}_2 \boldsymbol{\Sigma}_1) \\
 &\quad - \text{tr}(\widetilde{\mathbf{S}}_2 \boldsymbol{\Sigma}_1^2)) - 2n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1(\text{tr}(\widetilde{\mathbf{S}}_1 \boldsymbol{\Sigma}_1) - \text{tr}\boldsymbol{\Sigma}_1)] \\
 &\quad + \omega_2[\text{tr}(\widetilde{\mathbf{S}}_1 \widetilde{\mathbf{S}}_2^{-1})^2 - E(\text{tr}(\widetilde{\mathbf{S}}_1 \widetilde{\mathbf{S}}_2^{-1})^2|\widetilde{\mathbf{S}}_2) \\
 &\quad - 2(\text{tr}(\widetilde{\mathbf{S}}_1 \widetilde{\mathbf{S}}_2^{-1}) - E(\text{tr}(\widetilde{\mathbf{S}}_1 \widetilde{\mathbf{S}}_2^{-1})|\widetilde{\mathbf{S}}_2))].
 \end{aligned}$$

令 $\widetilde{\mathbf{B}}_1 = N_1^{-1} \sum_{i=1}^{N_1} \mathbf{w}_{1i} \mathbf{w}_{1i}^T$ 和 $\widetilde{\mathbf{B}}_2 = N_2^{-1} \sum_{i=1}^{N_2} \boldsymbol{\Gamma} \mathbf{w}_{2i} \mathbf{w}_{2i}^T \boldsymbol{\Gamma}^T$ 。仍然根据Zheng et al. (2015)^[53]中定理2.1的证明, 可得

$$\begin{aligned}
 E(T^*|\mathbf{S}_2) &= \omega_1[\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)^2 - E\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2) - \text{tr}(\boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1)) \\
 &\quad - 2n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2(\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) - \text{tr}\boldsymbol{\Sigma}_2) + \text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2 + n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1^2 \\
 &\quad + n_2^{-1}\text{tr}\boldsymbol{\Sigma}_2^2 + \beta_1 n_1^{-1}\text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1) + \beta_2 n_2^{-1}\text{tr}(\boldsymbol{\Sigma}_2 \circ \boldsymbol{\Sigma}_2)] \\
 &\quad + \omega_2[2(y_{n_1}c_1 - 1)(\text{tr}\widetilde{\mathbf{B}}_2^{-1} - E\text{tr}\widetilde{\mathbf{B}}_2^{-1}) + \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 - E\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 \\
 &\quad + p(y_{n_1}c_1^2 - 2c_1 + c_2) + 2y_{n_1}c_1\xi_1 - 2\xi_1 + \xi_2 \\
 &\quad + n_1^{-1}\text{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 + \beta_1 n_1^{-1}\text{tr}(\widetilde{\mathbf{S}}_2^{-1} \circ \widetilde{\mathbf{S}}_2^{-1}) + p] + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 T^* - E(T^*|\mathbf{S}_2) &= \omega_1[\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - E\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{S}}_2 \boldsymbol{\Sigma}_1) \\
 &\quad - \text{tr}(\widetilde{\mathbf{S}}_2 \boldsymbol{\Sigma}_1^2)) - 2n_1^{-1}\text{tr}\boldsymbol{\Sigma}_1(\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1) - \text{tr}\boldsymbol{\Sigma}_1)] \\
 &\quad + \omega_2[\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1})^2 - E(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1})^2|\widetilde{\mathbf{S}}_2) \\
 &\quad - 2(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1}) - E(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1})|\widetilde{\mathbf{S}}_2))] + o_p(1).
 \end{aligned}$$

根据§2.5.3节中假设a-b-c-d-f下面的讨论可知, $\|\widetilde{\mathbf{B}}_2\|$ 和 $\|\widetilde{\mathbf{B}}_2^{-1}\|$ 是几乎处处有界的。由于

$$\widetilde{\mathbf{S}}_2 = \frac{N_2}{n_2} (\widetilde{\mathbf{B}}_2 - N_2 \boldsymbol{\Gamma} \bar{\boldsymbol{\gamma}}_2 \bar{\boldsymbol{\gamma}}_2^T \boldsymbol{\Gamma}^T) \quad (2.20)$$

以及

$$\widetilde{\mathbf{S}}_2^{-1} = \frac{n_2}{N_2} \left(\widetilde{\mathbf{B}}_2^{-1} + \frac{1}{1 - N_2 \bar{\boldsymbol{\gamma}}_2^T \boldsymbol{\Gamma}^T \widetilde{\mathbf{B}}_2^{-1} \boldsymbol{\Gamma} \bar{\boldsymbol{\gamma}}_2} N_2 \widetilde{\mathbf{B}}_2^{-1} \boldsymbol{\Gamma} \bar{\boldsymbol{\gamma}}_2 \bar{\boldsymbol{\gamma}}_2^T \boldsymbol{\Gamma}^T \widetilde{\mathbf{B}}_2^{-1} \right), \quad (2.21)$$

再根据(2.86), (2.88), Zheng et al. (2015)^[53]中的引理5.2和引理5.3, 可得

$$\begin{aligned} \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{S}}_2 \boldsymbol{\Sigma}_1) - \text{tr}(\widetilde{\mathbf{S}}_2 \boldsymbol{\Sigma}_1^2) &= \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) - \text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2) + o_p(1), \\ \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1}) - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1}) | \widetilde{\mathbf{S}}_2) &= \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) | \widetilde{\mathbf{B}}_2) + o_p(1), \\ \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1})^2 - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}}_2^{-1})^2 | \widetilde{\mathbf{S}}_2) &= \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 | \widetilde{\mathbf{B}}_2) + o_p(1). \end{aligned}$$

由此可得

$$\begin{aligned} T^* - \text{E}(T^* | \mathbf{S}_2) &= \omega_1 [\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - \text{Etr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) \\ &\quad - \text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2)) - 2n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1 (\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1) - \text{tr} \boldsymbol{\Sigma}_1)] \\ &\quad + \omega_2 [\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) | \widetilde{\mathbf{B}}_2) \\ &\quad - 2(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 | \widetilde{\mathbf{B}}_2))] + o_p(1). \end{aligned}$$

令

$$\begin{aligned} \mu_1 &= \text{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2 + n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1^2 + n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2^2 \\ &\quad + \beta_1 n_1^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1) + \beta_2 n_2^{-1} \text{tr}(\boldsymbol{\Sigma}_2 \circ \boldsymbol{\Sigma}_2), \end{aligned} \quad (2.22)$$

$$\begin{aligned} \mu_2 &= p(y_{n_1} c_1^2 - 2c_1 + c_2) + p + 2y_{n_1} c_1 \xi_1 - 2\xi_1 + \xi_2 \\ &\quad + n_1^{-1} \text{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 + \beta_1 n_1^{-1} \text{tr}(\widetilde{\mathbf{S}}_2^{-1} \circ \widetilde{\mathbf{S}}_2^{-1}), \end{aligned} \quad (2.23)$$

$$\begin{aligned} T_A &= \omega_1 [\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)^2 - \text{Etr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2) - \text{tr}(\boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1)) \\ &\quad - 2n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2 (\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) - \text{tr} \boldsymbol{\Sigma}_2)] \\ &\quad + \omega_2 [2(y_{n_1} c_1 - 1)(\text{tr} \widetilde{\mathbf{B}}_2^{-1} - \text{Etr} \widetilde{\mathbf{B}}_2^{-1}) + \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 - \text{Etr}(\widetilde{\mathbf{B}}_2^{-1})^2], \end{aligned} \quad (2.24)$$

$$\begin{aligned} T_B &= \omega_1 [\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - \text{Etr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - 2(\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) - \text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2)) \\ &\quad - 2n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1 (\text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1) - \text{tr} \boldsymbol{\Sigma}_1)] \\ &\quad + \omega_2 [\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 | \widetilde{\mathbf{B}}_2) \\ &\quad - 2(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) - \text{E}(\text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) | \widetilde{\mathbf{B}}_2))]. \end{aligned} \quad (2.25)$$

综合上面的讨论可得

$$T = T_A + T_B + \omega_1 \mu_1 + \omega_2 \mu_2 + o_p(1).$$

根据Slutsky定理, 只需要考虑 $T_A + T_B$ 的渐近分布。

步骤2: 因为

$$T_A + T_B = \sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) T_A + \sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) T_B,$$

其中

$$\begin{aligned} & \sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) T_A \\ = & \omega_1 \left[\sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)^2 - 2 \sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2) \right. \\ & \left. - 2n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2 \sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) \right] \\ & + \omega_2 \left[2(y_{n_1} c_1 - 1) \sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr} \widetilde{\mathbf{B}}_2^{-1} + \sum_{j=1}^{N_2} (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 \right], \end{aligned}$$

和

$$\begin{aligned} & \sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) T_B \\ = & \omega_1 \left[\sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2 - 2 \sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) \right. \\ & \left. - 2n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1 \sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1) \right] \\ & + \omega_2 \left[\sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2 - 2 \sum_{i=1}^{N_1} (\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}) \right], \end{aligned}$$

因此, 根据鞅差序列的中心极限定理, 只需要计算

$$\sum_{j=1}^{N_2} \mathbf{E}_{j-1} [((\mathbf{E}_j - \mathbf{E}_{j-1}) T_A)^2] + \sum_{i=1}^{N_1} \mathbf{E}_{2,i-1} [((\mathbf{E}_{2,i} - \mathbf{E}_{2,i-1}) T_B)^2].$$

根据引理2.5.1中的(2.81)-(2.84), 引理2.5.4中的(2.97)-(2.100)和Zheng et al.(2015)^[53]

中的(2.5)式, 可得

$$\begin{aligned}
 \sigma_A &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1})T_A)^2] \\
 &= \omega_1^2 [4(n_2^{-1} \text{tr} \Sigma_2)^2 \eta_{33} + 4\eta_{44} + \eta_{55} + 8(n_2^{-1} \text{tr} \Sigma_2) \eta_{34} - 4(n_2^{-1} \text{tr} \Sigma_2) \eta_{35} - 4\eta_{45}] \\
 &\quad + 2\omega_1 \omega_2 [-4(y_{n_1} c_1 - 1)(n_2^{-1} \text{tr} \Sigma_2) \eta_{13} - 4(y_{n_1} c_1 - 1) \eta_{14} \\
 &\quad \quad + 2(y_{n_1} c_1 - 1) \eta_{15} - 2(n_2^{-1} \text{tr} \Sigma_2) \eta_{23} - 2\eta_{24} + \eta_{25}] \\
 &\quad + \omega_2^2 [4(y_{n_1} c_1 - 1)^2 \eta_{11} + \eta_{22} + 4(y_{n_1} c_1 - 1) \eta_{12}],
 \end{aligned}$$

当 $i, j \in \{1, 2\}$ 并且 $p \rightarrow \infty$ 时,

$$\begin{aligned}
 \eta_{ij} &\rightarrow -\frac{1}{2\pi^2} \oint_{C_1} \oint_{C_2} \frac{f_i(z_1) f_j(z_2)}{(\underline{m}_{y_2}(z_1) - \underline{m}_{y_2}(z_2))^2} d\underline{m}_{y_2}(z_1) d\underline{m}_{y_2}(z_2) \\
 &\quad - \frac{y_2 \beta_2}{4\pi^2} \oint_{C_1} \oint_{C_2} f_i(z_1) f_j(z_2) \left[\int \frac{t^2 dH(t)}{(\underline{m}_{y_2}(z_1)t + 1)^2 (\underline{m}_{y_2}(z_2)t + 1)^2} \right] d\underline{m}_{y_2}(z_1) d\underline{m}_{y_2}(z_2),
 \end{aligned} \tag{2.26}$$

其中 $f_1(x) = 1/x$, $f_2(x) = 1/x^2$, C_1 和 C_2 是复平面中包含 $F^{y_2, H}$ 的支撑集的闭围道, 并且 C_1 和 C_2 是不相交的,

$$\begin{aligned}
 \eta_{33} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \Sigma_1))^2] = \frac{1}{N_2} [2\text{tr} \Sigma_2^2 + \beta_2 \text{tr}(\Sigma_2 \circ \Sigma_2)], \\
 \eta_{44} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \Sigma_1^2))^2] \\
 &= \frac{1}{N_2} [2\text{tr}(\Sigma_1 \Sigma_2)^2 + \beta_2 \text{tr}(\Sigma_2^{1/2} \Sigma_1 \Sigma_2^{1/2} \circ \Sigma_2^{1/2} \Sigma_1 \Sigma_2^{1/2})], \\
 \eta_{55} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \Sigma_1^2))^2] \\
 &= \frac{4}{N_2^3} \text{tr}^2 \Sigma_2 [2\text{tr} \Sigma_2^2 + \beta_2 \text{tr}(\Sigma_2 \circ \Sigma_2)] + \frac{8}{N_2^2} \text{tr} \Sigma_2 [2\text{tr} \Sigma_2^3 + \beta_2 \text{tr}(\Sigma_2 \circ \Sigma_2^2)] \\
 &\quad + \frac{4}{N_2} [2\text{tr} \Sigma_2^4 + \beta_2 \text{tr}(\Sigma_2^2 \circ \Sigma_2^2)] + \frac{4}{N_2^2} \text{tr}^2 \Sigma_2^2 + o_p(1), \\
 \eta_{13} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr} \widetilde{\mathbf{B}}_2^{-1}) ((E_j - E_{j-1}) \text{tr}(\widetilde{\mathbf{B}}_2 \Sigma_1))] \\
 &= -\frac{2}{(1 - y_{N_2})} N_2^{-1} \text{tr} \Sigma_1 - \frac{\beta_2}{(1 - y_{N_2})} N_2^{-1} \text{tr}(\Sigma_1^{1/2} \Sigma_2^{-1} \Sigma_1^{1/2} \circ \Sigma_2) + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 \eta_{14} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr} \tilde{\mathbf{B}}_2^{-1}) ((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2))] \\
 &= -\frac{2}{(1-y_{N_2})} N_2^{-1} \text{tr} \boldsymbol{\Sigma}_1^2 - \frac{\beta_2}{(1-y_{N_2})} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{1/2}) + o_p(1), \\
 \eta_{15} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr} \tilde{\mathbf{B}}_2^{-1}) ((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2))] \\
 &= -\frac{4}{(1-y_{N_2})} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2) - \frac{2\beta_2}{(1-y_{N_2})} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^2) \\
 &\quad - \frac{2\beta_2}{(1-y_{N_2})} N_2^{-1} \text{tr} \boldsymbol{\Sigma}_2 N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2) + o_p(1), \\
 \eta_{23} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2^{-1})^2) ((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1))] \\
 &= -\frac{4}{(1-y_{N_2})^2} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^2 \boldsymbol{\Sigma}_2^{-1}) - \frac{4}{(1-y_{N_2})^3} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr} \boldsymbol{\Sigma}_1 \\
 &\quad - \frac{2\beta_2}{(1-y_{N_2})^2} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2) \\
 &\quad - \frac{2\beta_2}{(1-y_{N_2})^3} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2) + o_p(1), \\
 \eta_{24} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2^{-1})^2) ((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2))] \\
 &= -\frac{4}{(1-y_{N_2})^2} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^3 \boldsymbol{\Sigma}_2^{-1}) - \frac{4}{(1-y_{N_2})^3} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr} \boldsymbol{\Sigma}_1^2 \\
 &\quad - \frac{2\beta_2}{(1-y_{N_2})^2} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{1/2}) \\
 &\quad - \frac{2\beta_2}{(1-y_{N_2})^3} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{1/2}) + o_p(1), \\
 \eta_{25} &= \sum_{j=1}^{N_2} E_{j-1} [((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2^{-1})^2) ((E_j - E_{j-1}) \text{tr} (\tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2))] \\
 &= -\frac{8}{(1-y_{N_2})^2} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2) - \frac{8}{(1-y_{N_2})^3} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2) \\
 &\quad + \frac{4}{(1-y_{N_2})^2} (N_2^{-1} \text{tr} \boldsymbol{\Sigma}_1)^2 - \frac{4\beta_2}{(1-y_{N_2})^2} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^2) \\
 &\quad - \frac{4\beta_2}{(1-y_{N_2})^3} N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^2) \\
 &\quad - \frac{4\beta_2}{(1-y_{N_2})^2} N_2^{-1} \text{tr} \boldsymbol{\Sigma}_2 N_2^{-1} \text{tr} (\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2)
 \end{aligned}$$

$$\begin{aligned}
 & -\frac{4\beta_2}{(1-y_{N_2})^3}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\text{tr}\boldsymbol{\Sigma}_2N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2)+o_p(1), \\
 \eta_{34} &= \sum_{j=1}^{N_2}E_{j-1}[\{(E_j-E_{j-1})\text{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)\}\{(E_j-E_{j-1})\text{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2)\}] \\
 &= \frac{1}{N_2}[2\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^2)+\beta_2\text{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2})]+o_p(1), \\
 \eta_{35} &= \sum_{j=1}^{N_2}E_{j-1}[\{(E_j-E_{j-1})\text{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)\}\{(E_j-E_{j-1})\text{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2\}] \\
 &= \frac{2}{N_2^2}\text{tr}\boldsymbol{\Sigma}_2[2\text{tr}\boldsymbol{\Sigma}_2^2+\beta_2\text{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2)]+\frac{2}{N_2}[2\text{tr}\boldsymbol{\Sigma}_2^3+\beta_2\text{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2^2)]+o_p(1), \\
 \eta_{45} &= \sum_{j=1}^{N_2}E_{j-1}[\{(E_j-E_{j-1})\text{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2)\}\{(E_j-E_{j-1})\text{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2\}] \\
 &= \frac{2}{N_2^2}\text{tr}\boldsymbol{\Sigma}_2[2\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^2)+\beta_2\text{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2})] \\
 &\quad +\frac{2}{N_2}[2\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^3)+\beta_2\text{tr}(\boldsymbol{\Sigma}_2^2\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2})]+o_p(1).
 \end{aligned}$$

由引理2.5.2中的(2.91)-(2.94)可得

$$\begin{aligned}
 p^{-1}\text{tr}\widetilde{\mathbf{B}}_2^{-1} &= d_{1p}+o_p(1), & p^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 &= d_{2p}+o_p(1), \\
 p^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^3 &= d_{3p}+o_p(1), & p^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^4 &= d_{4p}+o_p(1)
 \end{aligned}$$

和

$$\begin{aligned}
 p^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\widetilde{\mathbf{B}}_2^{-1}) &= l_{1p}+o_p(1), \\
 p^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ(\widetilde{\mathbf{B}}_2^{-1})^2) &= l_{2p}+o_p(1), \\
 p^{-1}\text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ(\widetilde{\mathbf{B}}_2^{-1})^2) &= l_{3p}+o_p(1),
 \end{aligned}$$

其中

$$d_{1p} = \frac{1}{(1-y_{N_2})}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}), \quad (2.27)$$

$$d_{2p} = \frac{1}{(1-y_{N_2})^2}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})^2 + \frac{y_{N_2}}{(1-y_{N_2})^3}(p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}))^2, \quad (2.28)$$

$$d_{3p} = \frac{1}{(1-y_{N_2})^3}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})^3 + \frac{3y_{N_2}}{(1-y_{N_2})^4}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})^2 \quad (2.29)$$

$$\begin{aligned}
 & + \frac{2y_{N_2}^2}{(1-y_{N_2})^5} (p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}))^3, \\
 d_{4p} = & \frac{1}{(1-y_{N_2})^4} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1})^4 + \frac{4y_{N_2}}{(1-y_{N_2})^5} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1})^3 \quad (2.30) \\
 & + \frac{2y_{N_2}}{(1-y_{N_2})^5} (p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}))^2 + \frac{10y_{N_2}^2}{(1-y_{N_2})^6} (p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}))^2 p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1})^2 \\
 & + \frac{5y_{N_2}^3}{(1-y_{N_2})^7} (p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}))^4
 \end{aligned}$$

$$l_{1p} = \frac{1}{(1-y_{N_2})^2} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2}), \quad (2.31)$$

$$\begin{aligned}
 l_{2p} = & \frac{1}{(1-y_{N_2})^3} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2}) \quad (2.32) \\
 & + \frac{y_{N_2}}{(1-y_{N_2})^4} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2}),
 \end{aligned}$$

$$\begin{aligned}
 l_{3p} = & \frac{1}{(1-y_{N_2})^4} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2}) \quad (2.33) \\
 & + \frac{2y_{N_2}}{(1-y_{N_2})^5} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2}) \\
 & + \frac{y_{N_2}^2}{(1-y_{N_2})^6} (p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}))^2 p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2}).
 \end{aligned}$$

再根据引理2.5.1中的(2.81)-(2.85), 可得

$$\begin{aligned}
 \sigma_B & = \sum_{i=1}^{N_1} \mathbb{E}_{2,i-1} [((\mathbb{E}_{2,i} - \mathbb{E}_{2,i-1}) T_B)^2] \\
 & = \omega_1^2 [4(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1)^2 \theta_{11} + 4\theta_{33} + \theta_{55} + 8(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{13} - 4(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{15} - 4\theta_{35}] \\
 & \quad + 2\omega_1 \omega_2 [4(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{12} - 2(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{14} + 4\theta_{23} - 2\theta_{25} - 2\theta_{34} + \theta_{45}] \\
 & \quad + \omega_2^2 [4\theta_{22} + \theta_{44} - 4\theta_{24}],
 \end{aligned}$$

其中

$$\begin{aligned}
 \theta_{11} & = \sum_{i=1}^{N_1} \mathbb{E}_{2,i-1} [((\mathbb{E}_{2,i} - \mathbb{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1))^2] = \frac{1}{N_1} [2\text{tr} \boldsymbol{\Sigma}_1^2 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1)], \\
 \theta_{22} & = \sum_{i=1}^{N_1} \mathbb{E}_{2,i-1} [((\mathbb{E}_{2,i} - \mathbb{E}_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}))^2] \\
 & = \frac{1}{N_1} [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 + \beta_1 \text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \widetilde{\mathbf{B}}_2^{-1})] = y_{N_1} (2d_{2p} + \beta_1 l_{1p}) + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 \theta_{33} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1))^2] \\
 &= \frac{1}{N_1} [2\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^2)^2 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] \\
 &= \frac{2}{N_1 N_2} \text{tr}^2(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2) + \frac{2}{N_1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2)^2 + \frac{\beta_1}{N_1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{1/2}) + o_p(1), \\
 \theta_{44} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}))^2] \\
 &= \frac{4}{N_1^3} \text{tr}^2(\widetilde{\mathbf{B}}_2^{-1}) [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 + \beta_1 \text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \widetilde{\mathbf{B}}_2^{-1})] + \frac{4}{N_1^2} \text{tr}^2(\widetilde{\mathbf{B}}_2^{-1})^2 \\
 &\quad + \frac{8}{N_1^2} \text{tr}(\widetilde{\mathbf{B}}_2^{-1}) [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^3 + \beta_1 \text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ (\widetilde{\mathbf{B}}_2^{-1})^2)] \\
 &\quad + \frac{4}{N_1} [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^4 + \beta_1 \text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ (\widetilde{\mathbf{B}}_2^{-1})^2)] \\
 &= 4y_{N_1}^3 d_{1p}^2 (2d_{2p} + \beta_1 l_{1p}) + 4y_{N_1}^2 d_{2p}^2 + 8y_{N_1}^2 d_{1p} (2d_{3p} + \beta_1 l_{2p}) \\
 &\quad + 4y_{N_1} (2d_{4p} + \beta_1 l_{3p}) + o_p(1), \\
 \theta_{55} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2)^2] \\
 &= \frac{4}{N_1^3} \text{tr}^2 \boldsymbol{\Sigma}_1 [2\text{tr} \boldsymbol{\Sigma}_1^2 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1)] + \frac{8}{N_1^2} \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr} \boldsymbol{\Sigma}_1^3 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^2)] \\
 &\quad + \frac{4}{N_1} [2\text{tr} \boldsymbol{\Sigma}_1^4 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1^2 \circ \boldsymbol{\Sigma}_1^2)] + \frac{4}{N_1^2} \text{tr}^2 \boldsymbol{\Sigma}_1^2 + o_p(1), \\
 \theta_{12} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)) ((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1}))] \\
 &= \frac{1}{N_1} [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1)], \\
 \theta_{13} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)) ((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1))] \\
 &= \frac{1}{N_1} [2\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^3) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] \\
 &= \frac{1}{N_1} [2\text{tr}(\boldsymbol{\Sigma}_1^2 \boldsymbol{\Sigma}_2) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{1/2})] + o_p(1), \\
 \theta_{14} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)) ((E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2)]
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{2}{N_1^2} \text{tr}(\tilde{\mathbf{B}}_2^{-1}) [2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1)] \\
 &\quad + \frac{2}{N_1} [2\text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1)] + o_p(1), \\
 \theta_{15} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)) ((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2)] \\
 &= \frac{2}{N_1^2} \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr} \boldsymbol{\Sigma}_1^2 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_2)] + \frac{2}{N_1} [2\text{tr} \boldsymbol{\Sigma}_1^3 + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^2)] + o_p(1), \\
 \theta_{23} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \tilde{\mathbf{B}}_2^{-1})) ((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1))] \\
 &= \frac{1}{N_1} [2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)], \\
 \theta_{24} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \tilde{\mathbf{B}}_2^{-1})) ((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \tilde{\mathbf{B}}_2^{-1})^2)] \\
 &= \frac{2}{N_1^2} \text{tr}(\tilde{\mathbf{B}}_2^{-1}) [2\text{tr}(\tilde{\mathbf{B}}_2^{-1})^2 + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \tilde{\mathbf{B}}_2^{-1})] \\
 &\quad + \frac{2}{N_1} [2\text{tr}(\tilde{\mathbf{B}}_2^{-1})^3 + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ (\tilde{\mathbf{B}}_2^{-1})^2)] \\
 &= 2y_{N_1}^2 d_{1p} (2d_{2p} + \beta_1 l_{1p}) + 2y_{N_1} (2d_{3p} + \beta_1 l_{2p}) + o_p(1), \\
 \theta_{25} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \tilde{\mathbf{B}}_2^{-1})) ((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2)] \\
 &= \frac{2}{N_1^2} \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1)] \\
 &\quad + \frac{2}{N_1} [2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1^2) + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1^2)] + o_p(1), \\
 \theta_{34} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)) ((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \tilde{\mathbf{B}}_2^{-1})^2)] \\
 &= \frac{2}{N_1^2} \text{tr}(\tilde{\mathbf{B}}_2^{-1}) [2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] \\
 &\quad + \frac{2}{N_1} [2\text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] + o_p(1), \\
 \theta_{35} &= \sum_{i=1}^{N_1} E_{2,i-1} [((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)) ((E_{2,i} - E_{2,i-1}) \text{tr}(\tilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1)^2)] \\
 &= \frac{2}{N_1^2} \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr}(\tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^3) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)]
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{2}{N_1} [2\text{tr}(\widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1^4) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1^2 \circ \boldsymbol{\Sigma}_1 \widetilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] + o_p(1) \\
 = & \frac{2}{N_1^2} \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr}(\boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^2) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{1/2})] \\
 & + \frac{2}{N_1} [2\text{tr}(\boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^3) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1^2 \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{1/2})] + o_p(1), \\
 \theta_{45} = & \sum_{i=1}^{N_1} E_{2,i-1} [(E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{B}}_2^{-1})^2] (E_{2,i} - E_{2,i-1}) \text{tr}(\widetilde{\mathbf{B}}_1 \boldsymbol{\Sigma}_1^2) \\
 = & \frac{4}{N_1^3} \text{tr}(\widetilde{\mathbf{B}}_2^{-1}) \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1)] \\
 & + \frac{4}{N_1^2} \text{tr} \boldsymbol{\Sigma}_1 [2\text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1)] \\
 & + \frac{4}{N_1^2} \text{tr}(\widetilde{\mathbf{B}}_2^{-1}) [2\text{tr}(\widetilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1^2) + \beta_1 \text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1^2)] + \frac{4}{N_1^2} \text{tr}^2(\widetilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) \\
 & + \frac{4}{N_1} [2\text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1^2) + \beta_1 \text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1^2)] + o_p(1).
 \end{aligned}$$

经过计算和整理, 可得

$$\begin{aligned}
 \sigma_{11p} = & 4(n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2)^2 \eta_{33} + 4\eta_{44} + \eta_{55} + 8(n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2) \eta_{34} \tag{2.34} \\
 & - 4(n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2) \eta_{35} - 4\eta_{45} + 4(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1)^2 \theta_{11} + 4\theta_{33} \\
 & + \theta_{55} + 8(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{13} - 4(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{15} - 4\theta_{35} \\
 = & \frac{4}{N_1^2} \text{tr}^2 \boldsymbol{\Sigma}_1 + \frac{4}{N_2^2} \text{tr}^2 \boldsymbol{\Sigma}_2 + \frac{8}{N_1 N_2} \text{tr}^2(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2) \\
 & + \frac{4}{N_1} [2\text{tr}(\boldsymbol{\Sigma}_1 (\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1) \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1^{1/2} (\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1) \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{1/2})] \\
 & + \frac{4}{N_1} [2\text{tr}(\boldsymbol{\Sigma}_1^3 (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)) + \beta_1 \text{tr}(\boldsymbol{\Sigma}_1^2 \circ \boldsymbol{\Sigma}_1^{1/2} (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_1^{1/2})] \\
 & + \frac{4}{N_2} [2\text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2 (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_2) + \beta_2 \text{tr}(\boldsymbol{\Sigma}_2^{1/2} (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_2^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{1/2})] \\
 & + \frac{4}{N_2} [2\text{tr}((\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1) \boldsymbol{\Sigma}_2^3) + \beta_2 \text{tr}(\boldsymbol{\Sigma}_2^2 \circ \boldsymbol{\Sigma}_2^{1/2} (\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1) \boldsymbol{\Sigma}_2^{1/2})] + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 \sigma_{12p}^1 = & -4(y_{n_1} c_1 - 1)(n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2) \eta_{13} - 4(y_{n_1} c_1 - 1) \eta_{14} \tag{2.35} \\
 & + 2(y_{n_1} c_1 - 1) \eta_{15} - 2(n_2^{-1} \text{tr} \boldsymbol{\Sigma}_2) \eta_{23} - 2\eta_{24} + \eta_{25} \\
 = & 2(y_{n_1} c_1 - 1) \left[\frac{4y_{N_2}^2}{(1 - y_{N_2})} p^{-1} \text{tr} \boldsymbol{\Sigma}_1 p^{-1} \text{tr} \boldsymbol{\Sigma}_2 + \frac{4y_{N_2}}{(1 - y_{N_2})} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)) \right]
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{2\beta_2 y_{N_2}}{(1-y_{N_2})} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_2^{1/2}) \\
 & + \frac{8y_{N_2}^2}{(1-y_{N_2})^2} p^{-1} \text{tr} \boldsymbol{\Sigma}_2 p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^2 \boldsymbol{\Sigma}_2^{-1}) \\
 & + \frac{8y_{N_2}^3}{(1-y_{N_2})^3} p^{-1} \text{tr} \boldsymbol{\Sigma}_1 p^{-1} \text{tr} \boldsymbol{\Sigma}_2 p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) \\
 & + \frac{4y_{N_2}^2}{(1-y_{N_2})^2} (p^{-1} \text{tr} \boldsymbol{\Sigma}_1)^2 + \frac{8y_{N_2}}{(1-y_{N_2})^2} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)) \\
 & + \frac{8y_{N_2}^2}{(1-y_{N_2})^3} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)) \\
 & + \frac{4\beta_2 y_{N_2}}{(1-y_{N_2})^2} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_2^{1/2}) \\
 & + \frac{4\beta_2 y_{N_2}^2}{(1-y_{N_2})^3} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_2^{1/2} (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_2^{1/2}) + o_p(1),
 \end{aligned}$$

和

$$\begin{aligned}
 \sigma_{12p}^2 & = 4(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{12} - 2(n_1^{-1} \text{tr} \boldsymbol{\Sigma}_1) \theta_{14} + 4\theta_{23} - 2\theta_{25} - 2\theta_{34} + \theta_{45} \\
 & = 4(y_{N_1} d_{1p} - 1) N_1^{-1} [2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1^2) - 2\text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) \\
 & \quad + \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1^2) - \beta_1 \text{tr}(\tilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] \\
 & \quad + 4N_1^{-1} [2\text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1^2) - 2\text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) + \beta_1 \text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1^2) \\
 & \quad - \beta_1 \text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1)] + 4N_1^{-2} \text{tr}^2(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) + o_p(1).
 \end{aligned}$$

根据引理2.5.2中的(2.91)和(2.92)以及引理2.5.3中的(2.96)和(2.97), 可得

$$\begin{aligned}
 N_2^{-1} \text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1) & = \frac{1}{(1-y_{N_2})} N_2^{-1} \text{tr}(\boldsymbol{\Sigma}_1^2 \boldsymbol{\Sigma}_2^{-1}) + o_p(1), \\
 N_2^{-1} \text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1^2) & = \frac{1}{(1-y_{N_2})} N_2^{-1} \text{tr}(\boldsymbol{\Sigma}_1^3 \boldsymbol{\Sigma}_2^{-1}) + o_p(1), \\
 N_2^{-1} \text{tr}(\tilde{\mathbf{B}}_2^{-1} \boldsymbol{\Sigma}_1 \tilde{\mathbf{B}}_2 \boldsymbol{\Sigma}_1) & = \frac{1}{(1-y_{N_2})} N_2^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) \\
 & \quad - \frac{1}{(1-y_{N_2})} (N_2^{-1} \text{tr} \boldsymbol{\Sigma}_1)^2 + o_p(1), \\
 N_2^{-1} \text{tr}((\tilde{\mathbf{B}}_2^{-1})^2 \boldsymbol{\Sigma}_1^2) & = \frac{1}{(1-y_{N_2})^2} N_2^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^3 \boldsymbol{\Sigma}_2^{-1}) \\
 & \quad + \frac{1}{(1-y_{N_2})^3} N_2^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) N_2^{-1} \text{tr}(\boldsymbol{\Sigma}_1^3 \boldsymbol{\Sigma}_2^{-1}) + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 N_2^{-1}\text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) &= \frac{1}{(1-y_{N_2})^2}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}) \\
 &+ \frac{1}{(1-y_{N_2})^3}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}) \\
 &- \frac{2}{(1-y_{N_2})^2}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^2\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\text{tr}\boldsymbol{\Sigma}_1 \\
 &- \frac{1}{(1-y_{N_2})^3}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})(N_2^{-1}\text{tr}\boldsymbol{\Sigma}_1)^2 + o_p(1)
 \end{aligned}$$

和

$$\begin{aligned}
 N_2^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1^2) &= \frac{1}{(1-y_{N_2})}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^2) + o_p(1), \\
 N_2^{-1}\text{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) &= \frac{1}{(1-y_{N_2})}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2}) + o_p(1), \\
 N_2^{-1}\text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1^2) &= \frac{1}{(1-y_{N_2})^2}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^2) \\
 &+ \frac{1}{(1-y_{N_2})^3}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^2) + o_p(1), \\
 N_2^{-1}\text{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) &= \frac{1}{(1-y_{N_2})^2}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2}) \\
 &+ \frac{1}{(1-y_{N_2})^3}N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}) \\
 &\quad \times N_2^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2}) + o_p(1).
 \end{aligned}$$

由此可得

$$\begin{aligned}
 \sigma_{12p}^2 &= 4(y_{N_1}d_{1p} - 1)\left[\frac{2y_{N_1}y_{N_2}}{(1-y_{N_2})}(p^{-1}\text{tr}\boldsymbol{\Sigma}_1)^2 + \frac{2y_{N_1}}{(1-y_{N_2})}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})\right. \\
 &\quad \left. + \frac{\beta_1y_{N_1}}{(1-y_{N_2})}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_1^{1/2})\right] \quad (2.36) \\
 &+ 4\left[\frac{4y_{N_1}y_{N_2}}{(1-y_{N_2})^2}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1^2\boldsymbol{\Sigma}_2^{-1})p^{-1}\text{tr}\boldsymbol{\Sigma}_1\right. \\
 &+ \frac{2y_{N_1}y_{N_2}^2}{(1-y_{N_2})^3}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})(p^{-1}\text{tr}\boldsymbol{\Sigma}_1)^2 \\
 &+ \frac{2y_{N_1}}{(1-y_{N_2})^2}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)) \\
 &+ \frac{2y_{N_1}y_{N_2}}{(1-y_{N_2})^3}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})p^{-1}\text{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)) \\
 &+ \frac{\beta_1y_{N_1}}{(1-y_{N_2})^2}p^{-1}\text{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_1^{1/2})
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{\beta_1 y_{N_1} y_{N_2}}{(1 - y_{N_2})^3} p^{-1} \text{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1}) p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2} (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \boldsymbol{\Sigma}_1^{1/2})] \\
 & + \frac{4y_{N_1}^2}{(1 - y_{N_2})^2} (p^{-1} \text{tr}(\boldsymbol{\Sigma}_1^2 \boldsymbol{\Sigma}_2^{-1}))^2 + o_p(1).
 \end{aligned}$$

接下来, 经过整理可得

$$\sigma_{22p}^1 = 4(y_{n_1} c_1 - 1)^2 \eta_{11} + \eta_{22} + 4(y_{n_1} c_1 - 1) \eta_{12} + o_p(1) \quad (2.37)$$

和

$$\begin{aligned}
 \sigma_{22p}^2 & = 4\theta_{22} + \theta_{44} - 4\theta_{24} \quad (2.38) \\
 & = 4y_{N_1} (y_{N_1} d_{1p} - 1)^2 (2d_{2p} + \beta_1 l_{1p}) \\
 & \quad + 8y_{N_1} (y_{N_1} d_{1p} - 1) (2d_{3p} + \beta_1 l_{2p}) \\
 & \quad + 4y_{N_1} (2d_{4p} + \beta_1 l_{3p}) + 4y_{N_1}^2 d_{2p}^2 + o_p(1).
 \end{aligned}$$

最后, 根据(2.21)和(2.23), 可得

$$\begin{aligned}
 \mu_2 & = p(y_{n_1} c_1^2 - 2c_1 + c_2) + p + 2y_{n_1} c_1 \xi_1 - 2\xi_1 + \xi_2 \quad (2.39) \\
 & \quad + y_{n_1} d_{2p} + \beta_1 y_{n_1} l_{1p} + o_p(1).
 \end{aligned}$$

§2.5.2 定理和性质的证明

定理2.2.1的证明. 在这个证明过程中, 因为所有的量都是在原假设 H_0 下计算的, 所以在这些量的下标中添加一个0, 例如, 用 μ_{10} 代替 μ_1 . 在原假设 H_0 下, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, 由等式(2.22)可得

$$\mu_{10} = (y_{n_1} + y_{n_2}) p^{-1} \text{tr} \boldsymbol{\Sigma}^2 + (\beta_1 y_{n_1} + \beta_2 y_{n_2}) p^{-1} \text{tr}(\boldsymbol{\Sigma} \circ \boldsymbol{\Sigma}).$$

当 $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$ 时, $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}_1^{-1/2} \boldsymbol{\Sigma}_2^{1/2} = \mathbf{I}_p$ 和 $\mathbf{T}_p = \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T = \mathbf{I}_p$, 若将 $\widetilde{\mathbf{B}}_2$ 看作§2.5.3节中的 \mathbf{B}_n , 则 $\boldsymbol{\Gamma}$ 和 \mathbf{T}_p 满足假设c-d-f. \mathbf{T}_p 的经验谱分布 H_p 和极限谱分布 H 都为 δ_1 , 此时, $F^{y_{n_2}, H_p}$ 有如下的密度函数

$$\frac{1}{2\pi x y_{n_2}} \sqrt{(b_{y_{n_2}} - x)(x - a_{y_{n_2}})}, \quad a_{y_{n_2}} \leq x \leq b_{y_{n_2}}, \quad (2.40)$$

这是经典的Marčenko-Pastur分布, 支撑集为 $[a_{y_{n_2}}, b_{y_{n_2}}]$, 其中 $a_{y_{n_2}} = (1 - \sqrt{y_{n_2}})^2$, $b_{y_{n_2}} = (1 + \sqrt{y_{n_2}})^2$ 。由等式(2.18)可得

$$c_{10} = \frac{1}{(1 - y_{n_2})}, \quad c_{20} = \frac{1}{(1 - y_{n_2})^3}. \quad (2.41)$$

根据(2.19), (2.26)以及Zheng et al. (2015)^[53]中的性质A.1, 可得

$$\xi_{10} = \frac{y_{n_2}}{(1 - y_{n_2})^2} + \frac{\beta_2 y_{n_2}}{(1 - y_{n_2})}, \quad (2.42)$$

$$\xi_{20} = \frac{y_{n_2}^2 + 3y_{n_2}}{(1 - y_{n_2})^4} + \frac{\beta_2(-y_{n_2}^2 + 3y_{n_2})}{(1 - y_{n_2})^3}, \quad (2.43)$$

$$\eta_{110} = \frac{2y_2}{(1 - y_2)^4} + \frac{\beta_2 y_2}{(1 - y_2)^2}, \quad (2.44)$$

$$\eta_{120} = \frac{4y_2(1 + y_2)}{(1 - y_2)^6} + \frac{2\beta_2 y_2}{(1 - y_2)^4}, \quad (2.45)$$

$$\eta_{220} = \frac{4y_2(2y_2^2 + 5y_2 + 2)}{(1 - y_2)^8} + \frac{4\beta_2 y_2}{(1 - y_2)^6}. \quad (2.46)$$

由(2.27)-(2.33)可得

$$d_{10p} = \frac{1}{(1 - y_{N_2})} \rightarrow \frac{1}{(1 - y_2)} = d_{10}, \quad (2.47)$$

$$d_{20p} = \frac{1}{(1 - y_{N_2})^3} \rightarrow \frac{1}{(1 - y_2)^3} = d_{20}, \quad (2.48)$$

$$d_{30p} = \frac{1 + y_{N_2}}{(1 - y_{N_2})^5} \rightarrow \frac{1 + y_2}{(1 - y_2)^5} = d_{30}, \quad (2.49)$$

$$d_{40p} = \frac{y_{N_2}^2 + 3y_{N_2} + 1}{(1 - y_{N_2})^7} \rightarrow \frac{y_2^2 + 3y_2 + 1}{(1 - y_2)^7} = d_{40}, \quad (2.50)$$

$$l_{10p} = \frac{1}{(1 - y_{N_2})^2} \rightarrow \frac{1}{(1 - y_2)^2} = l_{10}, \quad (2.51)$$

$$l_{20p} = \frac{1}{(1 - y_{N_2})^4} \rightarrow \frac{1}{(1 - y_2)^4} = l_{20}, \quad (2.52)$$

$$l_{30p} = \frac{1}{(1 - y_{N_2})^6} \rightarrow \frac{1}{(1 - y_2)^6} = l_{30}. \quad (2.53)$$

因此, 由(2.39), (2.41), (2.42), (2.43), (2.48)和(2.51)可得

$$\begin{aligned} \mu_{20} = & p(y_{n_1} c_{10}^2 - 2c_{10} + c_{20}) + p + 2y_{n_1} c_{10} \xi_{10} - 2\xi_{10} + \xi_{20} \\ & + y_{n_1} d_{20p} + \beta_1 y_{n_1} l_{10p} + o_p(1) \end{aligned}$$

$$\begin{aligned}
 &= p \left[-\frac{2}{(1-y_{n_2})} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3} \right] + p \\
 &\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} \\
 &\quad + \frac{2\beta_2 y_{n_1} y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1-y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})^3} + o_p(1).
 \end{aligned}$$

令 $m_{10} = \int x dL(x)$ 和 $m_{20} = \int x^2 dL(x)$, 根据(2.34), (2.35), (2.36), (2.41)和(2.47), 可得

$$\begin{aligned}
 \sigma_{110p} &= 4[(y_{N_1} + y_{N_2})p^{-1} \text{tr} \Sigma^2]^2 + o_p(1) \rightarrow 4(y_1 + y_2)^2 m_{20}^2 = \sigma_{110}, \\
 \sigma_{120p}^1 &= \left[\frac{8(y_{N_1} + y_{N_2})y_{N_2}^2 + 4y_{N_2}^2}{(1-y_{N_2})^2} + \frac{8y_{N_2}^3}{(1-y_{N_2})^3} \right] (p^{-1} \text{tr} \Sigma)^2 + o_p(1), \\
 \sigma_{120p}^2 &= \left[\frac{8(y_{N_1} + y_{N_2})y_{N_1}y_{N_2} + 4y_{N_1}^2}{(1-y_{N_2})^2} + \frac{8y_{N_1}y_{N_2}}{(1-y_{N_2})^3} \right] (p^{-1} \text{tr} \Sigma)^2 + o_p(1),
 \end{aligned}$$

由此可得

$$\begin{aligned}
 \sigma_{120p} &= \sigma_{120p}^1 + \sigma_{120p}^2 \\
 &= \left[\frac{8y_{N_2}(y_{N_1} + y_{N_2})^2 + 4y_{N_1}^2 + 4y_{N_2}^2}{(1-y_{N_2})^2} + \frac{8y_{N_1}y_{N_2} + 8y_{N_2}^3}{(1-y_{N_2})^3} \right] (p^{-1} \text{tr} \Sigma)^2 + o_p(1) \\
 &\rightarrow \left[\frac{8y_2(y_1 + y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3} \right] m_{10}^2 = \sigma_{120}.
 \end{aligned}$$

根据(2.37), (2.38), (2.41), (2.44), (2.45), (2.46)和(2.47)-(2.53), 可得

$$\begin{aligned}
 \sigma_{220p} &= \sigma_{220p}^1 + \sigma_{220p}^2 \tag{2.54} \\
 &\rightarrow 4(y_1 d_{10} - 1)^2 (\eta_{110} + 2y_1 d_{20} + \beta_1 y_1 l_{10}) \\
 &\quad + 4(y_1 d_{10} - 1)(\eta_{120} + 4y_1 d_{30} + 2\beta_1 y_1 l_{20}) \\
 &\quad + \eta_{220} + 4y_1(2d_{40} + \beta_1 l_{30}) + 4y_1^2 d_{20}^2 \\
 &= \frac{8y_1^3 + 16y_1^2 y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2 y_2 + 64y_1 y_2^2}{(1-y_2)^6} \\
 &\quad + \frac{8y_1 y_2^4 + 56y_1 y_2^2 + 48y_2^3 + 8y_1 y_2}{(1-y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1-y_2)^8} \\
 &\quad + 4(\beta_1 y_1 + \beta_2 y_2) \left[\frac{(y_1 + y_2)^2}{(1-y_2)^4} + \frac{2y_2(y_1 + y_2)}{(1-y_2)^5} + \frac{y_2^2}{(1-y_2)^6} \right] = \sigma_{220}.
 \end{aligned}$$

因此, 在定理2.2.1的条件下,

$$T_A + T_B \xrightarrow{d} N(0, \omega_1^2 \sigma_{110} + 2\omega_1 \omega_2 \sigma_{120} + \omega_2^2 \sigma_{220}).$$

至此, 定理2.2.1得证。

定理2.2.2的证明. 在假设2.2.1-假设2.2.2和原假设 H_0 下, 令 $\Sigma_1 = \Sigma_2 = \Sigma$ 和 $\mathbf{S} = (n_1 + n_2)^{-1}(n_1 \mathbf{S}_1 + n_2 \mathbf{S}_2)$, 根据(2.13), (2.14)和(2.86), 可得

$$\begin{aligned} p^{-1} \text{tr} \mathbf{S} &= p^{-1} \text{tr} \Sigma + o_p(1), \\ p^{-1} \text{tr} \mathbf{S}^2 &= p^{-1} \text{tr} \Sigma^2 + (n_1 + n_2)^{-1} p^{-1} \text{tr}^2 \Sigma + o_p(1), \\ p^{-1} \text{tr}(\mathbf{S} \circ \mathbf{S}) &= p^{-1} \text{tr}(\Sigma \circ \Sigma) + o_p(1), \end{aligned}$$

这意味着

$$p^{-1} [\text{tr} \mathbf{S}^2 - (n_1 + n_2)^{-1} \text{tr}^2 \mathbf{S}] = p^{-1} \text{tr} \Sigma^2 + o_p(1).$$

根据(2.27)-(2.30)上面的等式, 可以发现 $d_{10}, d_{20}, d_{30}, d_{40}$ 分别是 $p^{-1} \text{tr} \widetilde{\mathbf{B}}_2^{-1}, p^{-1} \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2, p^{-1} \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^3, p^{-1} \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^4$ 的估计值的极限。当 $\Sigma_1 = \Sigma_2 = \Sigma$ 时, $\Gamma = \Sigma_1^{-1/2} \Sigma_2^{1/2} = \mathbf{I}_p$, 此时 $\text{tr} \widetilde{\mathbf{B}}_2^{-1}, \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^2, \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^3, \text{tr}(\widetilde{\mathbf{B}}_2^{-1})^4$ 是总体协方差矩阵为 \mathbf{I}_p 的样本协方差矩阵 $\widetilde{\mathbf{B}}_2$ 的线性谱统计量。因此, 由Zheng et al. (2015)^[53]中定理2.1的证明和性质A.1, 对于 $i = 1, 2, 3, 4$, 可得

$$p^{-1} (\text{tr}(\widetilde{\mathbf{B}}_2^{-1})^i - \text{Etr}(\widetilde{\mathbf{B}}_2^{-1})^i) = o_p(1), \quad (2.55)$$

$$p^{-1} \text{Etr}(\widetilde{\mathbf{B}}_2^{-1})^i = d_{i0p} + v_{i0p} + o(1), \quad (2.56)$$

其中

$$\begin{aligned} v_{10p} &= \frac{1}{p} \left[\frac{y_{n_2}}{(1-y_{n_2})^2} + \frac{\beta_2 y_{n_2}}{(1-y_{n_2})} \right], \\ v_{20p} &= \frac{1}{p} \left[\frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} - \frac{\beta_2 (y_{n_2}^2 - 3y_{n_2})}{(1-y_{n_2})^3} \right], \\ v_{30p} &= \frac{1}{p} \left[\frac{y_{n_2}^3 + 9y_{n_2}^2 + 6y_{n_2}}{(1-y_{n_2})^6} + \frac{6\beta_2 y_{n_2}}{(1-y_{n_2})^5} \right], \\ v_{40p} &= \frac{1}{p} \left[\frac{y_{n_2}^4 + 18y_{n_2}^3 + 35y_{n_2}^2 + 10y_{n_2}}{(1-y_{n_2})^8} + \frac{10\beta_2 (y_{n_2}^2 + y_{n_2})}{(1-y_{n_2})^7} \right]. \end{aligned}$$

因此, 对于 $i = 1, 2, 3, 4$, 用 $d_{i0p} + v_{i0p}$ 替换(2.54)中的 d_{i0} 时, 就能得到 $\hat{\sigma}_{220}$ 的表达式。综上, 定理2.2.2得证。

定理2.2.3的证明. 首先, 若 $f(\rho, x_1, x_2)$ 是二元正态分布 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$ 的密度函数, 其中 $0 < \rho < 1$, 则根据等式

$$\alpha = 1 - \int_{-g_\alpha}^{g_\alpha} \int_{-g_\alpha}^{g_\alpha} f(\rho, x_1, x_2) dx_1 dx_2$$

可以确定一个函数 $g(\rho) : \rho \rightarrow g_\alpha$, 并且 $g(\rho)$ 是一个连续函数。因此, 当定理2.2.1的假设条件成立时, 由定理2.2.1, 定理2.2.2和Slutsky定理可得

$$\begin{aligned} P(T_{\text{dr}} > t_\alpha) &= P(\max\{|T_{\text{d}} - \mu_0 - \hat{\mu}_{10}| / \sqrt{\hat{\sigma}_{110}}, |T_{\text{r}} - \mu_{20}| / \sqrt{\hat{\sigma}_{220}}\} > t_\alpha) \\ &= 1 - P(|T_{\text{d}} - \mu_0 - \hat{\mu}_{10}| / \sqrt{\hat{\sigma}_{110}} \leq t_\alpha, |T_{\text{r}} - \mu_{20}| / \sqrt{\hat{\sigma}_{220}} \leq t_\alpha) \\ &\rightarrow \alpha, \end{aligned}$$

这意味着检验 T_{dr} 的渐近水平为 α 。其次, 在定理2.2.1的条件和Cai et al. (2013)^[22]中的条件(C1), (C2)(或(C2*))和(C3)下,

$$\begin{aligned} P(T_{\text{drx}}^m > t_{\alpha/2}) &= P(\max\{T_{\text{dr}}, c_\alpha T_{\text{x}}\} > t_{\alpha/2}) \\ &\leq P(T_{\text{dr}} > t_{\alpha/2}) + P(T_{\text{x}} > q_{\alpha/2}) \\ &\rightarrow \alpha/2 + \alpha/2 = \alpha, \end{aligned}$$

这表明检验 T_{drx}^m 的渐近水平等于或小于 α 。最后, 在定理2.2.1的条件和Cai et al. (2013)^[22]中的条件(C1), (C2)(或(C2*))和(C3)下, 当阈值 $s(N_1, N_2, p) - 4 \log p \geq 0$ 时, 由(2.4)可得

$$P(T_{\text{x}} - 4 \log p + \log \log p \leq s(N_1, N_2, p) - 4 \log p + \log \log p) \rightarrow 1.$$

因此,

$$\begin{aligned} P(T_{\text{dr}} > t_\alpha) &\leq P(T_{\text{drx}}^w > t_\alpha) \\ &= P(T_{\text{dr}} + p^2 I(T_{\text{x}} > s(N_1, N_2, p)) > t_\alpha) \\ &\leq P(T_{\text{dr}} > t_\alpha) + P(p^2 I(T_{\text{x}} > s(N_1, N_2, p)) > 0) \end{aligned}$$

$$\begin{aligned}
 &= P(T_{\text{dr}} > t_\alpha) + 1 - P(p^2 I(T_x > s(N_1, N_2, p)) = 0) \\
 &= P(T_{\text{dr}} > t_\alpha) + 1 - P(T_x \leq s(N_1, N_2, p)) \\
 &\rightarrow \alpha + 1 - 1 = \alpha.
 \end{aligned}$$

由此可得检验 T_{drx}^w 的渐近水平为 α 。综上，定理2.2.3得证。

定理2.3.1的证明. 与定理2.2.1的证明类似，因为所有的量都是在 $(\Sigma_1, \Sigma_2) \in \Pi_1$ 下计算的，所以在这些量的下标中添加一个1。当 $(\Sigma_1, \Sigma_2) \in \Pi_1$ 时，

$$\Sigma_1 - \Sigma_2 = a_1/(p + a_1)\Sigma_1, \quad \Sigma_2^{-1} = \tau_p^{-1}\Sigma_1^{-1}. \quad (2.57)$$

由此可得

$$\text{tr}(\Sigma_1 - \Sigma_2)^2 = a_1^2/(p + a_1)^2 \text{tr}\Sigma_1^2 = o(1),$$

根据(2.22)，可得

$$\begin{aligned}
 \mu_{11} &= n_1^{-1} \text{tr}\Sigma_1^2 + n_2^{-1} \text{tr}\Sigma_2^2 + \beta_1 n_1^{-1} \text{tr}(\Sigma_1 \circ \Sigma_1) + \beta_2 n_2^{-1} \text{tr}(\Sigma_2 \circ \Sigma_2) + o(1) \\
 &= (y_{n_1} + y_{n_2} \tau_p^2) p^{-1} \text{tr}\Sigma_1^2 + (\beta_1 y_{n_1} + \beta_2 y_{n_2} \tau_p^2) p^{-1} \text{tr}(\Sigma_1 \circ \Sigma_1) + o(1).
 \end{aligned}$$

当 $\Sigma_2 = \tau_p \Sigma_1$ 时， $\Gamma_1 = \Sigma_1^{-1/2} \Sigma_2^{1/2} = \tau_p^{1/2} \mathbf{I}_p$ 和 $\mathbf{T}_{1p} = \Gamma_1 \Gamma_1^T = \tau_p \mathbf{I}_p$ 。若将 $\widetilde{\mathbf{B}}_2$ 看作§2.5.3节中的 \mathbf{B}_n ，则 Γ_1 和 \mathbf{T}_{1p} 也满足假设c-d-f。 \mathbf{T}_{1p} 的经验谱分布 H_{1p} 和极限谱分布 H_1 分别为 $\tau_p \delta_1$ 和 δ_1 。因此，对于 $i, j \in \{1, 2\}$ ，由(2.18)，(2.19)和(2.26)可得

$$c_{i1} = \tau_p^{-i} c_{i0}, \quad \xi_{i1} = \xi_{i0}, \quad \eta_{ij1} = \eta_{ij0}. \quad (2.58)$$

根据(2.57)和(2.27)-(2.33)，可得

$$d_{11p} = \frac{1}{(1 - y_{N_2}) \tau_p} \rightarrow \frac{1}{(1 - y_2)} = d_{10}, \quad (2.59)$$

$$d_{21p} = \frac{1}{(1 - y_{N_2})^3 \tau_p^2} \rightarrow \frac{1}{(1 - y_2)^3} = d_{20}, \quad (2.60)$$

$$d_{31p} = \frac{1 + y_{N_2}}{(1 - y_{N_2})^5 \tau_p^3} \rightarrow \frac{1 + y_2}{(1 - y_2)^5} = d_{30}, \quad (2.61)$$

$$d_{41p} = \frac{y_{N_2}^2 + 3y_{N_2} + 1}{(1 - y_{N_2})^7 \tau_p^4} \rightarrow \frac{y_2^2 + 3y_2 + 1}{(1 - y_2)^7} = d_{40}, \quad (2.62)$$

$$l_{11p} = \frac{1}{(1-y_{N_2})^2\tau_p^2} \rightarrow \frac{1}{(1-y_2)^2} = l_{10}, \quad (2.63)$$

$$l_{21p} = \frac{1}{(1-y_{N_2})^4\tau_p^3} \rightarrow \frac{1}{(1-y_2)^4} = l_{20}, \quad (2.64)$$

$$l_{31p} = \frac{1}{(1-y_{N_2})^6\tau_p^4} \rightarrow \frac{1}{(1-y_2)^6} = l_{30}. \quad (2.65)$$

因此, 由(2.39), (2.41), (2.42), (2.43), (2.58), (2.60)和(2.63)可得

$$\begin{aligned} \mu_{21} &= p(y_{n_1}c_{11}^2 - 2c_{11} + c_{21}) + p + 2y_{n_1}c_{11}\xi_{11} - 2\xi_{11} + \xi_{21} \\ &\quad + y_{n_1}d_{21p} + \beta_1y_{n_1}l_{11p} + o_p(1) \\ &= p\left[-\frac{2}{(1-y_{n_2})\tau_p} + \frac{y_{n_1}}{(1-y_{n_2})^2\tau_p^2} + \frac{1}{(1-y_{n_2})^3\tau_p^3}\right] + p \\ &\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} - \frac{2\beta_2y_{n_2}}{(1-y_{n_2})} \\ &\quad + \frac{2\beta_2y_{n_1}y_{n_2} + \beta_1y_{n_1} + \beta_2y_{n_2}}{(1-y_{n_2})^2} + \frac{2\beta_2y_{n_2}}{(1-y_{n_2})^3} + o_p(1). \end{aligned}$$

令 $m_{11} = \int x dL_1(x)$ 和 $m_{21} = \int x^2 dL_1(x)$, 根据(2.34), (2.35), (2.36), (2.57), (2.58), (2.41)和(2.59), 可得

$$\begin{aligned} \sigma_{111p} &= 4[(y_{N_1} + y_{N_2}\tau_p^2)p^{-1}\text{tr}\Sigma_1^2]^2 + o_p(1) \rightarrow 4(y_1 + y_2)^2 m_{21}^2 = \sigma_{111}, \\ \sigma_{121p}^1 &= \left[\frac{8(y_{N_1} + y_{N_2}\tau_p)y_{N_2}^2 + 4y_{N_2}^2}{(1-y_{N_2})^2} + \frac{8y_{N_2}^2 - 8y_{N_2}^2\tau_p}{(1-y_{N_2})^2} + \frac{8y_{N_2}^3}{(1-y_{N_2})^3}\right](p^{-1}\text{tr}\Sigma_1)^2 + o_p(1), \\ \sigma_{121p}^2 &= \left[\frac{8(y_{N_1} + y_{N_2}\tau_p)y_{N_1}y_{N_2} + 8y_{N_1}y_{N_2} - 8y_{N_1}y_{N_2}\tau_p}{(1-y_{N_2})^2\tau_p} + \frac{4y_{N_1}^2}{(1-y_{N_2})^2\tau_p^2} + \frac{8y_{N_1}y_{N_2}}{(1-y_{N_2})^3\tau_p}\right](p^{-1}\text{tr}\Sigma_1)^2 + o_p(1), \end{aligned}$$

由此可得

$$\begin{aligned} \sigma_{121p} &= \sigma_{121p}^1 + \sigma_{121p}^2 \\ &\rightarrow \left[\frac{8y_2(y_1 + y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3}\right]m_{11}^2 = \sigma_{121}. \end{aligned}$$

至于 σ_{221} , 根据(2.37), (2.38), (2.41), (2.44), (2.45), (2.46), (2.58)和(2.59)-(2.65),

可得

$$\begin{aligned}
 \sigma_{221p} &= \sigma_{221p}^1 + \sigma_{221p}^2 \\
 &\rightarrow 4(y_1 d_{10} - 1)^2(\eta_{110} + 2y_1 d_{20} + \beta_1 y_1 l_{10}) \\
 &\quad + 4(y_1 d_{10} - 1)(\eta_{120} + 4y_1 d_{30} + 2\beta_1 y_1 l_{20}) \\
 &\quad + \eta_{220} + 4y_1(2d_{40} + \beta_1 l_{30}) + 4y_1^2 d_{20}^2 \\
 &= \frac{8y_1^3 + 16y_1^2 y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2 y_2 + 64y_1 y_2^2}{(1-y_2)^6} \\
 &\quad + \frac{8y_1 y_2^4 + 56y_1 y_2^2 + 48y_2^3 + 8y_1 y_2}{(1-y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1-y_2)^8} \\
 &\quad + 4(\beta_1 y_1 + \beta_2 y_2) \left[\frac{(y_1 + y_2)^2}{(1-y_2)^4} + \frac{2y_2(y_1 + y_2)}{(1-y_2)^5} + \frac{y_2^2}{(1-y_2)^6} \right] = \sigma_{220}.
 \end{aligned}$$

因此, 在定理2.3.1的条件下,

$$T_A + T_B \xrightarrow{d} N(0, \omega_1^2 \sigma_{111} + 2\omega_1 \omega_2 \sigma_{121} + \omega_2^2 \sigma_{220}).$$

至此, 定理2.3.1得证。

性质2.3.1的证明. 在假设2.2.1-假设2.2.2下, 由(2.13)和(2.86)可得

$$p^{-1} \text{tr} \mathbf{S} = p^{-1} \text{tr} \boldsymbol{\Sigma}_w + o_p(1), \quad (2.66)$$

$$\begin{aligned}
 p^{-1} \text{tr} \mathbf{S}^2 &= p^{-1} \text{tr} \boldsymbol{\Sigma}_w^2 + \frac{n_1}{p(n_1 + n_2)^2} \text{tr}^2 \boldsymbol{\Sigma}_1 \\
 &\quad + \frac{n_2}{p(n_1 + n_2)^2} \text{tr}^2 \boldsymbol{\Sigma}_2 + o_p(1),
 \end{aligned} \quad (2.67)$$

$$p^{-1} \text{tr}(\mathbf{S} \circ \mathbf{S}) = p^{-1} \text{tr}(\boldsymbol{\Sigma}_w \circ \boldsymbol{\Sigma}_w) + o_p(1), \quad (2.68)$$

其中 $\boldsymbol{\Sigma}_w = n_1/(n_1 + n_2)\boldsymbol{\Sigma}_1 + n_2/(n_1 + n_2)\boldsymbol{\Sigma}_2$, 当 $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) \in \Pi_1$ 时,

$$\begin{aligned}
 \hat{\mu}_{10} &= \mu_{11} + o_p(1), & \hat{\sigma}_{110} &= \sigma_{111} + o_p(1), \\
 \hat{\sigma}_{120} &= \sigma_{121} + o_p(1), & \hat{\sigma}_{220} &= \sigma_{220} + o(1)
 \end{aligned}$$

和

$$\mu_{21} - \mu_{20} = 2a_1 \left[\frac{y_{n_1} + y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_2}}{(1-y_{n_2})^3} \right] + o(1).$$

因此, 根据定理2.3.1和Slutsky定理, 可得

$$\begin{aligned} & P\left(\frac{|T_d - \mu_0 - \hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}} > z_{1-\alpha/2}\right) \\ &= P\left(\frac{\sqrt{\sigma_{111}}}{\sqrt{\hat{\sigma}_{110}}} \cdot \frac{|T_d - \mu_0 - \mu_{11} + \mu_{11} - \hat{\mu}_{10}|}{\sqrt{\sigma_{111}}} > z_{1-\alpha/2}\right) \\ &\rightarrow 1 - \left[\Phi(z_{1-\alpha/2}) - \Phi(-z_{1-\alpha/2})\right] = \alpha \end{aligned}$$

和

$$\begin{aligned} & P\left(\frac{|T_r - \mu_{20}|}{\sqrt{\hat{\sigma}_{220}}} > z_{1-\alpha/2}\right) \\ &= P\left(\frac{\sqrt{\sigma_{220}}}{\sqrt{\hat{\sigma}_{220}}} \cdot \frac{|T_r - \mu_{21} + \mu_{21} - \mu_{20}|}{\sqrt{\sigma_{220}}} > z_{1-\alpha/2}\right) \\ &\rightarrow 1 - \left[\Phi(z_{1-\alpha/2} - \Delta_1) - \Phi(-z_{1-\alpha/2} - \Delta_1)\right] > \alpha, \end{aligned}$$

其中 $\Phi(\cdot)$ 是标准正态分布的分布函数,

$$\Delta_1 = \frac{2a_1}{\sqrt{\sigma_{220}}} \left[\frac{y_1 + y_2}{(1 - y_2)^2} + \frac{y_2}{(1 - y_2)^3} \right].$$

在定理2.3.1的条件下, 根据 t_α 的定义, 定理2.3.1和Slutsky定理, 可得

$$\begin{aligned} P(T_{dr} > t_\alpha) &= 1 - P\left(\frac{|T_d - \mu_0 - \hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}} \leq t_\alpha, \frac{|T_r - \mu_{20}|}{\sqrt{\hat{\sigma}_{220}}} \leq t_\alpha\right) \\ &= 1 - P\left(\frac{\sqrt{\sigma_{111}}}{\sqrt{\hat{\sigma}_{110}}} \cdot \frac{|T_d - \mu_0 - \mu_{11} + \mu_{11} - \hat{\mu}_{10}|}{\sqrt{\sigma_{111}}} \leq t_\alpha, \right. \\ &\quad \left. \frac{\sqrt{\sigma_{220}}}{\sqrt{\hat{\sigma}_{220}}} \cdot \frac{|T_r - \mu_{21} + \mu_{21} - \mu_{20}|}{\sqrt{\sigma_{220}}} \leq t_\alpha\right) \\ &\rightarrow 1 - \int_{-t'_\alpha - \Delta_1}^{t'_\alpha - \Delta_1} \int_{-t'_\alpha}^{t'_\alpha} f'(x_d, x_r) dx_d dx_r > \alpha, \end{aligned}$$

其中 $f'(x_d, x_r)$ 是 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix}\right)$ 的密度函数, $\rho_1 = \sigma_{121} / \sqrt{\sigma_{111}\sigma_{220}}$, t'_α 满足

$$\alpha = 1 - \int_{-t'_\alpha}^{t'_\alpha} \int_{-t'_\alpha}^{t'_\alpha} f'(x_d, x_r) dx_d dx_r.$$

根据 T_{drx}^w 的定义可知, $T_{drx}^w \geq T_{dr}$, 因此

$$P(T_{drx}^w > t_\alpha) \geq P(T_{dr} > t_\alpha),$$

由此可得

$$\lim_{p \rightarrow \infty} P(T_{\text{drx}}^w > t_\alpha) \geq \lim_{p \rightarrow \infty} P(T_{\text{dr}} > t_\alpha) > \alpha.$$

最后, 根据 T_{drx}^m 的定义, 可得

$$\begin{aligned} P(\max\{T_{\text{dr}}, c_\alpha T_x\} > t_{\alpha/2}) &= 1 - P(T_{\text{dr}} \leq t_{\alpha/2}, T_x \leq q_{\alpha/2}) \\ &\geq 1 - P(T_{\text{dr}} \leq t_{\alpha/2}) = P(T_{\text{dr}} > t_{\alpha/2}), \end{aligned}$$

由此可得

$$\lim_{p \rightarrow \infty} P(T_{\text{drx}}^m > t_{\alpha/2}) \geq 1 - \int_{-t'_{\alpha/2} - \Delta_1}^{t'_{\alpha/2} - \Delta_1} \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} f(x_d, x_r) dx_d dx_r > \alpha/2,$$

其中 $t'_{\alpha/2}$ 满足

$$\alpha/2 = 1 - \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} f(x_d, x_r) dx_d dx_r.$$

定理2.3.2的证明. 与定理2.2.1的证明类似, 因为所有的量都是在 $(\Sigma_1, \Sigma_2) \in \Pi_2$ 下计算的, 所以在这些量的下标中添加一个2。当 $(\Sigma_1, \Sigma_2) \in \Pi_2$ 时,

$$\Sigma_1 = \mathbf{I}_p, \quad \Sigma_2 - \Sigma_1 = a_2/p\mathbf{J}_p, \quad \Sigma_2^{-1} = \mathbf{I}_p - a_2/[p(1+a_2)]\mathbf{J}_p. \quad (2.69)$$

由此可得 $\text{tr}(\Sigma_1 - \Sigma_2)^2 = a_2^2$, 当 $\beta_2 = 0$ 时, 根据(2.22), 可得

$$\mu_{12} = a_2^2 + y_{n_1} + y_{n_2} + \beta_1 y_{n_1} + o(1).$$

当 $\Sigma_1 = \mathbf{I}_p$ 和 $\Sigma_2 = \Sigma_1 + a_2/p\mathbf{J}_p$ 时, $\Gamma_2 = \Sigma_1^{-1/2}\Sigma_2^{1/2} = \mathbf{I}_p + (\sqrt{1+a_2} - 1)/p\mathbf{J}_p$ 和 $\mathbf{T}_{2p} = \Gamma_2\Gamma_2^T = \Sigma_2$ 。若将 $\tilde{\mathbf{B}}_2$ 看作§2.5.3节中的 \mathbf{B}_n , 则 Γ_2 和 \mathbf{T}_{2p} 也满足假设c-d。 \mathbf{T}_{2p} 的特征值为 $1+a_2, 1, \dots, 1$, 符合Wang和Yao (2014)^[54]中考虑的离群模型。根据Wang和Yao (2014)^[54]中的定理2, 对于 $i = 1, 2$,

$$\begin{aligned} c_{i2} &= -\frac{1}{2\pi i p} \oint_{C_1} f_i\left(-\frac{1}{z} + \frac{y_{n_2}}{1+z}\right) \left(\frac{1}{y_{n_2}z} - \frac{(1+a_2)^2 z}{(1+(1+a_2)z)^2}\right) dz \\ &\quad - \frac{1}{2\pi i p} \oint_{C_1} f'_i\left(-\frac{1}{z} + \frac{y_{n_2}}{1+z}\right) \frac{a_2}{(1+(1+a_2)z)(1+z)} \left(\frac{1}{z} - \frac{y_{n_2}z}{(1+z)^2}\right) dz \\ &\quad + \left(1 - \frac{1}{p}\right) c_{i0} + \frac{1}{p} f_i(\phi_p(a_2)) I(a_2 > \sqrt{y_{n_2}}) + o\left(\frac{1}{p}\right), \end{aligned}$$

其中 $f_1(x) = 1/x$, $f_2(x) = 1/x^2$, $\phi_p(a_2) = 1 + a_2 + y_{n_2}(1 + a_2)/a_2$ 和 C_1 是逆时针方向的围道, 当将其限制在实轴上时, 包含区间 $[-1/(1 - \sqrt{y_{n_2}}), -1/(1 + \sqrt{y_{n_2}})]$ 。经过计算可得

$$c_{12} = \frac{1}{1 - y_{n_2}} - \frac{a_2}{p(1 - y_{n_2})(1 + a_2)} + o\left(\frac{1}{p}\right), \quad (2.70)$$

$$c_{22} = \frac{1}{(1 - y_{n_2})^3} - \frac{(1 + y_{n_2})a_2^2 + 2a_2}{p(1 - y_{n_2})^3(1 + a_2)^2} + o\left(\frac{1}{p}\right). \quad (2.71)$$

由于 \mathbf{T}_{2p} 的经验谱分布 H_{2p} 和极限谱分布 H_2 分别为 $(1 - p^{-1})\delta_1 + p^{-1}\delta_{1+a_2}$ 和 δ_1 , 当 $\beta_2 = 0$ 时, 由(2.19)和(2.26)可得

$$\xi_{12} = \frac{y_{n_2}}{(1 - y_{n_2})^2}, \quad \xi_{22} = \frac{y_{n_2}^2 + 3y_{n_2}}{(1 - y_{n_2})^4} \quad (2.72)$$

和

$$\eta_{112} = \frac{2y_2}{(1 - y_2)^4}, \quad \eta_{122} = \frac{4y_2(1 + y_2)}{(1 - y_2)^6}, \quad \eta_{222} = \frac{4y_2(2y_2^2 + 5y_2 + 2)}{(1 - y_2)^8}. \quad (2.73)$$

根据(2.69)和(2.27)-(2.33), 可得

$$d_{12p} = \frac{1}{(1 - y_{N_2})} + o(1) \rightarrow \frac{1}{(1 - y_2)} = d_{10}, \quad (2.74)$$

$$d_{22p} = \frac{1}{(1 - y_{N_2})^3} + o(1) \rightarrow \frac{1}{(1 - y_2)^3} = d_{20}, \quad (2.75)$$

$$d_{32p} = \frac{1 + y_{N_2}}{(1 - y_{N_2})^5} + o(1) \rightarrow \frac{1 + y_2}{(1 - y_2)^5} = d_{30}, \quad (2.76)$$

$$d_{42p} = \frac{y_{N_2}^2 + 3y_{N_2} + 1}{(1 - y_{N_2})^7} + o(1) \rightarrow \frac{y_2^2 + 3y_2 + 1}{(1 - y_2)^7} = d_{40}, \quad (2.77)$$

$$l_{12p} = \frac{1}{(1 - y_{N_2})^2} + o(1) \rightarrow \frac{1}{(1 - y_2)^2} = l_{10}, \quad (2.78)$$

$$l_{22p} = \frac{1}{(1 - y_{N_2})^4} + o(1) \rightarrow \frac{1}{(1 - y_2)^4} = l_{20}, \quad (2.79)$$

$$l_{32p} = \frac{1}{(1 - y_{N_2})^6} + o(1) \rightarrow \frac{1}{(1 - y_2)^6} = l_{30}. \quad (2.80)$$

因此, 由(2.39), (2.70), (2.71), (2.72), (2.75)和(2.78)可得

$$\begin{aligned} \mu_{22} &= p(y_{n_1}c_{12}^2 - 2c_{12} + c_{22}) + p + 2y_{n_1}c_{12}\xi_{12} - 2\xi_{12} + \xi_{22} \\ &\quad + y_{n_1}d_{22p} + \beta_1y_{n_1}l_{12p} + o_p(1) \end{aligned}$$

$$\begin{aligned}
 &= p \left[-\frac{2}{(1-y_{n_2})} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3} \right] + p \\
 &\quad + \frac{a_2^2 - 2a_2(1+a_2)(y_{n_1} + y_{n_2})}{(1-y_{n_2})^2(1+a_2)^2} - \frac{2a_2y_{n_2}}{(1-y_{n_2})^3(1+a_2)} \\
 &\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} + \frac{\beta_1y_{n_1}}{(1-y_{n_2})^2} + o_p(1).
 \end{aligned}$$

根据(2.34), (2.35), (2.36), (2.69), (2.70)和(2.74), 可得

$$\begin{aligned}
 \sigma_{112p} &= 4(y_{N_1} + y_{N_2})^2 + o_p(1) \rightarrow 4[(y_1 + y_2)]^2 = \sigma_{112}, \\
 \sigma_{122p}^1 &= \left[\frac{8(y_{N_1} + y_{N_2})y_{N_2}^2 + 4y_{N_2}^2}{(1-y_{N_2})^2} + \frac{8y_{N_2}^3}{(1-y_{N_2})^3} \right] + o_p(1), \\
 \sigma_{122p}^2 &= \left[\frac{8(y_{N_1} + y_{N_2})y_{N_1}y_{N_2} + 4y_{N_1}^2}{(1-y_{N_2})^2} + \frac{8y_{N_1}y_{N_2}}{(1-y_{N_2})^3} \right] + o_p(1),
 \end{aligned}$$

由此可得

$$\begin{aligned}
 \sigma_{122p} &= \sigma_{122p}^1 + \sigma_{122p}^2 \\
 &= \left[\frac{8y_{N_2}(y_{N_1} + y_{N_2})^2 + 4y_{N_1}^2 + 4y_{N_2}^2}{(1-y_{N_2})^2} + \frac{8y_{N_1}y_{N_2} + 8y_{N_2}^3}{(1-y_{N_2})^3} \right] + o_p(1) \\
 &\rightarrow \left[\frac{8y_2(y_1 + y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3} \right] = \sigma_{122}.
 \end{aligned}$$

根据(2.37), (2.38), (2.70), (2.73)和(2.74)-(2.80), 可得

$$\begin{aligned}
 \sigma_{222p} &= \sigma_{222p}^1 + \sigma_{222p}^2 \\
 &\rightarrow 4(y_1d_{10} - 1)^2(\eta_{112} + 2y_1d_{20} + \beta_1y_1l_{10}) \\
 &\quad + 4(y_1d_{10} - 1)(\eta_{122} + 4y_1d_{30} + 2\beta_1y_1l_{20}) \\
 &\quad + \eta_{222} + 4y_1(2d_{40} + \beta_1l_{30}) + 4y_1^2d_{20}^2 \\
 &= \frac{8y_1^3 + 16y_1^2y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2y_2 + 64y_1y_2^2}{(1-y_2)^6} \\
 &\quad + \frac{8y_1y_2^4 + 56y_1y_2^2 + 48y_2^3 + 8y_1y_2}{(1-y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1-y_2)^8} \\
 &\quad + 4\beta_1y_1 \left[\frac{(y_1 + y_2)^2}{(1-y_2)^4} + \frac{2y_2(y_1 + y_2)}{(1-y_2)^5} + \frac{y_2^2}{(1-y_2)^6} \right] = \sigma_{222}.
 \end{aligned}$$

因此, 在定理2.3.2的条件下,

$$T_A + T_B \xrightarrow{d} N(0, \omega_1^2 \sigma_{112} + 2\omega_1 \omega_2 \sigma_{122} + \omega_2^2 \sigma_{222}).$$

至此, 定理2.3.2得证。

性质2.3.2的证明. 在假设2.2.1-假设2.2.2下, 当 $(\Sigma_1, \Sigma_2) \in \Pi_2$ 和 $\beta_2 = 0$ 时, 根据(2.66), (2.67)和(2.68), 可得

$$\hat{\sigma}_{110} = \sigma_{112} + o_p(1), \quad \hat{\sigma}_{120} = \sigma_{122} + o_p(1), \quad \hat{\sigma}_{220} = \sigma_{222} + o(1),$$

$$\mu_{12} - \hat{\mu}_{10} = a_2^2 + o_p(1)$$

和

$$\mu_{22} - \mu_{20} = \frac{a_2^2 - 2a_2(1+a_2)(y_{n_1} + y_{n_2})}{(1-y_{n_2})^2(1+a_2)^2} - \frac{2a_2 y_{n_2}}{(1-y_{n_2})^3(1+a_2)}.$$

因此, 根据定理2.3.2和Slutsky定理, 可得

$$\begin{aligned} & P\left(\frac{|T_d - \mu_0 - \hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}} > z_{1-\alpha/2}\right) \\ &= P\left(\frac{\sqrt{\sigma_{112}}}{\sqrt{\hat{\sigma}_{110}}} \cdot \frac{|T_d - \mu_0 - \mu_{12} + \mu_{12} - \hat{\mu}_{10}|}{\sqrt{\sigma_{112}}} > z_{1-\alpha/2}\right) \\ &\rightarrow 1 - \left[\Phi(z_{1-\alpha/2} - \Delta_2) - \Phi(-z_{1-\alpha/2} - \Delta_2)\right] > \alpha \end{aligned}$$

和

$$\begin{aligned} & P\left(\frac{|T_r - \mu_{20}|}{\sqrt{\hat{\sigma}_{220}}} > z_{1-\alpha/2}\right) \\ &= P\left(\frac{\sqrt{\sigma_{222}}}{\sqrt{\hat{\sigma}_{220}}} \cdot \frac{|T_r - \mu_{22} + \mu_{22} - \mu_{20}|}{\sqrt{\sigma_{222}}} > z_{1-\alpha/2}\right) \\ &\rightarrow 1 - \left[\Phi(z_{1-\alpha/2} - \Delta_3) - \Phi(-z_{1-\alpha/2} - \Delta_3)\right] \geq \alpha, \end{aligned}$$

其中 $\Delta_2 = a_2^2/[2(y_1 + y_2)]$,

$$\Delta_3 = \frac{1}{\sqrt{\sigma_{222}}} \left[\frac{a_2^2 - 2a_2(1+a_2)(y_1 + y_2)}{(1-y_2)^2(1+a_2)^2} - \frac{2a_2 y_2}{(1-y_2)^3(1+a_2)} \right].$$

在定理2.3.2的条件下, 根据 t_α 的定义, 定理2.3.2和Slutsky定理, 可得

$$\begin{aligned}
 P(T_{\text{dr}} > t_\alpha) &= 1 - P\left(\frac{|T_d - \mu_0 - \hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}} \leq t_\alpha, \frac{|T_r - \mu_{20}|}{\sqrt{\hat{\sigma}_{220}}} \leq t_\alpha\right) \\
 &= 1 - P\left(\frac{\sqrt{\sigma_{112}}}{\sqrt{\hat{\sigma}_{110}}} \cdot \frac{|T_d - \mu_0 - \mu_{12} + \mu_{12} - \hat{\mu}_{10}|}{\sqrt{\sigma_{112}}} \leq t_\alpha, \right. \\
 &\quad \left. \frac{\sqrt{\sigma_{222}}}{\sqrt{\hat{\sigma}_{220}}} \cdot \frac{|T_r - \mu_{22} + \mu_{22} - \mu_{20}|}{\sqrt{\sigma_{222}}} \leq t_\alpha\right) \\
 &\rightarrow 1 - \int_{-t_\alpha^* - \Delta_3}^{t_\alpha^* - \Delta_3} \int_{-t_\alpha^* - \Delta_2}^{t_\alpha^* - \Delta_2} f^*(x_d, x_r) dx_d dx_r > \alpha,
 \end{aligned}$$

其中 $f^*(x_d, x_r)$ 是 $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{pmatrix}\right)$ 的密度函数, $\rho_2 = \sigma_{122}/\sqrt{\sigma_{112}\sigma_{222}}$, t_α^* 满足

$$\alpha = 1 - \int_{-t_\alpha^*}^{t_\alpha^*} \int_{-t_\alpha^*}^{t_\alpha^*} f^*(x_d, x_r) dx_d dx_r.$$

与性质2.3.1的证明类似, 可以得到(IV)和(V)。综上, 性质2.3.2得证。

性质2.3.3的证明. 首先给出结论(I)的证明。在结论(I)的条件下, 由(2.66), (2.67)和(2.68)可得

$$\begin{aligned}
 \hat{m}_{10} &= p^{-1} \text{tr} \Sigma_w + o_p(1), \\
 \hat{m}_{20} &= p^{-1} \text{tr} \Sigma_w^2 + \frac{y_{n_1}^2 y_{n_2}^2}{(y_{n_1} + y_{n_2})^3} (p^{-1} \text{tr} \Sigma_d)^2 + o_p(1),
 \end{aligned}$$

其中 $\Sigma_d = \Sigma_1 - \Sigma_2$, 这意味着临界值 t_α 依概率收敛到一个常数。根据Cai et al. (2013)^[221]中的(32)和(33), 可得

$$\begin{aligned}
 T_x &\stackrel{a.s.}{\geq} 0.5 \max_{1 \leq l_1 \leq l_2 \leq p} \frac{(\sigma_{1l_1 l_2} - \sigma_{2l_1 l_2})^2}{\hat{\theta}_{1l_1 l_2}/n_1 + \hat{\theta}_{2l_1 l_2}/n_2} - 4 \log p + 0.5 \log \log p \\
 &\stackrel{a.s.}{\geq} 0.5 \max_{1 \leq l_1 \leq l_2 \leq p} \frac{(\sigma_{1l_1 l_2} - \sigma_{2l_1 l_2})^2}{\theta_{1l_1 l_2}/n_1 + \theta_{2l_1 l_2}/n_2} - 4 \log p + 0.5 \log \log p.
 \end{aligned}$$

因此, 当 p 足够大时,

$$\begin{aligned}
 P(T_{\text{drx}}^w > t_\alpha) &= P(T_{\text{dr}} + p^2 I(T_x > s(N_1, N_2, p)) > t_\alpha) \\
 &\geq P(p^2 I(T_x > s(N_1, N_2, p)) > t_\alpha) = P(T_x > s(N_1, N_2, p)).
 \end{aligned}$$

由此可得, 当 $p \rightarrow \infty$ 时, $P(T_{\text{drx}}^w > t_\alpha) \rightarrow 1$ 。

接下来, 给出结论(II)的证明。根据 T_{drx}^m 的定义, 可得

$$\begin{aligned} P(T_{\text{drx}}^m > t_{\alpha/2}) &= P(\max\{T_{\text{dr}}, c_{\alpha}T_x\} > t_{\alpha/2}) \\ &= 1 - P(T_{\text{dr}} \leq t_{\alpha/2}, T_x \leq q_{\alpha/2}) \\ &\geq 1 - P(T_x \leq q_{\alpha/2}) = P(T_x > q_{\alpha/2}), \end{aligned}$$

当 $(\Sigma_1, \Sigma_2) \in \Pi_3$ 时, 根据Cai et al. (2013)^[221]中的定理2, 可得 $P(T_{\text{drx}}^m > t_{\alpha/2}) \rightarrow 1$ 。综上所述, 性质2.3.3得证。

§2.5.3 定理和性质证明过程中需要的引理

这一节给出§2.5.1节中所需要的引理。根据Bai和Silverstein (2004)^[551]中的符号, 在本节中, 样本的维数和样本量分布用 n 和 N 表示。样本 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ 来自 n 维总体 \mathbf{x} 。令 $\mathbf{x}_i = (x_{1i}, \dots, x_{ni})^T$ 和 $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_N)$, 定义

$$\begin{aligned} \mathbf{B}_n &= N^{-1} \mathbf{\Gamma} \mathbf{X}_n \mathbf{X}_n^T \mathbf{\Gamma}^T, \\ m_n(z) &= m_{F^{\mathbf{B}_n}}(z) = \int \frac{1}{\lambda - z} dF^{\mathbf{B}_n}(\lambda), \quad \Im(z) \neq 0, \end{aligned}$$

其中 \mathbf{X}_n^T 代表 \mathbf{X}_n 的转置, $\mathbf{\Gamma}$ 是可逆的 $n \times n$ 矩阵, $\Im(z)$ 代表复数 z 的虚部, $m_n(z)$ 代表 $F^{\mathbf{B}_n}$ 的Stieltjes变换。令 $\underline{\mathbf{B}}_n = N^{-1} \mathbf{X}_n^T \mathbf{\Gamma}^T \mathbf{\Gamma} \mathbf{X}_n$, 其Stieltjes变换定义如下

$$\underline{m}_n(z) = m_{F^{\underline{\mathbf{B}}_n}}(z) = -\frac{1 - c_n}{z} + c_n m_n(z),$$

其中 $c_n = n/N$ 。

- **假设a.** 元素 $\{x_{ji}, j = 1, \dots, n; i = 1, \dots, N\}$ 是独立同分布的, 对于某个序列 $\eta_N \rightarrow 0$, 满足

$$|x_{ji}| \leq \eta_N \sqrt{N}, \quad \mathbb{E}x_{ji} = 0, \quad \mathbb{E}x_{ji}^2 = 1, \quad \beta_x = \mathbb{E}x_{ji}^4 - 3.$$

- **假设b.** 维数 n 和样本量 N 成比例趋于无穷, 即 $c_n = n/N \rightarrow c > 0$ 。
- **假设c.** $\mathbf{T}_n = \mathbf{\Gamma} \mathbf{\Gamma}^T$ 的谱模是有界的, 并且 \mathbf{T}_n 的经验谱分布 H_n 收敛到极限谱分布 H 。
- **假设d.** \mathbf{T}_n 的逆矩阵 \mathbf{T}_n^{-1} 的谱模是有界的。

- 假设f. $\Gamma^T \Gamma$ 是对角的或者 $\beta_x = 0$ 。

根据Silverstein (1995)^[56]中的定理1.1., 在假设a-b-c下, 样本协方差矩阵 \mathbf{B}_n 有极限谱分布 $F^{c,H}$, $F^{c,H}$ 是以 (c, H) 为指标的Marčenko-Pastur分布, 其支撑集为

$$[(1 - \sqrt{c})^2 I(0 < c < 1) \liminf_n \lambda_{\min}(\mathbf{T}_n), (1 + \sqrt{c})^2 \limsup_n \lambda_{\max}(\mathbf{T}_n)].$$

此外, 当 $c > 1$ 时, $F^{c,H}$ 在原点具有质量为 $1 - 1/c$ 的点测度。定义 \underline{m}_c 为伴随极限谱分布

$$\underline{F}^{c,H} = (1 - c)\delta_0 + cF^{c,H}$$

的Stieltjes变换。则 \underline{m}_c 为方程

$$z = -\frac{1}{\underline{m}_c(z)} + c \int \frac{tdH(t)}{1 + t\underline{m}_c(z)}, \quad z \in \mathbb{C}^+ = \{z : \Im z > 0\}$$

在 $\{\underline{m}_c \in \mathbb{C} : (1 - c)/z + \underline{m}_c \in \mathbb{C}^+\}$ 中的唯一解。

令 $\gamma_j = (1/\sqrt{N})\Gamma \mathbf{x}_j$, $E_0(\cdot)$ 代表期望, $E_j(\cdot)$ 代表关于 $\gamma_1, \dots, \gamma_j$ 生成的 σ -域的条件期望。

引理2.5.1. 若 $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3$ 和 \mathbf{M}_4 是对称的、非负定的且谱模有界的矩阵, 则在假设a-b-c下,

$$\begin{aligned} & \sum_{j=1}^N E_{j-1} \{[(E_j - E_{j-1})\text{tr}(\mathbf{B}_n \mathbf{M}_1)]^2\} \\ &= \frac{1}{N} [2\text{tr}(\mathbf{T}_n \mathbf{M}_1)^2 + \beta_x \text{tr}(\Gamma^T \mathbf{M}_1 \Gamma \circ \Gamma^T \mathbf{M}_1 \Gamma)], \end{aligned} \quad (2.81)$$

$$\begin{aligned} & \sum_{j=1}^N E_{j-1} \{[(E_j - E_{j-1})\text{tr}(\mathbf{B}_n \mathbf{M}_3)^2]^2\} \\ &= \frac{4}{N^3} \text{tr}^2(\mathbf{T}_n \mathbf{M}_3) [2\text{tr}(\mathbf{T}_n \mathbf{M}_3)^2 + \beta_x \text{tr}(\Gamma^T \mathbf{M}_3 \Gamma \circ \Gamma^T \mathbf{M}_3 \Gamma)] \end{aligned} \quad (2.82)$$

$$\begin{aligned} & + \frac{8}{N^2} \text{tr}(\mathbf{T}_n \mathbf{M}_3) [2\text{tr}(\mathbf{T}_n \mathbf{M}_3)^3 + \beta_x \text{tr}(\Gamma^T \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \Gamma \circ \Gamma^T \mathbf{M}_3 \Gamma)] \\ & + \frac{4}{N} [2\text{tr}(\mathbf{T}_n \mathbf{M}_3)^4 + \beta_x \text{tr}(\Gamma^T \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \Gamma \circ \Gamma^T \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \Gamma)] \\ & + \frac{4}{N^2} \text{tr}^2(\mathbf{T}_n \mathbf{M}_3)^2 + o_p(1), \\ & \sum_{j=1}^N E_{j-1} \{[(E_j - E_{j-1})\text{tr}(\mathbf{B}_n \mathbf{M}_1)][(E_j - E_{j-1})\text{tr}(\mathbf{B}_n \mathbf{M}_2)]\} \end{aligned} \quad (2.83)$$

$$\begin{aligned}
 &= \frac{1}{N} [2\text{tr}(\mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n \mathbf{M}_2) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma})], \\
 &\quad \sum_{j=1}^N \mathbb{E}_{j-1} \{[(\mathbb{E}_j - \mathbb{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_1)] [(\mathbb{E}_j - \mathbb{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_3)^2]\} \quad (2.84)
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{2}{N^2} \text{tr}(\mathbf{T}_n \mathbf{M}_3) [2\text{tr}(\mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n \mathbf{M}_3) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{\Gamma})] \\
 &\quad + \frac{2}{N} [2\text{tr}(\mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \mathbf{\Gamma})] + o_p(1), \\
 &\quad \sum_{j=1}^N \mathbb{E}_{j-1} \{[(\mathbb{E}_j - \mathbb{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_3)^2] [(\mathbb{E}_j - \mathbb{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_4)^2]\} \quad (2.85)
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{4}{N^3} \text{tr}(\mathbf{T}_n \mathbf{M}_3) \text{tr}(\mathbf{T}_n \mathbf{M}_4) [2\text{tr}(\mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_4) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_4 \mathbf{\Gamma})] \\
 &\quad + \frac{4}{N^2} \text{tr}(\mathbf{T}_n \mathbf{M}_4) [2\text{tr}(\mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_4) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_4 \mathbf{\Gamma})] \\
 &\quad + \frac{4}{N^2} \text{tr}(\mathbf{T}_n \mathbf{M}_3) [2\text{tr}(\mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_4 \mathbf{T}_n \mathbf{M}_4) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_4 \mathbf{T}_n \mathbf{M}_4 \mathbf{\Gamma})] \\
 &\quad + \frac{4}{N} [2\text{tr}(\mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_4 \mathbf{T}_n \mathbf{M}_4) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_3 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_4 \mathbf{T}_n \mathbf{M}_4 \mathbf{\Gamma})] \\
 &\quad + \frac{4}{N^2} \text{tr}^2(\mathbf{T}_n \mathbf{M}_3 \mathbf{T}_n \mathbf{M}_4) + o_p(1).
 \end{aligned}$$

引理2.5.1的证明与Zheng et al. (2020)^[24]的补充材料中引理1的证明类似。此外, 根据Bai和Silverstein (2004)^[55]中的等式(1.15), 可得

$$\mathbb{E}[(\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 - \text{tr} \mathbf{A})(\mathbf{x}_1^T \mathbf{B} \mathbf{x}_1 - \text{tr} \mathbf{B})] = \beta_x \sum_{i=1}^n a_{ii} b_{ii} + \text{tr} \mathbf{A} \mathbf{B}^T + \text{tr} \mathbf{A} \mathbf{B}, \quad (2.86)$$

其中 $\mathbf{A} = (a_{ij})_{i,j=1}^n$ 和 $\mathbf{B} = (b_{ij})_{i,j=1}^n$ 为 $n \times n$ 矩阵。由等式(2.86)可得

$$\mathbb{E} \text{tr}(\mathbf{B}_n \mathbf{M}_1) = \text{tr}(\mathbf{T}_n \mathbf{M}_1), \quad (2.87)$$

$$\begin{aligned}
 \mathbb{E} \text{tr}(\mathbf{B}_n \mathbf{M}_3)^2 &= \frac{1}{N} \text{tr}^2(\mathbf{T}_n \mathbf{M}_3) + \text{tr}(\mathbf{T}_n \mathbf{M}_3)^2 \\
 &\quad + \frac{1}{N} [\text{tr}(\mathbf{T}_n \mathbf{M}_3)^2 + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_3 \mathbf{\Gamma})]. \quad (2.88)
 \end{aligned}$$

在给出引理2.5.2之前, 需要做一些准备工作。令 $\nu = \Im(z)$, 在下面的分析中, 假设 $\nu > 0$ 。模仿Bai和Silverstein (2004)^[55]中的符号, 令 $\mathbf{D}(z) = \mathbf{B}_n - z\mathbf{I}$, $\mathbf{D}_j(z) =$

$$\mathbf{D}(z) - \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T,$$

$$\begin{aligned} \varepsilon_j(z) &= \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j - \frac{1}{N} \operatorname{tr} \mathbf{T}_n \mathbf{D}_j^{-1}(z), \\ \delta_j(z) &= \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-2}(z) \boldsymbol{\gamma}_j - \frac{1}{N} \operatorname{tr} \mathbf{T}_n \mathbf{D}_j^{-2}(z) = \frac{d}{dz} \varepsilon_j(z), \\ \beta_j(z) &= \frac{1}{1 + \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j}, \\ \bar{\beta}_j(z) &= \frac{1}{1 + N^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_j^{-1}(z)}, \\ b_n(z) &= \frac{1}{1 + N^{-1} \operatorname{Etr} \mathbf{T}_n \mathbf{D}_1^{-1}(z)}. \end{aligned}$$

因此,

$$\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z) = -\mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \beta_j(z). \quad (2.89)$$

假设 $\mathbf{A} = (a_{ij})_{i,j=1}^n$ 是谱模有界的 $n \times n$ 矩阵, 由(2.86)可得

$$\mathbb{E}[(N^{-1} \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 - N^{-1} \operatorname{tr} \mathbf{A})^2] = \frac{1}{N^2} [\beta_x \sum_{i=1}^n a_{ii}^2 + \operatorname{tr} \mathbf{A} \mathbf{A}^T + \operatorname{tr} \mathbf{A}^2],$$

这意味着

$$N^{-1} \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 = N^{-1} \operatorname{tr} \mathbf{A} + o_p(1). \quad (2.90)$$

引理2.5.2. 若 \mathbf{M}_1 , \mathbf{M}_2 , \mathbf{M}_3 和 \mathbf{M}_4 是谱模有界的 $n \times n$ 矩阵, 在假设 a - b - c - d 下, 当 \mathbf{B}_n 可逆时,

$$N^{-1} \operatorname{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1) = \frac{1}{1 - c_n} N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) + o_p(1), \quad (2.91)$$

$$\begin{aligned} N^{-1} \operatorname{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2) &= \frac{1}{(1 - c_n)^2} N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) \\ &\quad + \frac{1}{(1 - c_n)^3} N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) + o_p(1), \end{aligned} \quad (2.92)$$

$$\begin{aligned} &N^{-1} \operatorname{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{M}_3) \\ &= \frac{1}{(1 - c_n)^3} N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3) \\ &\quad + \frac{1}{(1 - c_n)^4} N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \operatorname{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3) \end{aligned} \quad (2.93)$$

$$\begin{aligned}
 & + \frac{1}{(1-c_n)^4} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_3) \\
 & + \frac{1}{(1-c_n)^4} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) \\
 & + \frac{2}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) + o_p(1),
 \end{aligned}$$

和

$$\begin{aligned}
 & N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{M}_3 \mathbf{B}_n^{-1} \mathbf{M}_4) \tag{2.94} \\
 = & \frac{1}{(1-c_n)^4} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{1}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{1}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_3 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{1}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{1}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_4) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3) \\
 & + \frac{1}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{1}{(1-c_n)^5} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_4) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3) \\
 & + \frac{1}{(1-c_n)^6} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{1}{(1-c_n)^6} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_4) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_3) \\
 & + \frac{2}{(1-c_n)^6} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{2}{(1-c_n)^6} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_4) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_3) \\
 & + \frac{2}{(1-c_n)^6} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_4) \\
 & + \frac{2}{(1-c_n)^6} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_4) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) \\
 & + \frac{5}{(1-c_n)^7} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_4) + o_p(1).
 \end{aligned}$$

证明. 根据Bai和Silverstein (2004)^[55]中的(4.13),

$$\mathbf{D}^{-1}(z) = -(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} + b_n(z)\mathbf{A}(z) + \mathbf{B}(z) + \mathbf{C}(z), \tag{2.95}$$

其中

$$\begin{aligned}
 \mathbf{A}(z) &= \sum_{j=1}^N (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} (\boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T - N^{-1}\mathbf{T}_n) \mathbf{D}_j^{-1}(z), \\
 \mathbf{B}(z) &= \sum_{j=1}^N (\beta_j(z) - b_n(z)) (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z), \\
 \mathbf{C}(z) &= N^{-1} b_n(z) (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \sum_{j=1}^N (\mathbf{D}_j^{-1}(z) - \mathbf{D}^{-1}(z)) \\
 &= N^{-1} b_n(z) (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \sum_{j=1}^N \beta_j(z) \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z).
 \end{aligned}$$

在假设a-b-c-d下, $\|(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}\|$ 和 $\|\mathbf{D}_1^{-1}(z)\|$ 是几乎处处有界的。若 \mathbf{M} 是谱模有界的 $n \times n$ 非随机矩阵, 类似于Bai和Silverstein (2004)^[55]中的(4.15)和(4.16), 可得

$$\begin{aligned}
 N^{-1} \text{tr} \mathbf{A}(z) \mathbf{M} &= o_p(1), \\
 N^{-1} \text{tr} \mathbf{B}(z) \mathbf{M} &= o_p(1), \\
 N^{-1} \text{tr} \mathbf{C}(z) \mathbf{M} &= o_p(1),
 \end{aligned}$$

因此

$$N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 = -N^{-1} \text{tr} (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 + o_p(1).$$

当 $z = 0$ 时, 根据(2.56)和(2.47), 可得 $b_n(z) = 1 - c_n + o(1)$ 。因此

$$N^{-1} \text{tr} (\mathbf{B}_n^{-1} \mathbf{M}_1) = \frac{1}{1 - c_n} N^{-1} \text{tr} (\mathbf{T}_n^{-1} \mathbf{M}_1) + o_p(1),$$

这样就得到了(2.91)。接下来, 证明(2.92), 由(2.89), (2.90)和(2.95)可得

$$\begin{aligned}
 &N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \\
 &= -N^{-1} \text{tr} (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 + b_n(z) N^{-1} \text{tr} \mathbf{A}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 + o_p(1) \\
 &= -N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \\
 &\quad + b_n(z) N^{-1} \sum_{j=1}^N \text{tr} (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 (\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z)) \mathbf{M}_2 + o_p(1) \\
 &= -N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1
 \end{aligned}$$

$$\begin{aligned}
 & -b_n(z)N^{-1} \sum_{j=1}^N \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j + o_p(1) \\
 = & -N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \\
 & -b_n^2(z)N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n + o_p(1).
 \end{aligned}$$

因此, 当 $z = 0$ 时,

$$\begin{aligned}
 N^{-1} \text{tr} \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 &= (1 - c_n)^{-1} N^{-1} \text{tr} \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{T}_n^{-1} \mathbf{M}_1 \\
 &+ (1 - c_n) N^{-1} \text{tr} \mathbf{B}_n^{-1} \mathbf{M}_2 N^{-1} \text{tr} \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{T}_n + o_p(1).
 \end{aligned}$$

令 $\mathbf{M}_2 = \mathbf{T}_n$, 根据(2.91), 可得

$$N^{-1} \text{tr} \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{T}_n = \frac{1}{(1 - c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} \mathbf{M}_1 + o_p(1).$$

由此可得

$$\begin{aligned}
 N^{-1} \text{tr} (\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2) &= \frac{1}{(1 - c_n)^2} N^{-1} \text{tr} (\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) \\
 &+ \frac{1}{(1 - c_n)^3} N^{-1} \text{tr} (\mathbf{T}_n^{-1} \mathbf{M}_1) N^{-1} \text{tr} (\mathbf{T}_n^{-1} \mathbf{M}_2) + o_p(1),
 \end{aligned}$$

这样就得到了(2.92)。接下来, 证明(2.93)。由(2.89), (2.90)和(2.95)可得

$$\begin{aligned}
 & N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \\
 = & -N^{-1} \text{tr} (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \\
 & + b_n(z) N^{-1} \text{tr} \mathbf{A}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 + o_p(1) \\
 = & -N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \\
 & + b_n(z) N^{-1} \sum_{j=1}^N \beta_j^2(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & \quad \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \\
 & - b_n(z) N^{-1} \sum_{j=1}^N \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & \quad \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \\
 & - b_n(z) N^{-1} \sum_{j=1}^N \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j
 \end{aligned}$$

$$\begin{aligned}
 & \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j + o_p(1) \\
 = & -N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \\
 & + b_n^3(z) N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{T}_n \\
 & \quad \times N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \\
 & - b_n^2(z) N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n \\
 & \quad \times N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \\
 & - b_n^2(z) N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{T}_n \\
 & \quad \times N^{-1} \operatorname{tr} \mathbf{D}^{-1}(z) \mathbf{M}_3 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n + o_p(1).
 \end{aligned}$$

因此, 当 $z = 0$ 时,

$$\begin{aligned}
 & N^{-1} \operatorname{tr} (\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{M}_3) \\
 = & (1 - c_n)^{-1} N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{M}_3 \mathbf{T}_n^{-1} \mathbf{M}_1 \\
 & - (1 - c_n)^2 N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{T}_n N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{T}_n N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_3 \\
 & + (1 - c_n) N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{T}_n N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{M}_3 \\
 & + (1 - c_n) N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{T}_n N^{-1} \operatorname{tr} \mathbf{B}_n^{-1} \mathbf{M}_3 + o_p(1).
 \end{aligned}$$

令 $\mathbf{M}_3 = \mathbf{T}_n$, 由(2.91)和(2.92)可得

$$\begin{aligned}
 N^{-1} \operatorname{tr} (\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{T}_n) & = \frac{1}{(1 - c_n)^4} N^{-1} \operatorname{tr} \mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \\
 & + \frac{2}{(1 - c_n)^5} N^{-1} \operatorname{tr} \mathbf{T}_n^{-1} \mathbf{M}_1 N^{-1} \operatorname{tr} \mathbf{T}_n^{-1} \mathbf{M}_2 + o_p(1).
 \end{aligned}$$

由此可得(2.93)。最后, 证明(2.94)。再根据(2.89), (2.90)和(2.95), 可得

$$\begin{aligned}
 & N^{-1} \operatorname{tr} (\mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4) \\
 = & -N^{-1} \operatorname{tr} ((z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4) \\
 & + b_n(z) N^{-1} \operatorname{tr} (\mathbf{A}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4) + o_p(1) \\
 = & -N^{-1} \operatorname{tr} (\mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1) \\
 & - b_n(z) N^{-1} \sum_{j=1}^N \beta_j^3(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j
 \end{aligned}$$

$$\begin{aligned}
 & \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \\
 & + b_n(z) N^{-1} \sum_{j=1}^N \beta_j^2(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & \quad \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \\
 & + b_n(z) N^{-1} \sum_{j=1}^N \beta_j^2(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_4 \\
 & \quad \times (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & - b_n(z) N^{-1} \sum_{j=1}^N \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \mathbf{M}_4 \\
 & \quad \times (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & + b_n(z) N^{-1} \sum_{j=1}^N \beta_j^2(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & \quad \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \\
 & - b_n(z) N^{-1} \sum_{j=1}^N \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & \quad \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j \\
 & - b_n(z) N^{-1} \sum_{j=1}^N \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_1 \mathbf{D}_j^{-1}(z) \mathbf{M}_2 \mathbf{D}_j^{-1}(z) \mathbf{M}_3 \mathbf{D}_j^{-1}(z) \boldsymbol{\gamma}_j \\
 & \quad \times \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \boldsymbol{\gamma}_j + o_p(1) \\
 = & -N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{M}_1) \\
 & -b_n^4(z) N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{T}_n \\
 & \quad \times N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{T}_n N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \\
 & + b_n^3(z) N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{T}_n \\
 & \quad \times N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \\
 & + b_n^3(z) N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \\
 & \quad \times \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{T}_n N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \\
 & - b_n^2(z) N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_2 \mathbf{D}^{-1}(z) \mathbf{M}_3 \mathbf{D}^{-1}(z) \mathbf{M}_4 (z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} \mathbf{T}_n \\
 & \quad \times N^{-1} \text{tr} \mathbf{D}^{-1}(z) \mathbf{M}_1 \mathbf{D}^{-1}(z) \mathbf{T}_n
 \end{aligned}$$

$$\begin{aligned}
 & +b_n^3(z)N^{-1}\text{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\text{tr}\mathbf{D}^{-1}(z) \\
 & \quad \times\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\text{tr}\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n \\
 & -b_n^2(z)N^{-1}\text{tr}\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n \\
 & \quad \times N^{-1}\text{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_n \\
 & -b_n^2(z)N^{-1}\text{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{T}_n \\
 & \quad \times N^{-1}\text{tr}\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n + o_p(1).
 \end{aligned}$$

因此, 当 $z = 0$ 时,

$$\begin{aligned}
 & N^{-1}\text{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4) \\
 = & (1-c_n)^{-1}N^{-1}\text{tr}(\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4\mathbf{T}_n^{-1}\mathbf{M}_1) \\
 & +(1-c_n)^3N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_4 \\
 & -(1-c_n)^2N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4 \\
 & -(1-c_n)^2N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_4 \\
 & +(1-c_n)N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4 \\
 & -(1-c_n)^2N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_4 \\
 & +(1-c_n)N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4 \\
 & +(1-c_n)N^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\text{tr}\mathbf{B}_n^{-1}\mathbf{M}_4.
 \end{aligned}$$

令 $\mathbf{M}_4 = \mathbf{T}_n$, 根据(2.91), (2.92)和(2.93), 可得

$$\begin{aligned}
 & N^{-1}\text{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_n) \\
 = & \frac{1}{(1-c_n)^5}N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3 \\
 & +\frac{1}{(1-c_n)^6}N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_2N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_3 \\
 & +\frac{2}{(1-c_n)^6}N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_1N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3 \\
 & +\frac{2}{(1-c_n)^6}N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_3N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2 \\
 & +\frac{5}{(1-c_n)^7}N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_1N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_2N^{-1}\text{tr}\mathbf{T}_n^{-1}\mathbf{M}_3.
 \end{aligned}$$

由此可得(2.94)。至此, 引理2.5.2得证。 \square

引理2.5.3. 令 N' 为满足 $c'_n = n/N' \rightarrow c' > 0$ 的整数, \mathbf{M}_1 , \mathbf{M}_2 和 \mathbf{M}_3 是谱模有界的 $n \times n$ 矩阵. 在假设 $a-b-c-d$ 下, 当 \mathbf{B}_n 可逆时,

$$\begin{aligned} \frac{1}{N'} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n \mathbf{M}_2) &= \frac{1}{N'(1-c_n)} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n \mathbf{M}_2) \\ &\quad - \frac{1}{N'N(1-c_n)} \text{tr} \mathbf{M}_1 \text{tr} \mathbf{M}_2 + o_p(1) \end{aligned} \quad (2.96)$$

和

$$\begin{aligned} \frac{1}{N'} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n \mathbf{M}_3) &= \frac{1}{N'(1-c_n)^2} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_3) \\ &\quad + \frac{1}{N'N(1-c_n)^3} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_3) \\ &\quad - \frac{1}{N'N(1-c_n)^2} \text{tr}(\mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) \text{tr} \mathbf{M}_3 \\ &\quad - \frac{1}{N'N(1-c_n)^2} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{M}_3) \text{tr} \mathbf{M}_2 \\ &\quad - \frac{1}{N'N^2(1-c_n)^3} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) \text{tr} \mathbf{M}_2 \text{tr} \mathbf{M}_3 + o_p(1). \end{aligned}$$

证明. 根据 γ_j 的定义,

$$\mathbf{B}_n = N^{-1} \mathbf{\Gamma} \mathbf{X}_n \mathbf{X}_n^T \mathbf{\Gamma}^T = \sum_{j=1}^N \gamma_j \gamma_j^T,$$

令 $\mathbf{B}_{nj} = \mathbf{B}_n - \gamma_j \gamma_j^T$ 和 $\eta_j = (1 + \gamma_j^T \mathbf{B}_{nj}^{-1} \gamma_j)^{-1}$, 由(2.89), (2.90)和(2.91)可得

$$\begin{aligned} \frac{1}{N'} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n \mathbf{M}_2) &= \frac{1}{N'} \sum_{j=1}^N \gamma_j^T \mathbf{M}_2 \mathbf{B}_n^{-1} \mathbf{M}_1 \gamma_j \\ &= \frac{1}{N'} \sum_{j=1}^N \gamma_j^T \mathbf{M}_2 (\mathbf{B}_{nj}^{-1} - \eta_j \mathbf{B}_{nj}^{-1} \gamma_j \gamma_j^T \mathbf{B}_{nj}^{-1}) \mathbf{M}_1 \gamma_j \\ &= \frac{1}{N'} \sum_{j=1}^N [\gamma_j^T \mathbf{M}_2 \mathbf{B}_{nj}^{-1} \mathbf{M}_1 \gamma_j - \eta_j \gamma_j^T \mathbf{M}_2 \mathbf{B}_{nj}^{-1} \gamma_j \gamma_j^T \mathbf{B}_{nj}^{-1} \mathbf{M}_1 \gamma_j] \\ &= \frac{1}{N'} \left[\text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{T}_n \mathbf{M}_2) - \frac{\eta_1}{N} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{T}_n) \text{tr}(\mathbf{B}_n^{-1} \mathbf{T}_n \mathbf{M}_2) \right] + o_p(1) \\ &= \frac{1}{N'(1-c_n)} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n \mathbf{M}_2) - \frac{1}{N'N(1-c_n)} \text{tr} \mathbf{M}_1 \text{tr} \mathbf{M}_2 + o_p(1). \end{aligned}$$

类似地, 根据(2.89), (2.90), (2.91)和(2.92), 可得

$$\begin{aligned}
 & \frac{1}{N'} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{B}_n \mathbf{M}_3) = \frac{1}{N'} \sum_{j=1}^N \boldsymbol{\gamma}_j^T \mathbf{M}_3 \mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \boldsymbol{\gamma}_j \\
 &= \frac{1}{N'} \sum_{j=1}^N \boldsymbol{\gamma}_j^T \mathbf{M}_3 (\mathbf{B}_{nj}^{-1} - \eta_j \mathbf{B}_{nj}^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{B}_{nj}^{-1}) \mathbf{M}_1 (\mathbf{B}_{nj}^{-1} - \eta_j \mathbf{B}_{nj}^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{B}_{nj}^{-1}) \mathbf{M}_2 \boldsymbol{\gamma}_j \\
 &= \frac{1}{N'} \sum_{j=1}^N (\boldsymbol{\gamma}_j^T \mathbf{M}_3 \mathbf{B}_{nj}^{-1} \mathbf{M}_1 \mathbf{B}_{nj}^{-1} \mathbf{M}_2 \boldsymbol{\gamma}_j - \eta_j \boldsymbol{\gamma}_j^T \mathbf{M}_3 \mathbf{B}_{nj}^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{B}_{nj}^{-1} \mathbf{M}_1 \mathbf{B}_{nj}^{-1} \mathbf{M}_2 \boldsymbol{\gamma}_j \\
 &\quad - \eta_j \boldsymbol{\gamma}_j^T \mathbf{M}_3 \mathbf{B}_{nj}^{-1} \mathbf{M}_1 \mathbf{B}_{nj}^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{B}_{nj}^{-1} \mathbf{M}_2 \boldsymbol{\gamma}_j + \eta_j^2 \boldsymbol{\gamma}_j^T \mathbf{M}_3 \mathbf{B}_{nj}^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{B}_{nj}^{-1} \mathbf{M}_1 \mathbf{B}_{nj}^{-1} \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T \mathbf{B}_{nj}^{-1} \mathbf{M}_2 \boldsymbol{\gamma}_j) \\
 &= \frac{N}{N'} \left[N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_3) - \eta_1 N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{T}_n \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{T}_n) \right. \\
 &\quad \left. - \eta_1 N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{T}_n \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{T}_n) \right. \\
 &\quad \left. + \eta_1^2 N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{T}_n \mathbf{M}_3) N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_1 \mathbf{B}_n^{-1} \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{B}_n^{-1} \mathbf{M}_2 \mathbf{T}_n) \right] + o_p(1) \\
 &= \frac{1}{N'(1-c_n)^2} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_3) + \frac{1}{N'N(1-c_n)^3} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_3) \\
 &\quad - \frac{1}{N'N(1-c_n)^2} \text{tr}(\mathbf{M}_1 \mathbf{T}_n^{-1} \mathbf{M}_2) \text{tr} \mathbf{M}_3 - \frac{1}{N'N(1-c_n)^2} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1 \mathbf{M}_3) \text{tr} \mathbf{M}_2 \\
 &\quad - \frac{1}{N'N^2(1-c_n)^3} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) \text{tr} \mathbf{M}_2 \text{tr} \mathbf{M}_3 + o_p(1).
 \end{aligned}$$

□

引理2.5.4. 若 \mathbf{M}_1 和 \mathbf{M}_2 是对称的、非负定的且谱模有界的矩阵。在假设 $a-b-c-d$ 下, 当 \mathbf{B}_n 可逆时,

$$\sum_{j=1}^N E_{j-1} \{ [(\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\mathbf{B}_n^{-1})] [(\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_1)] \} \quad (2.97)$$

$$= -\frac{2}{(1-c_n)} N^{-1} \text{tr} \mathbf{M}_1 - \frac{\beta_x}{(1-c_n)} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \boldsymbol{\Gamma}^T \mathbf{M}_1 \boldsymbol{\Gamma}) + o_p(1),$$

$$\sum_{j=1}^N E_{j-1} \{ [(\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\mathbf{B}_n^{-2})] [(\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_1)] \} \quad (2.98)$$

$$\begin{aligned}
 &= -\frac{4}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) - \frac{4}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr} \mathbf{M}_1 \\
 &\quad - \frac{2\beta_x}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \boldsymbol{\Gamma}^T \mathbf{M}_1 \boldsymbol{\Gamma}) \\
 &\quad - \frac{2\beta_x}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \boldsymbol{\Gamma}^T \mathbf{M}_1 \boldsymbol{\Gamma}) + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 & \sum_{j=1}^N E_{j-1} \{[(E_j - E_{j-1})\text{tr}(\mathbf{B}_n^{-1})][(E_j - E_{j-1})\text{tr}(\mathbf{B}_n \mathbf{M}_2)^2]\} \quad (2.99) \\
 &= -\frac{4}{(1-c_n)} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2^2) - \frac{2\beta_x}{(1-c_n)} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \Gamma^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \Gamma) \\
 & \quad - \frac{2\beta_x}{(1-c_n)} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \Gamma^T \mathbf{M}_2 \Gamma) + o_p(1), \\
 & \sum_{j=1}^N E_{j-1} \{[(E_j - E_{j-1})\text{tr}(\mathbf{B}_n^{-2})][(E_j - E_{j-1})\text{tr}(\mathbf{B}_n \mathbf{M}_2)^2]\} \quad (2.100)
 \end{aligned}$$

$$\begin{aligned}
 &= -\frac{8}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2) - \frac{8}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2^2) \\
 & \quad + \frac{4}{(1-c_n)^2} (N^{-1} \text{tr} \mathbf{M}_2)^2 - \frac{4\beta_x}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \Gamma^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \Gamma) \\
 & \quad - \frac{4\beta_x}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \Gamma^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \Gamma) \\
 & \quad - \frac{4\beta_x}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \Gamma^T \mathbf{M}_2 \Gamma) \\
 & \quad - \frac{4\beta_x}{(1-c_n)^3} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \Gamma^T \mathbf{M}_2 \Gamma) + o_p(1).
 \end{aligned}$$

证明. 由(2.89)可得

$$\begin{aligned}
 (E_j - E_{j-1})\text{tr} \mathbf{D}^{-1}(z) &= \text{tr}[(E_j - E_{j-1})(\mathbf{D}_j^{-1}(z) - \beta_j(z)\mathbf{D}_j^{-1}(z)\gamma_j\gamma_j^T\mathbf{D}_j^{-1}(z))] \\
 &= -(E_j - E_{j-1})\beta_j(z)\gamma_j^T\mathbf{D}_j^{-2}(z)\gamma_j.
 \end{aligned}$$

因为

$$\begin{aligned}
 \beta_j(z) &= \bar{\beta}_j(z) - \beta_j(z)\bar{\beta}_j(z)\varepsilon_j(z) \\
 &= \bar{\beta}_j(z) - \bar{\beta}_j^2(z)\varepsilon_j(z) + \bar{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z),
 \end{aligned}$$

由此可得

$$\begin{aligned}
 & -(E_j - E_{j-1})\beta_j(z)\gamma_j^T\mathbf{D}_j^{-2}(z)\gamma_j \\
 &= -(E_j - E_{j-1})(\bar{\beta}_j(z)\delta_j(z) - \bar{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z) \\
 & \quad - \bar{\beta}_j^2(z)\varepsilon_j(z)N^{-1}\text{tr}(\mathbf{D}_j^{-2}(z)\mathbf{T}_n) + \bar{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z)\gamma_j^T\mathbf{D}_j^{-2}(z)\gamma_j) \\
 &= -E_j(\bar{\beta}_j(z)\delta_j(z) - \bar{\beta}_j^2(z)\varepsilon_j(z)N^{-1}\text{tr}(\mathbf{D}_j^{-2}(z)\mathbf{T}_n)) \\
 & \quad + (E_j - E_{j-1})\bar{\beta}_j^2(z)(\varepsilon_j(z)\delta_j(z) - \beta_j(z)\gamma_j^T\mathbf{D}_j^{-2}(z)\gamma_j\varepsilon_j^2(z)).
 \end{aligned}$$

根据Bai和Silverstein (2004)^[55]中569页的讨论, 可得

$$\begin{aligned} \mathbb{E} \left| \sum_{j=1}^N (\mathbf{E}_j - \mathbf{E}_{j-1}) \bar{\beta}_j^2(z) \varepsilon_j(z) \delta_j(z) \right|^2 &= \sum_{j=1}^N \mathbb{E} |(\mathbf{E}_j - \mathbf{E}_{j-1}) \bar{\beta}_j^2(z) \varepsilon_j(z) \delta_j(z)|^2 \\ &\leq 4 \sum_{j=1}^N \mathbb{E} |\bar{\beta}_j^2(z) \varepsilon_j(z) \delta_j(z)|^2 = o(1). \end{aligned}$$

这意味着

$$\sum_{j=1}^N (\mathbf{E}_j - \mathbf{E}_{j-1}) \bar{\beta}_j^2(z) \varepsilon_j(z) \delta_j(z) \xrightarrow{p} 0.$$

类似地,

$$\sum_{j=1}^N (\mathbf{E}_j - \mathbf{E}_{j-1}) \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-2}(z) \boldsymbol{\gamma}_j \varepsilon_j^2(z) \xrightarrow{p} 0.$$

因此

$$\begin{aligned} \text{tr}[\mathbf{D}^{-1}(z) - \mathbf{E} \mathbf{D}^{-1}(z)] &= \sum_{j=1}^N (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr} \mathbf{D}^{-1}(z) \\ &= - \sum_{j=1}^N (\mathbf{E}_j - \mathbf{E}_{j-1}) \beta_j(z) \boldsymbol{\gamma}_j^T \mathbf{D}_j^{-2}(z) \boldsymbol{\gamma}_j \\ &= - \sum_{j=1}^N \mathbb{E}_j (\bar{\beta}_j(z) \delta_j(z) - \bar{\beta}_j^2(z) \varepsilon_j(z) N^{-1} \text{tr}(\mathbf{D}_j^{-2}(z) \mathbf{T}_n)) + o_p(1). \end{aligned}$$

又因为

$$\begin{aligned} -\mathbb{E}_j \frac{d}{dz} \bar{\beta}_j(z) \varepsilon_j(z) &= -\mathbb{E}_j (\bar{\beta}_j(z) \delta_j(z) - \bar{\beta}_j^2(z) \varepsilon_j(z) N^{-1} \text{tr}(\mathbf{D}_j^{-2}(z) \mathbf{T}_n)), \\ (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_1) &= \boldsymbol{\gamma}_j^T \mathbf{M}_1 \boldsymbol{\gamma}_j - N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_1), \end{aligned}$$

根据鞅差序列的中心极限定理, 考虑下面的和

$$\begin{aligned} &\sum_{j=1}^N \mathbb{E}_{j-1} [-\bar{\beta}_j(z) \varepsilon_j(z) (\boldsymbol{\gamma}_j^T \mathbf{M}_1 \boldsymbol{\gamma}_j - N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_1))] \\ &= -\frac{1}{N^2} \sum_{j=1}^N \mathbb{E}_{j-1} [\bar{\beta}_j(z) (\mathbf{x}_j^T \boldsymbol{\Gamma}^T \mathbf{D}_j^{-1}(z) \boldsymbol{\Gamma} \mathbf{x}_j - \text{tr}(\mathbf{T}_n \mathbf{D}_j^{-1}(z))) (\mathbf{x}_j^T \boldsymbol{\Gamma}^T \mathbf{M}_1 \boldsymbol{\Gamma} \mathbf{x}_j - \text{tr}(\mathbf{T}_n \mathbf{M}_1))] \\ &= -\frac{1}{N^2} \sum_{j=1}^N \mathbb{E}_{j-1} [\bar{\beta}_j(z) (2\text{tr}(\mathbf{D}_j^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) + \beta_x \text{tr}(\boldsymbol{\Gamma}^T \mathbf{D}_j^{-1}(z) \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{M}_1 \boldsymbol{\Gamma}))] \end{aligned}$$

$$\begin{aligned}
 &= -\frac{1}{N^2} \sum_{j=1}^N E_{j-1} \left[b_n(z) (2\text{tr}(\mathbf{D}_j^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma})) \right] + o_p(1) \\
 &= -\frac{1}{N^2} \sum_{j=1}^N E_{j-1} \left[b_n(z) (2\text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{D}^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma})) \right] + o_p(1),
 \end{aligned}$$

其中最后一个等式是由Bai和Silverstein (2004)^[55]中的(2.3)式得到的。根据控制收敛定理，只需要计算

$$b_n(z) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) \quad \text{和} \quad b_n(z) N^{-1} \text{tr}(\mathbf{\Gamma}^T \mathbf{D}^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma})$$

的一阶导数和二阶导数的极限。由于

$$b'_n(z) = -b_n^2(z) N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n) + o_p(1),$$

由引理2.5.2, 当 $z \rightarrow 0$ 时,

$$\begin{aligned}
 &\frac{d}{dz} [b_n(z) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n)] \\
 &= -b_n^2(z) N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) \\
 &\quad + b_n(z) N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) + o_p(1) \\
 &= \frac{1}{(1-c_n)} N^{-1} \text{tr} \mathbf{M}_1 + o_p(1)
 \end{aligned}$$

和

$$\begin{aligned}
 &\frac{d^2}{dz^2} [b_n(z) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n)] \\
 &= 2b_n^3(z) (N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n))^2 N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) \\
 &\quad - 2b_n^2(z) N^{-1} \text{tr}(\mathbf{D}^{-3}(z) \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) \\
 &\quad - 2b_n^2(z) N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) \\
 &\quad + 2b_n(z) \text{tr}(\mathbf{D}^{-3}(z) \mathbf{T}_n \mathbf{M}_1 \mathbf{T}_n) + o_p(1) \\
 &= \frac{2}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) + \frac{2}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr} \mathbf{M}_1 + o_p(1).
 \end{aligned}$$

类似地,

$$\begin{aligned}
 & \frac{d}{dz} [b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma})] \\
 &= -b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \\
 & \quad + b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-2}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) + o_p(1) \\
 &= \frac{1}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) + o_p(1)
 \end{aligned}$$

和

$$\begin{aligned}
 & \frac{d^2}{dz^2} [b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma})] \\
 &= 2b_n^3(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n))^2N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \\
 & \quad - 2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \\
 & \quad - 2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-2}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \\
 & \quad + 2b_n^2(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-3}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \\
 &= \frac{2}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \\
 & \quad + \frac{2}{(1-c_n)^3}N^{-1}\text{tr}\mathbf{T}_n^{-1}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) + o_p(1).
 \end{aligned}$$

因此

$$\begin{aligned}
 & \sum_{j=1}^N \mathbf{E}_{j-1} \{[(\mathbf{E}_j - \mathbf{E}_{j-1})\text{tr}(\mathbf{B}_n^{-1})][(\mathbf{E}_j - \mathbf{E}_{j-1})\text{tr}(\mathbf{B}_n\mathbf{M}_1)]\} \\
 &= -\frac{1}{N} \sum_{j=1}^N \mathbf{E}_{j-1} \left[\frac{2}{(1-c_n)}N^{-1}\text{tr}\mathbf{M}_1 + \frac{\beta_x}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \right] + o_p(1) \\
 &= -\frac{2}{(1-c_n)}N^{-1}\text{tr}\mathbf{M}_1 - \frac{\beta_x}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) + o_p(1),
 \end{aligned}$$

和

$$\begin{aligned}
 & \sum_{j=1}^N \mathbf{E}_{j-1} \{[(\mathbf{E}_j - \mathbf{E}_{j-1})\text{tr}(\mathbf{B}_n^{-2})][(\mathbf{E}_j - \mathbf{E}_{j-1})\text{tr}(\mathbf{B}_n\mathbf{M}_1)]\} \\
 &= -\frac{1}{N} \sum_{j=1}^N \mathbf{E}_{j-1} \left[\frac{4}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1) \right. \\
 & \quad \left. + \frac{4}{(1-c_n)^3}N^{-1}\text{tr}\mathbf{T}_n^{-1}N^{-1}\text{tr}\mathbf{M}_1 + \frac{2\beta_x}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T\mathbf{M}_1\mathbf{\Gamma}) \right]
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{2\beta_x}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma}) \Big] + o_p(1) \\
 = & - \frac{4}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_1) - \frac{4}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr} \mathbf{M}_1 \\
 & - \frac{2\beta_x}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma}) \\
 & - \frac{2\beta_x}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T \mathbf{M}_1 \mathbf{\Gamma}) + o_p(1).
 \end{aligned}$$

关于 $\text{tr}[(\mathbf{B}_n \mathbf{M}_2)^2]$ 的鞅差, 有以下表达式

$$\begin{aligned}
 \text{tr}[(\mathbf{B}_n \mathbf{M}_2)^2] - \text{Etr}[(\mathbf{B}_n \mathbf{M}_2)^2] &= \sum_{j=1}^N (\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}[(\mathbf{B}_n \mathbf{M}_2)^2] \\
 = & \frac{1}{N^2} \sum_{j=1}^N \left[2(\mathbf{x}_j^T \mathbf{\Gamma}^T \mathbf{M}_2 \left(\sum_{i=1}^{j-1} \mathbf{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \mathbf{\Gamma}^T \right) \mathbf{M}_2 \mathbf{\Gamma} \mathbf{x}_j - \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \left(\sum_{i=1}^{j-1} \mathbf{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \mathbf{\Gamma}^T \right) \mathbf{M}_2 \mathbf{\Gamma} \right) \right. \\
 & + 2\text{tr}(\mathbf{T}_n \mathbf{M}_2) (\mathbf{x}_j^T \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma} \mathbf{x}_j - \text{tr}(\mathbf{T}_n \mathbf{M}_2)) \\
 & \left. + 2(N-j) (\mathbf{x}_j^T \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma} \mathbf{x}_j - \text{tr}(\mathbf{T}_n \mathbf{M}_2)^2) \right] + o_p(1).
 \end{aligned}$$

仍然根据鞅差序列的中心极限定理, 考虑下面的和

$$\begin{aligned}
 & \sum_{j=1}^N \mathbf{E}_{j-1} \left[(\mathbf{E}_j - \mathbf{E}_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_2)^2 (-\bar{\beta}_j(z) \varepsilon_j(z)) \right] \tag{2.101} \\
 = & -\frac{1}{N^3} \sum_{j=1}^N \mathbf{E}_{j-1} \left[2\bar{\beta}_j(z) (2\text{tr}(\mathbf{T}_n \mathbf{M}_2 \left(\sum_{i=1}^{j-1} \mathbf{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \mathbf{\Gamma}^T \right) \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z)) \right. \\
 & + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \left(\sum_{i=1}^{j-1} \mathbf{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \mathbf{\Gamma}^T \right) \mathbf{M}_2 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma})) \\
 & + 2\bar{\beta}_j(z) \text{tr}(\mathbf{T}_n \mathbf{M}_2) (2\text{tr}(\mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z)) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma})) \\
 & \left. + 2\bar{\beta}_j(z) (N-j) (2\text{tr}(\mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z)) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma})) \right] \\
 = & -\frac{1}{N^3} \sum_{j=1}^N \mathbf{E}_{j-1} \left[2\bar{\beta}_j(z) (2 \sum_{i=1}^{j-1} \mathbf{x}_i^T \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma} \mathbf{x}_i \right. \\
 & + \beta_x \sum_{i=1}^{j-1} \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma})) \\
 & + 2\bar{\beta}_j(z) \text{tr}(\mathbf{T}_n \mathbf{M}_2) (2\text{tr}(\mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z)) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma})) \\
 & \left. + 2\bar{\beta}_j(z) (N-j) (2\text{tr}(\mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z)) + \beta_x \text{tr}(\mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{D}_j^{-1}(z) \mathbf{\Gamma})) \right].
 \end{aligned}$$

令

$$\begin{aligned} \mathbf{D}_{ij}(z) &= \mathbf{D}(z) - \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^T - \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^T, \\ \beta_{ij}(z) &= \frac{1}{1 + \boldsymbol{\gamma}_i^T \mathbf{D}_{ij}^{-1}(z) \boldsymbol{\gamma}_i}, \\ b_1(z) &= \frac{1}{1 + N^{-1} \text{Etr} \mathbf{T}_n \mathbf{D}_{12}^{-1}(z)}, \end{aligned}$$

则有

$$\begin{aligned} & \sum_{i=1}^{j-1} \mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_j^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \boldsymbol{\Gamma} \mathbf{x}_i \\ &= \sum_{i=1}^{j-1} \mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n (\mathbf{D}_{ji}^{-1}(z) - \beta_{ji}(z) \mathbf{D}_{ji}^{-1}(z) \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i^T \mathbf{D}_{ji}^{-1}(z)) \mathbf{T}_n \mathbf{M}_2 \boldsymbol{\Gamma} \mathbf{x}_i \\ &= \sum_{i=1}^{j-1} (\mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_{ji}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \boldsymbol{\Gamma} \mathbf{x}_i - \beta_{ji}(z) N^{-1} (\mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{D}_{ji}^{-1}(z) \boldsymbol{\Gamma} \mathbf{x}_i)^2), \end{aligned}$$

和

$$\begin{aligned} & \sum_{i=1}^{j-1} \text{tr}(\boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{D}_j^{-1}(z) \boldsymbol{\Gamma}) \\ &= \sum_{i=1}^{j-1} (\text{tr}(\boldsymbol{\Gamma}^T \mathbf{D}_{ji}^{-1}(z) \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma}) \\ & \quad - \beta_{ji}(z) N^{-1} \text{tr}(\boldsymbol{\Gamma}^T \mathbf{D}_{ji}^{-1}(z) \boldsymbol{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{D}_{ji}^{-1}(z) \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma} \mathbf{x}_i \mathbf{x}_i^T \boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma})), \end{aligned}$$

由此可得

$$\begin{aligned} (2.101) &= -\frac{1}{N^2} \sum_{j=1}^N \text{E}_{j-1} \left[2b_n(z) (2(j-1)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) \right. \\ & \quad - 2(j-1)b_n(z) (N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n))^2 \\ & \quad + 2 \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) \\ & \quad + 2(N-j)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) \\ & \quad + \beta_n b_n(z) (2(j-1)N^{-1} \text{tr}(\boldsymbol{\Gamma}^T \mathbf{D}^{-1}(z) \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \boldsymbol{\Gamma}) \\ & \quad + 2 \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\boldsymbol{\Gamma}^T \mathbf{D}^{-1}(z) \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{M}_2 \boldsymbol{\Gamma}) \\ & \quad \left. + 2(N-j)N^{-1} \text{tr}(\boldsymbol{\Gamma}^T \mathbf{D}^{-1}(z) \boldsymbol{\Gamma} \circ \boldsymbol{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \boldsymbol{\Gamma}) \right] + o_p(1) \end{aligned}$$

$$\begin{aligned}
 &= -\frac{1}{N} \sum_{j=1}^N E_{j-1} \left[2b_n(z)(2N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n)) \right. \\
 &\quad - 2(j-1)N^{-1}b_n(z)(N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n))^2 \\
 &\quad + 2N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) \\
 &\quad + \beta_x b_n(z)(2N^{-1} \text{tr}(\mathbf{\Gamma}^T \mathbf{D}^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma})) \\
 &\quad \left. + 2N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{\Gamma}^T \mathbf{D}^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma}) \right] + o_p(1).
 \end{aligned}$$

和上面的讨论类似, 只需要计算

$$\begin{aligned}
 &b_n(z)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n), \quad b_n^2(z)(N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n))^2, \\
 &b_n(z)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n), \quad b_n(z)N^{-1} \text{tr}(\mathbf{\Gamma}^T \mathbf{D}^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma}),
 \end{aligned}$$

和

$$b_n(z)N^{-1} \text{tr}(\mathbf{\Gamma}^T \mathbf{D}^{-1}(z) \mathbf{\Gamma} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma})$$

的一阶导数和二阶导数的极限。由引理2.5.2, 当 $z \rightarrow 0$ 时,

$$\begin{aligned}
 &\frac{d}{dz} [b_n(z)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n)] \\
 &= -b_n^2(z)N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) \\
 &\quad + b_n(z)N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) + o_p(1) \\
 &= \frac{1}{(1-c_n)} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2^2) + o_p(1), \\
 &\frac{d}{dz} [b_n^2(z)(N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n))^2] \\
 &= -2b_n^3(z)N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n) (N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n))^2 \\
 &\quad + 2b_n^2(z)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) + o_p(1) \\
 &= \frac{2}{(1-c_n)} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{M}_2) + o_p(1), \\
 &\frac{d}{dz} [b_n(z)N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n)] \\
 &= -b_n^2(z)N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n) N^{-1} \text{tr}(\mathbf{D}^{-1}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) \\
 &\quad + b_n(z)N^{-1} \text{tr}(\mathbf{D}^{-2}(z) \mathbf{T}_n \mathbf{M}_2 \mathbf{T}_n) + o_p(1) \\
 &= \frac{1}{(1-c_n)} N^{-1} \text{tr}(\mathbf{M}_2) + o_p(1),
 \end{aligned}$$

$$\begin{aligned}
 & \frac{d}{dz} [b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})] \\
 = & -b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 & + b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) + o_p(1) \\
 = & \frac{1}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) + o_p(1)
 \end{aligned}$$

和

$$\begin{aligned}
 & \frac{d}{dz} [b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma})] \\
 = & -b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) \\
 & + b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-2}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) + o_p(1) \\
 = & \frac{1}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) + o_p(1).
 \end{aligned}$$

因此,

$$\begin{aligned}
 & \sum_{j=1}^N E_{j-1} \{ [(E_j - E_{j-1})\text{tr}(\mathbf{B}_n^{-1})][(E_j - E_{j-1})\text{tr}(\mathbf{B}_n\mathbf{M}_2)^2] \} \\
 = & -\frac{4}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n\mathbf{M}_2^2) - \frac{2\beta_x}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 & - \frac{2\beta_x}{(1-c_n)}N^{-1}\text{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) + o_p(1).
 \end{aligned}$$

接下来, 这些项的二阶导数如下

$$\begin{aligned}
 & \frac{d^2}{dz^2} [b_n(z)N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)] \\
 = & 2b_n^3(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n))^2N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & - 2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & - 2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & + 2b_n(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 = & \frac{2}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2) + \frac{2}{(1-c_n)^3}N^{-1}\text{tr}\mathbf{T}_n^{-1}N^{-1}\text{tr}(\mathbf{T}_n\mathbf{M}_2^2) + o_p(1), \\
 & \frac{d^2}{dz^2} [b_n^2(z)(N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n))^2] \\
 = & 6b_n^4(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n))^2(N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n))^2 \\
 & - 4b_n^3(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)(N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n))^2
 \end{aligned}$$

$$\begin{aligned}
 & -8b_n^3(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & +2b_n^2(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n))^2 \\
 & +4b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 = & \frac{2}{(1-c_n)^2}(N^{-1}\text{tr}\mathbf{M}_2)^2 + \frac{4}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\text{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2) \\
 & + \frac{4}{(1-c_n)^3}N^{-1}\text{tr}\mathbf{T}_n^{-1}N^{-1}\text{tr}\mathbf{M}_2N^{-1}\text{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2) + o_p(1), \\
 & \frac{d^2}{dz^2}[b_n(z)N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)] \\
 = & 2b_n^3(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n))^2N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & -2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & -2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 & +2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
 = & \frac{2}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2) + \frac{2}{(1-c_n)^3}N^{-1}\text{tr}\mathbf{T}_n^{-1}N^{-1}\text{tr}\mathbf{M}_2 + o_p(1), \\
 & \frac{d^2}{dz^2}[b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})] \\
 = & 2b_n^3(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n))^2N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 & -2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 & -2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-2}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 & +2b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-3}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 = & \frac{2}{(1-c_n)^2}N^{-1}\text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
 & + \frac{2}{(1-c_n)^3}N^{-1}\text{tr}\mathbf{T}_n^{-1}N^{-1}\text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) + o_p(1)
 \end{aligned}$$

和

$$\begin{aligned}
 & \frac{d^2}{dz^2}[b_n(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma})] \\
 = & 2b_n^3(z)(N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n))^2N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) \\
 & -2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) \\
 & -2b_n^2(z)N^{-1}\text{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-2}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma}) \\
 & +2b_n^2(z)N^{-1}\text{tr}(\mathbf{\Gamma}^T\mathbf{D}^{-3}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^T\mathbf{M}_2\mathbf{\Gamma})
 \end{aligned}$$

$$= \frac{2}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma}) \\ + \frac{2}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma}) + o_p(1).$$

由此可得

$$\sum_{j=1}^N E_{j-1} \{[(E_j - E_{j-1}) \text{tr}(\mathbf{B}_n^{-2})][(E_j - E_{j-1}) \text{tr}(\mathbf{B}_n \mathbf{M}_2)^2]\} \\ = -\frac{8}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2) - \frac{8}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2^2) \\ + \frac{4}{(1-c_n)^2} (N^{-1} \text{tr} \mathbf{M}_2)^2 - \frac{4\beta_x}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma}) \\ - \frac{4\beta_x}{(1-c_n)^3} N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{T}_n \mathbf{M}_2 \mathbf{\Gamma}) \\ - \frac{4\beta_x}{(1-c_n)^2} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma}) \\ - \frac{4\beta_x}{(1-c_n)^3} N^{-1} \text{tr}(\mathbf{T}_n \mathbf{M}_2) N^{-1} \text{tr} \mathbf{T}_n^{-1} N^{-1} \text{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^T \mathbf{M}_2 \mathbf{\Gamma}) + o_p(1).$$

至此，引理2.5.4得证。 □

§2.6 附录

附录包括两个部分：§2.6.1节给出了Cai et al. (2013)^[221]中的条件(C1), (C2)(或(C2*))和(C3)；§2.6.2节验证了数值模拟中考虑的4种设置满足Cai et al. (2013)^[221]中的条件(C1), (C2*)和(C3)。

§2.6.1 假定的条件

在给出Cai et al. (2013)^[221]中的条件(C1), (C2)(或(C2*))和(C3)之前，先介绍一些基本的符号。对于两个实数序列 $\{a_n\}$ 和 $\{b_n\}$ ，记 $a_n \asymp b_n$ ，若存在常数 $C > c > 0$ ，使得 n 充分大时， $c|b_n| \leq |a_n| \leq C|b_n|$ 。令 $N = \max\{N_1, N_2\}$ ， $\mathbf{R}_k = (\rho_{k\ell_1\ell_2})$ 为第 k 个总体的相关矩阵， $k = 1, 2$ 。

(C1). 假设存在一个正常数 α_0 和一个子集 $\Upsilon \subset \{1, 2, \dots, p\}$ ， $\text{card}(\Upsilon) = o(p)$ ，使得对于所有 $\gamma > 0$ ， $\max_{1 \leq \ell_2 \leq p, \ell_2 \notin \Upsilon} s_{\ell_2}(\alpha_0) = o(p^\gamma)$ ， $\ell_2 = 1, \dots, p$ ， s_{ℓ_2} 的定义如下

$$s_{\ell_2} = s_{\ell_2}(\alpha_0) := \text{card}\{\ell_1 : |\rho_{1\ell_1\ell_2}| \geq (\log p)^{-1-\alpha_0} \text{ 或 } |\rho_{2\ell_1\ell_2}| \geq (\log p)^{-1-\alpha_0}\}.$$

此外, 存在一个数列 $\Lambda_{p,r}$ 使得对于某个常数 $0 < r < 1$, $\text{card}(\Lambda(r)) \leq \Lambda_{p,r} = o(p)$, 其中

$$\Lambda(r) = \{1 \leq \ell_1 \leq p : |\rho_{1\ell_1\ell_2}| \geq r \text{ 或 } |\rho_{2\ell_1\ell_2}| \geq r \text{ 对于某个 } \ell_2 \neq \ell_1\}.$$

(C2). 假设 $\log p = o(N^{1/5})$ 和 $N_1 \asymp N_2$. 存在常数 $\eta > 0$ 和 $K > 0$ 使得

$$\begin{aligned} E \exp(\eta(\mathbf{x}_{1\ell} - \mu_{1\ell})^2 / \sigma_{1\ell\ell}) &\leq K, \\ E \exp(\eta(\mathbf{x}_{2\ell} - \mu_{2\ell})^2 / \sigma_{2\ell\ell}) &\leq K, \quad \ell = 1, \dots, p. \end{aligned}$$

此外, 对于某些常数 $\tau_1 > 0$ 和 $\tau_2 > 0$,

$$\min_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{\theta_{1\ell_1\ell_2}}{\sigma_{1\ell_1\ell_1}\sigma_{1\ell_2\ell_2}} \geq \tau_1 \quad \text{和} \quad \min_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{\theta_{2\ell_1\ell_2}}{\sigma_{2\ell_1\ell_1}\sigma_{2\ell_2\ell_2}} \geq \tau_2. \quad (2.102)$$

(C2*). 假设条件(2.102)成立, $N_1 \asymp N_2$, 并且对于某些常数 $\gamma_0, c_1 > 0$, $p \leq c_1 N^{\gamma_0}$. 此外, 对于某些常数 $\epsilon > 0$ 和 $K > 0$, 下列矩条件成立

$$\begin{aligned} E|(\mathbf{x}_{1\ell} - \mu_{1\ell}) / \sigma_{1\ell\ell}^{1/2}|^{4\gamma_0+4+\epsilon} &\leq K, \\ E|(\mathbf{x}_{2\ell} - \mu_{2\ell}) / \sigma_{2\ell\ell}^{1/2}|^{4\gamma_0+4+\epsilon} &\leq K, \quad \ell = 1, \dots, p. \end{aligned}$$

(C3). 对于任意的 $\ell_1, \ell_2, \ell_3, \ell_4 \in \{1, 2, \dots, p\}$ 和某些常数 $\kappa_1, \kappa_2 \geq \frac{1}{3}$,

$$\begin{aligned} E(\mathbf{x}_{1\ell_1} - \mu_{1\ell_1})(\mathbf{x}_{1\ell_2} - \mu_{1\ell_2})(\mathbf{x}_{1\ell_3} - \mu_{1\ell_3})(\mathbf{x}_{1\ell_4} - \mu_{1\ell_4}) \\ = \kappa_1(\sigma_{1\ell_1\ell_2}\sigma_{1\ell_3\ell_4} + \sigma_{1\ell_1\ell_3}\sigma_{1\ell_2\ell_4} + \sigma_{1\ell_1\ell_4}\sigma_{1\ell_2\ell_3}), \\ E(\mathbf{x}_{2\ell_1} - \mu_{2\ell_1})(\mathbf{x}_{2\ell_2} - \mu_{2\ell_2})(\mathbf{x}_{2\ell_3} - \mu_{2\ell_3})(\mathbf{x}_{2\ell_4} - \mu_{2\ell_4}) \\ = \kappa_2(\sigma_{2\ell_1\ell_2}\sigma_{2\ell_3\ell_4} + \sigma_{2\ell_1\ell_3}\sigma_{2\ell_2\ell_4} + \sigma_{2\ell_1\ell_4}\sigma_{2\ell_2\ell_3}). \end{aligned}$$

§2.6.2 验证数值模拟中考虑的模型设置

接下来, 验证数值模拟中的设置1-4满足§2.6.1中的(C1), (C2*)和(C3)。为了便于表示, 首先给出一些符号:

- 令 $\Sigma_{1k} = (\sigma_{1k\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$ 和 $\Sigma_{2k} = (\sigma_{2k\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$ 为设置 $k = 1, 2, 3, 4$ 中的总体协方差矩阵;

- 令 $\mathbf{R}_{1k} = (\rho_{1k\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$ 和 $\mathbf{R}_{2k} = (\rho_{2k\ell_1\ell_2})_{\ell_1, \ell_2=1}^p$ 为设置 $k = 1, 2, 3, 4$ 中的总体相关矩阵;
- 令 $\mathbf{r}_{1k\ell}$ 和 $\mathbf{r}_{2k\ell}$ 分别表示矩阵 $\Sigma_{1k}^{1/2}$ 和 $\Sigma_{2k}^{1/2}$ 的第 ℓ 列;
- 令 $\mathbf{w} = (w_1, \dots, w_p)$ 表示元素 i.i.d. 来自 $N(0, 1)$ 或 $\text{Gamma}(4, 2)$ -2 的向量;
- 在数值模拟中, 在原假设下 $\Sigma_{1k} = \Sigma_{2k}$ 并且样本 i.i.d. 来自 $\Sigma_{1k}^{1/2} \mathbf{w}$ 。

(1) 在原假设下, 设置 1-4 中的所有 Σ_{1k} 都满足 (C1)。对于固定的常数 $\alpha_0 > 0$, 令

$$s_{k\ell_2} = s_{k\ell_2}(\alpha_0) := \text{card}\{\ell_1 : |\rho_{1k\ell_1\ell_2}| \geq (\log p)^{-1-\alpha_0} \text{ 或 } |\rho_{2k\ell_1\ell_2}| \geq (\log p)^{-1-\alpha_0}\},$$

$$s'_{k\ell_2} = s'_{k\ell_2}(\alpha_0) := \text{card}\{\ell_1 : |\sigma_{1k\ell_1\ell_2}| \geq (\log p)^{-1-\alpha_0} \text{ 或 } |\sigma_{2k\ell_1\ell_2}| \geq (\log p)^{-1-\alpha_0}\}.$$

对于 $0 < r < 1$, 定义集合

$$\Lambda_k(r) = \{1 \leq \ell_1 \leq p : |\rho_{1k\ell_1\ell_2}| \geq r \text{ 或 } |\rho_{2k\ell_1\ell_2}| \geq r \text{ 对于某个 } \ell_2 \neq \ell_1\},$$

$$\Lambda'_k(r) = \{1 \leq \ell_1 \leq p : |\sigma_{1k\ell_1\ell_2}| \geq r \text{ 或 } |\sigma_{2k\ell_1\ell_2}| \geq r \text{ 对于某个 } \ell_2 \neq \ell_1\}.$$

因为 Σ_{1k} 的所有对角线元素都大于或等于 1, 所以 $\rho_{1k\ell_1\ell_2} \leq \sigma_{1k\ell_1\ell_2}$, 这表明 $s_{k\ell_2} \leq s'_{k\ell_2}$, $\text{card}(\Lambda_k(r)) \leq \text{card}(\Lambda'_k(r))$ 。因此, 只需证明 $s'_{k\ell_2}$ 和 $\text{card}(\Lambda'_k(r))$ 满足条件 (C1), $k = 1, 2, 3, 4$ 。

- 在设置 1 中, 当 $\ell_1 \neq \ell_2$ 时, $\sigma_{11\ell_1\ell_2} = 0$ 。对于固定的常数 $\alpha_0 > 0$ 和 $0 < r < 1$, $s'_{1\ell_2} = 1$ 和 $\text{card}(\Lambda'_1(r)) = 0$ 。因此, $s'_{1\ell_2}$ 和 $\text{card}(\Lambda'_1(r))$ 满足条件 (C1)。
- 在设置 2 中, 当 $\ell_1 \neq \ell_2$ 时, $\sigma_{12\ell_1\ell_2} = 0.5^{|\ell_1 - \ell_2|}$ 。当 $r = 0.5$ 时, $\text{card}(\Lambda'_2(r)) = 0$, 这意味着 $\text{card}(\Lambda'_2(r))$ 满足条件 (C1)。对于任意 $\gamma > 0$, 任意固定的常数 $\alpha_0 > 0$ 和足够大的 p , $0.5^{p^{\gamma/2}} \leq (\log p)^{-1-\alpha_0}$, 这表明 $\max_{1 \leq \ell_2 \leq p} s'_{2\ell_2} \leq 2p^{\gamma/2} + 1 = o(p^\gamma)$, 因此, $s'_{2\ell_2}$ 满足条件 (C1)。
- 在设置 3 中, 当 $\ell_1 \neq \ell_2$ 时, $\sigma_{13\ell_1\ell_2} = (0.1^{|\ell_1 - \ell_2|} + 0.2^{|\ell_1 - \ell_2|})/2$ 。与设置 2 中的讨论类似, 可以证明 $s'_{3\ell_2}$ 和 $\text{card}(\Lambda'_3(r))$ 满足条件 (C1)。

- 在设置4中, 当 $\ell_1 \neq \ell_2$ 时, $\sigma_{14\ell_1\ell_2} = (0.5 - |\ell_1 - \ell_2|/10)I(|\ell_1 - \ell_2| \leq 4)$ 。当 $r = 0.4$ 时, $\text{card}(\Lambda'_4(r)) = 0$, $\text{card}(\Lambda'_4(r))$ 显然满足条件(C1)。对于任意固定的常数 $\alpha_0 > 0$, $\max_{1 \leq \ell_2 \leq p} s'_{4\ell_2} \leq 9$, 这表明 $s'_{4\ell_2}$ 也满足条件(C1)。

(2) 对于 $k = 1, 2, 3, 4$, 第一组样本i.i.d.来自 $\Sigma_{1k}^{1/2} \mathbf{w}$, 第二组样本i.i.d.来自 $\Sigma_{2k}^{1/2} \mathbf{w}$ 。接下来, 证明无论 $\{w_1, \dots, w_p\}$ 来自 $N(0, 1)$ 还是 $\text{Gamma}(4, 2) - 2$, $\Sigma_{1k}^{1/2} \mathbf{w}$ 和 $\Sigma_{2k}^{1/2} \mathbf{w}$ 都满足条件(C2*)。

对于 $k = 1, 2, 3, 4$, 当条件(C2*)中的 $\gamma_0 = 1$, $c_1 = y_1$ 和 $\epsilon = 2$ 时,

$$\mathbb{E}|\mathbf{r}_{1k\ell}^T \mathbf{w} / \sigma_{1k\ell\ell}^{1/2}|^{4\gamma_0+4+\epsilon} = \mathbb{E}|\mathbf{r}_{1k\ell}^T \mathbf{w} / \sigma_{1k\ell\ell}^{1/2}|^{10} = \mathbb{E}(\mathbf{w}^T \mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T \mathbf{w})^5 / \sigma_{1k\ell\ell}^5.$$

根据Bai和Silverstein (1998)^[57]中的引理2.7和 $\mathbf{r}_{1k\ell}^T \mathbf{r}_{1k\ell} = \sigma_{1k\ell\ell}$, 可得

$$\begin{aligned} & \mathbb{E}(\mathbf{w}^T \mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T \mathbf{w})^5 \\ & \leq 16\mathbb{E}|\mathbf{w}^T \mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T \mathbf{w} - \text{tr}(\mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T)|^5 + 16|\text{tr}(\mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T)|^5 \\ & \leq 16K_5((\mathbb{E}|w_1|^4 \text{tr}(\mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T \mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T))^{5/2} + \mathbb{E}|w_1|^{10} \text{tr}(\mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T \mathbf{r}_{1k\ell} \mathbf{r}_{1k\ell}^T)^{5/2}) + 16(\mathbf{r}_{1k\ell}^T \mathbf{r}_{1k\ell})^5 \\ & = 16K_5((\mathbb{E}|w_1|^4)^{5/2} + \mathbb{E}|w_1|^{10})(\mathbf{r}_{1k\ell}^T \mathbf{r}_{1k\ell})^5 + 16(\mathbf{r}_{1k\ell}^T \mathbf{r}_{1k\ell})^5 \\ & = 16K_5((\mathbb{E}|w_1|^4)^{5/2} + \mathbb{E}|w_1|^{10})\sigma_{1k\ell\ell}^5 + 16\sigma_{1k\ell\ell}^5, \end{aligned}$$

其中 K_5 是一个常数。无论 w_1 来自 $N(0, 1)$ 还是 $\text{Gamma}(4, 2) - 2$, $\mathbb{E}|w_1|^4$ 和 $\mathbb{E}|w_1|^{10}$ 都是有界的。因此, 对于任意的 ℓ ,

$$\mathbb{E}|\mathbf{r}_{1k\ell}^T \mathbf{w} / \sigma_{1k\ell\ell}^{1/2}|^{4\gamma_0+4+\epsilon} \leq K.$$

同理, 对于 $k = 1, 2, 3, 4$ 和任意的 ℓ ,

$$\mathbb{E}|\mathbf{r}_{2k\ell}^T \mathbf{w} / \sigma_{2k\ell\ell}^{1/2}|^{4\gamma_0+4+\epsilon} \leq K.$$

令 $\theta_{1k\ell_1\ell_2} = \text{Var}(\mathbf{r}_{1k\ell_1}^T \mathbf{w} \mathbf{r}_{1k\ell_2}^T \mathbf{w})$, 根据等式(2.86), 可得

$$\begin{aligned} \text{Var}(\mathbf{r}_{1k\ell_1}^T \mathbf{w} \mathbf{r}_{1k\ell_2}^T \mathbf{w}) &= \mathbb{E}(\mathbf{w}^T \mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T \mathbf{w} - \mathbf{r}_{1k\ell_2}^T \mathbf{r}_{1k\ell_1})^2 \\ &= \text{tr}(\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T \mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T) + \text{tr}(\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T \mathbf{r}_{1k\ell_2} \mathbf{r}_{1k\ell_1}^T) + \beta_w \sum_{i=1}^p (\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T)_{ii}^2 \\ &= \sigma_{1k\ell_1\ell_2}^2 + \sigma_{1k\ell_1\ell_1} \sigma_{1k\ell_2\ell_2} + \beta_w \sum_{i=1}^p (\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T)_{ii}^2, \end{aligned}$$

其中 $(\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T)_{ii}$ 代表矩阵 $\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T$ 的第 i 个对角线元素, $\beta_w = Ew_1^4 - 3 \geq 0$ 。因此,

$$\min_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{\theta_{1k\ell_1\ell_2}}{\sigma_{1k\ell_1\ell_1} \sigma_{1k\ell_2\ell_2}} \geq 1.$$

同理, 令 $\theta_{2k\ell_1\ell_2} = \text{Var}(\mathbf{r}_{2k\ell_1}^T \mathbf{w} \mathbf{r}_{2k\ell_2}^T \mathbf{w})$, 也可以得到

$$\min_{1 \leq \ell_1 \leq \ell_2 \leq p} \frac{\theta_{2k\ell_1\ell_2}}{\sigma_{2k\ell_1\ell_1} \sigma_{2k\ell_2\ell_2}} \geq 1.$$

(3) 接下来, 证明当 \mathbf{w} 中的元素i.i.d.来自 $N(0, 1)$ 时, $\Sigma_{1k}^{1/2} \mathbf{w}$ 满足条件(C3); 当 \mathbf{w} 中的元素i.i.d.来自 $\text{Gamma}(4, 2) - 2$ 时, $\Sigma_{1k}^{1/2} \mathbf{w}$ 不满足条件(C3)。

对任意的 $\ell_1, \ell_2, \ell_3, \ell_4 \in \{1, 2, \dots, p\}$, 根据等式(2.86), 可得

$$\begin{aligned} & E(\mathbf{r}_{1k\ell_1}^T \mathbf{w} \mathbf{r}_{1k\ell_2}^T \mathbf{w} \mathbf{r}_{1k\ell_3}^T \mathbf{w} \mathbf{r}_{1k\ell_4}^T \mathbf{w}) \\ &= E(\mathbf{w}^T \mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T \mathbf{w} \mathbf{w}^T \mathbf{r}_{1k\ell_3} \mathbf{r}_{1k\ell_4}^T \mathbf{w}) \\ &= \text{tr}(\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T) \text{tr}(\mathbf{r}_{1k\ell_3} \mathbf{r}_{1k\ell_4}^T) + \text{tr}(\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T \mathbf{r}_{1k\ell_3} \mathbf{r}_{1k\ell_4}^T) \\ & \quad + \text{tr}(\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T \mathbf{r}_{1k\ell_4} \mathbf{r}_{1k\ell_3}^T) + \beta_w \sum_{i=1}^p (\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T)_{ii} (\mathbf{r}_{1k\ell_3} \mathbf{r}_{1k\ell_4}^T)_{ii} \\ &= \sigma_{1k\ell_1\ell_2} \sigma_{1k\ell_3\ell_4} + \sigma_{1k\ell_2\ell_3} \sigma_{1k\ell_1\ell_4} + \sigma_{1k\ell_1\ell_3} \sigma_{1k\ell_2\ell_4} + \beta_w \sum_{i=1}^p (\mathbf{r}_{1k\ell_1} \mathbf{r}_{1k\ell_2}^T)_{ii} (\mathbf{r}_{1k\ell_3} \mathbf{r}_{1k\ell_4}^T)_{ii}. \end{aligned}$$

由上式可得, 当 w_1 来自标准正态分布 $N(0, 1)$ 时, $\beta_w = 0$, $\Sigma_{1k}^{1/2} \mathbf{w}$ 满足条件(C3); 当 w_1 来自 $\text{Gamma}(4, 2) - 2$ 时, $\beta_w = 1.5$, $\Sigma_{1k}^{1/2} \mathbf{w}$ 不满足条件(C3)。

然而, 根据Cai et al. (2013)^[22]中的性质1, 在原假设 H_0 和条件(C2*)下,

$$P(T_x - 4 \log p + \log \log p \geq c_\alpha) \leq -\log(1 - \alpha) + o(1),$$

其中 c_α 是第I类极值分布的 $1 - \alpha$ 分位点。这表明即使样本不满足条件(C1)和(C3), 检验 T_x 也可以有效地控制犯第一类错误的概率。这也许是当 \mathbf{w} 中的元素i.i.d.来自 $\text{Gamma}(4, 2) - 2$ 时, $\Sigma_{1k}^{1/2} \mathbf{w}$ 不满足条件(C3), 检验 T_{drx}^w 和检验 T_{drx}^m 的经验水平仍然令人满意的原因。

第三章 检验数据是否来自平稳向量自回归过程

本章考虑检验一组数据是否来自系数未知的平稳向量自回归过程。在原假设下,可以将这组数据看作是来自主平稳自回归过程的独立同分布的样本。因此,本章首先估计这个过程中的未知系数,然后根据该过程的协方差矩阵结构和系数的估计值,提出了一个新的检验统计量并且在原假设下给出了这个统计量的渐近分布。最后通过数值模拟验证了这个检验方法的正确性和有效性。

§3.1 检验问题

令 $\mathbf{X} = (x_{it})_{N \times T} = (\mathbf{x}_1, \dots, \mathbf{x}_T)$ 是观测到的时间序列构成的矩阵, $\mathbf{E} = (e_{it})_{N \times T} = (\mathbf{e}_1, \dots, \mathbf{e}_T)$ 为元素是独立同分布的,均值为0且方差为 σ_e^2 的白噪声矩阵。本章考虑的假设检验问题可以写为

$$H_{0p} : \phi(B)\mathbf{x}_t = \mathbf{e}_t \quad \text{versus} \quad H_{1p} : \text{非}H_{0p}, \quad (3.1)$$

其中 $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, B 是滞后算子满足 $B^j \mathbf{x}_t = \mathbf{x}_{t-j}$, $j = 0, 1, \dots$, 并且AR特征方程 $1 - \phi_1 x - \dots - \phi_p x^p = 0$ 每一个根的模都大于1。在原假设 H_{0p} 下, \mathbf{X} 的行与行之间是相互独立的且每行的协方差矩阵是相同的。

若 $\mathbf{Y} = \mathbf{X}^T = (\mathbf{y}_1, \dots, \mathbf{y}_N)$, 在原假设 H_{0p} 下, $\mathbf{y}_1, \dots, \mathbf{y}_N$ 可以被看作是来自主平稳AR(p)过程 \mathbf{y} 的 N 个样本。 \mathbf{y} 的协方差矩阵为 $\sigma_e^2 \Sigma_T$, 其中

$$\begin{aligned} \Sigma_T &= \gamma_0 (\rho_{|i-j|})_{i,j=1}^T, \\ \gamma_0 &= \frac{1}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p}, \end{aligned}$$

$\rho_1, \rho_2, \dots, \rho_p$ 满足下列Yule-Walker方程组(Cryer和Chan (2008)^[58], p.55):

$$\left. \begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \phi_3 \rho_2 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \phi_3 \rho_1 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \phi_3 \rho_{p-3} + \dots + \phi_p \end{aligned} \right\} \quad (3.2)$$

以及

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} + \cdots + \phi_p \rho_{k-p}, \quad k \geq p+1.$$

§3.2 主要结果

由于 $\text{Cov}(\boldsymbol{\Sigma}_T^{-1/2} \mathbf{y}) = \sigma_e^2 \mathbf{I}_T$, 其中 \mathbf{I}_T 为 $T \times T$ 单位阵, 可以通过检验 \mathbf{y} 的协方差矩阵结构来检验原假设 H_{0p} . 但与一般的球形检验问题不同的是矩阵 $\boldsymbol{\Sigma}_T$ 不是已知的, 而是由未知系数 ϕ_1, \dots, ϕ_p 唯一确定的. 因此, 本节首先估计这些未知系数, 然后利用系数的估计值进行检验.

§3.2.1 系数的最小二乘估计

若 $\mathbf{y}_i = (y_{1i}, \dots, y_{Ti})^T$ 是一个平稳 $\text{AR}(p)$ 过程, 则

$$y_{ti} = \phi_1 y_{t-1,i} + \phi_2 y_{t-2,i} + \cdots + \phi_p y_{t-p,i} + e_{ti}, \quad t = p+1, \dots, T,$$

可将上式看成是以 $y_{t-1,i}, y_{t-2,i}, \dots, y_{t-p,i}$ 为预测变量, y_{ti} 为响应变量的回归模型. 令条件平方和函数为

$$S_c(\boldsymbol{\phi}) = \sum_{i=1}^N \sum_{t=p+1}^T (y_{ti} - \phi_1 y_{t-1,i} - \phi_2 y_{t-2,i} - \cdots - \phi_p y_{t-p,i})^2,$$

其中 $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_p)^T$ 是由回归系数构成的 p 维向量. 根据最小二乘准则, $\boldsymbol{\phi}$ 的估计值 $\hat{\boldsymbol{\phi}}$ 是通过最小化 $S_c(\boldsymbol{\phi})$ 得到的, 即

$$\hat{\boldsymbol{\phi}} = \text{Arg}\{\min_{\boldsymbol{\phi}} S_c(\boldsymbol{\phi})\} = \text{Arg}\{\min_{\boldsymbol{\phi}} \sum_{i=1}^N \sum_{t=p+1}^T (y_{ti} - \phi_1 y_{t-1,i} - \phi_2 y_{t-2,i} - \cdots - \phi_p y_{t-p,i})^2\}.$$

因此,

$$\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\Gamma}}_p^{-1} \hat{\boldsymbol{\gamma}}_p,$$

其中

$$\begin{aligned} \hat{\boldsymbol{\Gamma}}_p &= \left((NT)^{-1} \sum_{i=1}^N \sum_{t=p+1}^T y_{t-k,i} y_{t-\ell,i} \right)_{k,\ell=1}^p, \\ \hat{\boldsymbol{\gamma}}_p &= \left((NT)^{-1} \sum_{i=1}^N \sum_{t=p+1}^T y_{t-\ell,i} y_{ti} \right)_{\ell=1}^p. \end{aligned}$$

为了得到 $\hat{\boldsymbol{\phi}}$ 的渐近正态性, 需要下面两个假设条件.

假设3.2.1. 维数 T 和样本量 N 成比例趋于无穷, 即

$$y_N = T/N \rightarrow y \in (0, \infty).$$

假设3.2.2. 白噪声矩阵 $\mathbf{E} = (e_{it})_{N \times T}$ 中的元素是独立同分布的, 满足

$$\mathbb{E}e_{it} = 0, \mathbb{E}e_{it}^2 = \sigma_e^2, \mathbb{E}e_{it}^4 = (\beta_e + 3)\sigma_e^4.$$

若平稳性条件满足, 即AR特征方程 $1 - \phi_1 x - \dots - \phi_p x^p = 0$ 每一个根的模都大于1, 则 \mathbf{y} 也可以表示为一般线性过程. 此时 $\mathbf{y} = \mathbf{Q}\mathbf{e}$, 其中 $\mathbf{e} = (e_T, e_{T-1}, e_{T-2}, \dots)^T$ 是一个 ∞ 维白噪声向量, \mathbf{Q} 是一个 $T \times \infty$ 矩阵, 具有以下形式

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 & \psi_1 & \psi_2 & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ 0 & 0 & 1 & \cdots & \psi_{T-3} & \psi_{T-2} & \psi_{T-1} & \cdots \\ 0 & 1 & \psi_1 & \cdots & \psi_{T-2} & \psi_{T-1} & \psi_T & \cdots \\ 1 & \psi_1 & \psi_2 & \cdots & \psi_{T-1} & \psi_T & \psi_{T+1} & \cdots \end{pmatrix},$$

其中 ψ 系数由下式决定

$$\left. \begin{aligned} \psi_0 &= 1 \\ \psi_1 &= \phi_1 \\ \psi_2 &= \phi_2 \psi_0 + \phi_1 \psi_1 \\ &\vdots \\ \psi_j &= \phi_p \psi_{j-p} + \phi_{p-1} \psi_{j-p+1} + \cdots + \phi_1 \psi_{j-1} \\ &\vdots \end{aligned} \right\}.$$

定理3.2.1. 在原假设 H_{0p} 下, 若假设3.2.1-假设3.2.2成立, 则

$$\sqrt{NT}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \xrightarrow{d} N(\mathbf{0}_p, \boldsymbol{\Sigma}_p^{-1}),$$

其中 $\mathbf{0}_p$ 为元素都是0的 p 维向量, $\boldsymbol{\Sigma}_p = \gamma_0(\rho_{|i-j|})_{i,j=1}^p$ 为 $p \times p$ 协方差矩阵。

根据定理3.2.1和假设3.2.1, 对于 $j = 1, \dots, p$,

$$\hat{\phi}_j - \phi_j = O_p(T^{-1}), \tag{3.3}$$

其中 $\hat{\phi}_j$ 代表 $\hat{\boldsymbol{\phi}}$ 的第 j 个分量。

§3.2.2 检验方法

根据Verbyla (1985)^[59]中的等式(3), Σ_T^{-1} 的表达式为:

$$\Sigma_T^{-1} = \mathbf{I}_T + \sum_{j=1}^p \phi_j^2 \mathbf{E}_j - \sum_{j=1}^p \phi_j \mathbf{F}_j + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} \phi_k \phi_{k+j} \mathbf{G}_{k,k+j}, \quad (3.4)$$

其中 \mathbf{E}_j 是将单位阵中前 j 个和后 j 个对角线元素变为0的矩阵, \mathbf{F}_j 是第上 j 个和第下 j 个次对角线元素为1, 其余元素为0的矩阵, $\mathbf{G}_{k,k+j} = \mathbf{E}_k \mathbf{F}_j \mathbf{E}_k$ 。很自然地, 矩阵 Σ_T^{-1} 的估计为

$$\hat{\Sigma}_T^{-1} = \mathbf{I}_T + \sum_{j=1}^p \hat{\phi}_j^2 \mathbf{E}_j - \sum_{j=1}^p \hat{\phi}_j \mathbf{F}_j + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} \hat{\phi}_k \hat{\phi}_{k+j} \mathbf{G}_{k,k+j}. \quad (3.5)$$

受Wang和Yao (2013)^[4]中修正后的John检验的启发, 本节提出以下统计量来检验 \mathbf{y} 的协方差矩阵结构

$$\hat{T}_{Np} = \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} / [T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1})] - \mathbf{I}_T)^2,$$

其中 $\mathbf{B}_N = N^{-1} \sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^T$ 。

为了得到统计量 \hat{T}_{Np} 在原假设 H_{0p} 下的渐近分布, 除了假设3.2.1-假设3.2.2, 还需要下面两个假设条件。

假设3.2.3. 矩阵 $\mathbf{Q} = (q_{tj})_{t=1, \dots, T, j=1, \dots, \infty}$ 满足下列条件: 存在一个数列 $\{k_T\}$, $k_T \geq T$, 使得

$$\sum_{t=1}^T \sum_{j=k_T+1}^{\infty} |q_{tj}|^2 = o(1).$$

假设3.2.4. $\beta_e = 0$ 或 $\|\mathbf{Q}_1^T \Sigma_T^{-1} \mathbf{Q}_1 - \text{diag}(\mathbf{Q}_1^T \Sigma_T^{-1} \mathbf{Q}_1)\| = o(1)$, 其中 \mathbf{Q}_1 是 \mathbf{Q} 的前 k_T 列构成的矩阵, $\text{diag}(\mathbf{Q}_1^T \Sigma_T^{-1} \mathbf{Q}_1)$ 是由矩阵 $\mathbf{Q}_1^T \Sigma_T^{-1} \mathbf{Q}_1$ 的对角线元素构成的对角矩阵。

注释3.2.1. 若 \mathbf{y} 是一个平稳AR(1)过程, 则相应的矩阵 \mathbf{Q} 有以下形式

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 & \phi & \phi^2 & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ 0 & 0 & 1 & \cdots & \phi^{T-3} & \phi^{T-2} & \phi^{T-1} & \cdots \\ 0 & 1 & \phi & \cdots & \phi^{T-2} & \phi^{T-1} & \phi^T & \cdots \\ 1 & \phi & \phi^2 & \cdots & \phi^{T-1} & \phi^T & \phi^{T+1} & \cdots \end{pmatrix},$$

其中系数 ϕ 满足 $|\phi| < 1$ 。令 $k_T = 2T$ ，则

$$\sum_{t=1}^T \sum_{j=2T+1}^{\infty} |q_{tj}|^2 = \frac{\phi^{2T+2}(1 - \phi^{2T})}{(1 - \phi^2)^2} = o(1).$$

此时，假设3.2.3成立。

下面的定理给出了统计量 \hat{T}_{Np} 在原假设 H_{0p} 下的渐近分布。

定理3.2.2. 在原假设 H_{0p} 和假设3.2.1-假设3.2.4下，若 $Ee_{it}^8 < \infty$ ， $\text{tr}(\Sigma_T^2) = o(T^2)$ ，则

$$\frac{\hat{T}_{Np} - Ty_N - (\beta_e + 1)y_N}{2y_N} \xrightarrow{d} N(0, 1). \quad (3.6)$$

注释3.2.2. 根据定理3.2.2的证明，可以得到，如果 ϕ^* 是 ϕ 的另外一个 T 相合估计，即对于 $j = 1, \dots, p$ ，

$$\phi_j^* - \phi_j = O_p(T^{-1}),$$

其中 ϕ_j^* 代表 ϕ^* 的第 j 个分量，那么

$$\frac{T_{Np}^* - Ty_N - (\beta_e + 1)y_N}{2y_N} \xrightarrow{d} N(0, 1),$$

其中 T_{Np}^* 是用 ϕ^* 替换 \hat{T}_{Np} 中的 $\hat{\phi}$ 得到的统计量。这表明即使统计量 \hat{T}_{Np} 中使用的不是最小二乘估计 $\hat{\phi}$ ，而是 ϕ 的其他 T 相合估计，那么它依然具有(3.6)所示的渐近正态性。

根据定理3.2.2，对于给定的显著水平 α ，基于统计量 \hat{T}_{Np} 的检验的拒绝域为

$$\{\mathbf{y}_1, \dots, \mathbf{y}_N : \hat{T}_{Np} > Ty_N + (\beta_e + 1)y_N + 2y_N z_{1-\alpha}\}, \quad (3.7)$$

其中 $z_{1-\alpha}$ 是标准正态分布 $N(0, 1)$ 的 $1 - \alpha$ 分位点。在下文中，将基于统计量 \hat{T}_{Np} 的检验方法简称为检验 \hat{T}_{Np} 。

注意到(3.7)中的临界值涉及参数 β_e ，在实际应用中， β_e 通常是未知的，因此需要对其进行估计。根据Bai和Silverstein (2004)^[55]中的等式(1.15)，在原假设 H_{0p} 下

$$\text{Var}(\mathbf{y}^T \Sigma_T^{-1} \mathbf{y}) = \sigma_e^4 \left[2T + \beta_e \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell)^2 \right],$$

其中 \mathbf{q}_ℓ 代表矩阵 \mathbf{Q} 的第 ℓ 列。因此，可以构造 β_e 的一个矩估计。首先用样本方差

$$\hat{V}_e = (N-1)^{-1} \sum_{i=1}^N (\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i - N^{-1} \sum_{i=1}^N \mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i)^2$$

估计 $\text{Var}(\mathbf{y}^T \Sigma_T^{-1} \mathbf{y})$ 。由于 $\text{Cov}(\Sigma_T^{-1/2} \mathbf{y}) = \sigma_e^2 \mathbf{I}_T$ ，用 $T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1})$ 估计 σ_e^2 。综上， β_e 的估计为

$$\hat{\beta}_e = \frac{\hat{V}_e / [T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1})]^2 - 2T}{\sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell)^2}, \quad (3.8)$$

其中 $\hat{\mathbf{q}}_\ell$ 代表矩阵 $\hat{\mathbf{Q}}$ 的第 ℓ 列，而 $\hat{\mathbf{Q}}$ 是用 $\hat{\phi}$ 替换矩阵 \mathbf{Q} 中的 ϕ 得到的矩阵，即

$$\hat{\mathbf{Q}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 & \hat{\psi}_1 & \hat{\psi}_2 & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ 0 & 0 & 1 & \cdots & \hat{\psi}_{T-3} & \hat{\psi}_{T-2} & \hat{\psi}_{T-1} & \cdots \\ 0 & 1 & \hat{\psi}_1 & \cdots & \hat{\psi}_{T-2} & \hat{\psi}_{T-1} & \hat{\psi}_T & \cdots \\ 1 & \hat{\psi}_1 & \hat{\psi}_2 & \cdots & \hat{\psi}_{T-1} & \hat{\psi}_T & \hat{\psi}_{T+1} & \cdots \end{pmatrix},$$

其中 $\hat{\psi}$ 系数由下式决定

$$\left. \begin{aligned} \hat{\psi}_0 &= 1 \\ \hat{\psi}_1 &= \hat{\phi}_1 \\ \hat{\psi}_2 &= \hat{\phi}_2 \hat{\psi}_0 + \hat{\phi}_1 \hat{\psi}_1 \\ &\vdots \\ \hat{\psi}_j &= \hat{\phi}_p \hat{\psi}_{j-p} + \hat{\phi}_{p-1} \hat{\psi}_{j-p+1} + \cdots + \hat{\phi}_1 \hat{\psi}_{j-1} \\ &\vdots \end{aligned} \right\}.$$

下面的性质说明在原假设 H_{0p} 下， $\hat{\beta}_e$ 是 β_e 的一个相合估计。

性质3.2.1. 在原假设 H_{0p} 和假设3.2.1-假设3.2.2下，若 $\text{E}e_{it}^8 < \infty$ ，则

$$\hat{\beta}_e = \beta_e + o_p(1).$$

注释3.2.3. 注意到 $\hat{\beta}_e$ 的表达式中存在无穷级数。在实际应用中，可以用一个较大的整数代替 ∞ 。根据注释3.2.1，在下一节的数值模拟中，用

$$\hat{\beta}_e^* = \frac{\hat{V}_e / [T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1})]^2 - 2T}{\sum_{\ell=1}^{2T} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell)^2} \quad (3.9)$$

估计 β_e 。

§3.3 数值模拟

本节通过数值模拟评估新的检验方法在有限样本情形下的表现。令白噪声矩阵 $\mathbf{E} = (e_{it})_{N \times T}$ 中的元素是来自高斯分布 $N(0, 1)$ 或伽马分布 $\text{Gamma}(4, 2)$ -2 的 i.i.d. 样本, 在这两种情况下, 参数 β_e 分别为 0 和 1.5。样本的维数 T 取 50, 100, 200, 500, 样本量 N 取 100, 200, 300。显著水平 α 设置为 0.05, 独立重复试验次数为 5,000。

为了模拟检验 \hat{T}_{N1} 和检验 \hat{T}_{N2} 分别在原假设 H_{01} 和原假设 H_{02} 下的经验水平, 考虑来自下面两个模型的样本。

- **模型1.** $H_{01} : \mathbf{x}_t = \phi \mathbf{x}_{t-1} + \mathbf{e}_t$, 其中 $\phi \in \{-0.6, -0.3, 0, 0.3, 0.6\}$.
- **模型2.** $H_{02} : \mathbf{x}_t = \phi_1 \mathbf{x}_{t-1} + \phi_2 \mathbf{x}_{t-2} + \mathbf{e}_t$, 其中 $\phi_1 \in \{0.3, 0.6\}$, $\phi_2 \in \{0.2, 0.3\}$.

表3.1和表3.2分别记录了在不同的 (T, N) 值下, 检验 \hat{T}_{N1} 和检验 \hat{T}_{N2} 的经验水平。结果显示无论白噪声矩阵 \mathbf{E} 中的元素来自正态分布还是伽马分布, 随着 T 或 N 的增大, 新的检验方法犯第一类错误的概率都接近于事先给定的显著水平 0.05。这表明新提出的检验方法对非正态分布的样本也是适用的。

接下来, 为了展示(3.9)中的估计量 $\hat{\beta}_e^*$ 在模型1和模型2下的有限样本表现, 基于 5,000 次重复试验的结果, 表3.3-表3.6中给出了 $\hat{\beta}_e^*$ 的均值和方差。数值结果表明, 该估计量的表现是令人满意的, 尤其是在维数 T 和样本量 N 较大的情况下。

为了模拟检验 \hat{T}_{N1} 的经验功效, 考虑来自下面两个模型的样本。

- **模型3.** $H_{11} : \mathbf{x}_t = \phi_1 \mathbf{x}_{t-1} + \phi_2 \mathbf{x}_{t-2} + \mathbf{e}_t$, 其中 $\phi_1 \in \{-0.6, -0.3, 0, 0.3, 0.6\}$, $\phi_2 \in \{0.1, 0.15\}$.
- **模型4.** $H_{11} : \mathbf{x}_t = \phi \mathbf{x}_{t-1} + \mathbf{e}_t + \theta \mathbf{e}_{t-1}$, 其中 $\phi \in \{-0.6, -0.3, 0, 0.3, 0.6\}$, $\theta \in \{0.4, 0.5\}$.

在模型3中, 样本来自 VAR(2) 模型, 系数 ϕ_1 和模型1中的系数 ϕ 取值相同。在模型4中, 样本来自 VARMA(1,1) 模型, 系数 ϕ 也和模型1中的系数 ϕ 取值相同。表3.7和表3.8记录了当样本来自模型3时检验 \hat{T}_{N1} 的经验功效。表3.9和表3.10记录了当样本来自模型4时检验 \hat{T}_{N1} 的经验功效。

从表3.7和表3.8中可以看到, 当样本来自 VAR(2) 模型时, 对于不同的 ϕ_1 , 检验 \hat{T}_{N1} 的经验功效的变化趋势类似。对于给定的 ϕ_1 和一组 (T, N) 值, 检验 \hat{T}_{N1} 的经

表 3.1 检验 \hat{T}_{N1} 在模型1下的经验水平(百分比)。

ϕ	N	$N(0, 1)$				Gamma(4,2)-2			
		$T = 50$	100	200	500	$T = 50$	100	200	500
-0.6	100	4.58	5.66	4.54	4.80	5.06	5.04	4.74	5.16
	200	5.34	4.94	4.82	4.94	5.00	5.58	4.64	4.84
	300	4.64	4.96	4.96	5.06	5.10	4.94	5.26	4.72
-0.3	100	4.62	4.84	5.00	5.18	5.62	4.64	4.86	4.70
	200	4.74	4.84	4.96	4.82	5.70	4.70	5.54	4.46
	300	5.30	5.08	5.38	4.96	5.08	5.18	5.26	4.64
0	100	4.94	5.00	5.18	4.78	4.64	5.24	4.92	4.76
	200	4.44	4.76	5.20	4.98	5.08	4.82	4.72	5.06
	300	4.82	4.86	5.26	4.86	5.42	5.86	5.34	5.32
0.3	100	4.52	4.92	5.18	5.06	5.74	4.74	4.86	4.74
	200	4.70	4.90	5.04	4.92	5.62	4.66	5.40	4.68
	300	5.30	5.12	5.50	4.94	4.92	5.30	5.30	4.48
0.6	100	4.82	5.72	4.50	4.96	5.46	5.20	4.44	5.30
	200	4.86	4.92	4.84	4.80	4.54	5.88	4.58	4.66
	300	4.84	4.62	5.22	5.20	4.88	4.92	5.28	4.84

验功效随着 ϕ_2 的增大而增大,这是符合直观的,因为 ϕ_2 越大,备择假设 H_{11} 和原假设 H_{01} 的差别越大。从表3.9和表3.10中的模拟结果可知,当样本来自VARMA(1,1)模型时,检验 \hat{T}_{N1} 的经验功效并不是单纯地随着 θ 的增大而增大。当 $\phi + \theta$ 的值接近0时,检验 \hat{T}_{N1} 的表现很差;当 $\phi + \theta$ 的值远离0时,检验 \hat{T}_{N1} 的表现会变好。此外,无论样本来自VAR(2)模型还是VARMA(1,1)模型,对于固定的 N ,检验 \hat{T}_{N1} 的经验功效虽然随着 T 增大而增大,但是数值不会变化太大。反之,对于固定的 T ,当 N 增大时,检验 \hat{T}_{N1} 的经验功效会明显地增大。

总的来说,通过这一节的模拟可知,新提出的检验方法对正态分布和非正态分布的样本都是适用的,并且在大部分情况下,新的检验方法都具有较高的经验功效。

表 3.2 检验 \hat{T}_{N_2} 在模型2下的经验水平(百分比)。

N(0, 1)									
		$\phi_1 = 0.3$				$\phi_1 = 0.6$			
ϕ_2	N	$T = 50$	100	200	500	$T = 50$	100	200	500
0.2	100	4.40	5.12	4.20	5.16	4.78	5.04	4.54	5.04
	200	4.60	5.04	4.98	4.48	4.50	4.64	4.60	5.04
	300	4.16	4.78	4.78	4.64	4.28	4.76	4.86	4.82
0.3	100	4.50	5.10	5.30	5.16	4.56	5.28	4.30	5.08
	200	4.82	4.72	4.90	5.36	5.04	4.90	5.14	5.66
	300	4.54	4.96	4.90	5.12	4.68	4.76	5.24	4.78

Gamma(4,2)-2									
		$\phi_1 = 0.3$				$\phi_1 = 0.6$			
ϕ_2	N	$T = 50$	100	200	500	$T = 50$	100	200	500
0.2	100	4.84	5.40	5.30	5.04	5.92	4.52	4.94	5.26
	200	5.18	5.26	5.06	5.08	5.68	5.10	5.30	4.76
	300	5.38	4.52	5.18	5.08	5.10	5.00	5.14	5.38
0.3	100	4.90	5.18	5.32	4.92	4.92	5.10	5.08	4.82
	200	5.08	5.18	5.18	4.92	4.90	5.18	4.98	4.70
	300	5.18	4.94	5.36	4.84	4.66	5.02	4.86	4.80

表 3.3 当样本来自高斯分布时, 估计量 $\hat{\beta}_e^*$ 在模型1下的数值结果, 此时 $\beta_e = 0$ 。

ϕ	N	$T = 50$	100	200	500
		均值(方差)	均值(方差)	均值(方差)	均值(方差)
-0.6	100	-0.0018 (0.0816)	0.0012 (0.0814)	0.0074 (0.0811)	0.0032 (0.0825)
	200	0.0016 (0.0413)	0.0046 (0.0399)	0.0064 (0.0412)	0.0007 (0.0384)
	300	0.0041 (0.0281)	0.0018 (0.0263)	0.0023 (0.0272)	-0.0002 (0.0273)
-0.3	100	0.0060 (0.0843)	0.0025 (0.0813)	0.0045 (0.0817)	0.0067 (0.0814)
	200	0.0025 (0.0428)	0.0044 (0.0412)	0.0071 (0.0403)	0.0016 (0.0407)
	300	0.0004 (0.0278)	0.0030 (0.0265)	0.0041 (0.0267)	0.0005 (0.0278)
0	100	0.0053 (0.0855)	0.0070 (0.0847)	0.0033 (0.0821)	0.0051 (0.0832)
	200	0.0070 (0.0429)	0.0037 (0.0417)	0.0027 (0.0409)	0.0001 (0.0417)
	300	0.0040 (0.0281)	0.0034 (0.0272)	-0.0004 (0.0275)	0.0002 (0.0280)
0.3	100	0.0060 (0.0846)	0.0016 (0.0812)	0.0046 (0.0813)	0.0069 (0.0814)
	200	0.0026 (0.0429)	0.0040 (0.0410)	0.0071 (0.0403)	0.0017 (0.0407)
	300	0.0003 (0.0277)	0.0029 (0.0263)	0.0042 (0.0267)	0.0005 (0.0278)
0.6	100	-0.0025 (0.0816)	0.0029 (0.0816)	0.0074 (0.0810)	0.0034 (0.0825)
	200	0.0015 (0.0410)	0.0048 (0.0400)	0.0064 (0.0410)	0.0008 (0.0383)
	300	0.0044 (0.0280)	0.0019 (0.0263)	0.0023 (0.0270)	-0.0003 (0.0271)

表 3.4 当样本来自伽马分布时, 估计量 $\hat{\beta}_e^*$ 在模型1下的数值结果, 此时 $\beta_e = 1.5$ 。

ϕ	N	$T = 50$	100	200	500
		均值(方差)	均值(方差)	均值(方差)	均值(方差)
-0.6	100	1.4799 (0.3535)	1.4857 (0.2867)	1.4997 (0.2695)	1.4962 (0.2591)
	200	1.4815 (0.1758)	1.4967 (0.1450)	1.5003 (0.1366)	1.4968 (0.1273)
	300	1.4850 (0.1223)	1.4962 (0.1017)	1.4989 (0.0922)	1.4936 (0.0832)
-0.3	100	1.4902 (0.3496)	1.4898 (0.2923)	1.4994 (0.2648)	1.4956 (0.2605)
	200	1.4891 (0.1693)	1.4963 (0.1483)	1.4989 (0.1358)	1.4913 (0.1300)
	300	1.4897 (0.1153)	1.4944 (0.0993)	1.4987 (0.0901)	1.4942 (0.0868)
0	100	1.4936 (0.3436)	1.4904 (0.2933)	1.4970 (0.2645)	1.4964 (0.2571)
	200	1.4899 (0.1687)	1.4946 (0.1462)	1.4984 (0.1401)	1.4936 (0.1261)
	300	1.4879 (0.1125)	1.4964 (0.0990)	1.4949 (0.0907)	1.4939 (0.0839)
0.3	100	1.4901 (0.3505)	1.4895 (0.2914)	1.4994 (0.2651)	1.4957 (0.2605)
	200	1.4889 (0.1696)	1.4965 (0.1478)	1.4987 (0.1358)	1.4915 (0.1299)
	300	1.4898 (0.1156)	1.4944 (0.0991)	1.4985 (0.0903)	1.4944 (0.0868)
0.6	100	1.4819 (0.3564)	1.4843 (0.2871)	1.4995 (0.2705)	1.4964 (0.2592)
	200	1.4824 (0.1761)	1.4963 (0.1453)	1.5001 (0.1365)	1.4972 (0.1275)
	300	1.4855 (0.1223)	1.4953 (0.1019)	1.4996 (0.0924)	1.4940 (0.0833)

表 3.5 当样本来自高斯分布时, 估计量 $\hat{\beta}_e^*$ 在模型2下的数值结果, 此时 $\beta_e = 0$ 。

$\phi_1 = 0.3$					
ϕ_2	N	$T = 50$	100	200	500
		均值(方差)	均值(方差)	均值(方差)	均值(方差)
0.2	100	0.0090 (0.0857)	0.0001 (0.0821)	0.0049 (0.0808)	0.0063 (0.0864)
	200	0.0037 (0.0412)	0.0030 (0.0409)	0.0055 (0.0406)	0.0021 (0.0416)
	300	0.0043 (0.0280)	0.0012 (0.0276)	0.0021 (0.0271)	-0.0015 (0.0278)
0.3	100	0.0062 (0.0857)	0.0033 (0.0850)	0.0016 (0.0792)	0.0059 (0.0828)
	200	0.0026 (0.0425)	0.0037 (0.0407)	0.0039 (0.0402)	0.0026 (0.0410)
	300	0.0042 (0.0287)	0.0034 (0.0276)	0.0006 (0.0266)	0.0012 (0.0264)
$\phi_1 = 0.6$					
ϕ_2	N	$T = 50$	100	200	500
		均值(方差)	均值(方差)	均值(方差)	均值(方差)
0.2	100	0.0007 (0.0839)	0.0027 (0.0830)	0.0024 (0.0784)	0.0023 (0.0811)
	200	-0.0007 (0.0409)	0.0030 (0.0427)	0.0052 (0.0396)	0.0011 (0.0406)
	300	0.0026 (0.0283)	0.0039 (0.0284)	0.0023 (0.0272)	-0.0002 (0.0268)
0.3	100	-0.0004 (0.0874)	0.0013 (0.0860)	0.0031 (0.0807)	0.0006 (0.0828)
	200	0.0022 (0.0439)	0.0008 (0.0427)	0.0030 (0.0414)	-0.0012 (0.0403)
	300	0.0015 (0.0287)	-0.0028 (0.0285)	0.0032 (0.0267)	-0.0021 (0.0272)

表 3.6 当样本来自伽马分布时, 估计量 $\hat{\beta}_e^*$ 在模型2下的数值结果, 此时 $\beta_e = 1.5$ 。

$\phi_1 = 0.3$					
ϕ_2	N	$T = 50$	100	200	500
		均值(方差)	均值(方差)	均值(方差)	均值(方差)
0.2	100	1.4816 (0.3505)	1.4924 (0.2871)	1.4970 (0.2688)	1.5010 (0.2603)
	200	1.4847 (0.1762)	1.4948 (0.1506)	1.5003 (0.1352)	1.4951 (0.1287)
	300	1.4914 (0.1160)	1.4958 (0.1026)	1.5011 (0.0908)	1.4937 (0.0872)
0.3	100	1.4889 (0.3501)	1.4904 (0.2898)	1.5045 (0.2799)	1.5046 (0.2600)
	200	1.4918 (0.1774)	1.4998 (0.1493)	1.5042 (0.1361)	1.5008 (0.1319)
	300	1.4909 (0.1205)	1.5002 (0.1023)	1.5018 (0.0921)	1.4988 (0.0879)
$\phi_1 = 0.6$					
ϕ_2	N	$T = 50$	100	200	500
		均值(方差)	均值(方差)	均值(方差)	均值(方差)
0.2	100	1.4800 (0.3411)	1.4784 (0.2946)	1.4992 (0.2685)	1.5055 (0.2718)
	200	1.4841 (0.1806)	1.4867 (0.1483)	1.4958 (0.1367)	1.4991 (0.1307)
	300	1.4925 (0.1220)	1.4934 (0.0993)	1.4956 (0.0901)	1.4973 (0.0854)
0.3	100	1.4816 (0.3511)	1.4925 (0.3029)	1.5018 (0.2737)	1.5017 (0.2554)
	200	1.4910 (0.1739)	1.4961 (0.1535)	1.5006 (0.1359)	1.4980 (0.1306)
	300	1.4968 (0.1182)	1.4997 (0.1013)	1.5004 (0.0919)	1.4956 (0.0837)

表 3.7 当样本来自高斯分布时，检验 \hat{T}_{N1} 在模型3下的经验功效(百分比)。

ϕ_1	N	$T=50$	$\phi_2 = 0.1$			$\phi_2 = 0.15$			
			100	200	500	$T=50$	100	200	500
-0.6	100	23.80	24.32	24.72	24.62	65.64	69.16	70.22	70.76
	200	56.72	61.26	62.26	62.48	98.78	99.24	99.58	99.70
	300	84.62	88.64	89.98	90.72	100	100	100	100
-0.3	100	23.58	24.20	24.40	25.20	65.72	69.10	69.64	71.36
	200	57.16	59.88	60.34	62.90	98.70	99.40	99.62	99.90
	300	85.00	88.22	89.22	90.28	100	100	100	100
0	100	23.98	25.16	26.26	24.36	67.08	69.70	70.38	72.60
	200	56.88	60.44	62.68	63.20	98.68	99.38	99.58	99.84
	300	85.06	87.78	89.92	90.36	100	100	100	100
0.3	100	23.56	24.20	24.64	24.88	66.06	69.76	69.60	71.20
	200	56.96	60.66	60.36	62.70	98.98	99.36	99.62	99.76
	300	85.24	88.18	89.16	90.38	100	100	100	100
0.6	100	23.88	24.74	25.02	24.86	64.86	68.66	69.92	71.10
	200	57.40	61.94	62.66	61.66	98.66	99.20	99.56	99.70
	300	85.18	88.82	89.56	90.82	99.98	100	100	100

表 3.8 当样本来自伽马分布时，检验 \hat{T}_{N1} 在模型3下的经验功效(百分比)。

ϕ_1	N	$\phi_2 = 0.1$				$\phi_2 = 0.15$			
		$T = 50$	100	200	500	$T = 50$	100	200	500
-0.6	100	23.70	24.78	24.52	24.70	63.76	67.14	69.36	72.70
	200	56.86	59.98	61.42	61.86	98.60	99.54	99.54	99.62
	300	84.44	87.02	88.92	89.94	99.98	100	100	100
-0.3	100	24.58	23.16	25.94	24.80	63.56	68.68	68.90	70.54
	200	56.96	60.18	62.16	64.08	98.46	99.38	99.68	99.66
	300	84.00	87.56	89.02	90.46	100	100	100	100
0	100	24.18	25.44	25.42	26.24	64.22	67.12	70.60	70.78
	200	56.72	58.94	62.24	64.22	98.74	99.28	99.64	99.78
	300	85.06	86.82	89.96	90.52	100	100	100	100
0.3	100	23.64	23.30	26.18	24.90	64.14	68.32	68.86	71.66
	200	56.28	60.08	62.26	63.86	98.56	99.44	99.64	99.50
	300	84.36	87.38	89.08	90.58	100	100	100	100
0.6	100	23.58	25.26	24.82	24.70	63.96	67.86	69.66	71.96
	200	57.38	59.40	61.56	61.30	98.28	99.42	99.46	99.72
	300	84.18	86.72	89.18	90.08	100	100	100	100

表 3.9 当样本来自正态分布时，检验 \hat{T}_{M1} 在模型4下的经验功效(百分比)。

		$\theta = 0.4$				$\theta = 0.5$			
ϕ	N	$T = 50$	100	200	500	$T = 50$	100	200	500
-0.6	100	21.20	22.70	22.98	22.92	10.02	9.82	10.24	9.86
	200	52.90	55.48	58.40	58.82	18.96	18.20	18.78	20.04
	300	80.98	84.06	87.08	87.42	29.72	30.74	31.98	31.28
-0.3	100	6.84	6.72	6.50	6.82	22.20	24.18	23.96	24.70
	200	9.52	8.88	9.84	10.00	55.92	59.94	59.80	61.30
	300	12.70	12.98	12.38	12.40	83.78	88.00	88.18	89.78
0	100	57.42	59.82	62.42	61.88	97.90	98.22	98.90	98.94
	200	97.42	98.60	98.64	98.64	100	100	100	100
	300	100	100	100	100	100	100	100	100
0.3	100	99.40	99.62	99.84	99.72	100	100	100	100
	200	100	100	100	100	100	100	100	100
	300	100	100	100	100	100	100	100	100
0.6	100	100	100	100	100	100	100	100	100
	200	100	100	100	100	100	100	100	100
	300	100	100	100	100	100	100	100	100

表 3.10 当样本来自伽马分布时, 检验 \hat{T}_{M1} 在模型4下的经验功效(百分比)。

		$\theta = 0.4$				$\theta = 0.5$			
ϕ	N	$T = 50$	100	200	500	$T = 50$	100	200	500
-0.6	100	22.88	22.66	24.06	24.46	10.88	10.30	11.00	10.82
	200	53.06	56.92	57.20	58.86	18.28	18.88	19.16	19.58
	300	80.02	84.88	85.80	87.26	29.88	30.32	30.48	30.88
-0.3	100	7.50	6.86	6.62	6.70	22.80	23.92	24.58	24.46
	200	10.12	9.28	9.74	9.48	55.50	58.38	59.16	61.24
	300	12.92	11.80	12.44	11.88	82.72	86.72	88.00	88.78
0	100	56.32	58.98	62.56	61.76	97.12	98.46	98.60	99.02
	200	96.86	98.40	98.70	99.02	100	100	100	100
	300	100	99.98	100	100	100	100	100	100
0.3	100	99.20	99.64	99.62	99.82	100	100	100	100
	200	100	100	100	100	100	100	100	100
	300	100	100	100	100	100	100	100	100
0.6	100	100	100	100	100	100	100	100	100
	200	100	100	100	100	100	100	100	100
	300	100	100	100	100	100	100	100	100

§3.4 理论证明

本节给出定理3.2.1, 定理3.2.2和性质3.2.1的证明。

§3.4.1 定理3.2.1的证明

首先, 给出在定理3.2.1的证明中所需要的两个引理。

引理3.4.1. [Shumway和Stoffer (2006)^[60], p.508, 定理A.2] 对于 $n = 1, 2, \dots$ 和 $m = 1, 2, \dots$, 若 \mathbf{x}_n 和 \mathbf{y}_{mn} 为 $k \times 1$ 维的随机向量, 使得

(i) 对于每个 m , 当 $n \rightarrow \infty$ 时, $\mathbf{y}_{mn} \xrightarrow{d} \mathbf{y}_m$;

(ii) 当 $m \rightarrow \infty$ 时, $\mathbf{y}_m \xrightarrow{d} \mathbf{y}$;

(iii) 当 $m, n \rightarrow \infty$ 时, $E\{|\mathbf{x}_n - \mathbf{y}_{mn}|^2\} \rightarrow 0$,

则 $\mathbf{x}_n \xrightarrow{d} \mathbf{y}$ 。

引理3.4.2. 令 $\mathbf{x} = (x_1, x_2, \dots, x_T)^T$ 是均值为 0 且自协方差函数为 $\gamma(\cdot)$ 的严平稳 M 相依随机变量序列, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ 是 \mathbf{x} 的一组 *i.i.d.* 样本,

$$V_M = \sum_{u=-M}^M \gamma(\cdot).$$

若 $V_M \neq 0$, 则在假设3.2.1下,

$$(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T x_{ti} \xrightarrow{d} N(0, V_M).$$

证明. 这个引理的证明类似于Shumway和Stoffer (2006)^[60]中定理A.4的证明。首先构造一系列随机变量 y_{mT} 近似

$$(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T x_{ti},$$

然后验证引理3.4.1中的条件(i), (ii)和(iii)。对于 $m > 2M$, 令

$$\begin{aligned} y_{mT} &= (NT)^{-1/2} \sum_{i=1}^N [(x_{1i} + \dots + x_{m-M,i}) + (x_{m+1,i} + \dots + x_{2m-M,i}) \\ &\quad + (x_{2m+1,i} + \dots + x_{3m-M,i}) + \dots + (x_{(r-1)m+1,i} + \dots + x_{rm-M,i})] \\ &= (NT)^{-1/2} \sum_{i=1}^N (z_{1i} + z_{2i} + \dots + z_{ri}), \end{aligned}$$

其中 $r = [T/m]$, $[T/m]$ 是小于或等于 T/m 的最大整数。由于 $\mathbf{x} = (x_1, x_2, \dots, x_T)^T$ 是严平稳 M 相依随机变量序列, 因此 $\{z_{ji}, 1 \leq j \leq r, 1 \leq i \leq N\}$ 是独立同分布的, 其均值为 0, 方差为

$$S_{m-M} = \sum_{|u| < m-M} (m-M-|u|)\gamma(u).$$

接下来, 依次验证引理 3.4.1 中的条件 (i), (ii) 和 (iii)。

(i): 首先将 y_{mT} 改写成

$$y_{mT} = (NT)^{-1/2} \sum_{i=1}^N \sum_{j=1}^r z_{ji} = (T/r)^{-1/2} (rN)^{-1/2} \sum_{i=1}^N \sum_{j=1}^r z_{ji}.$$

因为 $(T/r)^{-1/2} \rightarrow m^{-1/2}$ 和

$$(rN)^{-1/2} \sum_{i=1}^N \sum_{j=1}^r z_{ji} \xrightarrow{d} N(0, S_{m-M}),$$

当 $T \rightarrow \infty$ 时, 由 Slutsky 定理可得

$$y_{mT} \xrightarrow{d} y_m \sim N(0, S_{m-M}/m).$$

(ii): 根据控制收敛定理, 当 $m \rightarrow \infty$ 时, $S_{m-M}/m \rightarrow V_M$ 。因此, 当 $m \rightarrow \infty$ 时, $y_m \xrightarrow{d} y \sim N(0, V_M)$ 。

(iii): 因为

$$\begin{aligned} (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T x_{ti} - y_{mT} &= (NT)^{-1/2} \sum_{i=1}^N [(x_{m-M+1,i} + \dots + x_{mi}) \\ &\quad + (x_{2m-M+1,i} + \dots + x_{2m,i}) \\ &\quad \vdots \\ &\quad + (x_{(r-1)m-M+1,i} + \dots + x_{(r-1)m,i}) \\ &\quad + (x_{rm-M+1,i} + \dots + x_{Ti})] \\ &= (NT)^{-1/2} \sum_{i=1}^N (\omega_{1i} + \omega_{2i} + \dots + \omega_{ri}), \end{aligned}$$

由此可得

$$E \left| (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T x_{ti} - y_{mT} \right|^2 = T^{-1} [(r-1)S_M + \text{Var}(\omega_r)],$$

其中

$$\text{Var}(\omega_r) = \sum_{|u| < n - [n/m]m + M} (n - [n/m]m + M - |u|)\gamma(u).$$

因此, 当 $T \rightarrow \infty$ 时,

$$\mathbb{E} \left| (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T x_{ti} - y_{mT} \right|^2 \rightarrow m^{-1} S_M.$$

又因为当 $m \rightarrow \infty$ 时, $m^{-1} S_M \rightarrow 0$, 这意味着条件(iii)是成立的。

至此, 引理3.4.2得证。 \square

定理3.2.1的证明分为以下三个步骤, 每个步骤的框架如下:

步骤1的框架. 首先证明对于 $k, \ell = 1, \dots, p$, $\hat{\Gamma}_{pk\ell} \xrightarrow{p} \sigma_e^2 \Sigma_{pk\ell}$, 其中 $\hat{\Gamma}_{pk\ell}$ 和 $\Sigma_{pk\ell}$ 分别代表矩阵 $\hat{\Gamma}_p$ 和 Σ_p 的第 (k, ℓ) 个元素。

步骤2的框架. 引入 ϕ 的一个新的估计量 $\tilde{\phi}$, 其定义如下

$$\tilde{\phi} = \tilde{\Gamma}_p^{-1} \tilde{\gamma}_p,$$

其中

$$\begin{aligned} \tilde{\Gamma}_p &= \left((NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T y_{t-k,i} y_{t-\ell,i} \right)_{k,\ell=1}^p, \\ \tilde{\gamma}_p &= \left((NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T y_{t-\ell,i} y_{ti} \right)_{\ell=1}^p. \end{aligned}$$

在这一步中, 将证明

$$\sqrt{NT}(\hat{\phi} - \tilde{\phi}) = o_p(1).$$

步骤3的框架. 根据步骤2, 为了得到 $\hat{\phi}$ 的渐近正态性, 只需证明

$$\sqrt{NT}(\tilde{\phi} - \phi) \xrightarrow{d} N(\mathbf{0}_p, \Sigma_p^{-1}). \quad (3.10)$$

接下来, 分别给出这三个步骤的详细过程。

步骤1: 根据 $\hat{\Gamma}_p$ 的定义, 可得

$$\hat{\Gamma}_{pk\ell} = (NT)^{-1} \sum_{i=1}^N \sum_{t=p+1}^T y_{t-k,i} y_{t-\ell,i}.$$

由于 $\mathbf{y}_1, \dots, \mathbf{y}_N$ 是来自平稳AR(p)过程 \mathbf{y} 的i.i.d.样本, 因此

$$E(\hat{\Gamma}_{pk\ell}) = (T-p)T^{-1}\sigma_e^2\gamma_0\rho_{|k-\ell|},$$

并且 $\mathbf{y}_i = \mathbf{Q}\mathbf{e}_i$, 其中 $\mathbf{e}_i = (e_{Ti}, e_{T-1,i}, e_{T-2,i}, \dots)^T$ 是一个 ∞ 维向量。对于 $k = 0, 1, \dots, p$, 令

$$\mathbf{y}_{ik} = \mathbf{Q}_k\mathbf{e}_i,$$

其中

$$\mathbf{Q}_k = \underbrace{\begin{pmatrix} 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & \psi_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \\ 0 & \cdots & 0 & 0 & 1 & \psi_1 & \cdots & \psi_{T-p-2} & \psi_{T-p-1} & \cdots \\ 0 & \cdots & 0 & 1 & \psi_1 & \psi_2 & \cdots & \psi_{T-p-1} & \psi_{T-p} & \cdots \end{pmatrix}}_k.$$

是一个 $(T-p) \times \infty$ 矩阵。由此可得

$$\hat{\Gamma}_{pk\ell} = (NT)^{-1} \sum_{i=1}^N \mathbf{y}_{ik}^T \mathbf{y}_{i\ell} = (NT)^{-1} \sum_{i=1}^N \mathbf{e}_i^T \mathbf{Q}_k^T \mathbf{Q}_\ell \mathbf{e}_i.$$

对于任意的 $k \in \{0, 1, \dots, p\}$, $\mathbf{Q}_k \mathbf{Q}_k^T = \sigma_e^2 \boldsymbol{\Sigma}_{T-p}$, 因此

$$\begin{aligned} & E(\hat{\Gamma}_{pk\ell} - E\hat{\Gamma}_{pk\ell})^2 \\ &= N^{-2}T^{-2}E\left[\sum_{i=1}^N (\mathbf{e}_i^T \mathbf{Q}_k^T \mathbf{Q}_\ell \mathbf{e}_i - \text{tr}(\mathbf{Q}_\ell \mathbf{Q}_k^T))\right]^2 \\ &= N^{-1}T^{-2}E(\mathbf{e}_1^T \mathbf{Q}_k^T \mathbf{Q}_\ell \mathbf{e}_1 - \text{tr}(\mathbf{Q}_\ell \mathbf{Q}_k^T))^2 \\ &= \sigma_e^4 N^{-1}T^{-2}[\text{tr}(\mathbf{Q}_k^T \mathbf{Q}_\ell \mathbf{Q}_k^T \mathbf{Q}_\ell) + \text{tr}(\mathbf{Q}_k^T \mathbf{Q}_\ell \mathbf{Q}_\ell^T \mathbf{Q}_k) + \beta_e \text{tr}(\mathbf{Q}_k^T \mathbf{Q}_\ell \circ \mathbf{Q}_k^T \mathbf{Q}_\ell)] \\ &\leq (2 + |\beta_e|)\sigma_e^4 N^{-1}T^{-2} \text{tr}(\boldsymbol{\Sigma}_{T-p}^2) = o(1). \end{aligned}$$

根据切比雪夫不等式(Chebyshev's inequality)可得

$$(NT)^{-1} \sum_{i=1}^N \sum_{t=p+1}^T y_{t-k,i} y_{t-\ell,i} = (T-p)T^{-1}\sigma_e^2\gamma_0\rho_{|k-\ell|} + o_p(1).$$

又因为 $(T-p)T^{-1}\sigma_e^2\gamma_0\rho_{|k-\ell|} = \sigma_e^2\gamma_0\rho_{|k-\ell|} + o(1)$, 所以有

$$(NT)^{-1} \sum_{i=1}^N \sum_{t=p+1}^T y_{t-k,i} y_{t-\ell,i} = \sigma_e^2\gamma_0\rho_{|k-\ell|} + o_p(1).$$

步骤2: 根据 $\hat{\phi}$ 和 $\tilde{\phi}$ 的定义,

$$\begin{aligned}\sqrt{NT}(\hat{\phi} - \tilde{\phi}) &= \sqrt{NT}(\hat{\Gamma}_p^{-1}\hat{\gamma}_p - \tilde{\Gamma}_p^{-1}\tilde{\gamma}_p) \\ &= \sqrt{NT}\hat{\Gamma}_p^{-1}(\hat{\gamma}_p - \tilde{\gamma}_p) + \sqrt{NT}(\hat{\Gamma}_p^{-1} - \tilde{\Gamma}_p^{-1})\tilde{\gamma}_p.\end{aligned}\quad (3.11)$$

类似步骤1中的证明, 对于 $k, \ell = 1, \dots, p$, 可以得到

$$\tilde{\Gamma}_{pk\ell} \xrightarrow{p} \sigma_e^2 \gamma_0 \rho_{|k-\ell|}, \quad \hat{\gamma}_{p\ell} \xrightarrow{p} \sigma_e^2 \gamma_0 \rho_\ell, \quad \tilde{\gamma}_{p\ell} \xrightarrow{p} \sigma_e^2 \gamma_0 \rho_\ell.$$

向量 $\sqrt{NT}(\hat{\gamma}_p - \tilde{\gamma}_p)$ 的第 ℓ 个分量为 $-(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^p y_{t-\ell,i} y_{ti}$, 模仿步骤1中的证明可得

$$-(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^p y_{t-\ell,i} y_{ti} = -py_N^{-1/2} \sigma_e^2 \gamma_0 \rho_\ell + o_p(1).$$

根据等式(3.2), 可得

$$\sqrt{NT}\hat{\Gamma}_p^{-1}(\hat{\gamma}_p - \tilde{\gamma}_p) = -py_N^{-1/2} \phi + o_p(1).\quad (3.12)$$

又因为

$$\sqrt{NT}(\hat{\Gamma}_p^{-1} - \tilde{\Gamma}_p^{-1}) = \sqrt{NT}\hat{\Gamma}_p^{-1}(\tilde{\Gamma}_p - \hat{\Gamma}_p)\tilde{\Gamma}_p^{-1}.$$

矩阵 $\sqrt{NT}(\tilde{\Gamma}_p - \hat{\Gamma}_p)$ 的第 (k, ℓ) 个元素为 $(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^p y_{t-k,i} y_{t-\ell,i}$, 仍然模仿步骤1中的证明, 可以得到

$$(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^p y_{t-k,i} y_{t-\ell,i} = py_N^{-1/2} \sigma_e^2 \gamma_0 \rho_{|k-\ell|} + o_p(1).$$

因此,

$$\sqrt{NT}(\hat{\Gamma}_p^{-1} - \tilde{\Gamma}_p^{-1})\tilde{\gamma}_p = py_N^{-1/2} \phi + o_p(1).\quad (3.13)$$

由(3.11), (3.12)和(3.13)可得

$$\sqrt{NT}(\hat{\phi} - \tilde{\phi}) = o_p(1).$$

步骤3: 对于 $i = 1, 2, \dots, N$,

$$y_{ti} = \phi_1 y_{t-1,i} + \phi_2 y_{t-2,i} + \dots + \phi_p y_{t-p,i} + e_{ti}, \quad t = 1, \dots, T,$$

将这些等式代入 $\tilde{\gamma}_p$ 中可得

$$\tilde{\phi} = \phi + \tilde{\Gamma}_p^{-1}(NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T \boldsymbol{\varepsilon}_{ti},$$

其中 $\boldsymbol{\varepsilon}_{ti} = e_{ti}(y_{t-1,i}, y_{t-2,i}, \dots, y_{t-p,i})^T$ 。由此可得

$$(NT)^{1/2}(\tilde{\phi} - \phi) = \tilde{\Gamma}_p^{-1}(NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T \boldsymbol{\varepsilon}_{ti}. \quad (3.14)$$

由于 e_{ti} 和 $y_{t-1,i}, y_{t-2,i}, y_{t-3,i}, \dots$ 是独立的, 因此

$$\mathbf{E}\boldsymbol{\varepsilon}_{ti} = \mathbf{0}, \quad \mathbf{E}(\boldsymbol{\varepsilon}_{ti}\boldsymbol{\varepsilon}_{ti}^T) = \sigma_e^4 \gamma_0 \boldsymbol{\Sigma}_p, \quad \mathbf{E}(\boldsymbol{\varepsilon}_{t+h,i}\boldsymbol{\varepsilon}_{ti}^T) = \mathbf{0},$$

其中 $h = \pm 1, \pm 2, \dots$ 。

根据表达式(3.14), 为了得到 $\sqrt{NT}(\tilde{\phi} - \phi)$ 的渐近分布, 只需考虑下列和的中心极限定理

$$S_T = (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T \boldsymbol{\lambda}^T \boldsymbol{\varepsilon}_{ti},$$

其中 $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p)^T$ 是任意常数向量。接下来, 利用引理3.4.1和引理3.4.2导出 S_T 的渐近分布。对于 $i = 1, 2, \dots, N$, 定义

$$y_{ii}^m = \sum_{\ell=0}^m \psi_{\ell} e_{t-\ell,i}$$

和 $(m+p)$ 相依过程 $\boldsymbol{\varepsilon}_{ii}^m = e_{ii}(y_{t-1,i}^m, y_{t-2,i}^m, \dots, y_{t-p,i}^m)^T$, 则 S_T 的近似值为

$$S_{mT} = (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T \boldsymbol{\lambda}^T \boldsymbol{\varepsilon}_{ii}^m.$$

首先, 对于固定的 m , 根据引理3.4.2, 可得 $S_{mT} \xrightarrow{d} S_m$, 其中 S_m 服从均值为0, 方差为 $\boldsymbol{\lambda}^T \boldsymbol{\Gamma}_p^{(m)} \boldsymbol{\lambda}$ 的正态分布, 这里 $\boldsymbol{\Gamma}_p^{(m)}$ 是 $\boldsymbol{\varepsilon}_{ii}^m$ 的协方差矩阵。其次, 当 $m \rightarrow \infty$ 时, $\boldsymbol{\lambda}^T \boldsymbol{\Gamma}_p^{(m)} \boldsymbol{\lambda} \rightarrow \sigma_e^4 \boldsymbol{\lambda}^T \boldsymbol{\Sigma}_p \boldsymbol{\lambda}$ 。因此, S_m 依分布收敛到均值为0, 方差为 $\sigma_e^4 \boldsymbol{\lambda}^T \boldsymbol{\Sigma}_p \boldsymbol{\lambda}$ 的正态分布, 即 S_{mT} 满足引理3.4.1中的条件(ii)。对于 $t = 1, 2, \dots, T$ 和 $i = 1, 2, \dots, N$,

$$y_{ti} - y_{ii}^m = \sum_{\ell=m+1}^{\infty} \psi_{\ell} e_{t-\ell,i},$$

因此, 当 $T, m \rightarrow \infty$ 时,

$$\begin{aligned} \mathbb{E}[(S_{mT} - S_T)^2] &= (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T \lambda^T \mathbb{E}[(\boldsymbol{\varepsilon}_{ti} - \boldsymbol{\varepsilon}_{ti}^m)(\boldsymbol{\varepsilon}_{ti} - \boldsymbol{\varepsilon}_{ti}^m)^T] \lambda \\ &= T^{-1} \sum_{t=1}^T \lambda^T \mathbb{E}[(\boldsymbol{\varepsilon}_{t1} - \boldsymbol{\varepsilon}_{t1}^m)(\boldsymbol{\varepsilon}_{t1} - \boldsymbol{\varepsilon}_{t1}^m)^T] \lambda \rightarrow 0. \end{aligned}$$

综上, 根据引理3.4.1, 可得 $S_T \xrightarrow{d} N(0, \sigma_e^4 \lambda^T \boldsymbol{\Sigma}_p \lambda)$ 。再根据(3.14)和Slutsky定理, 可得

$$\sqrt{NT}(\tilde{\boldsymbol{\phi}} - \boldsymbol{\phi}) \xrightarrow{d} N(\mathbf{0}_p, \boldsymbol{\Sigma}_p^{-1}).$$

至此, 定理3.2.1得证。

§3.4.2 定理3.2.2的证明

定理3.2.2的证明分为以下两个步骤, 每个步骤的框架如下:

步骤1的框架. 令

$$T_{Np} = \text{tr}\{\mathbf{B}_N \boldsymbol{\Sigma}_T^{-1} / [T^{-1} \text{tr}(\mathbf{B}_N \boldsymbol{\Sigma}_T^{-1})] - \mathbf{I}_T\}^2.$$

在这一步中, 将证明

$$\frac{T_{Np} - T y_N - (\beta_e + 1) y_N}{2 y_N} \xrightarrow{d} N(0, 1).$$

步骤2的框架. 根据步骤1和Slutsky定理, 为了得到 \hat{T}_{Np} 的渐近正态性, 只需验证 $\hat{T}_{Np} - T_{Np} = o_p(1)$ 。

接下来, 分别给出这两个步骤的详细过程。

步骤1: 在下面的证明中, 不失一般性, 假设 $\sigma_e^2 = 1$ 。回顾定理3.2.1的证明中的步骤1可知 $\mathbf{y}_i = \mathbf{Q} \mathbf{e}_i$, 其中 \mathbf{e}_i 是一个 ∞ 维向量。根据假设3.2.3, 分别将 \mathbf{Q} 和 \mathbf{e}_i 截断为 $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2)$ 和 $\mathbf{e}_i = \begin{pmatrix} \mathbf{e}_{1i} \\ \mathbf{e}_{2i} \end{pmatrix}$, 其中 \mathbf{Q}_1 是 $T \times k_T$ 矩阵, \mathbf{Q}_2 是 $T \times \infty$ 矩阵, \mathbf{e}_{1i} 是 k_T 维向

量和 \mathbf{e}_{2i} 是 ∞ 维向量。因此,

$$\begin{aligned}\mathbf{B}_N &= N^{-1} \sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^T = N^{-1} \sum_{i=1}^N (\mathbf{Q}_1 \mathbf{e}_{1i} + \mathbf{Q}_2 \mathbf{e}_{2i})(\mathbf{e}_{1i}^T \mathbf{Q}_1^T + \mathbf{e}_{2i}^T \mathbf{Q}_2^T) \\ &= N^{-1} \sum_{i=1}^N \mathbf{Q}_1 \mathbf{e}_{1i} \mathbf{e}_{1i}^T \mathbf{Q}_1^T + N^{-1} \sum_{i=1}^N \mathbf{Q}_1 \mathbf{e}_{1i} \mathbf{e}_{2i}^T \mathbf{Q}_2^T + N^{-1} \sum_{i=1}^N \mathbf{Q}_2 \mathbf{e}_{2i} \mathbf{e}_{1i}^T \mathbf{Q}_1^T \\ &\quad + N^{-1} \sum_{i=1}^N \mathbf{Q}_2 \mathbf{e}_{2i} \mathbf{e}_{2i}^T \mathbf{Q}_2^T \equiv \mathbf{B}_{11N} + \mathbf{B}_{12N} + \mathbf{B}_{21N} + \mathbf{B}_{22N}.\end{aligned}$$

由此可得

$$\begin{aligned}\text{tr}(\mathbf{B}_N \boldsymbol{\Sigma}_T^{-1}) &= \text{tr}(\mathbf{B}_{11N} \boldsymbol{\Sigma}_T^{-1}) + 2\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1}) + \text{tr}(\mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1}), \\ \text{tr}(\mathbf{B}_N \boldsymbol{\Sigma}_T^{-1})^2 &= \text{tr}(\mathbf{B}_{11N} \boldsymbol{\Sigma}_T^{-1})^2 + 4\text{tr}(\mathbf{B}_{11N} \boldsymbol{\Sigma}_T^{-1} \mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1}) + 2\text{tr}(\mathbf{B}_{11N} \boldsymbol{\Sigma}_T^{-1} \mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1}) \\ &\quad + 2\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1})^2 + 2\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1} \mathbf{B}_{21N} \boldsymbol{\Sigma}_T^{-1}) \\ &\quad + 4\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1} \mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1}) + \text{tr}(\mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1})^2.\end{aligned}$$

接下来, 将证明

$$\text{tr}(\mathbf{B}_N \boldsymbol{\Sigma}_T^{-1}) = \text{tr}(\mathbf{B}_{11N} \boldsymbol{\Sigma}_T^{-1}) + o_p(1), \quad (3.15)$$

$$\text{tr}(\mathbf{B}_N \boldsymbol{\Sigma}_T^{-1})^2 = \text{tr}(\mathbf{B}_{11N} \boldsymbol{\Sigma}_T^{-1})^2 + o_p(1). \quad (3.16)$$

为了得到(3.15), 需要证明

$$\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1}) = o_p(1), \quad \text{tr}(\mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1}) = o_p(1).$$

由于

$$\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1}) = N^{-1} \sum_{i=1}^N \mathbf{e}_{2i}^T \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1} \mathbf{Q}_1 \mathbf{e}_{1i}, \quad \text{tr}(\mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1}) = N^{-1} \sum_{i=1}^N \mathbf{e}_{2i}^T \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1} \mathbf{Q}_2 \mathbf{e}_{2i},$$

因此,

$$\mathbb{E}(\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1})) = 0, \quad \mathbb{E}(\text{tr}(\mathbf{B}_{22N} \boldsymbol{\Sigma}_T^{-1})) = \text{tr}(\mathbf{Q}_2 \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1}),$$

$$\text{Var}(\text{tr}(\mathbf{B}_{12N} \boldsymbol{\Sigma}_T^{-1})) = N^{-1} \text{tr}(\mathbf{Q}_1 \mathbf{Q}_1^T \boldsymbol{\Sigma}_T^{-1} \mathbf{Q}_2 \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1}).$$

根据 $\boldsymbol{\Sigma}_T = \mathbf{Q}_1 \mathbf{Q}_1^T + \mathbf{Q}_2 \mathbf{Q}_2^T$ 和Bushell和Trustrum (1990)^[61]中的(1.1), 可得

$$\text{tr}(\mathbf{Q}_1 \mathbf{Q}_1^T \boldsymbol{\Sigma}_T^{-1} \mathbf{Q}_2 \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1}) \leq \text{tr}(\mathbf{Q}_2 \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1}) \leq \lambda_{\max}(\boldsymbol{\Sigma}_T^{-1}) \text{tr}(\mathbf{Q}_2 \mathbf{Q}_2^T) = o(1). \quad (3.17)$$

因此, 由切比雪夫不等式可得

$$\text{tr}(\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1), \quad \text{tr}(\mathbf{B}_{22N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1).$$

接下来, 证明 $\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1)$ 。由于

$$\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1}) = N^{-2} \sum_{i=1}^N \sum_{j=1}^N \mathbf{e}_{1i}^T \mathbf{Q}_1^T \boldsymbol{\Sigma}_T^{-1} \mathbf{Q}_1 \mathbf{e}_{1j} \mathbf{e}_{2j}^T \mathbf{Q}_2^T \boldsymbol{\Sigma}_T^{-1} \mathbf{Q}_2 \mathbf{e}_{1i},$$

因此 $E(\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1})) = 0$ 。再根据 $\boldsymbol{\Sigma}_T = \mathbf{Q}_1\mathbf{Q}_1^T + \mathbf{Q}_2\mathbf{Q}_2^T$ 和Bushell和Trustrum (1990)^[61]中的(1.1) 可得

$$\text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1}\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1}\mathbf{Q}_2\mathbf{Q}_2^T\boldsymbol{\Sigma}_T^{-1}) = o(1), \quad (3.18)$$

$$\text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1}\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1}\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1}\mathbf{Q}_2\mathbf{Q}_2^T\boldsymbol{\Sigma}_T^{-1}) = o(1). \quad (3.19)$$

由此可得 $\text{Var}(\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1})) = o(1)$ 。仍然由切比雪夫不等式可得

$$\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1).$$

类似地, 还可以得到

$$\text{tr}(\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1})^2 = o_p(1), \quad \text{tr}(\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{22N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1),$$

$$\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{22N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1), \quad \text{tr}(\mathbf{B}_{12N}\boldsymbol{\Sigma}_T^{-1}\mathbf{B}_{21N}\boldsymbol{\Sigma}_T^{-1}) = o_p(1), \quad \text{tr}(\mathbf{B}_{22N}\boldsymbol{\Sigma}_T^{-1})^2 = o_p(1).$$

这样就得到了(3.16)。

令 $E_0(\cdot)$ 代表期望, $E_i(\cdot)$ 代表关于 $\mathbf{e}_{11}, \dots, \mathbf{e}_{1i}$ 生成的 σ -域的条件期望。经过计算可得

$$\begin{aligned} & \sum_{i=1}^N E_{i-1} \{[(E_i - E_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})]^2\} \\ &= (2 + \beta_e)N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^2 + o(1), \\ & \sum_{i=1}^N E_{i-1} \{[(E_i - E_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})][E_i - E_{i-1}]\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})^2\} \\ &= 2(2 + \beta_e)N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^3 + 2(2 + \beta_e)N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^2 + o_p(1), \\ & \sum_{i=1}^N E_{i-1} \{[(E_i - E_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})]^2\} \\ &= 4(2 + \beta_e)N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^4 + 8(2 + \beta_e)N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^3 \\ & \quad + 4(2 + \beta_e)N^{-2} \text{tr}^2(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})N^{-1} \text{tr}(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^2 + 4N^{-2} \text{tr}^2(\mathbf{Q}_1\mathbf{Q}_1^T\boldsymbol{\Sigma}_T^{-1})^2 + o_p(1). \end{aligned}$$

由(3.17), (3.18)和(3.19)可得

$$\begin{aligned} \sum_{i=1}^N \mathbf{E}_{i-1} \{[(\mathbf{E}_i - \mathbf{E}_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})]^2\} &= (2 + \beta_e)y_N + o(1), \\ \sum_{i=1}^N \mathbf{E}_{i-1} \{[(\mathbf{E}_i - \mathbf{E}_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})][(\mathbf{E}_i - \mathbf{E}_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})^2]\} &= 2(2 + \beta_e)(y_N + y_N^2) + o_p(1), \\ \sum_{i=1}^N \mathbf{E}_{i-1} \{[(\mathbf{E}_i - \mathbf{E}_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})^2]^2\} &= 4(2 + \beta_e)(y_N + 2y_N^2 + y_N^3) + 4y_N^2 + o_p(1). \end{aligned}$$

根据Chen et al. (2010)^[62]中性质A.3 的证明, 可得

$$\sum_{i=1}^N \mathbf{E} \left\{ [(\mathbf{E}_i - \mathbf{E}_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})]^4 \right\} = o(1), \quad \sum_{i=1}^N \mathbf{E} \left\{ [(\mathbf{E}_i - \mathbf{E}_{i-1})\text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})^2]^4 \right\} = o(1).$$

因此, 根据鞅差序列的中心极限定理和Slutsky定理可得

$$\begin{pmatrix} \text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1}) - \mu_1 \\ \text{tr}(\mathbf{B}_{11N}\boldsymbol{\Sigma}_T^{-1})^2 - \mu_2 \end{pmatrix} \xrightarrow{d} N \left(\mathbf{0}_2, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right),$$

其中

$$\begin{aligned} \mu_1 &= T, \quad \mu_2 = T(1 + y_N) + (1 + \beta_e)y_N, \quad \sigma_{11} = (2 + \beta_e)y, \\ \sigma_{12} = \sigma_{21} &= 2(2 + \beta_e)(y + y^2), \quad \sigma_{22} = 4(2 + \beta_e)(y + 2y^2 + y^3) + 4y^2. \end{aligned}$$

结合(3.15)和(3.16)可得

$$\begin{pmatrix} \text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1}) - \mu_1 \\ \text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})^2 - \mu_2 \end{pmatrix} \xrightarrow{d} N \left(\mathbf{0}_2, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right).$$

最后, 应用Delta方法, 可以得到

$$\frac{T_{Np} - Ty_N - (\beta_e + 1)y_N}{2y_N} \xrightarrow{d} N(0, 1).$$

步骤2: 根据 \hat{T}_{Np} 和 T_{Np} 的定义, 可得

$$\begin{aligned} \hat{T}_{Np} - T_{Np} &= \frac{\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})^2}{[T^{-1}\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})]^2} - \frac{\text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})^2}{[T^{-1}\text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})]^2} \\ &= \frac{\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})^2}{[T^{-1}\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})]^2} - \frac{\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})^2}{[T^{-1}\text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})]^2} + \frac{\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})^2 - \text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})^2}{[T^{-1}\text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})]^2}. \end{aligned}$$

由于 $\mathbf{Q}\mathbf{Q}^T = \boldsymbol{\Sigma}_T$ 和 $Ee_{it}^8 < \infty$, 模仿Zheng et al. (2019)^[63]中定理2.1的证明可得

$$T^{-1}\text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1}) = 1 + o_p(1), \quad (3.20)$$

$$T^{-1}\text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})^2 = 1 + y_N + o_p(1). \quad (3.21)$$

因此, 为了得到 $\hat{T}_{Np} - T_{Np} = o_p(1)$, 只需验证

$$\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1}) - \text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1}) = o_p(1), \quad (3.22)$$

$$\text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1})^2 - \text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1})^2 = o_p(1). \quad (3.23)$$

根据 $\boldsymbol{\Sigma}_T^{-1}$ 和 $\hat{\boldsymbol{\Sigma}}_T^{-1}$ 的表达式(3.4)和(3.5), 可得

$$\hat{\boldsymbol{\Sigma}}_T^{-1} - \boldsymbol{\Sigma}_T^{-1} = \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2)\mathbf{E}_j - \sum_{j=1}^p (\hat{\phi}_j - \phi_j)\mathbf{F}_j + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j})\mathbf{G}_{k,k+j}, \quad (3.24)$$

由此可得

$$\begin{aligned} \text{tr}(\mathbf{B}_N\hat{\boldsymbol{\Sigma}}_T^{-1}) - \text{tr}(\mathbf{B}_N\boldsymbol{\Sigma}_T^{-1}) &= \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2)\text{tr}(\mathbf{B}_N\mathbf{E}_j) - \sum_{j=1}^p (\hat{\phi}_j - \phi_j)\text{tr}(\mathbf{B}_N\mathbf{F}_j) \\ &\quad + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j})\text{tr}(\mathbf{B}_N\mathbf{G}_{k,k+j}). \end{aligned}$$

根据矩阵 \mathbf{E}_j , \mathbf{F}_j 和 $\mathbf{G}_{k,k+j}$ 的定义, 经过简单的计算可得

$$|\text{tr}(\boldsymbol{\Sigma}_T\mathbf{E}_j)| \leq T\gamma_0, \quad \text{tr}(\boldsymbol{\Sigma}_T\mathbf{E}_j)^2 \leq \text{tr}\boldsymbol{\Sigma}_T^2, \quad (3.25)$$

$$|\text{tr}(\boldsymbol{\Sigma}_T\mathbf{F}_j)| \leq 2T\gamma_0, \quad \text{tr}(\boldsymbol{\Sigma}_T\mathbf{F}_j)^2 \leq 4\text{tr}\boldsymbol{\Sigma}_T^2, \quad (3.26)$$

$$|\text{tr}(\boldsymbol{\Sigma}_T\mathbf{G}_{k,k+j})| \leq 2T\gamma_0, \quad \text{tr}(\boldsymbol{\Sigma}_T\mathbf{G}_{k,k+j})^2 \leq 4\text{tr}\boldsymbol{\Sigma}_T^2. \quad (3.27)$$

因此,

$$T^{-1}\text{tr}(\mathbf{B}_N\mathbf{E}_j) = T^{-1}\text{tr}(\boldsymbol{\Sigma}_T\mathbf{E}_j) + o_p(1),$$

$$T^{-1}\text{tr}(\mathbf{B}_N\mathbf{F}_j) = T^{-1}\text{tr}(\boldsymbol{\Sigma}_T\mathbf{F}_j) + o_p(1),$$

$$T^{-1}\text{tr}(\mathbf{B}_N\mathbf{G}_{k,k+j}) = T^{-1}\text{tr}(\boldsymbol{\Sigma}_T\mathbf{G}_{k,k+j}) + o_p(1).$$

结合(3.3), 可以得到

$$\begin{aligned}
 \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}(\mathbf{B}_N \Sigma_T^{-1}) &= \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) \text{tr}(\Sigma_T \mathbf{E}_j) - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) \text{tr}(\Sigma_T \mathbf{F}_j) \\
 &\quad + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) \text{tr}(\Sigma_T \mathbf{G}_{k,k+j}) + o_p(1) \\
 &= \text{tr}(\Sigma_T \hat{\Sigma}_T^{-1}) - T + o_p(1).
 \end{aligned}$$

再由方程组(3.2)和 γ_0 的定义, 当 $p+1 \leq t \leq T-p$ 时, 矩阵 $\Sigma_T \hat{\Sigma}_T^{-1}$ 的第 t 个对角线元素为

$$\begin{aligned}
 (\Sigma_T \hat{\Sigma}_T^{-1})_{tt} &= 1 + \gamma_0 \{ -\rho_1(\hat{\phi}_1 - \phi_1) - \rho_2(\hat{\phi}_2 - \phi_2) - \cdots - \rho_p(\hat{\phi}_p - \phi_p) \\
 &\quad + \hat{\phi}_1 [(\hat{\phi}_1 - \phi_1) + \rho_1(\hat{\phi}_2 - \phi_2) + \rho_2(\hat{\phi}_3 - \phi_3) \cdots + \rho_{p-1}(\hat{\phi}_p - \phi_p)] \\
 &\quad + \hat{\phi}_2 [\rho_1(\hat{\phi}_1 - \phi_1) + (\hat{\phi}_2 - \phi_2) + \rho_1(\hat{\phi}_3 - \phi_3) \cdots + \rho_{p-2}(\hat{\phi}_p - \phi_p)] \\
 &\quad \vdots \\
 &\quad + \hat{\phi}_p [\rho_{p-1}(\hat{\phi}_1 - \phi_1) + \rho_{p-2}(\hat{\phi}_2 - \phi_2) + \rho_{p-3}(\hat{\phi}_3 - \phi_3) \cdots + (\hat{\phi}_p - \phi_p)] \} \\
 &= 1 + \gamma_0 \{ (\hat{\phi}_1 - \phi_1) [(\hat{\phi}_1 - \phi_1) + \rho_1(\hat{\phi}_2 - \phi_2) + \cdots + \rho_{p-1}(\hat{\phi}_p - \phi_p)] \\
 &\quad + (\hat{\phi}_2 - \phi_2) [\rho_1(\hat{\phi}_1 - \phi_1) + (\hat{\phi}_2 - \phi_2) + \cdots + \rho_{p-2}(\hat{\phi}_p - \phi_p)] \\
 &\quad \vdots \\
 &\quad + (\hat{\phi}_p - \phi_p) [\rho_{p-1}(\hat{\phi}_1 - \phi_1) + \rho_{p-2}(\hat{\phi}_2 - \phi_2) + \cdots + (\hat{\phi}_p - \phi_p)] \},
 \end{aligned}$$

当 $t < p+1$ 或 $t > T-p$ 时, $(\Sigma_T \hat{\Sigma}_T^{-1})_{tt} = 1 + o_p(1)$ 。因此,

$$\text{tr}(\Sigma_T \hat{\Sigma}_T^{-1}) = T + o_p(1). \tag{3.28}$$

由此可得

$$\text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}(\mathbf{B}_N \Sigma_T^{-1}) = o_p(1).$$

为了验证(3.23), 只需证明

$$\begin{aligned}
 \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) &= o_p(1), \\
 \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) - \text{tr}(\mathbf{B}_N \Sigma_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) &= o_p(1).
 \end{aligned}$$

由(3.24)可得

$$\begin{aligned}
 & \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}(\mathbf{B}_N \Sigma_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) \\
 = & \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \mathbf{E}_j) - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \mathbf{F}_j) \\
 & + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \mathbf{G}_{k,k+j}).
 \end{aligned}$$

为了简化符号,在接下来的证明中,用 \mathbf{M}_1 和 \mathbf{M}_2 表示 \mathbf{E}_j , \mathbf{F}_j 或 $\mathbf{G}_{k,k+j}$ 。仍然由(3.24)可得

$$\begin{aligned}
 & T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \mathbf{M}_1) - T^{-1} \text{tr}(\mathbf{B}_N \Sigma_T^{-1} \mathbf{B}_N \mathbf{M}_1) \\
 = & \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) T^{-1} \text{tr}(\mathbf{B}_N \mathbf{M}_1 \mathbf{B}_N \mathbf{E}_j) - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) T^{-1} \text{tr}(\mathbf{B}_N \mathbf{M}_1 \mathbf{B}_N \mathbf{F}_j) \\
 & + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) T^{-1} \text{tr}(\mathbf{B}_N \mathbf{M}_1 \mathbf{B}_N \mathbf{G}_{k,k+j}).
 \end{aligned}$$

注意到 $Ee_{ii}^8 < \infty$ 和 $\text{tr} \Sigma_T^2 = o(T^2)$, 根据(3.25)-(3.27)和切比雪夫不等式, 可以得到 $T^{-2} \text{tr}(\mathbf{B}_N \mathbf{M}_1 \mathbf{B}_N \mathbf{M}_2) = o_p(1)$ 。这意味着

$$T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \mathbf{M}_1) - T^{-1} \text{tr}(\mathbf{B}_N \Sigma_T^{-1} \mathbf{B}_N \mathbf{M}_1) = o_p(1).$$

类似于(3.21)的证明,可以得到

$$T^{-1} \text{tr}(\mathbf{B}_N \Sigma_T^{-1} \mathbf{B}_N \mathbf{M}_1) = (1 + y_N) T^{-1} \text{tr}(\Sigma_T \mathbf{M}_1) + o_p(1), \quad (3.29)$$

由此可得

$$\begin{aligned}
 & \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) \\
 = & (1 + y_N) \left[\sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) \text{tr}(\Sigma_T \mathbf{E}_j) - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) \text{tr}(\Sigma_T \mathbf{F}_j) \right. \\
 & \left. + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) \text{tr}(\Sigma_T \mathbf{G}_{k,k+j}) \right] \\
 = & (1 + y_N) [\text{tr}(\Sigma_T \hat{\Sigma}_T^{-1}) - T] = o_p(1).
 \end{aligned}$$

同理, 根据(3.24), (3.28)和(3.29), 可得

$$\text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) - \text{tr}(\mathbf{B}_N \Sigma_T^{-1} \mathbf{B}_N \Sigma_T^{-1}) = o_p(1).$$

至此, 定理3.2.2得证。

§3.4.3 性质3.2.1的证明

首先将 $\hat{\beta}_e$ 改写成

$$\hat{\beta}_e = \frac{T^{-1} \hat{V}_e / [T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1})]^2 - 2}{T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_{\ell}^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_{\ell})^2}.$$

由(3.20)和(3.22)可得

$$T^{-1} \text{tr}(\mathbf{B}_N \hat{\Sigma}_T^{-1}) = \sigma_e^2 + o_p(1).$$

因此, 为了得到 $\hat{\beta}_e$ 的相合性, 只需要证明

$$T^{-1} \hat{V}_e = \sigma_e^4 \left[2 + \beta_e T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^T \Sigma_T^{-1} \mathbf{q}_{\ell})^2 \right] + o_p(1), \quad (3.30)$$

$$T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_{\ell}^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_{\ell})^2 = T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^T \Sigma_T^{-1} \mathbf{q}_{\ell})^2 + o_p(1). \quad (3.31)$$

令

$$V_e = (N-1)^{-1} \sum_{i=1}^N \left(\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i - N^{-1} \sum_{i=1}^N \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i \right)^2,$$

将(3.30)的证明分为以下两步

$$T^{-1} \hat{V}_e = T^{-1} V_e + o_p(1), \quad (3.32)$$

$$T^{-1} V_e = \sigma_e^4 \left[2 + \beta_e T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^T \Sigma_T^{-1} \mathbf{q}_{\ell})^2 \right] + o_p(1). \quad (3.33)$$

注意到

$$\begin{aligned} T^{-1} \hat{V}_e - T^{-1} V_e &= T^{-1} (N-1)^{-1} \sum_{i=1}^N [(\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i)^2 - (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2] \\ &\quad - T^{-1} (N-1)^{-1} N \left[(N^{-1} \sum_{i=1}^N \mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i)^2 - (N^{-1} \sum_{i=1}^N \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2 \right] \\ &= T^{-1} (N-1)^{-1} \sum_{i=1}^N [(\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i)^2 - (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2] \\ &\quad - T^{-1} (N-1)^{-1} N [\text{tr}^2(\mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}^2(\mathbf{B}_N \Sigma_T^{-1})]. \end{aligned}$$

仍然由(3.20)和(3.22)可得

$$T^{-1}[\text{tr}^2(\mathbf{B}_N \hat{\Sigma}_T^{-1}) - \text{tr}^2(\mathbf{B}_N \Sigma_T^{-1})] = o_p(1).$$

此外, 还有

$$\begin{aligned} & T^{-1}(N-1)^{-1} \sum_{i=1}^N [(\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i)^2 - (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2] \\ = & T^{-1}(N-1)^{-1} \sum_{i=1}^N (\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i - \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2 \\ & + 2T^{-1}(N-1)^{-1} \sum_{i=1}^N \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i (\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i - \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i). \end{aligned}$$

根据(3.24), 可得

$$\begin{aligned} \mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i - \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i &= \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) \mathbf{y}_i^T \mathbf{E}_j \mathbf{y}_i - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) \mathbf{y}_i^T \mathbf{F}_j \mathbf{y}_i \\ &+ \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) \mathbf{y}_i^T \mathbf{G}_{k,k+j} \mathbf{y}_i. \end{aligned}$$

为了简化符号, 在接下来的证明中, 用 \mathbf{M}_1 和 \mathbf{M}_2 表示 \mathbf{E}_j , \mathbf{F}_j 或 $\mathbf{G}_{k,k+j}$ 。由于 $Ee_{it}^8 < \infty$, 根据(3.25)-(3.27)和切比雪夫不等式, 可以得到

$$T^{-2}(N-1)^{-1} \sum_{i=1}^N \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i \mathbf{y}_i^T \mathbf{M}_1 \mathbf{y}_i = \sigma_e^4 T^{-1} \text{tr}(\Sigma_T \mathbf{M}_1) + o_p(1). \quad (3.34)$$

这意味着

$$\begin{aligned} & T^{-1}(N-1)^{-1} \sum_{i=1}^N \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i (\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i - \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i) \\ = & \sigma_e^4 \left[\sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) \text{tr}(\Sigma_T \mathbf{E}_j) - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) \text{tr}(\Sigma_T \mathbf{F}_j) \right. \\ & \left. + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) \text{tr}(\Sigma_T \mathbf{G}_{k,k+j}) \right] \\ = & \sigma_e^4 [\text{tr}(\Sigma_T \hat{\Sigma}_T^{-1}) - T] + o_p(1) = o_p(1). \end{aligned}$$

类似于(3.34)的证明, 可以得到

$$T^{-3}(N-1)^{-1} \sum_{i=1}^N \mathbf{y}_i^T \mathbf{M}_1 \mathbf{y}_i \mathbf{y}_i^T \mathbf{M}_2 \mathbf{y}_i = o_p(1), \quad (3.35)$$

由此可得

$$T^{-1}(N-1)^{-1} \sum_{i=1}^N (\mathbf{y}_i^T \hat{\Sigma}_T^{-1} \mathbf{y}_i - \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2 = o_p(1).$$

这样就完成了(3.32)的证明。接下来, 验证(3.33)。经过简单的计算可得

$$\begin{aligned} T^{-1}V_e &= (TN)^{-1} \sum_{i=1}^N (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i)^2 - (TN)^{-1}(N-1)^{-1} \sum_{i \neq j} \mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i \mathbf{y}_j^T \Sigma_T^{-1} \mathbf{y}_j \\ &= (TN)^{-1} \sum_{i=1}^N (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i - T\sigma_e^2)^2 \\ &\quad - (TN)^{-1}(N-1)^{-1} \sum_{i \neq j} (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i - T\sigma_e^2)(\mathbf{y}_j^T \Sigma_T^{-1} \mathbf{y}_j - T\sigma_e^2). \end{aligned}$$

根据Bai和Silverstein (2004)^[55]中等式(1.15)和切比雪夫不等式可得

$$(TN)^{-1} \sum_{i=1}^N (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i - T\sigma_e^2)^2 = \sigma_e^4 \left[2 + \beta_e T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell)^2 \right] + o_p(1)$$

和

$$(TN)^{-1}(N-1)^{-1} \sum_{i \neq j} (\mathbf{y}_i^T \Sigma_T^{-1} \mathbf{y}_i - T\sigma_e^2)(\mathbf{y}_j^T \Sigma_T^{-1} \mathbf{y}_j - T\sigma_e^2) = o_p(1),$$

由此可得(3.33)。

令

$$\gamma_0 = \sum_{j=0}^{\infty} \psi_j^2, \quad \gamma_\ell = \sum_{j=0}^{\infty} \hat{\psi}_j \psi_j, \quad \gamma_r = \sum_{j=0}^{\infty} \psi_j \hat{\psi}_j, \quad \gamma_{\ell r} = \sum_{j=0}^{\infty} \hat{\psi}_j \hat{\psi}_j.$$

由于 $\hat{\boldsymbol{\phi}} = (\hat{\phi}_1, \dots, \hat{\phi}_p)^T$ 是 $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)^T$ 的 T 相合估计, 因此

$$\gamma_\ell = \gamma_0 + o_p(1), \quad \gamma_r = \gamma_0 + o_p(1), \quad \gamma_{\ell r} = \gamma_0 + o_p(1).$$

注意到, 由柯西-施瓦茨不等式(Cauchy-Schwarz inequality)可得

$$\begin{aligned} & T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell)^2 - T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 \\ &= T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell - \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)(\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell + \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell) \\ &\leq \sqrt{T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell - \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell + \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2}. \end{aligned}$$

由于 \mathbf{E}_j 和 \mathbf{F}_j 是谱模有界的并且 $\mathbf{q}_\ell^T \mathbf{q}_\ell \leq \gamma_0$ 和 $\hat{\mathbf{q}}_\ell^T \hat{\mathbf{q}}_\ell \leq \gamma_{lr}$, 因此

$$\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell = O_p(\lambda_{\max}(\hat{\Sigma}_T^{-1}) \hat{\mathbf{q}}_\ell^T \hat{\mathbf{q}}_\ell) = O_p(1), \quad \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell = O_p(\lambda_{\max}(\hat{\Sigma}_T^{-1}) \mathbf{q}_\ell^T \mathbf{q}_\ell) = O_p(1).$$

再由(3.28)可得

$$T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell + \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 = O_p(1).$$

再根据柯西-施瓦茨不等式, 可得

$$\begin{aligned} & T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell - \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 \\ &= T^{-1} \sum_{\ell=1}^{\infty} [(\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell)^T \hat{\Sigma}_T^{-1} (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell) + 2\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell)]^2 \\ &\leq 2T^{-1} \sum_{\ell=1}^{\infty} [(\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell)^T \hat{\Sigma}_T^{-1} (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell)]^2 + 8T^{-1} \sum_{\ell=1}^{\infty} \mathbf{q}_\ell^T \hat{\Sigma}_T^{-2} \mathbf{q}_\ell (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell)^T (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell). \end{aligned}$$

又因为

$$\begin{aligned} T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell)^T (\hat{\mathbf{q}}_\ell - \mathbf{q}_\ell) &= T^{-1} [\text{tr}(\hat{\mathbf{Q}}^T \hat{\mathbf{Q}}) - \text{tr}(\hat{\mathbf{Q}}^T \mathbf{Q}) - \text{tr}(\mathbf{Q}^T \hat{\mathbf{Q}}) + \text{tr}(\mathbf{Q}^T \mathbf{Q})] \\ &= \gamma_{lr} - \gamma_\ell - \gamma_r + \gamma_0 = o_p(1), \end{aligned}$$

由此可得

$$T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell - \mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 = o_p(1).$$

因此,

$$T^{-1} \sum_{\ell=1}^{\infty} (\hat{\mathbf{q}}_\ell^T \hat{\Sigma}_T^{-1} \hat{\mathbf{q}}_\ell)^2 - T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 = o_p(1).$$

仍然根据柯西-施瓦茨不等式, 可得

$$\begin{aligned} & T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell)^2 - T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell)^2 \\ &= T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell - \mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell) (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell + \mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell) \\ &\leq \sqrt{T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell - \mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell)^2 T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_\ell^T \hat{\Sigma}_T^{-1} \mathbf{q}_\ell + \mathbf{q}_\ell^T \Sigma_T^{-1} \mathbf{q}_\ell)^2}. \end{aligned}$$

与前面的讨论类似, 可以证明

$$T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^{\top} \hat{\Sigma}_T^{-1} \mathbf{q}_{\ell} + \mathbf{q}_{\ell}^{\top} \Sigma_T^{-1} \mathbf{q}_{\ell})^2 = O_p(1).$$

由(3.24)可得

$$\begin{aligned} \mathbf{q}_{\ell}^{\top} \hat{\Sigma}_T^{-1} \mathbf{q}_{\ell} - \mathbf{q}_{\ell}^{\top} \Sigma_T^{-1} \mathbf{q}_{\ell} &= \sum_{j=1}^p (\hat{\phi}_j^2 - \phi_j^2) \mathbf{q}_{\ell}^{\top} \mathbf{E}_j \mathbf{q}_{\ell} - \sum_{j=1}^p (\hat{\phi}_j - \phi_j) \mathbf{q}_{\ell}^{\top} \mathbf{F}_j \mathbf{q}_{\ell} \\ &\quad + \sum_{j=1}^{p-1} \sum_{k=1}^{p-j} (\hat{\phi}_k \hat{\phi}_{k+j} - \phi_k \phi_{k+j}) \mathbf{q}_{\ell}^{\top} \mathbf{G}_{k,k+j} \mathbf{q}_{\ell}. \end{aligned}$$

再根据(3.25)-(3.27), 可得

$$\begin{aligned} T^{-3} \sum_{\ell=1}^{\infty} \mathbf{q}_{\ell}^{\top} \mathbf{M}_1 \mathbf{q}_{\ell} \mathbf{q}_{\ell}^{\top} \mathbf{M}_2 \mathbf{q}_{\ell} &\leq \sqrt{T^{-3} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^{\top} \mathbf{M}_1 \mathbf{q}_{\ell})^2 T^{-3} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^{\top} \mathbf{M}_2 \mathbf{q}_{\ell})^2} \\ &\leq \sqrt{T^{-3} \text{tr}(\Sigma_T \mathbf{M}_1)^2 T^{-3} \text{tr}(\Sigma_T \mathbf{M}_2)^2} = o(1). \end{aligned}$$

这意味着

$$\sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^{\top} \hat{\Sigma}_T^{-1} \mathbf{q}_{\ell} - \mathbf{q}_{\ell}^{\top} \Sigma_T^{-1} \mathbf{q}_{\ell})^2 = o_p(1),$$

由此可得

$$T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^{\top} \hat{\Sigma}_T^{-1} \mathbf{q}_{\ell})^2 - T^{-1} \sum_{\ell=1}^{\infty} (\mathbf{q}_{\ell}^{\top} \Sigma_T^{-1} \mathbf{q}_{\ell})^2 = o_p(1).$$

因此, (3.31)成立。

至此, 性质3.2.1得证。

第四章 总结和未来研究工作

§4.1 总结

本文主要考虑两个高维假设检验问题。第一个是检验两组高维数据的总体协方差矩阵是否相等；第二个是通过检验一组高维数据的总体协方差矩阵结构，来检验这组数据是否来自平稳向量自回归过程。对于上述两个高维假设检验问题，本文基于样本协方差矩阵，分别提出了新的检验方法。新的检验方法不仅适用于高维数据，而且无需假定样本的分布。接下来，对本文所做的工作做一个简单的总结。

在第二章中，为了检验两组高维数据的总体协方差矩阵是否相等，本文基于新的目标函数提出了检验统计量 T_r 。利用统计量 T_r 和原有的统计量 T_d ，本文给出了第一个检验方法 T_{dr} ， T_{dr} 的表现接近 T_d 和 T_r 中表现较好的那一个。由于检验 T_d ，检验 T_r 和检验 T_{dr} 仅适用于稠密备择假设，对于稀疏备择假设，它们有时会失效。通过结合 T_d ， T_r 和极值统计量 T_x ，本文又提出了另外两个检验方法 T_{drx}^w 和 T_{drx}^m ，这两个方法在一般的备择假设下都可以保持较高的检验功效。在原假设和三个特殊的备择假设下，本文对这三个检验方法的理论性质进行了广泛的研究，从而阐明了将多个统计量用不同方法组合的优缺点。大量的数值模拟结果显示，新的检验方法的功效会等于或优于现有的方法。尤其是在既不稠密也不稀疏的备择假设下，检验方法 T_{drx}^w 和 T_{drx}^m 具有明显的优越性。

在第三章中，本文检验一组时间序列数据是否来自VAR模型(3.1)。在原假设 H_{0p} 下，可以将这组数据看作是来自平稳AR(p)过程的独立同分布的样本，而平稳AR(p)过程的协方差矩阵等于噪声的方差乘以一个由回归系数唯一确定的矩阵。因此，可以通过检验这组数据的总体协方差矩阵结构来检验 H_{0p} 。首先，本文给出了回归系数向量 ϕ 的最小二乘估计 $\hat{\phi}$ 和 $\hat{\phi}$ 的渐近正态性。然后，利用估计值 $\hat{\phi}$ 构造了检验统计量 \hat{T}_{Np} ，并且通过 \hat{T}_{Np} 在原假设 H_{0p} 下的渐近分布给出了一个检验方法，这个方法对正态分布和非正态分布的样本都是适用的。数值模拟结果显示，在大部分的备择假设下，新的检验方法都具有较高的经验功效，这表明新的检验方法适用范围比较广泛，表现比较稳定。

§4.2 未来研究工作

虽然本文对上述两个高维假设检验问题进行了研究，取得了阶段性的成果，但是仍有许多不足之处。因此，后续的未来研究工作如下：

在两组高维数据的协方差矩阵检验问题中，本文提出的检验方法要求一组数据的样本量大于数据的维数，这个要求缩小了这些方法的使用范围。在下一步研究中，我们的目标是修正第二章中提出的统计量，进而得到新的检验方法，这些方法在两组数据的样本量均可以小于维数的假设下，也能具有较高的检验功效。除此之外，本文的第二章只考虑了检验两组数据的协方差矩阵，然而在许多实际应用中，需要检验多组数据的协方差矩阵是否相等。因此，我们应该思考如何将第二章中的检验方法推广，进而提出有效的方法来检验多组数据的协方差矩阵是否相等。

在第三章中，本文考虑的是检验一组时间序列数据是否来自特殊的VAR模型，该模型中的系数矩阵是对角线元素一致的对角矩阵。检验数据是否来自系数矩阵更一般的VMA模型，VAR模型或VARMA模型是具有挑战性的工作。如何对高维时间序列数据进行统计推断，这将是今后要研究的一个方向。

参考文献

- [1] Mauchly J W. Significance test for sphericity of a normal n -variate distribution[J]. The Annals of Mathematical Statistics, 1940, 11(2): 204-209.
- [2] John S. Some optimal multivariate tests[J]. Biometrika, 1971, 58(1): 123-127.
- [3] Ledoit O, Wolf M. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size[J]. The Annals of Statistics, 2002, 30(4): 1081-1102.
- [4] Wang Q W, Yao J F. On the sphericity test with large-dimensional observations[J]. Electronic Journal of Statistics, 2013, 7: 2164-2192.
- [5] Fisher R A. The statistical utilization of multiple measurements[J]. The Annals of Eugenics, 1938, 8: 376-386.
- [6] Wald A. On a statistical problem arising in the classification of an individual into one of two groups[J]. The Annals of Mathematical Statistics, 1944, 15(2): 145-162.
- [7] Friedman J H. Regularized discriminant analysis[J]. Journal of the American Statistical Association, 1989, 84(405): 165-175.
- [8] Guo Y Q, Hastie T, Tibshirani R. Regularized linear discriminant analysis and its application in microarrays[J]. Biostatistics, 2007, 8(1): 86-100.
- [9] Xu P, Brock G N, Parrish R S. Modified linear discriminant analysis approaches for classification of high-dimensional microarray data[J]. Computational Statistics & Data Analysis, 2009, 53(5): 1674-1687.
- [10] Witten D M, Tibshirani R. Penalized classification using Fisher's linear discriminant[J]. Journal of the Royal Statistical Society: Series B, 2011, 73(5): 753-772.
- [11] Bai Z D, Jiang D D, Yao J F, et al. Corrections to LRT on large-dimensional covariance matrix by RMT[J]. The Annals of Statistics, 2009, 37(6B): 3822-3840.
- [12] Zheng S R. Central limit theorems for linear spectral statistics of large dimensional F-matrices[J]. Annales de l'Institut Henri Poincaré Probabilités Et Statistiques, 2012, 48(2): 444-476.
- [13] Jiang T F, Yang F. Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions[J]. The Annals of Statistics, 2013, 41(4): 2029-2074.
- [14] Zhang Q Y, Hu J, Bai Z D. Invariant test based on the modified correction to LRT for the equality of two high-dimensional covariance matrices[J]. Electronic Journal of Statistics, 2019, 13(1): 850-881.
- [15] Wilks S S. Certain generalizations in the analysis of variance[J]. Biometrika, 1932, 24(3): 471-494.

- [16] Pillai K C S. On some distribution problems in multivariate analysis[D]: [PhD thesis]. Chapel Hill: University of North Carolina, 1954.
- [17] Zhang Q Y, Hu J, Bai Z D. Modified Pillai's trace statistics for two high-dimensional sample covariance matrices[J]. *Journal of Statistical Planning and Inference*, 2020, 207: 255-275.
- [18] Bai Z D, Hu J, Pan G M, et al. Convergence of the empirical spectral distribution function of Beta matrices[J]. *Bernoulli*, 2015, 21(3): 1538-1574.
- [19] Schott J R. A test for the equality of covariance matrices when the dimension is large relative to the sample sizes[J]. *Computational Statistics & Data Analysis*, 2007, 51(12): 6535-6542.
- [20] Srivastava M S, Yanagihara H. Testing the equality of several covariance matrices with fewer observations than the dimension[J]. *Journal of Multivariate Analysis*, 2010, 101(6): 1319-1329.
- [21] Li J, Chen S X. Two sample tests for high-dimensional covariance matrices[J]. *The Annals of Statistics*, 2012, 40(2): 908-940.
- [22] Cai T, Liu W D, Xia Y. Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings[J]. *Journal of the American Statistical Association*, 2013, 108(501): 265-277.
- [23] Yang Q, Pan G M. Weighted statistic in detecting faint and sparse alternatives for high-dimensional covariance matrices[J]. *Journal of the American Statistical Association*, 2017, 112(517): 188-200.
- [24] Zheng S R, Lin R T, Guo J H, et al. Testing homogeneity of high-dimensional covariance matrices[J]. *Statistica Sinica*, 2020, 30(1): 35-53.
- [25] Jin B S, Miao B Q, Ye W Y, et al. The application of spectral distribution of product of two random matrices in the factor analysis[J]. *Science in China Series A: Mathematics*, 2007, 50(9): 1303-1315.
- [26] Forni M, Lippi M. Aggregation of linear dynamic microeconomic models[J]. *Journal of Mathematical Economics*, 1999, 31(1): 131-158.
- [27] Forni M, Hallin M, Lippi M, et al. The generalized dynamic factor model: identification and estimation[J]. *The Review of Economics and Statistics*, 2000, 82(4): 540-554.
- [28] Forni M, Hallin M, Lippi M, et al. The generalized dynamic factor model consistency and rates[J]. *Journal of Econometrics*, 2004, 119(2): 231-255.
- [29] Bai Z D, Zhou W. Large sample covariance matrices without column independence structure[J]. *Statistica Sinica*, 2008, 18(2): 425-442.

- [30] Jin B S, Wang C, Miao B Q, et al. Limiting spectral distribution of large-dimensional sample covariance matrices generated by VARMA[J]. *Journal of Multivariate Analysis*, 2009, 100(9): 2112-2125.
- [31] Wang C, Jin B S, Miao B Q. On limiting spectral distribution of large sample covariance matrices by VARMA(p,q)[J]. *Journal of Time Series Analysis*, 2011, 32(5): 539-546.
- [32] Pfaffel O, Schlemm E. Eigenvalue distribution of large sample covariance matrices of linear processes[J]. *Probability & Mathematical Statistics*, 2011, 31(2): 313-329.
- [33] Yao J F. A note on a marčenko-pastur type theorem for time series[J]. *Statistics & Probability Letters*, 2012, 82(1): 22-28.
- [34] Jin B S, Wang C, Bai Z D, et al. Limiting spectral distribution of a symmetrized auto-cross covariance matrix[J]. *The Annals of Applied Probability*, 2014, 24(3): 1199-1225.
- [35] Liu H Y, Aue A, Paul D. On the marčenko-pastur law for linear time series[J]. *The Annals of Statistics*, 2015, 43(2): 675-712.
- [36] Bhattacharjee M, Bose A. Large sample behaviour of high dimensional autocovariance matrices[J]. *The Annals of Statistics*, 2016, 44(2): 598-628.
- [37] Bhattacharjee M, Bose A. Joint convergence of sample autocovariance matrices when $p/n \rightarrow 0$ with application[J]. *The Annals of Statistics*, 2019, 47(6): 3470-3503.
- [38] 张阳春. 几类高维随机矩阵模型的极限谱行为及其统计应用[D]: [博士学位论文]. 哈尔滨: 哈尔滨工业大学数学学院, 2020.
- [39] Sugiura N, Nagao H. Unbiasedness of some test criteria for the equality of one or two covariance matrices[J]. *The Annals of Mathematical Statistics*, 1968, 39(5): 1686-1692.
- [40] Gupta D S, Giri N. Properties of tests concerning covariance matrices of normal distributions[J]. *The Annals of Statistics*, 1973, 1(6): 1222-1224.
- [41] Perlman M D. Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations[J]. *The Annals of Statistics*, 1980, 8(2): 247-263.
- [42] Gupta A K, Tang J. Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate gaussian models[J]. *Biometrika*, 1984, 71(3): 555-559.
- [43] O'Brien P C. Robust procedures for testing equality of covariance matrices[J]. *Biometrics*, 1992, 48(3): 819-827.
- [44] Anderson T W. An introduction to multivariate statistical analysis[M]. 3rd ed. New York: John Wiley & Sons, 2003.
- [45] Chang J Y, Zhou W, Zhou W X, et al. Comparing large covariance matrices under weak conditions on the dependence structure and its application to gene clustering[J]. *Biomet-*

- rics, 2017, 73(1): 31-41.
- [46] Zhu L X, Lei J, Devlin B, et al. Testing high dimensional differential matrices, with application to detecting schizophrenia risk genes[J]. *The Annals of Applied Statistics*, 2017, 11(3): 1810-1831.
- [47] Han F, Xu S, Zhou W X. On Gaussian comparison inequality and its application to spectral analysis of large random matrices[J]. *Bernoulli*, 2018, 24(3): 1787-1833.
- [48] Chen J, Wang X Y, Zheng S R, et al. Tests for high dimensional covariance matrices[J]. *Random Matrices: Theory and Applications*, 2019, 9(3): 1-26.
- [49] 何冰, 薄晓玲. 基于随机 F -矩阵的高维双样本协方差矩阵相等性检验[J]. *吉林大学学报(理学版)*, 2019, 57(1): 65-71.
- [50] Westberg M. Combining independent statistical tests[J]. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 1985, 34(3): 287-296.
- [51] Heard N A, Rubin-Delanchy P. Choosing between methods for combining p -values[J]. *Biometrika*, 2018, 105(1), 239-246.
- [52] Billingsley P. *Probability and measure*[M]. 3rd ed. New York: John Wiley & Sons, 1995.
- [53] Zheng S R, Bai Z D, Yao J F. Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing[J]. *The Annals of Statistics*, 2015, 43(2): 546-591.
- [54] Wang Q W, Yao J F. A note on the CLT of the LSS for sample covariance matrix from a spiked population model[J]. *Journal of Multivariate Analysis*, 2014, 130: 194-207.
- [55] Bai Z D, Silverstein J W. CLT for linear spectral statistics of large-dimensional sample covariance matrices[J]. *The Annals of Probability*, 2004, 32(1): 553-605.
- [56] Silverstein J W. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices[J]. *Journal of Multivariate Analysis*, 1995, 55(2): 331-339.
- [57] Bai Z D, Silverstein J W. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices[J]. *The Annals of Probability*, 1998, 26(1): 316-345.
- [58] Cryer J D, Chan K-S. *Time series analysis with applications in R*[M]. 2nd ed. New York: Springer, 2008.
- [59] Verbyla A P. A note on the inverse covariance matrix of the autoregressive process[J]. *Australian Journal of Statistics*, 1985, 27(2): 221-224.
- [60] Shumway R H, Stoffer D S. *Time series analysis and its applications with R examples*[M]. 2nd ed. New York: Springer, 2006.
- [61] Bushell P J, Trustrum G B. Trace inequalities for positive definite matrix power product-

- s[J]. *Linear Algebra & Its Applications*, 1990, 132: 173-178.
- [62] Chen S X, Zhang L X, Zhong P S. Tests for high-dimensional covariance matrices[J]. *Journal of the American Statistical Association*, 2010, 105(490): 810-819.
- [63] Zheng S R, Chen Z, Cui H J, et al. Hypothesis testing on linear structures of high-dimensional covariance matrix[J]. *The Annals of Statistics*, 2019, 47(6): 3300-3334.

致谢

随着博士论文的撰写工作接近尾声，这意味着我在师大十一年的学习生活即将结束。回想过去十一年的点点滴滴，我得到了许多师长、同学和朋友的关怀与帮助，在此谨向他们表示最诚挚的谢意！

首先，特别感谢我的导师郑术蓉老师，我的每一步成长都离不开郑老师的悉心指导。在我刚开始读研时，郑老师告诉我一定要好好学习基础课，尤其是概率论和数理统计。于是在两年的硕士研究生阶段，我格外注重基础课的学习，这为我以后的博士生涯打下了坚实的基础。当我步入博士研究生学习阶段，在郑老师的帮助和指导下，我慢慢学会怎么找问题、怎么解决问题、怎么编写程序以及如何写文章，虽然这个过程并不是一帆风顺，但郑老师总会安慰我和鼓励我，让我每次都有勇气可以直面困难。此外，郑老师也经常鼓励我出去开会、做报告，推荐我去别的学校交流访问，这不仅让我认识了一些优秀的老师和青年学者，还开拓了我的眼界。正是郑老师的因材施教和言传身教，这些年来，我在学习态度、学习习惯和为人处世等方面都有很大的提高。

感谢在学业上给予我指导的白志东老师。每次去参加白老师的讨论班，总会被他老人家渊博的学识，思考问题的角度和强大的记忆力震撼到。每次有想不通的问题去问白老师时，他总是放下手头的工作，耐心地为我解答。白老师似乎有种“魔力”，每次和他老人家讨论完问题，总会有种茅塞顿开的感觉。

感谢统计系的史宁中老师，郭建华老师，高巍老师，朱文圣老师，孙法省老师，刘秉辉老师等多位老师，还有已故的陶剑老师长期以来的谆谆教诲，从他们的身上我感受到了师大的校训“勤奋创新，为人师表”。感谢南方科技大学的田国梁老师为我提供交流的机会。感谢美国MD安德森癌症中心的林瑞涛老师对我文章的指导。

感谢统计系的张超师兄，王东营师姐，王昊师兄，陆婧师姐，陈晶师妹，周怡杉师妹，梁春辉师妹，赵凯鸽师弟，赖佳钰师妹，李铭师妹等的帮助和陪伴，尤其是陈晶师妹，每次有事求她帮忙，她都毫不犹豫地帮助我。希望各位师弟师妹的科研之路更加顺利。

感谢我的室友兼同学王潇逸，孟静波，荣鑫淼，王素敏，朱红月，还有同学刘芳，感谢你们的鼓励和陪伴。在这条枯燥且艰辛的读博路上，因为有了你们，我从未感觉到孤单。

特别感谢我的父母，感谢他们默默付出，替我承受生活的艰辛。感谢他们无条件的包容、支持和鼓励，没有他们无私的爱，就不可能有现在的我。

最后，祝愿所有帮助和关心过我的老师、同学、朋友和家人，身体健康，平安顺遂。

邹婷婷

2021年3月

在学期间公开发表论文情况

文章名称	发表刊物	刊发时间	刊物级别	第几作者
Two-sample Tests for High-Dimensional Covariance Matrices Using both Difference and Ratio	Electronic Journal of Statistics	2021年	SCI	1
CLT for Linear Spectral Statistics of Large Dimensional Sample Covariance Matrices with Dependent Data	Statistical Papers	Minor Revision	SCI	1
High-Dimensional Sample Correlation Matrices	The Annals of Statistics	Submitted	SCI	3(联合通讯作者)
Testing High-Dimensional Correlation Matrix Structures		Manuscript		1
One-sample Test for High-Dimensional Vector Autoregressive Structure		Manuscript		1