基于空间中位数的超高维位置参数统计推断

Guanghui Cheng[a]、Liuhua Peng[b]、ChangliangZou[c]

[a]广州国际金融研究院、广州大学，[b]墨尔本大学数学与统计学院，[c] 南开大学统计与数据科学学院

arXiv:2301.03126v1 [stat.ME] 9 Jan 2023

摘要

受机器学习中广泛使用的几何中位数均值估计量启发，本文研究一般多元模型下基于样本空间中位数的超高维位置参数统计推断问题，包括同时置信区间的构造、全局检验，以及带错误发现率控制的多重检验。为实现这些目标，我们推导了样本空间中位数的一种新型 Bahadur表示，并给出了余项的最大范数界，同时建立了样本空间中位数在超矩形类上的高斯近似。此外，我们提出了一种乘子自助法算法，用于近似样本空间中位数的分布。当维数以样本量的指数速度增长时，这些近似仍然有效，从而促进了空间中位数在超高维领域的应用。我们还通过模拟以及对一项微阵列研究中的基因组数据集进行分析，进一步说明了所提出方法的表现。

关键词：自助法近似；高斯近似；高维；空间中位数；错误发现率控制

1 引言

几何中位数均值（GMOM）已被广泛用于多元均值的稳健估计，并在机器学习中得到广泛应用（Minsker 2015年，Hsu & Sabato 2016，Prasad et al. 2020）。GMOM 的基本思路是先将数据划分为互不相交的子样本，并分别计算各子样本的经验均值。然后，将得到的经验均值的空间中位数（又称几何中位数）作为 GMOM 估计量。以往关于 GMOM 的研究主要聚焦于在特定重尾假设下建立其非渐近误差界

其分布性质对于统计推断至关重要，但目前仍未得到了解。

在许多科学领域中，人们都遇到了这样一类高维数据：随着观测数量趋于无穷，数据维度也趋于无穷。越来越多的证据表明，由于重尾性的存在，使用多元正态分布对高维数据建模存在问题，且不足以刻画尾部依赖。例如，在许多基因表达数据中，即使经过对数变换，微阵列表达的分布仍被发现是非正态且具有厚尾的（Purdom & Holmes 2005，Wang、Peng 和 Li（2015））。再如，对于美国股票收益数据，椭圆分布，尤其是多元 t分布和对称多元正态逆高斯分布，作为模型远优于多元正态分布（McNeil et al. 2005）。在这类情况下，估计位置参数时，样本空间中位数比样本均值更受青睐。上述讨论有力地促使我们研究高维情形下的空间中位数，尤其是其分布性质，以及如何将其应用于高维位置参数的统计推断。

设 X_1、.....、 X_n 是来自总体 X 的一列独立同分布（i.i.d.） p 维随机向量，其累积分布函数为 F_X ，且 $\mathbb{R}^p$。本文研究一个一般的多元模型，其中 X 具有如下随机表示:

$$X = \boldsymbol{\theta} + \nu\Gamma U\,, \tag{1}$$

其中， $\boldsymbol{\theta}$ 为位置参数， ν 为非负单变量随机变量， U 为具有独立分量的p维随机向量。模型（1）涵盖了许多常用的多元模型和分布族，包括独立分量模型（Yao 等人（2015））和椭圆分布族（Fang 等人（1990））。更详细的讨论请参见第 2 节。

空间中位数是单变量中位数向多元分布的扩展，最初是为位置参数的稳健推断而提出的（Haldane 1948，Weber 1929）。样本空间中位数 $\hat{}\,\boldsymbol{\theta}_n \in \mathbb{R}^p$ 使经验准则函数 $L_n(\boldsymbol{\beta}) = \sum_{i=1}^{n}(\|X_i - \boldsymbol{\beta}\| - \|X_i\|)$达到最小值,

其中 $\|\cdot\|$ 表示欧几里得范数。等价地，

$$\hat{\boldsymbol{\theta}}_n = \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} L_n(\boldsymbol{\beta}) = \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} (\|X_i - \boldsymbol{\beta}\| - \|X_i\|) . \tag{2}$$

ˆ 函数 $L_n(\boldsymbol{\beta})$ 是凸的；当观测值 $\{X_i\}_{i=1}^n$ 在 $\mathbb{R}^p$ 中不集中于一条直线上时， $\boldsymbol{\theta}_n$ 是唯一的，其中 $p > 2$ （Milasevic & Ducharme 1987）。当维数 p 固定时，空间中位数已在文献中得到充分研究。关于这一主题的精彩综述，可参见 Oja（2010）第 6.2 章。

在高维情形下，随着观测值数量 $n \to \infty$增加，维数 p 趋于无穷，已有一些研究探讨了样本空间中位数的渐近性质 ˆ。Zou 等人（2014）针对具有相同形状矩阵的椭圆分布，给出了 $\boldsymbol{\theta}_n$ 的展开；Cheng 等人（2019）则将这一结果推广到一般形状矩阵。近期，Li & Xu（2022）在更强的条件下改进了 Cheng 等人（2019）的展开，得到了阶更小的余项，并建立了平方欧氏距离 $\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\|^2$的中心极限定理。在 Zou 等人（2014）和 Cheng 等人（2019）的研究中，均要求 $p = O(n^2)$。此外，Li & Xu（2022）还要求 p以与 n相同的速率趋于无穷。然而，在基因组学和蛋白质组学等现代领域中，数据维数可能随样本量呈指数增长，进入“超高维”范围（Fan & Lv 2008）。以多项式维数为限制条件的既有研究，限制了空间中位数在超高维情形下的应用。此外，既有结果均基于椭圆分布。因此，建立空间中位数在超高维及超越椭圆分布情形下的渐近性质并探究其应用，具有重要意义。

本文首先在超越椭圆分布的一般模型（1）下，为样本空间中位数建立了适用于超矩形的高斯近似与自助法近似；当维数随样本量呈指数级增长时，这些近似仍然有效。它们构成了超高维情形下基于样本空间中位数对位置参数进行统计推断的理论基础。本文建立了关于位置参数的一致同时置信区间（SCIs）和全局检验。我们还研究了基于 $\boldsymbol{\theta}_n$对 $\boldsymbol{\theta}$ 的每个分量进行多重检验。受关于 $\boldsymbol{\theta}$的同时推断启发，我们

定义了样本空间中位数相对于样本均值的高维渐近相对效率。最重要的是，我们的理论结果保证了所提出推断方法在 p呈指数级增长时的有效性。模拟和真实数据分析验证了所提方法的优势。

本文的主要贡献概括如下。首先，我们基于样本空间中位数 $\boldsymbol{\theta}_n$，为位置参数 $\boldsymbol{\theta}$ 建立了新的同时置信区间，这在现有文献中尚属首次。自助法近似的一致性保证了在超高维情形下，同时置信区间覆盖位置参数所有分量的概率趋近于名义置信水平。我们还提出了一种基于样本空间中位数最大范数的新型超高维位置参数检验。对于呈指数级增长的 p，所提出的检验不仅在渐近意义下保持名义显著性水平，而且与基于 L_2范数的方法相比，在稀疏备择下具有更高的检验功效（Li & Xu 2022，Wang、Peng 和 Li 2015）。另一项主要推断工作是研究位置参数各分量的多重检验；结合基于样本空间中位数的 Benjamini-Hochberg 程序，可以很好地控制错误发现率（FDR），这拓展了现有基于样本均值的方法（Liu 和 Shao 2014）。在所有推断方法中，对于重尾分布，基于样本空间中位数的程序优于基于样本均值的程序。

其次，本文是首篇为超高维情形下的样本空间中位数提供高斯近似与自助法近似的研究。过去十年间，高维样本均值的高斯近似与自助法近似受到了广泛关注。Chernozhukov 等（2013）和 Chernozhukov、Chetverikov 和 Kato（2017）分别在 Kolmogorov 距离和超矩形上，为中心化独立随机向量之和的最大值建立了高斯近似与自助法近似。相关研究还可参见 Chen（2018）、Chernozhukov 等人（2019）和 Chernozhukov 等（2020）。与形式简单的样本均值相比，样本空间中位数在理论上的难点在于它没有显式形式。本文通过推导样本空间中位数的一种新型 Bahadur表示，并为余项给出最大范数界，解决了这一问题；该结果将 Zou 等人（2014）、Cheng 等人（2019）和 Li & Xu（2022）在椭圆分布及多项式维数情形下的结果加以推广。此外，我们的结果

在合理条件下可应用于 GMOM，从而提升 GMOM 的实际应用价值。

第三，我们为样本空间中位数提出一种新颖的乘子自助法。与通常用于M估计量的做法不同（Imaizumi & Otsu 2021），我们不对损失函数乘以乘子，而是将乘子应用于中心化后的 X_i。具体而言，$\hat{}\,\boldsymbol{\theta}_n$ 的自助法版本定义为 $\tilde{\boldsymbol{\theta}}_n = \text{argmin}_{\boldsymbol{\beta}\in\mathbb{R}^d} \sum_{i=1}^n \|Z_i(X_i - \hat{}\,\boldsymbol{\theta}_n) - \boldsymbol{\beta}\|$，其中 Z_1、……、Z_n 为乘子。得益于这一新颖的构造，即使在超高维情形下，乘子自助法仍具有一致性。然而，这不同于针对样本均值的乘子自助法，因为后者具有显式形式（Chernozhukov 等，2013；Chernozhukov、Chetverikov 和 Kato，2017）。

本文其余部分安排如下。第 2 节介绍模型和假设。第 3 节建立样本空间中位数分布的高斯近似与自助法近似。第 4 节介绍基于样本空间中位数对位置参数进行统计推断。第 5 节报告数值结果，包括模拟和真实数据分析。主要结果的预备引理及其证明见补充材料的附录 A。预备引理的证明和额外模拟见补充材料的附录 B 和 C。

记号： 记 $|x|_\infty = \max(|x_1|, \ldots, |x_d|)$ 为 $x = (x_1, \ldots, x_d)^\top$ 的最大范数。若 $a_n \leqslant Cb_n$，则记为 $a_n \lesssim b_n$，其中 C 为正数常数；$a_n \asymp b_n$ 表示 $a_n \lesssim b_n$ 和 $b_n \lesssim a_n$。对于 $\alpha > 0$，令 $\psi_\alpha(x) = \exp(x^\alpha) - 1$ 为定义在 $[0, \infty)$ 上的函数。随机变量 X 的 Orlicz 范数 $\|\cdot\|_{\psi_\alpha}$ 定义为 $\|X\|_{\psi_\alpha} = \inf\{t > 0, \mathbb{E}\{\psi_\alpha(|X|/t)\} \leqslant 1\}$.。我们用 $\text{tr}(\cdot)$ 表示方阵的迹算子。此外，记 I_p 为 $p \times p$ 单位矩阵。对于 $a, b \in \mathbb{R}$，记 $a \wedge b = \min(a, b)$。

2 模型与假设

本文考虑分布 F_X 的一般多元模型，使得 X_i 具有如下随机表示：

$$X_i = \boldsymbol{\theta} + \nu_i \Gamma U_i\,, \tag{3}$$

其中， $\boldsymbol{\theta}$ 是位置参数， Γ 是非随机且可逆的 $p \times p$ 矩阵， U_i 是一个具有独立标准化分量的 p 维随机向量，而 ν_i 是一个非负单变量随机变量，且与 U_i 的空间符号相互独立。 X_i 的分布通过形状矩阵 $\Omega = \Gamma\Gamma^\top$ 依赖于 Γ 。

注 1. 模型 (3) 涵盖了许多常用的多元模型和分布族。首先，独立分量模型（Yao 等人（2015））符合 (3)，其中 ν_i 为非负常数。其次，取$U_i \sim N(0, I_p)$ 和 $\nu_i = \xi_i/\|U_i\|$，并令其为与 U_i独立的某个非负随机变量 ξ_i，模型 (3) 也涵盖椭圆分布。在这种情况下， ν_i 与 U_i的空间符号独立，但与 U_i不独立。独立分量模型在高维数据分析、信号处理和机器学习领域得到了广泛应用（Hyv¨arinen 等人，2001）。此外，椭圆分布族涵盖了许多非高斯分布，例如多元 t 分布、多元逻辑斯蒂分布等。在研究样本空间中位数的文献中，这类分布被广泛采用（Cheng 等人，2019，Li 和 Xu（2022），Zou 等人（2014））。就 GMOM 而言，如果数据来自独立分量模型，子样本均值显然满足模型 (3)。此外，椭圆分布的某些子族在卷积下封闭，因此子样本均值也服从模型 (3)。在这些情况下，我们的结果可直接应用于广义中位数的中位数估计量。

对于 $i = 1, \ldots, n$ 和 $k = 1, 2, 3, 4$，记为

$$W_i = S(X_i - \boldsymbol{\theta}) \quad \text{and} \quad R_i = \|X_i - \boldsymbol{\theta}\| \tag{4}$$

分别表示 $X_i - \boldsymbol{\theta}$ 的空间符号和半径，其中 $S(X) = \|X\|^{-1} X \mathbb{I}(X \neq 0)$ 是多元符号函数， $\mathbb{I}(\cdot)$ 是示性函数。因此， $\hat{\boldsymbol{\theta}}_n$ 满足 $\sum_{i=1}^n S(X_i - \hat{\boldsymbol{\theta}}_n) = 0$.

记 $U_i = (U_{i,1}, \ldots, U_{i,p})^\top$，我们提出以下三个条件。

条件 C.1. $U_{i,1}, \ldots, U_{i,p}$ 是独立同分布的对称随机变量，且 $\mathbb{E}(U_{i,j}) = 0$, $\mathbb{E}(U_{i,j}^2) = 1$，以及 $\|U_{i,j}\|_{\psi_\alpha} \leqslant c_0$ ，其中常数为 $c_0 > 0$ 和 $1 \leqslant \alpha \leqslant 2$。

条件 C.2。 当 p 足够大时， $k = 1$ 的矩 $\zeta_k = \mathbb{E}(R_i^{-k})$ 存在。此外，存在两个正数常数 b 和 $\bar{B}$，使得 $b \leqslant \limsup_p \mathbb{E}(R_i/\sqrt{p})^{-k} \leqslant \bar{B}$ 对于 $k = 1$， 2， 3， 4。

条件 C.3。 形状矩阵 $\Omega = (\omega_{j\ell})_{p\times p}$ 满足 $\mathrm{tr}(\Omega) = p$ ，且属于以下类:

$$\mathcal{U}(a_0(p), \underline{m}, \bar{M}) = \left\{ \Omega : \underline{m} \leqslant \omega_{jj} \leqslant \bar{M},\ \sum_{\ell=1}^{p} |\omega_{j\ell}| \leqslant a_0(p),\ \text{ for all } j = 1, \ldots, p \right\},$$

其中 $m \leqslant M$ 为有界正数常数。

注释 2. 在条件 C.1 中，对称性假设是为了确保模型 (3) 中的 $\boldsymbol{\theta}$ 与总体空间中位数重合，而总体空间中位数使 $L(\boldsymbol{\beta}) = \mathbb{E}(\|X - \boldsymbol{\beta}\| - \|X\|)$最小化。显然，对于满足 $U_i \sim N(0, I_p)$ 的椭圆分布，条件 C.1 成立。条件 $\|U_{i,j}\|_{\psi_\alpha} \leqslant c_0$ 意味着 $U_{i,j}$ 服从次指数分布。值得强调的是，只需对主要定理的证明稍作修改，就可以通过以下假设替代条件 C.1，从而放宽对 $U_{i,1}, \ldots, U_{i,p}$的独立同分布条件: $U_{i,1}, \ldots, U_{i,p}$ 是相互独立的对称随机变量，满足 $\mathbb{E}(U_{i,j}) = 0$, $\mathbb{E}(U_{i,j}^2) = 1$ ，对所有 $j = 1, \ldots, p$均成立，并且 $\sup_{1\leqslant j\leqslant p} \|U_{i,j}\|_{\psi_\alpha} \leqslant c_0$ ，其中 $c_0 > 0$ 和 $1 \leqslant \alpha \leqslant 2$为某些常数。

注释 3. 条件 $b \leqslant \limsup_p \mathbb{E}(R_i/\sqrt{p})^{-k} \leqslant \bar{B}$ 表明，对于 k =1、2、3、4， $\zeta_k \asymp p^{-k/2}$。引入该条件是为了避免 X_i 过度集中于 $\boldsymbol{\theta}$附近。对于椭圆分布，它是 Zou 等人（2014）提出的假设 1 的推广，许多常见分布都满足该条件。对于独立分量模型，条件 C.2 等价于 $b \leqslant \limsup_p \mathbb{E}(\|\Gamma U_i\|/\sqrt{p})^{-k} \leqslant \bar{B}$.。根据附录 A 中的引理 A2， $\mathbb{E}(\|\Gamma U_i\|^k) = p^{k/2}\{1 + o(1)\}$ 对于 $k = 1$、2、3、4。于是，由柯西-施瓦茨不等式可得 $\mathbb{E}(\|\Gamma U_i\|^{-k}) \geqslant \{\mathbb{E}(\|\Gamma U_i\|^k)\}^{-1} = p^{-k/2}\{1 + o(1)\}$, ，由此可知 $\mathbb{E}(\|\Gamma U_i\|^{-k}) \gtrsim p^{-k/2}$。此外，记 Γ_j 为 Γ的第 j行，则根据调和平均数与平方平均数的不等式,

$$p^2\|\Gamma U_i\|^{-4} = \left\{ \frac{p}{(\Gamma_1 U_i)^2 + \cdots + (\Gamma_p U_i)^2} \right\} \leqslant \frac{(\Gamma_1 U_i)^{-4} + \cdots + (\Gamma_p U_i)^{-4}}{p}.$$

由此可知， $\mathbb{E}(\|\Gamma U_i\|^{-4}) \lesssim p^{-2}$ 如果 $\mathbb{E}\{(\Gamma_1 U_i)^{-4}\}$，……， $\mathbb{E}\{(\Gamma_p U_i)^{-4}\}$ 一致有界，那么根据 $\mathbb{E}(\|\Gamma U_i\|^{-k}) \lesssim p^{-k/2}$ 詹森不等式。因此，只要 $\Gamma_1 U_i$，……， $\Gamma_p U_i$ 不会过度集中在 0 附近，独立分量模型就满足条件 C.2。另请参见 Cardot 等人（2013）关于类似条件的讨论。

注4。 注意到，形状矩阵 Ω 仅在相差一个标量倍数的意义下定义，因此使用条件 $\mathrm{tr}(\Omega) = p$ 对 Ω 进行规范化，使模型 (3) 可识别。类 $\mathcal{U}(a_0(p),\ m,\ M)$ 涵盖了广泛的对称方阵，且在高维分析文献中被普遍采用。例如，Bickel & Levina (2008)引入了一个类似的矩阵类。条件 $m \leqslant \omega_{jj} \leqslant M$ 要求对角元素有界。 $a_0(p)$的阶（将在后文说明）控制 Ω 的非对角元素的阶。

—

3 高斯分布与自助法近似

3.1 Bahadur表示与高斯近似

ˆ本节针对 $\boldsymbol{\theta}_n$ 建立高斯近似，当 p 随 n 指数级发散时，该近似仍然成立。以下引理给出了 $\boldsymbol{\theta}_n$ 的 Bahadur表示，并为定理1中的高斯近似结果奠定了基础。

引理 1. （*Bahadur*表示）假设条件*C.1*、*C.2*和*C.3*在某个正数常数 $\delta \leqslant 1/2$ 下成立，且 $a_0(p) \asymp p^{1-\delta}$。若 $\log p = o(n^{1/3})$ 且 $\log n = o(p^{1/3\wedge\delta}),$ ，则

$$n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} W_i + C_n\,,$$

其中 $|C_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$。

注5. 据我们所知，引理 1 首次给出了样本空间中位数的 Bahadur表示，并对余项给出了最大范数界。在 Zou 等人（2014）和 Cheng 等人（2019）中，得到了相同的展开式，其中余项 C_n 满足 $\|C_n\| = o_p(\zeta_1^{-1})$ ；Li & Xu（2022）通过在线性项中以 $n^{-1}\sum_{i=1}^{n} R_i^{-1}$ 替换 ζ_1 ，将他们的结果改进为 $\|C_n\| = o_p(1)$ ，但所需条件更为严格，即 p 和 n 同阶。值得注意的是，先前的结果（Cheng 等人 2019、Li & Xu 2022、Zou 等人 2014）均是在椭圆分布下推导得到的。

设 $\mathcal{A}^{\mathrm{re}} = \{\prod_{j=1}^{p}[a_j, b_j]\colon -\infty \leqslant a_j \leqslant b_j \leqslant \infty, j = 1, \ldots, p\}$ 是.....中的矩形类

$\mathbb{R}^p$。借助引理 1 中的 Bahadur表示，我们为 $\boldsymbol{\theta}_n$ 在超矩形上的情形建立了以下高斯分布近似结果。

定理1. （高斯近似）假设条件*C.1*、*C.2*和*C.3*成立，其中 $a_0(p) \asymp p^{1-\delta}$，且某个正数常数 $\delta \leqslant 1/2$ 满足条件。若 $\log p = o(n^{1/5})$ 且 $\log n = o(p^{1/3 \wedge \delta})$, ，则

$$\rho_n(\mathcal{A}^{\mathrm{re}}) = \sup_{A \in \mathcal{A}^{\mathrm{re}}} \left| \mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\} - \mathbb{P}\left(G \in A\right) \right| \to 0$$

当 $n \to \infty$ 时， $G \sim N(0,\ \zeta_1^{-2}\mathbb{B})$，其中 $\mathbb{B} = \mathbb{E}(W_1 W_1^{\top})$。

$\boldsymbol{\theta}_n$^ 的高斯近似表明，对于超矩形 $A \in \mathcal{A}^{\mathrm{re}}$，概率 $\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\}$ 可由协方差矩阵为 $\zeta_1^{-2}\mathbb{B}$ 的中心化高斯随机向量的相应概率近似。定理1允许 p 指数发散，这适用于超高维情形。与 Li & Xu（2022）中 $\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\|^2$ 的渐近正态性相比（其中假设 p 与 n 同阶），定理1中的高斯近似结果对 n 和 p 的速率要求弱得多。

注 6. 令 $\mathbb{B}_{j\ell}$ 为 $\mathbb{B}$的第 $(j,\ \ell)$个元素。根据附录 A 中的引理A4（iii），对于所有 $1 \leqslant j,\ \ell \leqslant p$，都有 $\zeta_1^{-2}\mathbb{B}_{j\ell} = \zeta_1^{-2} p^{-1} \omega_{j,\ \ell} + O(p^{-\delta/2})$ 。因此，定理1中 G 的协方差矩阵渐近正比于形状矩阵 Ω。

注7. 由于样本空间中位数是一种特殊的M估计量，Imaizumi & Otsu（2021）关于M估计量的高斯近似可能适用于高维情形下的空间中位数。然而，值得强调的是，Imaizumi & Otsu（2021）中的结果并不适用于我们的框架。具体而言，Imaizumi & Otsu（2021）中的假设1（A3）假设存在常数 $C > 0$和 $\alpha \in (0,\ 2)$使得对于所有 $\varepsilon \in (0,\ 1)$，$\log \mathcal{N}(\varepsilon,\ \Theta,\ \|\cdot\|) \leqslant C\varepsilon^{-\alpha}$成立，其中 Θ为参数空间，而 $\mathcal{N}(\varepsilon,\ \Theta,\ \|\cdot\|)$是 Θ 在欧几里得范数 $\|\cdot\|$下的 ε-覆盖数（van der Vaart & Wellner，1996）。当 Θ 是 $\mathbb{R}^p$的紧子集时， $\mathcal{N}(\varepsilon,\ \Theta,\ \|\cdot\|)$的阶为 $O(\varepsilon^{-p})$。在这种情况下，当 $p \to \infty$时，$\log \mathcal{N}(\varepsilon,\ \Theta,\ \|\cdot\|) \leqslant C\varepsilon^{-\alpha}$无法满足。因此，我们的理论发现独立于 Imaizumi & Otsu（2021）中的结果。

由于上确界范数对应的 Kolmogorov 距离是 $\mathcal{A}^{\mathrm{re}}$ 中最大超矩形在 $\mathbb{R}^p$ 中的子集，因此定理1立即推出以下推论。

推论 1. 在定理1所假设的条件下，当 $n \to \infty$ 时，

$$\rho_n = \sup_{t\in\mathbb{R}} \left| \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}(|G|_\infty \leqslant t) \right| \to 0.$$

3.2 乘子自助法近似

定理1使我们能够用 G hit 超矩形的分布来近似 $n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$ 的分布，其中 $G \sim N(0,\ \zeta_1^{-2}\mathbb{B})$。然而，由于量 ζ_1 和矩阵 $\mathbb{B}$ 取决于底层分布 F_X，因而未知，因此无法将其直接用于对 $\boldsymbol{\theta}$ 的统计推断。为解决这一问题，我们提出一种易于实现的自助法，用于近似 $n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$ 的分布。

设 $Z_1, \ldots, Z_n$ 是一列均值为零、方差为单位值的独立同分布随机变量。将样本空间中位数的自助法版本定义为

$$\tilde{\boldsymbol{\theta}}_n = \underset{\boldsymbol{\beta}\in\mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n} \|Z_i(X_i - \hat{\boldsymbol{\theta}}_n) - \boldsymbol{\beta}\|. \tag{5}$$

然后，利用给定 $X_1, \ldots, X_n$ 时 $n^{1/2}\tilde{\boldsymbol{\theta}}_n$ 的条件分布来近似 $n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$ 的分布。该算法称为乘子自助法，$Z_1, \ldots, Z_n$ 为乘子权重。

关于附录 B 中引理 A5 的证明，乘子权重$Z_1, \ldots, Z_n$最好有界并满足 $\mathbb{E}(Z_i^{-2}) < \infty$。因此，我们选择拉ademacher变量作为乘子（Chernozhukov 等，2019），即 $\mathbb{P}(Z_i = 1) = \mathbb{P}(Z_i = -1) = 1/2$。

定理 2.（自助法近似）在定理*1*所假设的条件下，

$$\rho_n^{\mathrm{MB}}(\mathcal{A}^{\mathrm{re}}) = \sup_{A\in\mathcal{A}^{\mathrm{re}}} \left| \mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\} - \mathbb{P}^*(n^{1/2}\tilde{\boldsymbol{\theta}}_n \in A) \right| \to 0$$

当 $n \to \infty$ 时依概率成立，其中 $\mathbb{P}^*$ 表示给定 $X_1, \ldots, X_n$ 时的条件概率。

在与定理1中相同的 n 和 p 发散速率条件下，定理 2 验证了：给定 $X_1, \ldots, X_n$，自助样本空间˜ ˆ中位数 $\boldsymbol{\theta}_n$ 的分布在超矩形上始终一致地近似 $\boldsymbol{\theta}_n$ 的分布。

注 8. 定理 2 的证明并非易事，且无法直接从现有

˜结果中推出，因为 $\boldsymbol{\theta}_n$ 没有显式形式，这与文献中已分析的针对高维样本均值的乘子自助法不同。˜ ˆ 证明中的关键步骤是为 $\boldsymbol{\theta}_n$ 获得类似于引理 1 中 $\boldsymbol{\theta}_n$ 的 Bahadur表示。具体而言，我们在附录 A 的引理 A5 中证明了 $n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n Z_i W_i + \tilde{C}_n$ 满足 $|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$ 。

下一个推论是定理 2 的直接推论。

推论 2。 在定理 2 的假设条件下，当 $n \to \infty$ 时，

$$\rho_n^{\mathrm{MB}} = \sup_{t\in\mathbb{R}} \left| \mathbb{P}\{n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant t\} - \mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant t) \right| \to 0 \ \text{ in probability.}$$

4 统计推断

样本空间中位数的高斯分布近似和乘子自助法近似，为超高维总体位置参数提供了多种统计推断方法。本节介绍以下统计推断：总体位置参数的同时置信区间（SCIs）和全局检验、针对 $\boldsymbol{\theta}$ 每个分量的多重检验，以及与样本均值相比，样本空间中位数在高维下的渐近相对效率。

4.1 同时置信区间

我们关注如何为 $\boldsymbol{\theta} = (\theta_1、\ldots\ldots、\theta_p)^\top$ 的所有分量构造同时置信区间。推论 2 启发我们采用以下方法，为 $\boldsymbol{\theta}$ 构造同时置信区间。给定名义置信水平 $1-\tau$，定义集合 $\mathcal{C}_\tau$ 为

$$\mathcal{C}_\tau = \left\{ \boldsymbol{\theta} \in \mathbb{R}^p, n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty < q_{1-\tau}^B \right\},$$

其中， $q_{1-\tau}^B$是给定 $X_1、\ldots\ldots、X_n$时 $n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty$的第 $(1-\tau)$分位数。记ˆ $\boldsymbol{\theta}_n = (\hat{\theta}_{n,\,1}、\ldots\ldots、\hat{\theta}_{n,\,p})^\top$，则对于 $j = 1、\ldots\ldots、p$，置信区间为$[\theta_{n,\,j}^-,\ \theta_{n,\,j}^+]$，其中

$$\theta_{n,j}^- = \hat{\theta}_{n,j} - n^{-1/2}q_{1-\tau}^B \ \text{ and } \ \theta_{n,j}^+ = \hat{\theta}_{n,j} + n^{-1/2}q_{1-\tau}^B.$$

下一个定理表明，在超高维情形下， $\mathcal{C}_\tau$ 渐近地保持名义同时置信水平 $1-\tau$。

定理3. 假设定理 *2* 的条件成立，则当 $n \to \infty$ 时， $\mathbb{P}(\boldsymbol{\theta} \in \mathcal{C}_\tau) \to 1-\tau$。等价地，对于所有 $1 \leqslant j \leqslant p) \to 1-\tau$，当 $n \to \infty$ 时，$\mathbb{P}(\theta_j \in [\theta^-_{n,\ j}, \theta^+_{n,\ j}]$。

注 9. 与固定维度情形不同， $n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty$ 和 $n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty$ 是发散数量的变量的最大值，其分位数通常也会随着 $p \to \infty$ 发散。因此，定理3并非推论 2 的直接推论。为从理论上确定 $\mathcal{C}_\tau$的一致性，我们利用发散随机序列的反集中不等式证明：以趋于 1 的概率， $q^B_{1-\tau}$ 被 $n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty$ 的两个分位数所界定，这两个分位数的分位数水平都足够接近 $1-\tau$。

注 10. 样本均值 $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ 的高斯近似（Chernozhukov 等，2013；Chernozhukov、Chetverikov 和 Kato，2017；Chernozhukov 等人（2019））表明，如果 $\log p = o(n^{1/5}),$

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(n^{1/2}|\bar{X}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}(|G_0|_\infty \leqslant t)\right| \to 0 \tag{6}$$

当 $n \to \infty$ 时，在一些适度条件下成立，其中 $G_0 \sim N(0,\ \Sigma)$ 满足 $\Sigma = \mathbb{E}(XX^\top)$。定义 $X^*_i = Z_i(X_i - X_n)$，其中 $i = 1,$ ……，$n,$ 其中 $Z_1,$ ……，Z_n 为 Rademacher 权重。记 $\bar{X}^*_n = n^{-1}\sum_{i=1}^n X^*_i$，Chernozhukov 等人（2019）已证明

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(n^{1/2}|\bar{X}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}^*(n^{1/2}|\bar{X}^*_n|_\infty \leqslant t)\right| \to 0 \tag{7}$$

当 $\log p = o(n^{1/5})$ 时，以概率收敛于 $n \to \infty$。根据 **(7)**，定义

$$\mathcal{C}'_\tau = \left\{\boldsymbol{\theta} \in \mathbb{R}^p, n^{1/2}|\bar{X}_n - \boldsymbol{\theta}|_\infty < q^{B\prime}_{1-\tau}\right\},$$

其中， $q^{B\prime}_{1-\tau}$是给定 X_1、……、 X_n时 $n^{1/2}|\bar{X}^*_n|_\infty$的 $(1-\tau)$分位数。于是， $\mathcal{C}'_\tau$也是 $\boldsymbol{\theta}$的渐近 $1-\tau$ 同时置信区间。根据第 4.4 节中的讨论，在重尾分布下， $\mathcal{C}_\tau$ 优于 $\mathcal{C}'_\tau$（区间相对较短）。关于这一点的有限样本论证，请参见第 5.1 节。

4.2 高维位置参数的全局检验

本节提出一种针对高维位置参数进行全局检验的新方法。设 $\boldsymbol{\theta}_0$ 为一个已知的 p维向量，我们关注以下检验问题:

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{versus} \quad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \tag{8}$$

定理 1 和 2 启发我们提出一种最大范数型检验统计量。定义

$$T_n = n^{1/2} |\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0|_\infty \tag{9}$$

作为检验统计量，当 T_n 大于临界值时，拒绝 H_0 。我们可以使用乘子自助法来近似 T_n 在 H_0 下的分布。具体而言，给定名义显著性水平 τ，若 $T_n > q^B_{1-\tau}$，则拒绝零假设。定理3保证，基于 T_n 的检验在超高维情形下渐近地维持名义显著性水平，即当 $\log p = o(n^{1/5})$ 时，$\mathbb{P}(T_n > q^B_{1-\tau} \mid H_0) \to \tau$，其中 $n \to \infty$ 。

注11。 可以基于 X_n 构造针对 (8) 的另一种检验，其检验统计量定义为 $T_{\text{Mean}} = n^{1/2} |\bar{X}_n - \boldsymbol{\theta}_0|_\infty$。若 $T_{\text{Mean}} > q^{B\prime}_{1-\tau}$，则拒绝零假设。基于 T_n 的检验可视为基于 T_{Mean} 的检验的非参数扩展。正如第 4.4 节所讨论的，在重尾分布下，对 $\boldsymbol{\theta}$ 进行同时推断时，$\hat{\bar{\boldsymbol{\theta}}}_n$ 比 X_n 更有效，因此我们预计，在这些情况下，基于 T_n 的检验比基于 T_{Mean} 的检验功效更高。补充材料附录 C 中的模拟结果印证了这一点。

下一个定理总结了基于 T_n 的所提检验的渐近功效。

定理 4. 假设定理 *2* 的条件成立。对于任意给定的 $0 < \tau < 1$，若 $|\boldsymbol{\theta} - \boldsymbol{\theta}_0|_\infty \geqslant C$ 对数 $^{1/2}(\tau^{-1}) n^{-1/2}$ 对数$^{1/2}(np)$，其中常数 $C > 0$ 足够大，则当 $n \to \infty$ 时，$\mathbb{P}(T_n > q^B_{1-\tau} \mid H_1) \to 1$ 。

定理 4 表明，对于固定的显著性水平 τ，当 $n^{1/2}|\boldsymbol{\theta} - \boldsymbol{\theta}_0|$ 的最大元素的绝对值远大于 $\log^{1/2}(\tau^{-1}) \log^{1/2}(np)$ 时，基于 T_n 的检验具有一致性。

注释 12。 Wang、Peng 和 Li（2015）针对（8）提出了一种基于 $T_{\mathrm{WPL}}=\sum_{i=1}^{n}\sum_{j=1}^{i-1}W_i^{\top}W_i$、带有 $\boldsymbol{\theta}_0=0$ 的 L_2-范数型检验（WPL 检验）。Wang、Peng 和 Li（2015）以及 Li & Xu（2022）指出，WPL 检验的信号取决于 $\|\boldsymbol{\theta}\|$ 的大小，其中 $\|\boldsymbol{\theta}\|$ 是 $\boldsymbol{\theta}$ 的 L_2-范数。相比之下，基于 T_n 的检验的检验功效取决于 $|\boldsymbol{\theta}|_{\infty}$。因此，在稀疏备择下，当 $\boldsymbol{\theta}$ 仅包含有限个非零分量且其最大元素具有一定阶的大小时，预计所提出的基于 T_n 的检验具有更高的检验功效。在这种情况下， $\|\boldsymbol{\theta}\|$ 不够大，无法拒绝 WPL 检验。数值论证见补充材料中的附录 C 和第 5.3 节。

4.3 大规模检验中的假发现率控制与多重检验

基于错误发现率（FDR）控制的多重检验已应用于许多实际问题，例如在基因组学研究中检测差异表达基因。本节中，我们采用 Benjamini 和 Hochberg（B-H）错误发现率控制方法，研究基于空间中位数对 $\boldsymbol{\theta}$ 的每个分量进行多重检验。对于 $j=1,\ldots,p$，我们感兴趣的是检验

$$H_{0j}:\theta_j=\theta_{0,j}\quad\text{versus}\quad H_{1j}:\theta_j\neq\theta_{0,j}$$

同时进行检验，其中 $\theta_{0,1},\ldots,\theta_{0,p}$ 为给定值。

定义检验统计量如下

$$T_{n,j}=n^{1/2}(\hat{\theta}_{n,j}-\theta_{0,j})/s_{n,j}$$

对于 $j=1,\ldots,p$，其中 $s_{n,j}^2=\hat{\zeta}_1^{-2}\hat{\mathbb{B}}_{jj}$ with $\hat{\zeta}_1=n^{-1}\sum_{i=1}^{n}\|X_i-\hat{\boldsymbol{\theta}}_n\|^{-1}$，且 $\hat{\mathbb{B}}_{jj}$ 是 $\hat{\mathbb{B}}=n^{-1}\sum_{i=1}^{n}\|X_i-\hat{\boldsymbol{\theta}}_n\|^{-2}(X_i-\hat{\boldsymbol{\theta}}_n)(X_i-\hat{\boldsymbol{\theta}}_n)^{\top}$的第 j个对角元素。

根据附录 A 中定理 5 的证明，$T_{n,j}$ 在 H_{0j}下依分布收敛于标准正态分布，其中 $j=1,\ldots,p$。因此，我们利用标准正态分布估计边际 p 值。对于 $j=1,\ldots,p$，定义 H_{0j} 的 p 值为 $P_j=2-2\Phi(|T_{n,j}|)$。记 $P_{(1)}\leqslant\cdots\leqslant P_{(p)}$ 为按序排列的 p 值，并定义

$$\hat{k}=\max\left\{j=0,\ldots,p:P_{(j)}\leqslant\tau j/p\right\}$$

对于预先指定的显著性水平 τ。然后，B-H程序拒绝满足 $P_j \leqslant P_{(\hat{k})}$ 的零假设。记 $\mathcal{H}_R = \{j : P_j \leqslant P_{(\hat{k})}\}$ 为指标 j 的集合，其中 H_{0j} 被 B-H 方法拒绝；令 $|\mathcal{H}_R|$ 为 $\mathcal{H}_R$ 的基数，即被拒绝的零假设总数。

令 $\mathcal{H}_0 \subset \{1, \ldots, p\}$ 为与真实零假设 H_{0j}对应的指标集合 j。B-H 方法的错误发现比例（FDP）和错误发现率（FDR）定义如下：

$$\mathrm{FDP}_M = \frac{|\mathcal{H}_0 \cap \mathcal{H}_R|}{|\mathcal{H}_R| \vee 1} \quad \text{and} \quad \mathrm{FDR}_M = \mathbb{E}(\mathrm{FDP}_M).$$

鉴于 $T_{n,1}, \ldots, T_{n,p}$之间存在依赖关系，我们对 W_i任意两个分量之间的弱依赖施加如下条件。定义 $(r_{j\ell})_{p\times p} = \{\mathsf{diag}(\mathbb{B})\}^{-1/2}\mathbb{B}\{\mathsf{diag}(\mathbb{B})\}^{-1/2}$为相关矩阵，其中 $\mathsf{diag}(\mathbb{B})$ 为 $\mathbb{B}$的对角矩阵。

条件 C.4. 设 $\max_{1\leqslant j,\ell\leqslant p}|r_{j\ell}| \leqslant r$ ，其中 $0 < r < 1$为某个常数。此外，$\sum_{j=1}^{p}\mathbb{I}(r_{j\ell} = 0) = O(p^{\eta})$ 对某个常数 $0 < \eta < (1-r)/(1+r)$成立。

Liu 和 Shao（2014）以及 Belloni 等人（2018年）也假设了类似的条件。令 $p_0 = |\mathcal{H}_0|$为真实零假设的数量， $\mathbb{B}_{jj}$ 为 $\mathbb{B}$的第 j个对角元素。

定理 5. 假设条件 *C.4* 和定理*1*的条件成立。此外，存在 $\mathcal{H} \subset \{1、\ldots\ldots、 p\}$，使得 $\mathcal{H} = \big\{j : \zeta_1\mathbb{B}_{jj}^{-1/2}n^{1/2}|\theta_j - \theta_{0,j}| \geqslant 2\,$对数$^{1/2}(p)\big\}$ 且 $|\mathcal{H}| \geqslant$对数对数 $p \to \infty$，当 $p \to \infty$ 时。假设对于某个$0 < \varpi < 1$，错误零假设的数量 $p_1 \leqslant p^{\varpi}$。那么，当 $n \to \infty$时，假发现率 $_M/(\tau p_0/p) \to 1$。

定理 5 表明，基于 $P_1, \ldots, P_p$的 B-H程序渐近地控制假发现率，并将 Liu 和 Shao（2014）中的定理 4.1 推广到了基于空间中位数的检验统计量。

4.4 高维渐近相对效率

作为位置参数 $\boldsymbol{\theta}$ 的两个候选估计量，研究样本空间中位数 $\boldsymbol{\theta}_n$ 相对于样本均值 X_n 的渐近相对效率（ARE）($\boldsymbol{\theta}_n$, X_n)颇具意义。当 p 固定时，对于球形多元正态分布，Brown（1983）表明， $\boldsymbol{\theta}_n$ 相对于 X_n 的渐近效率（记为 ARE($\boldsymbol{\theta}_n$）超过通常的一元

情形 $2/\pi$。此外，ARE($\boldsymbol{\theta}_n$， X_n) 会随着维数增加而增加，并且当 p 趋于足够大时趋近于 1（Magyar 和 Tyler（2011））。然而，当 $p \to \infty$ 时，由于 $\boldsymbol{\theta}_n$ 和 X_n 不存在明显的“最终”极限分布，ARE 并不容易量化 $^{-}$。受第 4.1 节和第 4.2 节讨论的启发，我们比较 $\boldsymbol{\theta}_n$ 和 X_n 在关于 $\boldsymbol{\theta}$ 的同时推断中的效率 $|\boldsymbol{\theta} - \boldsymbol{\theta}|_\infty$

$^{-}$ $\hat{}$ 和 $|X_n - \boldsymbol{\theta}|_\infty$，这些效率由其变化决定。根据推论 1 和（6），我们将 $^{-}$ 相对于 X_n 在关于 $\boldsymbol{\theta}$ 的同时推断中的高维渐近相对效率定义为

$$\mathrm{ARE}(\hat{\boldsymbol{\theta}}_n, \bar{X}_n) = \mathrm{Var}(|G_0|_\infty)/\mathrm{Var}(|G|_\infty)\,, \tag{10}$$

$^{-}$ $\hat{}$ $\hat{}^{-}$，其近似于 $\mathrm{Var}(|X_n - \boldsymbol{\theta}|_\infty)/\mathrm{Var}(|\boldsymbol{\theta}_n - \boldsymbol{\theta}|_\infty)$。如果 $\lim_{p\to\infty} \mathrm{ARE}(\boldsymbol{\theta}_n,$ $X_n) > 1$，则称在高维性下对 $\boldsymbol{\theta}$ 进行同时推断时，$\hat{}^{-}\boldsymbol{\theta}_n$ 比 X_n 更有效。

如注 6 所述 $G \sim N(0,\ \zeta_1^{-2}\mathbb{B})$，对所有 $1 \leqslant j,\ \ell \leqslant p,\ \zeta_1^{-2}\mathbb{B}_{j\ell} = \zeta_1^{-2}p^{-1}\omega_{i\ell}$。此外，我们可以证明$\Sigma_{j\ell} = \mathbb{E}(\nu_i^2)\omega_{j\ell} + O(p^{-1/2})$，这与补充材料附录 B 中引理 A3 的证明类似，其中 $\Sigma_{j\ell}$ 是 Σ 的第 ℓ) 个元素 (j, 。因此，协方差矩阵 Σ 和 $\zeta_1^{-2}\mathbb{B}$ 渐近地都与 Ω 成正比，并且 ARE($\hat{}\boldsymbol{\theta}_n$， $\bar{X}_n$) 约为 $\mathbb{E}(\nu_i^2)\zeta_1^2 p$。

由于实际中很少知道 Σ和 $\zeta_1^{-2}\mathbb{B}$，我们使用自助法近似来估计 $\mathrm{Var}(|G_0|_\infty)/\mathrm{Var}(|G|_\infty)$ 的值。结合推论 2 和 (7)，我们建议使用

$$\mathrm{Var}^*(|\bar{X}_n^*|_\infty)/\mathrm{Var}^*(|\tilde{\boldsymbol{\theta}}_n|_\infty),$$

$\hat{}^{-}$来估计渐近相对效率($\boldsymbol{\theta}_n$， X_n)。

例 1。 假设 X_1， $\ldots X_n$ 是来自 $N(\boldsymbol{\theta},\ I_p)$的独立同分布样本，则 ν_i^2服从自由度为 p 的卡方分布。因此， $\mathbb{E}(\nu_i^2) = p$ 且 $\mathbb{E}(\nu_i^{-1}) = \Gamma(p/2 - 1/2)/\{2^{1/2}\Gamma(p/2)\}$,, 其中 $\Gamma(\cdot)$ 是伽马函数。。因此，渐近相对效率为 ARE($\hat{}\boldsymbol{\theta}_n$， $\bar{X}_n$) $= p\{\Gamma(p/2-1/2)\}^2/\{2^{1/2}\Gamma(p/2)\}^2.\hat{}^{-}$。利用斯特林公式，$\lim_{p\to\infty} \mathrm{ARE}(\boldsymbol{\theta}_n,\ X_n) = 1.$。因此，对于高维高斯数据，在对 $\boldsymbol{\theta}$进行同时推断时，样本空间中位数与样本均值具有相同的渐近效率。

例 2. 当数据来自自由度为$v > 2$ 、形状矩阵为 $\Omega = I_p$的多元 t-分布时， $\nu_i^2/p \sim F_{p,v}$, 其中 $F_{p,v}$ 是参数为

p和 v的分布。于是， $\mathbb{E}(\nu_i^2) = pv/(v-2)$ 和 $\mathbb{E}(\nu_i^{-1}) = \Gamma(v/2+1/2)\Gamma(p/2-1/2)/\{v^{1/2}\Gamma(v/2)\Gamma(p/2)\}$。因此，渐近相对效率为 ARE($\hat{\boldsymbol{\theta}}_n$, $\bar{X}_n$) $= (v-2)^{-1}p\{\Gamma(v/2+1/2)\Gamma(p/2-1/2)\}^2/\{\Gamma(v/2)\Gamma(p/2)\}^2$.显然，当 p足够大时，ARE($\boldsymbol{\theta}_n$, X) > 1。此外，$\lim_{p\to\infty}$ ARE($\boldsymbol{\theta}_n$, X_n) $= 2(v-2)^{-1}\{\Gamma(v/2+1/2)\}^2/\{\Gamma(v/2)\}^2 > 1$. 因此，对于高维 t-分布，在对 $\boldsymbol{\theta}$进行同时推断时，样本空间中位数的渐近效率高于样本均值。

图1展示了不同模型下，在一系列维度和样本量条件下模拟得到的 Var($|X_n|_\infty$)/Var($|\boldsymbol{\theta}_n|_\infty$) 值。对于高斯数据，相对效率随着 p的增加而持续提高，并在 p 增大时趋近于 1。对于从多元t分布模拟得到的数据，在所有 n 和 p的组合下，相对效率均大于 1。这表明，对于 t分布，样本空间中位数的效率高于样本均值。在模拟中考察的不同协方差结构下，结果均一致。

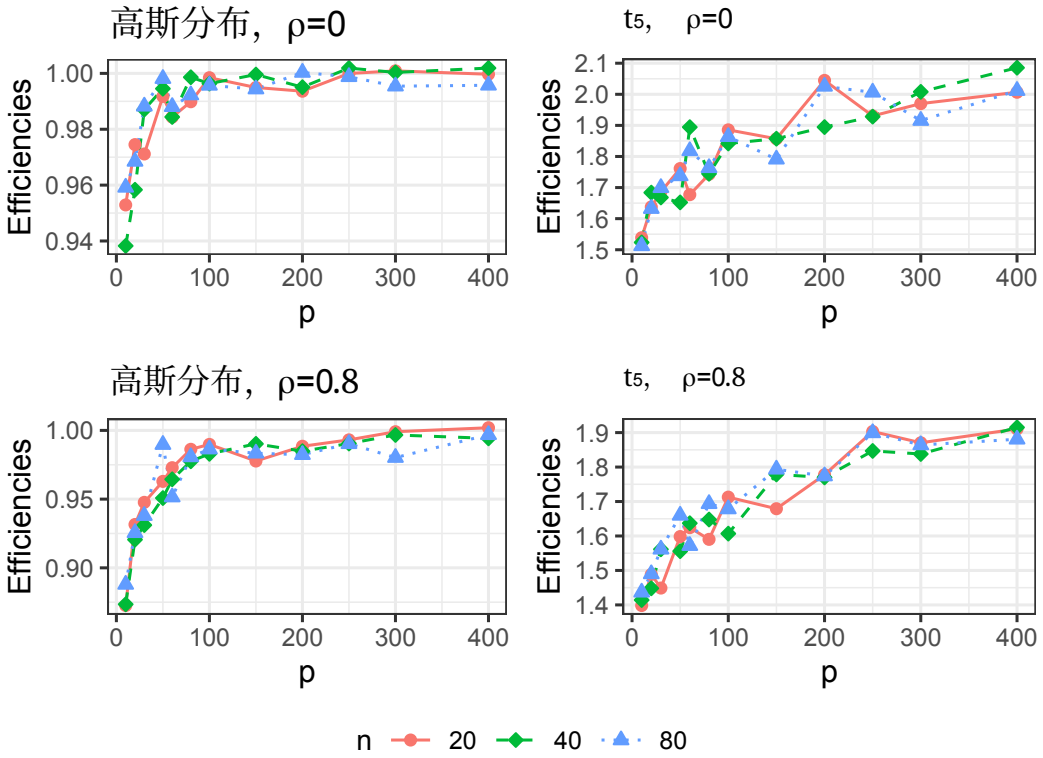


图1：基于5000次重复模拟，比较 $|\boldsymbol{\theta}_n|_\infty$ 与 $|X_n|_\infty$ 的有限样本相对效率。数据生成自多元正态分布（高斯分布）和自由度为5的 t分布（t_5）。形状矩阵$\Omega = (\rho^{|j-\ell|})_{p\times p}$，其中 $\rho = 0$和0.8。

5 数值研究

本节报告同时置信区间和采用错误发现率控制的多重检验的蒙特卡洛模拟结果，并进行真实数据分析，以展示所提出方法的性能。有关全局检验的更多模拟结果见补充材料的附录 C。所有模拟结果均基于 2500 次重复模拟。自助法的实现中，自助迭代次数设为$B = 400$。

5.1 同时置信区间的模拟

ˆ 我们首先考察基于 $\boldsymbol{\theta}_n$ 的同时置信区间的表现，并将其与基于¯ X_n的同时置信区间进行比较。样本量 n 取 100 或 200，对于每种样本量，考虑维度 $p = 100$ 和 1000。我们考虑两类常用的椭圆分布：(I) 多元正态分布 $N(\boldsymbol{\theta},\ \Sigma)$；(II) 自由度为 3、均值向量为 $\boldsymbol{\theta}$、协方差矩阵为 Σ 的多元 t 分布。此外，我们还纳入以下独立分量模型：(III) $X_i = \boldsymbol{\theta} + \Sigma^{1/2} Z_i,$，其中 Z_i 的每个分量均独立同分布于标准拉普拉斯分布。我们设定 $\Sigma = (\rho^{|j-\ell|})$，其中 $\rho = 0,$ 0.2，0.5 和 0.8。为节省篇幅，此处仅给出 $\rho = 0$ 和 0.8 的结果。$\rho \in \{0.2,\ 0.5\}$ 的结果相似，详见补充材料。对于 $\boldsymbol{\theta}$，我们考虑稀疏和稠密两种情形：(i) $\boldsymbol{\theta}_1 = (2,\ -2,\ 3,\ 0,\ \ldots,\ 0)$；(ii) $\boldsymbol{\theta}_2 = (0.2,\ \ldots,\ 0.2_{\lfloor p/4 \rfloor}, 0, \ldots, 0)$。此处 $\lfloor \cdot \rfloor$ 表示向下取整函数。

ˆ 表 1 报告了基于 $\boldsymbol{\theta}_n$ 的同时置信区间的覆盖概率和中位数长度，基于 X_n 的同时置信区间的¯结果列于括号内。对于来自椭圆¯分布的模型 I 和模型 II，我们观察到，基于 $\boldsymbol{\theta}_n$ 和 X_n 的同时置信区间在 ρ、$\boldsymbol{\theta}$、n 和 p 取不同值时均能达到令人满意的覆盖概率。对于从多元 ˆ正态分布模拟的数据，基于 $\boldsymbol{\theta}_n$ 的同时置信区间的中位数长度与基于¯ˆ ¯的同时置信区间的中位数长度非常接近。这些结果表明，在第 ˆ 4.4 节讨论的高维高斯模型下，对于 $\boldsymbol{\theta}$ 的同时推断，$\boldsymbol{\theta}_n$ 与 X_n 具有相似的渐近效率。对于多元 t分布，基于 $\boldsymbol{\theta}_n$ 的同时置信区间明显窄于基于 ¯ ˆ X_n的同时置信区间。这些结果表明，基于 $\boldsymbol{\theta}_n$ 的同时置信区间比同时置信区间

¯基于 X_n 的多元 t 分布的同时置信区间更有效，该分布具有重尾特征。这与第 4.4 节中的渐近分析一致。此外，模型 III 的结果表明，尽管模型 III 不属于椭圆分布族，但基于空间中位数的同时置信区间具有稳健性，其表现与基于样本均值的同时置信区间相似。我们还注意到，对于每个模型，当 n 增大或 p 减小时，同时置信区间的中位数长度会缩短。

ˆ 表 1：基于 $\boldsymbol{\theta}_n$的同时置信区间的覆盖概率（%）和中位数长度；基于 X_n 的同时置信区间结果见括号内。

				$\boldsymbol{\theta}=\boldsymbol{\theta}_1$				$\boldsymbol{\theta}=\boldsymbol{\theta}_2$			
				覆盖概率		中位数长度		覆盖概率		中位数长度	
模型	ρ	n	p	90%	95%	90%	95%	90%	95%	90%	95%
I	0	100	100	89.6 (89.9)	94.4 (94.4)	0.65 (0.65)	0.69 (0.69)	88.9（88.8）	94.1（93.9）	0.65（0.65）	0.69（0.69）
			1000	89.5 (89.6)	94.7 (94.4)	0.77 (0.77)	0.80 (0.80)	89.5（89.5）	94.0（94.0）	0.77（0.77）	0.81（0.80）
		200	100	89.8 (89.8)	95.1 (95.1)	0.46 (0.46)	0.49 (0.49)	88.6 (88.8)	94.4 (94.7)	0.46 (0.46)	0.49 (0.49)
			1000	89.7 (89.7)	94.4 (94.6)	0.55 (0.55)	0.57 (0.57)	89.1 (89.2)	94.7 (94.6)	0.55 (0.55)	0.57 (0.57)
	0.8	100	100	89.1 (88.7)	94.6 (94.6)	0.64 (0.63)	0.68 (0.67)	88.4 (88.6)	93.7 (94.1)	0.64 (0.63)	0.68 (0.67)
			1000	88.4 (88.4)	93.8 (93.7)	0.76 (0.76)	0.80 (0.79)	89.0 (89.2)	94.6 (94.6)	0.76 (0.76)	0.80 (0.79)
		200	100	90.5（90.1）	95.2（94.9）	0.45（0.45）	0.48（0.48）	89.6 (89.6)	94.0 (94.1)	0.45 (0.45)	0.48 (0.48)
			1000	90.4 (90.4)	94.5 (94.4)	0.54 (0.54)	0.56 (0.56)	88.4 (88.5)	93.6 (93.8)	0.54 (0.54)	0.56 (0.56)
II	0	100	100	89.7 (88.6)	94.7 (93.7)	0.71 (1.05)	0.75 (1.11)	88.8 (88.8)	94.5 (94.2)	0.71 (1.05)	0.75 (1.11)
			1000	89.4 (91.0)	95.8 (95.0)	0.84 (1.25)	0.88 (1.30)	89.1（89.0）	94.4（94.5）	0.84（1.25）	0.88（1.31）
		200	100	88.6 (89.1)	94.2 (95.1)	0.50 (0.76)	0.53 (0.81)	89.5（89.7）	94.4（94.8）	0.50（0.76）	0.53（0.80）
			1000	89.6 (88.7)	94.8 (94.6)	0.59 (0.90)	0.62 (0.94)	90.1 (89.5)	94.8 (93.9)	0.59 (0.90)	0.62 (0.94)
	0.8	100	100	89.1 (90.7)	94.4 (94.9)	0.69 (1.02)	0.74 (1.09)	89.4 (89.7)	94.2 (94.4)	0.69 (1.02)	0.74 (1.09)
			1000	89.3 (89.1)	94.6 (94.4)	0.83 (1.23)	0.87 (1.29)	89.8（88.8）	94.7（94.4）	0.83（1.23）	0.87（1.29）
		200	100	87.6 (87.7)	93.4 (93.6)	0.49 (0.73)	0.52 (0.78)	90.3（90.1）	94.9（95.2）	0.49（0.73）	0.52（0.78）
			1000	88.7 (89.7)	94.7 (94.6)	0.59 (0.88)	0.61 (0.92)	90.2 (90.8)	94.7 (95.7)	0.59 (0.89)	0.61 (0.93)
III	0	100	100	89.8 (89.4)	94.6 (94.5)	0.65 (0.66)	0.69 (0.70)	89.1 (89.0)	94.4 (94.4)	0.65 (0.66)	0.69 (0.70)
			1000	88.3 (88.2)	93.6 (93.7)	0.78 (0.78)	0.82 (0.82)	89.1 (89.0)	94.2 (93.8)	0.78 (0.78)	0.82 (0.82)
		200	100	90.6（91.1）	95.0（95.0）	0.46（0.46）	0.49（0.49）	90.6 (90.1)	95.2 (95.2)	0.46 (0.46)	0.49 (0.49)
			1000	90.1 (90.4)	95.0 (94.6)	0.55 (0.55)	0.57 (0.58)	88.7 (89)	93.6 (93.8)	0.55 (0.55)	0.57 (0.58)
	0.8	100	100	90.4（89.7）	95.0（94.8）	0.63（0.63）	0.68（0.68）	89.0 (88.9)	95.0 (94.9)	0.63 (0.63)	0.67 (0.68)
			1000	88.7 (88.9)	93.8 (94.0)	0.77 (0.77)	0.80 (0.80)	89.0（89.0）	94.6（94.3）	0.76（0.76）	0.80（0.80）
		200	100	88.8 (89.1)	94.2 (94.0)	0.45 (0.45)	0.48 (0.48)	90.2 (89.7)	94.8 (95.0)	0.45 (0.45)	0.48 (0.48)
			1000	90.0 (90.3)	95.0 (95.0)	0.54 (0.54)	0.57 (0.57)	88.8（89.1）	94.2（94.1）	0.54（0.54）	0.57（0.57）

5.2 关于错误发现率控制的多重检验模拟

本节考察第 4.3 节介绍的基于样本空间中位数的 B-H 方法的性能，并将其与基于样本均值的 B-H 程序进行比较，后者的 p 值根据 $N(0,\ 1)$ 在 Liu 和 Shao（2014）中的方法计算。我们对所有 $j=1,\ \ldots\ldots,\ p$设定 $\theta_{0,\ j}=0$。数据由模型 I 和模型 II 生成，其中 $p=1000$。对于 $\boldsymbol{\theta}=(\theta_1,\ \ldots\ldots,\ \theta_p)^{\top}$，令$\theta_j=2(\log p/n)^{1/2}$，其中 $1\leqslant j\leqslant p_1$，且 $\theta_j=0$，其中 $(p_1+1)\leqslant j\leqslant p$，其中 $p_1=0.1p$。

表 2 报告了基于样本空间中位数的 B-H 程序（经验FDR$_M$和检验功效$_M$）与基于样本均值的 B-H 程序（经验FDR$_A$ 和检验功效$_A$）的经验假发现率和检验功效（Liu 和 Shao

2014），其名义水平为 $\alpha = 0.1$ 和 0。2。结果表明，两种方法都能很好地控制假发现率。对于多元正态分布，基于空间中位数和样本均值的 B-H 程序表现相近。然而，就经验检验功效而言，在重尾的多元 t-分布下，基于样本空间中位数的 B-H 方法优于基于样本均值的 B-H 程序。

表 2：通过 B-H 程序，对 Liu 和 Shao (2014) 中基于空间中位数（假发现率$_M$ 和检验功效$_M$）及基于样本均值（假发现率$_A$ 和检验功效$_A$）的方法进行的经验假发现率和检验功效。

			$\alpha = 0.1$				$\alpha = 0.2$			
模型	ρ	n	假发现率$_M$	假发现率$_A$	检验功效$_M$	检验功效$_A$	假发现率$_M$	假发现率$_A$	检验功效$_M$	检验功效$_A$
I	0	50	0.124	0.124	0.996	0.996	0.224	0.222	0.999	0.999
		100	0.107	0.106	0.997	0.997	0.202	0.201	0.999	0.999
	0.2	50	0.125	0.124	0.996	0.996	0.224	0.223	0.999	0.999
		100	0.107	0.106	0.997	0.997	0.202	0.201	0.999	0.999
	0.5	50	0.125	0.124	0.996	0.996	0.225	0.223	0.999	0.999
		100	0.107	0.105	0.997	0.997	0.202	0.201	0.999	0.999
	0.8	50	0.127	0.124	0.996	0.996	0.227	0.223	0.999	0.999
		100	0.108	0.105	0.997	0.997	0.204	0.199	0.999	0.999
II	0	50	0.117	0.099	0.984	0.728	0.215	0.193	0.992	0.805
		100	0.103	0.088	0.987	0.710	0.197	0.179	0.994	0.795
	0.2	50	0.117	0.098	0.984	0.727	0.215	0.194	0.992	0.805
		100	0.103	0.087	0.987	0.709	0.198	0.179	0.994	0.795
	0.5	50	0.118	0.099	0.984	0.727	0.216	0.194	0.992	0.803
		100	0.103	0.087	0.987	0.708	0.198	0.178	0.994	0.794
	0.8	50	0.120	0.098	0.984	0.724	0.218	0.192	0.992	0.800
		100	0.104	0.087	0.987	0.705	0.199	0.177	0.994	0.791

5.3 真实数据分析

2 型糖尿病是一种人体对胰岛素的正常作用产生抵抗，并逐渐丧失足量胰岛素生成能力的疾病。由于骨骼肌是胰岛素刺激葡萄糖清除的主要组织，骨骼肌胰岛素抵抗通常被视为全身胰岛素抵抗的关键组成部分，因此对 2 型糖尿病的发病机制至关重要。为研究胰岛素对骨骼肌基因表达的影响，研究人员使用 Affymetrix Hu95A 芯片，对 15 名糖尿病患者在胰岛素治疗前后采集的肌肉活检样本开展了一项微阵列研究（Wu 等人，2007）。本文关注的是基因表达改变，即由治疗引起的基因表达水平变化，

数据可从 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE22309 获取。研究人员使用 limma R 软件包中的 normalizeQuantiles 函数进行分位数标准化。遵循 Wang、Peng 和 Li（2015年）的做法，我们聚焦于 2547 个经过整理的基因集，这些基因集均至少包含 15 个基因，来源于 GSEA 在线通路数据库中的 C2 集合。当多个探针对应同一个基因时，取其平均值来汇总基因表达值。

我们基于 T_n 对 2519 个基因集实施了中位数检验。这等价于检验基因表达水平变化向量的中位数是否等于 0。自助迭代次数为 $B = 10^5$。经 Bonferroni 校正后，有 1242 个基因集在 5% 显著性水平上被判定为显著。作为比较，我们对相同的基因集应用了 WPL 检验（Wang、Peng 和 Li 2015）和 CQ 检验（Chen 和 Qin 2010）。WPL 检验选出了 1060 个显著基因集；CQ 检验则判定 630 个基因集显著。在 CQ 检验选出的 630 个基因集中，我们提出的方法也识别出其中 605 个，WPL 检验则识别出其中 629 个。Wang、Peng 和 Li（2015）指出，某些基因表达水平具有厚尾特征，因为其峰度远高于正态分布的峰度 3。因此，基于空间中位数的方法（中位数检验和 WPL 检验）预计比基于矩的方法（CQ 检验）更加稳健且有效。此外，在 WPL 检验识别出的 1060 个基因集中，我们提出的方法判定其中 958 个显著。

正如注释 12 所述，基于 T_n 的中位数检验在检测强稀疏信号方面比 WPL 检验更有检验功效。为此，我们考察以下三个基因集：(1) ZHAN MULTIPLE MYELOMA UP；

(2) MIKKELSEN MEF HCP WITH H3K27ME3；
(3) JAZAG TGFB1 SIGNALING VIA SMAD4 UP。

这三个基因集的 WPL 检验 p 值分别为 0.41、0.31 和 0.27。然而，对于这三个基因集，采用 1.0×10^{-5} 次自助迭代后，中位数检验的 p 值均小于 $B = 10^5$。图 2 展示了这三个基因集的基因表达水平变化所对应的空间中位数向量的同时置信区间。未覆盖 0 的置信区间以红色标示。很明显，空间中位数中仅有一两个较大的值

导致中位数检验拒绝原假设，而其他维度上的信号不足以使 WPL 检验拒绝原假设。

最后，我们采用基于空间中位数的B-H程序，对三个基因集进行多重检验并控制错误发现率，以检测差异表达基因（DEG），这是基因组分析中最重要的目标之一。表3报告了在名义水平 $\alpha = 0.1$ 下，各基因集中检测到的差异表达基因（DEG），以及相应的边际 p 值 $P_j = 2 - 2\Phi(|T_{n,\ j}|)$ 和所选基因在同时置信区间中的置信区间。可以看出，对于所有选中的基因，边际 p 值都非常小，且相应的置信区间均不包含 0。

表3：采用基于空间中位数的B-H程序，在 $\alpha = 0.1$条件下对三个基因集检测到的差异表达基因（DEG）；“p-value”指边际 p-value $P_j = 2-2\Phi(|T_{n,\ j}|)$，“CI”指所选基因在同时置信区间中的置信区间。

基因集	DEG	p 值	CI
ZHAN_MULTIPLE_MYELOMA_UP	CDKN1A	0.00082	(0.234, 0.550)
MIKKELSEN_MEF_HCP_WITH_H3K27ME3	MYOD1	< 0.00001	(0.433, 0.791)
JAZAG经SMAD4上调的TGFB1信号通路	HDAC4	0.00058	(0.254, 0.644)

6 讨论

本文在椭圆分布之外的一般模型下，为超高维样本空间中位数建立了单样本和双样本高斯近似及自助法近似。研究我们的结果能否推广到其他一些分布族很有意义，我们将此留待未来研究。此外，与基于 L_2范数的检验相比，所提出的基于样本空间中位数最大值的检验在稀疏备择下具有更高的检验功效。众所周知，基于 L_2范数的检验在稠密备择下具有更高的检验功效。因此，考虑将基于最大范数的检验与 L_2范数检验相结合很有意义，这种方法有望在稀疏和稠密备择下均具有较高的检验功效。我们也将此留待未来研究。

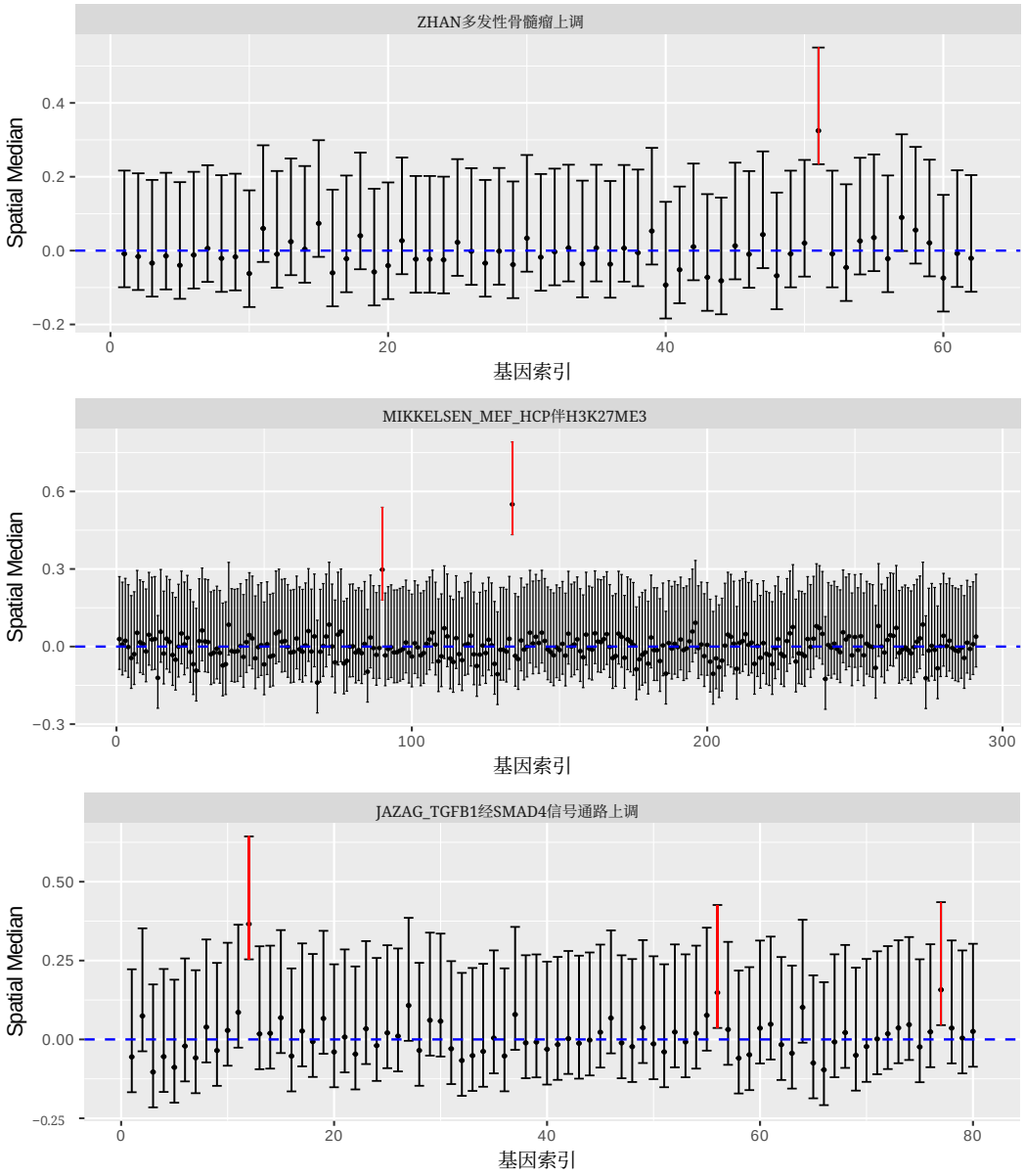


图 2：三个基因集的空间中位数同时置信区间（SCIs）。

补充材料

补充材料包含所有技术证明以及一些额外的数值结果。

参考文献

Belloni, A., Chernozhukov, V., Chetverikov, D., Hansen, C. & Kato, K.（2018年），‘高维计量经济学与广义 GMM’，$arXiv$ p.1806.01888。

Bickel, P. J. & Levina, E.（2008），‘通过阈值化进行协方差正则化’，《统计年鉴》**36**，2577–2604。

Brown, B. （1983），《空间中位数的统计用途》，《皇家统计学会杂志》B **45**，25–3 0.

Cardot, H., C´enac, P. & Zitt, P.-A.（2013），《在希尔伯特空间中使用平均随机梯度算法高效快速地估计几何中位数》，$Bernoulli$ **19**，18–43。

Chen, S. X. & Qin, Y.（2010），《用于高维数据的双样本检验及其在基因集检验中的应用》，《统计年鉴》**38**，808–835。

Chen, X.（2018年），《高维 U-统计量的高斯近似与自助法近似及其应用》，《统计年鉴》**46**，642–678。

Cheng, G., Liu, B., Peng, L., Zhang, B. & Zheng, S.（2019），《检验两个高维空间符号协方差矩阵的相等性》，《斯堪的纳维亚统计学杂志》**46**，257–271。

Chernozhukov, V.、Chetverikov, D. 和 Kato, K.（2013），《高维随机向量和的最大值的高斯分布近似与乘子自助法》，《统计年鉴》**41**，2786– 2819。

Chernozhukov, V.、Chetverikov, D. 和 Kato, K.（2017），《高维中的中心极限定理与自助法》，《概率年刊》**45**，2309–2352。

Chernozhukov, V., Chetverikov, D. & Kato, K. (2019), ‘改进的高维中心极限定理与自助法近似’,$arXiv$ 第 1912.10529 页。

Chernozhukov, V., Chetverikov, D., Kato, K. & Koike, Y. (2020), ‘近乎最优的高维中心极限定理与自助法近似’, $arXiv$ 第 2012.09513 页。Fan, J. & Lv, J. (2008), ‘超高维特征空间的确定性独立性筛选’,《*J. R. Statist. Soc.*》B **70**, 849–911. Fang, K. W., Kotz, S. & Ng, K. W. (1990), 对称多元分布及相关分布, 博卡拉顿, FL：CRC出版社。Haldane, J. B. S. (1948), ‘关于多元分布中位数的注记’, $Biometrika$ **35**, 414– 417. Hsu, D. & Sabato, S. (2016), ‘厚尾情形下的损失最小化与参数估计’, 机器学习研究杂志 **17**, 1–40. Hyv¨arinen, P., Karhunen, J. & Oja, E. (2001), 独立成分分析, 纽约：Wiley出版社。Imaizumi, M. & Otsu, T. (2021), ‘关于M估计量的高斯近似’, $arXiv$第 2012.15678v2 页。Koike, Y. (2021), ‘超矩形高维中心极限定理中的维数依赖性札记’, 日本统计与数据科学杂志 **1**, 257–297. Li, W. & Xu, Y. (2022), ‘椭圆分布下高维空间中位数的渐近性质及其应用’, 多元分析杂志 **190**, 104975. Liu, W. & Shao, Q.-M. (2014), ‘具有错误发现率控制的大规模t检验中的相变与正则化自助法’,《统计年鉴》 **42**, 2003–2025. Magyar, A. & Tyler, D. E. (2011), ‘椭圆对称分布下空间中位数的渐近效率’, $Sankhya\ B$ **73**, 165–192. McNeil, A. J., Frey, R. & Embrechts, P. (2005), 定量风险管理：概念、技术与工具, 普林斯顿, NJ：普林斯顿大学出版社。Milasevic, P. & Ducharme, G. R. (1987), ‘空间中位数的唯一性’,《统计年鉴》**15**, 1332–1333.

Minsker, S.（2015年），《Banach空间中的几何中位数与稳健估计》，*Bernoulli* **21**，2308–2335。

Oja, H.（2010），《基于空间符号和秩的多元非参数方法：*R*语言实践》，统计学讲义，Springer，纽约。

Prasad, A.、Suggala, A. S.、Balakrishnan, S. 与 Ravikumar, P.（2020年），《通过稳健梯度估计实现稳健估计》，《英国皇家统计学会杂志*B*辑》**82**，601–627。

Purdom, E. 与 Holmes, S. P.（2005），《基因表达数据的误差分布》，遗传学与分子生物学中的统计应用**4**，1–35。

van der Vaart, A. W. 与 Wellner, J. A.（1996），《弱收敛与经验过程：统计学应用》，Springer。

Wang, L., Peng, B. & Li, R. (2015年), '针对均值向量的高维非参数多元检验',美国统计学会杂志**110**, 1658–1669.

Weber, A.(1929), 工业区位（*Alfred Weber* 的工业区位理论），美国伊利诺伊州芝加哥：芝加哥大学出版社。

Wu, X., Wang, J., Cui, X., Maianu, L., Rhees, B., Rosinski, J., So, W. V., Willi, S. M., Osier, M. V., Hill, H. S., Page, G. P., Allison, D. B., Maritin, M. & Garvey, W. T. (2007), '胰岛素对人类骨骼肌中基因表达及生化通路的影响',*Endocrine* **31**,5–17.

Yao, J., Zheng, S. & Bai, Z. (2015年), 样本协方差矩阵与高维数据分析,剑桥大学出版社,剑桥。

Zou, C.、Peng, L.、Feng, L. 和 Wang, Z.（2014），《基于符号的多元高维球形性检验》，*Biometrika* **101**，229–236。

《基于空间中位数的超高维位置参数统计推断》补充材料

附录 A：技术证明

我们首先介绍并回顾一些记号。对于一个 $d_1 \times d_2$ 矩阵 $M = (m_{j\ell})_{d_1 \times d_2}$，其矩阵$\varrho$-范数为 $\|M\|_\varrho$ =上确界$\{\|Mx\|_\varrho$: $\|x\|_\varrho = 1\}$。具体而言， M的 1-、2- 和 8-范数分别为$\|M\|_1$ =最大值$_{1\leqslant\ell\leqslant d_2} \sum_{j=1}^{d_1} |m_{j\ell}|$、 $\|M\|_2 = \{\lambda_{\max}(M^\top M)\}^{1/2}$,和 $\|M\|_\infty$ =最大值$_{1\leqslant j\leqslant d_1} \sum_{\ell=1}^{d_2} |m_{j\ell}|$。 M的 Frobenius 范数为 $\|M\|_F = \{\sum_{j=1}^{d_1}\sum_{\ell=1}^{d_2} m_{j\ell}^2\}^{1/2}$。

定义一个随机 $p \times p$矩阵 $Q = n^{-1}\sum_{i=1}^n R_i^{-1}W_iW_i^\top$，使得$\mathbb{E}(Q) = \mathbb{E}(R_i^{-1}W_iW_i^\top)$，并将 $Q_{j\ell}$ 记为 Q的第 (j,ℓ)个元素。将 $\mathbb{E}^*(\cdot)$ 和 $\mathsf{Var}^*(\cdot)$ 分别记为在给定 $X_1,\ldots,X_n$条件下的期望和方差。回顾，$W_{i,j}$是 W_i 中第 j个元素，其中$i = 1,\ldots,n$ 且 $j = 1,\ldots,p$；$\omega_{j\ell}$是 Ω 的第 (j,ℓ)个元素； Γ_j是 Γ的第 j行。最后，我们用 $C_1, C_2, C_3, \ldots$ 表示各种正常数，除非另有说明。

A.1 预备引理

本节给出若干预备引理，其证明见在线附录 B。

引理 A1. （范数集中）假设条件 *C.1* 和 *C.3* 成立，且 $a_0(p) \asymp p^{1-\delta}$ 对某个正数常数 $\delta \leqslant 1/2$ 成立。那么，当 p 充分大时，存在正数常数 c_1 和 c_2，使得

$$\mathbb{P}\left\{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_1\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\right\} \geqslant 1 - c_1\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

and

$$\mathbb{P}\left\{(1-\epsilon)\mathrm{tr}(\Omega) \leqslant \|\Gamma U_1\|^2 \leqslant (1+\epsilon)\mathrm{tr}(\Omega)\right\} \geqslant 1 - c_1\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

对于任意固定的 $0 < \epsilon < 1$。

引理 A2。 假设条件*C.1*、*C.2*和*C.3*在 $a_0(p) \asymp p^{1-\delta}$ 成立，其中 $\delta \leqslant 1/2$为某个正数常数。那么，对于任意 $i = 1, \ldots, n$,

(i) $\mathbb{E}(\|U_i\|^4) = p\mathbb{E}(U_{i,\ j}^4) + p(p-1)$,

$$\begin{aligned}
\mathbb{E}(\|U_i\|^6) &= p\mathbb{E}(U_{i,j}^6) + 3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2), \\
\mathbb{E}(\|U_i\|^8) &= p\mathbb{E}(U_{i,j}^8) + 4p(p-1)\mathbb{E}(U_{i,j_1}^6) + 3p(p-1)\{\mathbb{E}(U_{i,j_1}^4)\}^2 \\
&\quad + 3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)(p-3)\,.
\end{aligned}$$

此外，对于任意正整数k， $\mathbb{E}(\|U_i\|^{2k}) = p^k + O(p^{k-1})$ 和 $\mathbb{E}(\|U\|^k) = p^{k/2} + O(p^{k/2-1})$ 。

(ii) $\mathbb{E}(\|\Gamma U_i\|^4) = p^2 + O(p^{2-\delta})$, $\mathbb{E}(\|\Gamma U_i\|^6) = p^3 + O(p^{3-\delta})$。此外， $\mathbb{E}(\|\Gamma U_i\|) = p^{1/2} + O(p^{1/2-\delta})$ 和 $\mathbb{E}(\|\Gamma U_i\|^3) = p^{3/2} + O(p^{3/2-\delta})$。

(iii) $\mathbb{E}\{\|\Gamma S(U_i)\|^2\} = 1 + O(p^{-1/2})$ 和 $\mathbb{E}\{\|\Gamma S(U_i)\|^4\} = 1 + O(p^{-1/3})$。

(iv) 当 $k = 1$ 时， $\mathbb{E}(\nu_i^{-k}) \lesssim \zeta_k p^{k/2}$ ， 2,3.

引理 A3. 假设条件*C.1*、*C.2*和*C.3*与某个正数常数$\delta \leqslant 1/2$ 下的 $a_0(p) \asymp p^{1-\delta}$ 成立。定义一个随机 $p \times p$ 矩阵 $Q = n^{-1}\sum_{i=1}^n R_i^{-1} W_i W_i^\top$， 并令 $Q_{j\ell}$ 为 Q的 (j, ℓ)个元素。那么，

(i) $|Q_{j\ell}| \lesssim \zeta_1 p^{-1}|\omega_{j\ell}| + O_p(\zeta_1 n^{-1/2} p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$。

(ii) $Q_{j\ell} = Q_{0,\ j\ell} + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$, 其中 $Q_{0,\ j\ell}$ 是第 $(j$、 $\ell)$ 个元素

$$Q_0 = n^{-1} p^{-1/2} \sum_{i=1}^n \nu_i^{-1} \{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top.$$

此外， Q_0 满足

$$\mathrm{tr}[\mathbb{E}(Q_0^2) - \{\mathbb{E}(Q_0)\}^2] = O(n^{-1}p^{-1}).$$

引理 A4. 假设条件*C.1*、*C.2*和*C.3*在某个正数常数$\delta \leqslant 1/2$ 下成立，其中 $a_0(p) \asymp p^{1-\delta}$ 。那么，

(i) $\mathbb{E}\{(\zeta_1^{-1} W_{i,\ j})^4\} \lesssim \bar{M}^2$ 且 $\mathbb{E}\{(\zeta_1^{-1} W_{i,\ j})^2\} \gtrsim m$ 对所有 $i = 1$、……、 n 和 $j = 1$、……、 p.

(ii) $\|\zeta_1^{-1} W_{i,\ j}\|_{\psi_\alpha} \lesssim \bar{B}$ 对所有 $i = 1, \ldots, n$ 和 $j = 1, \ldots, p$。

(iii) $\mathbb{E}(W_{i,\,j}^2) = p^{-1}\omega_{jj} + O(p^{-1-\delta/2})$ 对于 $j = 1, \ldots, p$ 和 $\mathbb{E}(W_{i,\,j}^2) = p^{-1}\omega_{j\ell} + O(p^{-1-\delta/2})$ 对于$1 \leqslant j \neq \ell \leqslant p$。

(iv) 若 $\log p = o(n^{1/3})$,

$$\left| n^{-1/2} \sum_{i=1}^{n} \zeta_1^{-1} W_i \right|_\infty = O_p\{\log^{1/2}(np)\} \quad \text{and} \quad \left| n^{-1} \sum_{i=1}^{n} (\zeta_1^{-1} W_i)^2 \right|_\infty = O_p(1)\,.$$

引理 A5. 假设定理 *2* 的条件成立，则

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1} \sum_{i=1}^{n} Z_i W_i + \tilde{C}_n\,, \tag{S.11}$$

其中 $|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$。

下述引理是 Nazarov 不等式，其证明见 Chernozhukov、Chetverikov 和 Kato（2017）。

引理 A6 （Nazarov 不等式）。 设 $Y_0 = (Y_{0,1}, \ldots, Y_{0,p})^\top$ 是 $\mathbb{R}^p$ 中的一个中心化高斯随机向量，且对于所有 $j = 1, \ldots, p$ 和某个常数 $b > 0$，有 $\mathbb{E}(Y_{0,j}^2) \geqslant b$，则对任意 $y \in \mathbb{R}^p$ 和 $a > 0$,

$$\mathbb{P}(Y_0 \leqslant y + a) - \mathbb{P}(Y_0 \leqslant y) \lesssim a\log^{1/2}(p)\,.$$

A.2 主要结果的证明

引理 *1* 的证明。 由于 $\boldsymbol{\theta}$ 是位置参数，不妨设 $\boldsymbol{\theta} = 0$ 。于是，对于 $i = 1, \ldots, n$，样本空间中位数 $\hat{\boldsymbol{\theta}}_n$ 满足

$$\sum_{i=1}^{n} S(X_i - \hat{\boldsymbol{\theta}}_n) = \sum_{i=1}^{n} \frac{X_i - \hat{\boldsymbol{\theta}}_n}{\|X_i - \hat{\boldsymbol{\theta}}_n\|} = \sum_{i=1}^{n} \frac{W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n}{\|W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n\|} = 0\,,$$

这等价于

$$n^{-1} \sum_{i=1}^{n} (W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1 - 2R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n + R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2)^{-1/2} = 0$$

因为 $W_i^\top W_i = 1$。

在条件 C.2 下， $\zeta_k = \mathbb{E}(R_i^{-k}) = O(p^{-k/2})$ 对于 $k = 1$，2，3，4。此外，引理

A3 表明 $Q_{j\ell} = Q_{0,\ j\ell} + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$, 其中 $Q_{0,\ j\ell}$ 是 $Q_0 = n^{-1}p^{-1/2}\sum_{i=1}^n \nu_i\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top$. 的第 $(j,\ \ell)$ 个元素。此外， Q_0 满足 tr $[\mathbb{E}(Q_0^2) - \{\mathbb{E}(Q_0)\}^2] = O(n^{-1}p^{-1})$. 因此，仿照 Cheng 等人（2019）引理 1.2 的证明中的类似程序，我们可以证明

$$\|\hat{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1}n^{-1/2}).$$

于是，对于 $i = 1$、.....、 n，有 $|R_i^{-1}W_i^T\hat{\boldsymbol{\theta}}_n| \leqslant R_i^{-1}\|\hat{\boldsymbol{\theta}}_n\| = O_p(n^{-1/2})$ 以及 $R_i^{-2}||\hat{\boldsymbol{\theta}}_n||^2 = O_p(n^{-1})$。根据一阶泰勒展开，上式可改写为

$$n^{-1}\sum_{i=1}^n (W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1 + R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i}) = 0\,, \tag{S.12}$$

其中 $\delta_{1i} = O_p\{(R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2)^2\} = O_p(n^{-1})$. 根据马尔可夫不等式，对于任意 $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\max_{1\leqslant i\leqslant n} R_i^{-1} \geqslant \varepsilon\zeta_1 n^{1/4}\right) &= \mathbb{P}\left(\max_{1\leqslant i\leqslant n} R_i^{-4} \geqslant \varepsilon^4\zeta_1^4 n\right)\\ &\leqslant \mathbb{E}\left(\max_{1\leqslant i\leqslant n} R_i^{-4}\right)/(\varepsilon^4\zeta_1^4 n) \leqslant n\mathbb{E}(R_i^{-4})/(\varepsilon^4\zeta_1^4 n) \lesssim \varepsilon^{-4}\,,\end{aligned}$$

其中，最后一个不等式源于条件 C.2。因此，$\max_{1\leqslant i\leqslant n} R_i^{-2} = O_p(\zeta_1^2 n^{1/2}),$，从而 $\max_{1\leqslant i\leqslant n}\delta_{1i} = O_p(\|\hat{\boldsymbol{\theta}}_n\|^2\max_{1\leqslant i\leqslant n} R_i^{-2}) = O_p(n^{-1/2})$。将 (S.12) 改写为

$$\begin{aligned}&n^{-1}\sum_{i=1}^n (1 - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i})W_i + n^{-1}\sum_{i=1}^n R_i^{-1}(W_i^\top\hat{\boldsymbol{\theta}}_n)W_i\\ =\ &n^{-1}\sum_{i=1}^n R_i^{-1}(1 - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i})\hat{\boldsymbol{\theta}}_n + n^{-1}\sum_{i=1}^n R_i^{-2}(W_i^\top\hat{\boldsymbol{\theta}}_n)\hat{\boldsymbol{\theta}}_n\,,\end{aligned}$$

这意味着

$$\begin{aligned}&n^{-1}\sum_{i=1}^n (1 - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i})W_i + n^{-1}\sum_{i=1}^n R_i^{-1}(W_i^\top\hat{\boldsymbol{\theta}}_n)W_i\\ =\ &n^{-1}\sum_{i=1}^n R_i^{-1}(1 + \delta_{1i} + \delta_{2i})\hat{\boldsymbol{\theta}}_n\,,\end{aligned} \tag{S.13}$$

其中 $\delta_{2i} = R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 = O_p(\delta_{1i}^{1/2})$ 满足 $\max_{1\leqslant i\leqslant n}\delta_{2i} = O_p(n^{-1/4})$。这是

不难验证 $n^{-1}\sum_{i=1}^n R_i^{-1}(W_i^\top \hat{\boldsymbol{\theta}}_n)W_i = n^{-1}\sum_{i=1}^n R_i^{-1}W_iW_i^\top \hat{\boldsymbol{\theta}}_n = Q\hat{\boldsymbol{\theta}}_n$。由引理 A3,

$$|Q_{j\ell}| \lesssim \zeta_1 p^{-1}|\omega_{j\ell}| + O_p(\zeta_1 n^{-1/2}p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2}),$$

由此可得

$$|Q\hat{\boldsymbol{\theta}}_n|_\infty \leqslant \|Q\|_1 |\hat{\boldsymbol{\theta}}_n|_\infty \lesssim \zeta_1 p^{-1}\|\Omega\|_1 |\hat{\boldsymbol{\theta}}_n|_\infty + O_p(\zeta_1 n^{-1/2} + \zeta_1 p^{-1/6} + \zeta_1 p^{-\delta/2})|\hat{\boldsymbol{\theta}}_n|_\infty.$$

根据引理 A4，我们有 $|n^{-1}\sum_{i=1}^n \zeta_1^{-1}W_i|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}$。于是,

$$\left|\zeta_1^{-1}n^{-1}\sum_{i=1}^n \delta_{1i}W_i\right|_\infty^2 \leqslant \left|n^{-1}\sum_{i=1}^n(\zeta_1^{-1}W_i)^2\right|_\infty \left(n^{-1}\sum_{i=1}^n \delta_{1i}^2\right) \lesssim O_p(n^{-2}).$$

此外，我们有 $|\zeta_1^{-1}n^{-1}\sum_{i=1}^n R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 W_i|_\infty \lesssim O_p(n^{-1})$。结合等式 (S.13) 以及 $\zeta_1^{-1}n^{-1}\sum_{i=1}^n R_i^{-1} = 1 + O_p(n^{-1/2})$, 这一事实，可得

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_n|_\infty &\lesssim \left|\zeta_1^{-1}n^{-1}\sum_{i=1}^n W_i\right|_\infty + \zeta_1^{-1}|Q\hat{\boldsymbol{\theta}}_n|_\infty \\
&\lesssim p^{-1}a_0(p)|\hat{\boldsymbol{\theta}}_n|_\infty + O_p(n^{-1/2} + p^{-(1/6\wedge\delta/2)})|\hat{\boldsymbol{\theta}}_n|_\infty + O_p\{n^{-1/2}\log^{1/2}(np)\}.
\end{aligned}$$

因此，当 $a_0(p) \asymp p^{1-\delta}$ 时，我们得出 $|\hat{\boldsymbol{\theta}}_n|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}$ 。此外，我们有$|\zeta_1^{-1}Q\hat{\boldsymbol{\theta}}_n|_\infty = O_p\{n^{-1/2}p^{-(1/6\wedge\delta/2)}\log^{1/2}(np) + n^{-1}\log^{1/2}(np)\}$，且 $n^{-1}\sum_{i=1}^n R_i^{-1}(1+\delta_{1i}+\delta_{2i}) = \zeta_1\{1+O_p(n^{-1/4})\}$。最后，我们可以写成

$$n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n W_i + C_n\,,$$

其中 C_n 满足 $|C_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$. □

定理*1*的证明。令 $L_{n,p} = n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)$。那么，对于任意序列$\eta_n \to \infty$ 和任意 $t \in \mathbb{R}^p$,

$$\mathbb{P}\left\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\right\} = \mathbb{P}\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n W_i + C_n \leqslant t\right)$$

$$\leqslant \quad \mathbb{P}\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} W_i \leqslant t + \eta_n L_{n,p}\right) + \mathbb{P}(|C_n|_\infty > \eta_n L_{n,p})\,.$$

根据引理 A4, $\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^4\} \lesssim \bar{M}^2$, $\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \gtrsim m$, 以及 $\|\zeta_1^{-1}W_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$ 对所有 $i = 1, \ldots, n$ 和 $j = 1, \ldots, p$ 均成立。根据 Koike (2021) 中关于独立部分和的高斯近似, 令 $G \sim N(0, \zeta_1^{-2}\mathbb{B})$ 满足 $\mathbb{B} = \mathbb{E}(W_1 W_1^\top)$, 我们有

$$\begin{aligned}
&\mathbb{P}\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} W_i \leqslant t + \eta_n L_{n,p}\right) \leqslant \mathbb{P}(G \leqslant t + \eta_n L_{np}) + O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right) \\
\leqslant \quad &\mathbb{P}(G \leqslant t) + O\{\eta_n L_{np}\log^{1/2}(p)\} + O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right),
\end{aligned}$$

其中, 最后一个不等式来自引理 A6 中的 Nazarov 不等式。还值得注意的是, Chernozhukov 等人（2019）将阶 $O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right)$ 改进为 $O\left(\left\{n^{-1}\log^5(np)\right\}^{1/4}\right)$。因此,

$$\begin{aligned}
&\mathbb{P}\left\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\right\} \\
\leqslant \quad &\mathbb{P}(G \leqslant t) + O\{\eta_n L_{np}\log^{1/2}(p)\} + O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right) \\
&+\mathbb{P}(|C_n|_\infty > \eta_n L_{n,p})\,.
\end{aligned}$$

另一方面，我们还有

$$\begin{aligned}
&\mathbb{P}\left\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\right\} \\
\geqslant \quad &\mathbb{P}(G \leqslant t) - O\{\eta_n L_{np}\log^{1/2}(p)\} - O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right) \\
&-\mathbb{P}(|C_n|_\infty > \eta_n L_{n,p})\,,
\end{aligned}$$

其中 $\mathbb{P}(|C_n|_\infty > \eta_n L_{n,\ p}) \to 0$, 根据引理 1, $n \to \infty$ 。

于是, 若 $\log p = o(n^{1/5})$ 和 $\log\ n = o(p^{1/3\wedge\delta})$, 满足足够缓慢的 $\eta_n \to \infty$, 则有

$$\sup_{t\in\mathbb{R}^p}\left|\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\} - \mathbb{P}(G \leqslant t)\right| \to 0\,.$$

我们得到 imm 直接由 Chernozhukov、Chetverikov 和 Kato（2017）中的推论 5.1 可知 t

$$\rho_n(\mathcal{A}^{\mathrm{re}}) = \sup_{A\in\mathcal{A}^{\mathrm{re}}} \left|\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\} - \mathbb{P}\left(G \in A\right)\right| \to 0\,,$$

由此得出本定理的结论。 □

~ ˆ ~ ~定理 *2* 的证明。令 $X_i = X_i - \boldsymbol{\theta}_n$ 和 $R_i = \|X_i\|$ 满足 $i = 1, \ldots, n$。根据引理 A5,

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i W_i + \tilde{C}_n\,,$$

其中 $\tilde{C}_n$ 满足 $|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$。

记 $\bar{W}_n = n^{-1}\sum_{i=1}^{n} W_i$ 并改写

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i(W_i - \bar{W}_n) + \left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i\right)\bar{W}_n + \tilde{C}_n,$$

其中

$$\left|\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i\right)\bar{W}_n\right|_\infty \leqslant \zeta_1^{-1}\left|n^{-1/2}\sum_{i=1}^{n} Z_i\right|\left|\bar{W}_n\right|_\infty \lesssim n^{-1/2}\log^{1/2}(np)$$

根据引理A4（iii）。

显然 ， $\mathbb{E}^*\left\{n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i(W_i - \bar{W}_n)\right\} = 0$。令ˆ$\mathbb{B} = n^{-1}\sum_{i=1}^{n} W_i W_i^\top$， 则

$$\mathrm{Var}^*\left\{n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i(W_i - \bar{W}_n)\right\} = \zeta_1^{-1}\hat{\mathbb{B}} - \zeta_1^{-2}\bar{W}_n\bar{W}_n^\top.$$

ˆ ˆ ¯分别表示 $\mathbb{B}_{j\ell}$ 和 $\mathbb{B}_{j\ell}$ 的第 ℓ)个元素，其中 $(j$, $\mathbb{B}$ 和 $\mathbb{B}$分别为。此外，记 $W_{n,j}$¯为 W_n 的第 j个元素。定义

$$\Delta_n = \max_{1\leqslant j,\ell\leqslant p}\left|\zeta_1^{-2}\hat{\mathbb{B}}_{j\ell} - \zeta_1^{-2}\bar{W}_{n,j}\bar{W}_{n,\ell} - \zeta_1^{-2}\mathbb{B}_{j\ell}\right|,$$

then

$$\Delta_n \leqslant \Delta_{n1} + \max_{1\leqslant j,\ell\leqslant p} \left|\zeta_1^{-2}\bar{W}_{n,j}\bar{W}_{n,\ell}\right| \lesssim \Delta_{n1} + n^{-1}\log(np),$$

其中

$$\Delta_{n1} = \max_{1\leqslant j,\ell\leqslant p} |\zeta_1^{-2}\hat{\mathbb{B}}_{j\ell} - \zeta_1^{-2}\mathbb{B}_{j\ell}| = \max_{1\leqslant j,\ell\leqslant p} \left| n^{-1}\zeta_1^{-2}\sum_{i=1}^{n}\{W_{i,j}W_{i,\ell} - \mathbb{E}(W_{i,j}W_{i,\ell})\}\right|.$$

根据 ψ_α 范数的性质，可得

$$\begin{aligned}\left\|\max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p} |\zeta_1^{-2}W_{i,j}W_{i,\ell}|\right\|_{\psi_{\alpha/2}} &\lesssim \left\|\max_{1\leqslant i\leqslant n,1\leqslant j\leqslant p} |\zeta_1^{-2}W_{i,j}|^2\right\|_{\psi_{\alpha/2}} \\ &= \zeta_1^{-2}\left\|\max_{1\leqslant i\leqslant n,1\leqslant j\leqslant p} |W_{i,j}|\right\|_{\psi_\alpha}^2 \lesssim \log^2(np)\,.\end{aligned}$$

令 $J_n = \max_{1\leqslant i\leqslant n;1\leqslant j,\ \ell\leqslant p}\zeta_1^{-2}|W_{i,\ j}W_{i,\ \ell} - \mathbb{E}(W_{i,\ j}W_{i,\ \ell})|$，并且

$$\begin{aligned}\sigma_n^2 &= \max_{1\leqslant j,\ell\leqslant p}\zeta_1^{-2}\sum_{i=1}^{n}\mathbb{E}\{W_{i,j}W_{i,\ell} - \mathbb{E}(W_{i,j}W_{i,\ell})\}^2 \\ &\lesssim \max_{1\leqslant j,\ell\leqslant p}\zeta_1^{-2}\sum_{i=1}^{n}\mathbb{E}\{|W_{i,j}W_{i,\ell}|^2\} \lesssim n\,.\end{aligned}$$

还可得

$$\|J_n\|_{\psi_{\alpha/2}} \lesssim \zeta_1^{-2}\left\|\max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p}|W_{ij}W_{i,\ell}|\right\|_{\psi_{\alpha/2}} + \max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p}\zeta_1^{-2}\mathbb{E}(|W_{i,j}W_{i,\ell}|) \lesssim \log^2(np)\,.$$

根据 Chernozhukov、Chetverikov 和 Kato（2017）中的引理 E.1，可得

$$\begin{aligned}\mathbb{E}(\Delta_{n1}) &\lesssim n^{-1}\left[\sigma_n\log^{1/2}(p) + \{\mathbb{E}(J_n^2)\}^{1/2}\log p\right] \\ &\lesssim n^{-1}\left\{n^{1/2}\log^{1/2}(p) + \log^{1/\alpha+1}(np)\right\} \\ &\lesssim n^{-1/2}\log^{1/2}(np)\,.\end{aligned}$$

然后应用 Chernozhukov、Chetverikov 和 Kato（2017）中的引理 E.2 与 $\eta = 1$ 以及

$\beta = \alpha/2$, 我们得到

$$\mathbb{P}(\Delta_{n1} \geqslant 2\mathbb{E}(\Delta_n) + t) \lesssim \exp\left(-C_1 n t^2\right) + 3\exp\left\{-C_2\{tn\log^{-2/\alpha}(np)\}^{\alpha/2}\right\}.$$

因此，存在一个依赖于 δ 的常数 C_1 ，使得

$$\mathbb{P}\left\{\Delta_{n1} > C_1 n^{-1/2}\log^{1/2}(np)\right\} \lesssim p^{-\delta} \to 0\,.$$

根据 Chernozhukov、Chetverikov 和 Kato（2017）以及 Koike (2021) 中的乘子自助法定理和高斯分布比较,

$$\begin{aligned} & \sup_{t\in\mathbb{R}^p}\left|\mathbb{P}^*\left\{n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n Z_i(W_i - \bar{W}_n) \leqslant t\right\} - \mathbb{P}(G \leqslant t)\right| \\ \lesssim\ & \Delta_n^{1/2}\log(p) + \{n^{-1}\log^5(np)\}^{1/4}\,, \end{aligned}$$

关于 $\{\Delta_n \lesssim n^{-1/2}\log^{1/2}(np)\},$ ，该结论以概率 $1 - p^{-\delta}$成立。

最后，类似于定理1的证明，我们可以证明，在条件C.2和C.3以及 $a_0(p) \asymp p^{1-\delta}$成立的情况下，若$\log p = o(n^{1/5})$ 和$\log\ n = o(p^{1/3\wedge\delta}),$ ，则有

$$\sup_{A\in\mathcal{A}^{\mathrm{re}}}\left|\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\} - \mathbb{P}^*(n^{1/2}\tilde{\boldsymbol{\theta}}_n \in A)\right| \to 0$$

依概率成立，从而完成该定理的证明。 □

定理3的证明。定理1和定理2表明，存在一个正序列$\beta_{n,\ p} \to 0$，当 $n,\ \ p \to \infty$ 时，使得

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}(|G|_\infty \leqslant t)\right| \leqslant \beta_{n,p}/2$$

and

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant t)\right| \leqslant \beta_{n,p}$$

当 $n \to \infty$时，以趋近于 1 的概率成立。令 $q_{1-\alpha}$ 为 $(1-\alpha)$分位数，即

$n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty$，即 $q_{1-\alpha} = \inf\{u \in \mathbb{R} : \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > u) \leqslant \alpha\}$。那么，

$$\mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant q_{1-\alpha+\beta_{n,p}}) \geqslant \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha+\beta_{n,p}}) - \beta_{n,p} \geqslant 1 - \alpha\,,$$

当 $n \to \infty$ 趋于某值时，该结论以趋于 1 的概率成立。另一方面，以相同的概率有

$$\begin{aligned}
& \mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) \\
\leqslant\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) + \beta_{n,p} \\
=\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6}) + \beta_{n,p} \\
& + \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) \\
& - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6}) \\
<\ & 1 - \alpha - 2\beta_{n,p} + \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) \\
& - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6})\,,
\end{aligned}$$

其中，$\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,\ p}}) - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,\ p}} - n^{-1/6})$ 可被界定为

$$\begin{aligned}
& \Big|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) \\
& \quad - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6})\Big| \\
\leqslant\ & \Big|\mathbb{P}(|G|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) - \mathbb{P}(|G|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6})\Big| \\
& + \Big|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) - \mathbb{P}(|G|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}})\Big| \\
& + \Big|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6}) \\
& \qquad - \mathbb{P}(|G|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6})\Big| \\
\leqslant\ & \Big|\mathbb{P}(|G|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) - \mathbb{P}(|G|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}} - n^{-1/6})\Big| + \beta_{n,p} \\
\leqslant\ & C_1\left\{n^{-1}\log^5(np)\right\}^{1/6} + \beta_{n,p},
\end{aligned}$$

其中 C_1 为某个正数常数，最后一个不等式由 Nazarov 不等式得出。

取 $C_1\left\{n^{-1}\log^5(np)\right\}^{1/6} \leqslant \beta_{n,p},$，得到

$$\mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant q_{1-\alpha-3\beta_{n,p}}) < 1-\alpha$$

以趋近于 1 的概率成立。因此，

$$\mathbb{P}(q_{1-\alpha-3\beta_{n,p}} < q^B_{1-\alpha} \leqslant q_{1-\alpha+\beta_{n,p}}) \to 1, \quad \text{as } n,p \to \infty\,.$$

因此，

$$\begin{aligned} & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q^B_{1-\alpha}) \\ \leqslant\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q_{1-\alpha-3\beta_{n,p}}) + \mathbb{P}(q^B_{1-\alpha} \leqslant q_{1-\alpha-3\beta_{n,p}}) \\ \leqslant\ & \alpha + 3\beta_{n,p} + o(1) \end{aligned} \tag{S.14}$$

and

$$\begin{aligned} & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q^B_{1-\alpha}) \\ \geqslant\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q_{1-\alpha+\beta_{n,p}}) - \mathbb{P}(q^B_{1-\alpha} > q_{1-\alpha+\beta_{n,p}}) \\ \geqslant\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q_{1-\alpha+\beta_{n,p}} - n^{-1/6}) - o(1) \\ & + \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q_{1-\alpha+\beta_{n,p}}) \\ & - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > q_{1-\alpha+\beta_{n,p}} - n^{-1/6}) \\ \geqslant\ & \alpha - 2\beta_{n,p} - C_2\left\{n^{-1}\log^5(np)\right\}^{1/6} \geqslant \alpha - 3\beta_{n,p}. \end{aligned}$$

对于某个正数常数 C_2，其中倒数第二个不等式由 Nazarov 不等式推出，最后一个不等式则是通过选取 $\beta_{n,\ p} \geqslant C_2\left\{n^{-1}\log^5(np)\right\}^{1/6}$ 得到的。最后，由于$\beta_{n,\ p} \to 0$,

$$\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \geqslant q^B_{1-\alpha}) - \alpha \to 0,$$

这就完成了该定理的证明。 □

定理 4 的证明。不失一般性，假设 $\boldsymbol{\theta}_0 = 0$。将检验统计量重写为$T_n = n^{1/2}|\hat{\boldsymbol{\theta}}_n|_\infty$，并令 $T^c_n = n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty$；在 H_0 下，它与 T_n 具有相同的分布。

于是，有

$$T_n \geqslant n^{1/2}|\boldsymbol{\theta}|_\infty - T_n^c\,.$$

因此，基于 T_n 的检验功效满足

$$\begin{aligned}\mathbb{P}(T_n > q_{1-\alpha}^B \mid H_1) &\geqslant \mathbb{P}(n^{1/2}|\boldsymbol{\theta}|_\infty - T_n^c \geqslant q_{1-\alpha}^B \mid H_1)\\ &= \mathbb{P}(T_n^c \leqslant n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\alpha}^B \mid H_1)\end{aligned}$$

在定理 2 的条件下，存在一个正序列 $\beta_{n,p} \to 0$ ，当 $n, p \to \infty$时，满足

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(T_n^c > t \mid H_1) - \mathbb{P}(|G|_\infty > t \mid H_1)| \leqslant \beta_{n,p}, \tag{S.15}$$

其中 $G \sim N(0, \zeta_1^{-2}\mathbb{B})$。令 $q_{1-\alpha}$为 T_n^C的 $(1-\alpha)$分位数， $q_{1-\alpha}^G$为 $|G|_\infty$的 $(1-\alpha)$分位数。在方程 (S.15) 中取 $t = q_{1-\alpha+2\beta_{n,p}}^G$ ，可得 $|\mathbb{P}(T_n^c > q_{1-\alpha+2\beta_{n,p}}^G |H_1) - \alpha + 2\beta_{n,p}| \leqslant \beta_{n,p}$ 且 $\mathbb{P}(T_n^c > q_{1-\alpha+2\beta_{n,p}}^G \mid H_1) \leqslant \alpha - \beta_{n,p}$，这意味着 $q_{1-\alpha+\beta_{n,p}} \leqslant q_{1-\alpha+2\beta_{n,p}}^G$。

注意， $q_{1-\alpha}^B$是给定 X_1、.....、 X_n时 $n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty$的第 $(1-\alpha)$分位数。通过采用与公式（ S.14）的证明中类似的程序，我们可以证明

$$\mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\alpha}^B \mid H_1) \leqslant \mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\alpha+\beta_{n,p}} \mid H_1) + o(1)\,. \tag{S.16}$$

由此可得

$$\mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\alpha}^B \mid H_1) \leqslant \mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\alpha+2\beta_{n,p}}^G \mid H_1) + o(1)\,.$$

对于 $|G|_\infty$，我们知道 $\||G|_\infty\|_{\psi_2} \lesssim \log^{1/2}(np)$。此外，对于任意 $t > 0$,

$$\mathbb{P}(|G|_\infty > t) \leqslant 2\exp\{-C_1(t/\||G|_\infty\|_{\psi_2})^2\} \leqslant 2\exp\{-C_2 t^2 \log^{-1}(np)\}\,.$$

取 $t = C_2^{-1/2}\log^{1/2}(2/(\alpha - 2\beta_{n,p}))\log^{1/2}(np),$ ，可得

$$\mathbb{P}(|G|_\infty > C_2^{-1/2}\log^{1/2}(2/(\alpha - 2\beta_{n,p}))\log^{1/2}(np)) \leqslant \alpha - 2\beta_{n,p}\,,$$

由此可得

$$q^G_{1-\alpha+2\beta_{n,p}} \leqslant C_2^{-1/2}\log^{1/2}(2/(\alpha-2\beta_{n,p}))\log^{1/2}(np).$$

那么，若 $|\boldsymbol{\theta}|_\infty \geqslant Cn^{-1/2}\log^{1/2}(\alpha^{-1})\log^{1/2}(np)$ 对于足够大的常数 C 成立，则当 C_3 足够大时，有

$$\begin{aligned}
& \mathbb{P}(T_n > q^B_{1-\alpha} \mid H_1) \\
\geqslant\ & \mathbb{P}(T_n^c \leqslant n^{1/2}|\boldsymbol{\theta}|_\infty - q^G_{1-\alpha+2\beta_{n,p}} \mid H_1) + o(1) \\
\geqslant\ & \mathbb{P}\{T_n^c \leqslant C_3\log^{1/2}(np)\log^{1/2}(\alpha^{-1}) \mid H_1\} + o(1) \\
\geqslant\ & \mathbb{P}\{|G|_\infty \leqslant C_3\log^{1/2}(np)\log^{1/2}(\alpha^{-1}) \mid H_1\} - \beta_{n,p} + o(1) \\
\geqslant\ & 1 - 2\alpha^{C_2C_3^2} - \beta_{n,p} + o(1).
\end{aligned}$$

我们完成了该定理的证明。 □

定理 5 的证明。 回顾，$\hat{}\ \zeta_1 = n^{-1}\sum_{i=1}^n \|X_i - \hat{}\boldsymbol{\theta}_n\|^{-1}$。引理 A5 的证明中已证明

$$\|X_i - \hat{\boldsymbol{\theta}}_n\|^{-1} = R_i^{-1}\left(1 + R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \tilde{\delta}_{1i}\right),$$

其中 $\tilde{\delta}_{1i}$ 满足 $\tilde{\delta}_{1i} = O_p(n^{-1})$，且 $\max_{1\leqslant i\leqslant n}\tilde{\delta}_{1i} = O_p(n^{-1/2})$。因此，

$$\begin{aligned}
\hat{\zeta}_1 &= n^{-1}\sum_{i=1}^n R_i^{-1}\left(1 + R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \tilde{\delta}_{1i}\right) \\
&= n^{-1}\sum_{i=1}^n R_i^{-1}(1 + \tilde{\delta}_{3i}),
\end{aligned}$$

其中，$\tilde{\delta}_{3i}$ 满足 $\tilde{\delta}_{3i} = O_p(n^{-1/2})$，且 $\max_{1\leqslant i\leqslant n}\tilde{\delta}_{3i} = O_p(n^{-1/4})$。根据 $n^{-1}\sum_{i=1}^n R_i^{-1} = \zeta_1 + O_p(\zeta_1 n^{-1/2})$, 这一事实，我们得出

$$\hat{\zeta}_1/\zeta_1 - 1 = O_p(n^{-1/2}).$$

˜ ˆˆ令 $W_i = (X_i - \boldsymbol{\theta}_n)/\|X_i - \boldsymbol{\theta}_n\|$，其中 $i = 1, \ldots\ldots, n$。根据引理 A5 的证明 ，

$$\tilde{W}_i = (W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1 + \tilde{\delta}_{2i}) = W_i + W_i\tilde{\delta}_{2i} - R_i^{-1}\hat{\boldsymbol{\theta}}_n(1 + \tilde{\delta}_{2i}),$$

其中， $\tilde{\delta}_{2i}$ 满足 $\tilde{\delta}_{2i} = O_p(n^{-1/2})$ 且 $\max_{1\leqslant i\leqslant n}\tilde{\delta}_{2i} = O_p(n^{-1/4})$。令$\tilde{W}_{i,j}$为 W_i的第 j个分量˜，则

$$\begin{aligned}
\hat{\mathbb{B}}_{jj} &= n^{-1}\sum_{i=1}^{n}\tilde{W}_{i,j}^2 \\
&= n^{-1}\sum_{i=1}^{n}W_{i,j}^2\{1 + O_p(\tilde{\delta}_{2i})\} + n^{-1}\sum_{i=1}^{n}R_i^{-1}W_{i,j}\hat{\theta}_{n,j}\{1 + O_p(\tilde{\delta}_{2i})\} \\
&\quad + n^{-1}\sum_{i=1}^{n}R_i^{-2}\hat{\theta}_{n,j}^2\{1 + O_p(\tilde{\delta}_{2i})\},
\end{aligned}$$

其中 $\max_{1\leqslant j\leqslant p} n^{-1}\sum_{i=1}^{n}W_{i,j}^2\{1 + O_p(\tilde{\delta}_{2i})\}/\mathbb{B}_{jj} = 1 + O_p(n^{-1/4})$,

$$\begin{aligned}
&\max_{1\leqslant j\leqslant p}\left\|n^{-1}\sum_{i=1}^{n}R_i^{-2}\hat{\theta}_{n,j}^2\{1 + O_p(\tilde{\delta}_{2i})\}\right\| \\
\lesssim\ &\left|n^{-1}\sum_{i=1}^{n}R_i^{-2}\right|\max_{1\leqslant j\leqslant p}|\hat{\theta}_{n,j}^2| \\
=\ &O_p\{\zeta_1^2 n^{-1}\log^{1/2}(np)\}
\end{aligned}$$

and

$$\begin{aligned}
&\max_{1\leqslant j\leqslant p}\left|n^{-1}\sum_{i=1}^{n}R_i^{-1}W_{i,j}\hat{\theta}_{n,j}\{1 + O_p(\tilde{\delta}_{2i})\}\right| \\
\lesssim\ &\left(n^{-1}\sum_{i=1}^{n}R_i^{-2}\right)^{1/2}\max_{1\leqslant j\leqslant p}\left\{\left(n^{-1}\sum_{i=1}^{n}W_{i,j}^2\right)^{1/2}\right\}\max_{1\leqslant j\leqslant p}|\hat{\theta}_{n,j}| \\
=\ &O_p\{\zeta_1 p^{-1/2}n^{-1/2}\log^{1/2}(np)\}.
\end{aligned}$$

由此可得

$$\max_{1\leqslant j\leqslant p}\hat{\mathbb{B}}_{jj}/\mathbb{B}_{jj} = 1 + O_p\{n^{-1/4}\log^{1/2}(np)\}.$$

因此，Belloni 等人（2018年）的条件 A (ii) 由 $s_{n,\ j}$ 满足。显然，Belloni 等人（2018年）的条件 A (i) 由余项 C_n 满足。因此，根据 Belloni 等人（2018年）的定理 2.4，对于任意 $1 \leqslant j \leqslant p$，若 $\log p = o(n^{1/5})$ 且 $\log\ n = o(p^{1/3 \wedge \delta}),$，则有

$$\sup_{0 \leqslant x \leqslant 2^{1/2} \log^{1/2}(np)} \left| \mathbb{P}\left\{ n^{1/2}(\hat{\theta}_{n,j} - \theta_j)/s_{n,j} > x \right\} - \{1 - \Phi(x)\} \right| \to 0\,.$$

令 $\bar{T}_j = n^{-1/2} \sum_{i=1}^n W_{i,j} / \{ n^{-1} \sum_{i=1}^n W_{i,j}^2 - (n^{-1} \sum_{i=1}^n W_{i,j})^2 \}^{1/2}$。根据 Liu 和 Shao (2014) 的公式 (13)，对于任意序列 $d_n \to \infty$ 和 $d_n = o(p)$，当 $n \to \infty$ 时，在满足条件 C.4 的情况下，

$$\sup_{0 \leqslant t \leqslant \mathcal{G}_\kappa^{-1}(d_n/p)} \left| \frac{\sum_{j \in \mathcal{H}_0} \mathbb{I}\{|\bar{T}_j| \geqslant t\}}{p_0 \mathcal{G}_\kappa(t)} - 1 \right| = o_p(1)\,,$$

其中，$\mathcal{G}_\kappa(t)$ 是某个满足以下条件的函数：对所有 $t \in \mathbb{R}$，均有 $\mathcal{G}_\kappa(t) \geqslant \mathcal{G}(t) = 2\{1 - \Phi(t)\}$；并且 $\mathcal{G}_\kappa(t) = \mathcal{G}(t)\{1 + o(1)\}$ 关于 $0 \leqslant t \leqslant 2^{1/2} \log^{1/2}(p)$ 一致成立。因此，当 n 足够大，且 $|\mathcal{H}| \to \infty$ 和 $|\mathcal{H}| > 2/\alpha$，成立时，我们有

$$\alpha |\mathcal{H}|/p \geqslant 2/p = 2 \exp\{-(2^{1/2} \log^{1/2} p)^2/2\} \geqslant 2\{1 - \Phi(2^{1/2} \log^{1/2} p)\} = \mathcal{G}(2^{1/2} \log^{1/2} p)\,.$$

由此可得 $\mathcal{G}^{-1}(\alpha |\mathcal{H}|/p) \leqslant 2^{1/2} \log^{1/2} p,$，进而得到，

$$\sup_{0 \leqslant t \leqslant \mathcal{G}^{-1}(\alpha |\mathcal{H}|/p)} \left| \frac{\sum_{j \in \mathcal{H}_0} \mathbb{I}\{|\bar{T}_j| \geqslant t\}}{p_0 \mathcal{G}(t)} - 1 \right| = o_p(1)\,.$$

令 $T_j' = n^{1/2}(\hat{\theta}_{n,\ j} - \theta_j)/s_{n,j},$经过一些仔细计算，我们得到 $\max_{1 \leqslant j \leqslant p} |T_j' - \bar{T}_j| = o_p\{\log^{-1/2}(p)\}$。按照 Belloni 等人（2018年）第 84 页中类似的程序，可得

$$\sup_{0 \leqslant t \leqslant \mathcal{G}^{-1}(\alpha |\mathcal{H}|/p)} \left| \frac{\sum_{j \in \mathcal{H}_0} \mathbb{I}\{|T_j'| \geqslant t\}}{p_0 \mathcal{G}(t)} - 1 \right| = o_p(1)\,. \tag{S.17}$$

采用 P_1，……，P_p 的 B-H 方法等价于以下程序：当且仅当 $P_j \leqslant \hat{t}_0$ 时，拒绝 H_{0j}，其中

$$\hat{t}_0 = \sup\left\{ 0 \leqslant t \leqslant 1 : t \leqslant \frac{\alpha \max\{\sum_{j=1}^p \mathbb{I}(P_j \leqslant t)\}}{p} \right\}.$$

于是我们有 $\hat{t}_0 = \frac{\alpha \text{最大值}\{\sum_{j=1}^p \mathbb{I}(P_j \leqslant \hat{\quad} t_0)\ 1\}}{p},$，以及 $\alpha |\mathcal{H}|/p \geqslant \mathcal{G}(2^{1/2} \log^{1/2} p)$。设 $t = \mathcal{G}^{-1}(\alpha |\mathcal{H}|/p),$

于是 $t \leqslant 2^{1/2}\log^{1/2} p$ 以概率趋于 1。因此，我们有

$$\begin{aligned} \mathcal{G}(t) = \frac{\alpha|\mathcal{H}|}{p} &\leqslant \frac{\alpha \max\{\sum_{j=1}^{p} \mathbb{I}(|T_i| \geqslant 2^{1/2}\log^{1/2} p), 1\}}{p} \\ &\leqslant \frac{\alpha \max\{\sum_{j=1}^{p} \mathbb{I}(|T_i| \geqslant t), 1\}}{p}, \end{aligned}$$

其中，第二个不等式由 Belloni 等人（2018年）的（B.29）式推出。这意味着 $\mathbb{P}(\hat{t}_0 \geqslant \alpha|\mathcal{H}|/p) \to 1$ with $\hat{t}_0 = \mathcal{G}(\hat{t})$，结合（S.17），可得

$$\frac{\sum_{j\in\mathcal{H}_0} \mathbb{I}\{|T_j'| \geqslant \hat{t}\}}{p_0\mathcal{G}(\hat{t})} = \frac{\sum_{j\in\mathcal{H}_0} \mathbb{I}(|T_j| \geqslant \hat{t})}{p_0\mathcal{G}(\hat{t})} \to 1,$$

这等价于

$$\frac{\sum_{j\in\mathcal{H}_0} \mathbb{I}(P_j \leqslant \hat{t}_0)}{p_0\hat{t}_0} \to 1.$$

最后，

$$\mathrm{FDR}_M = \frac{\sum_{j\in\mathcal{H}_0} \mathbb{I}(P_j \leqslant \hat{t}_0)}{\max\{\sum_{j=1}^{p} \mathbb{I}(P_j \leqslant \hat{t}_0), 1\}} = \frac{\sum_{j\in\mathcal{H}_0} \mathbb{I}(P_j \leqslant \hat{t}_0)}{p\hat{t}_0/\alpha} \to \frac{\alpha p_0}{p}$$

由于 $n \to \infty$，这就完成了该定理的证明。 □

附录 B：预备引理的证明

本节给出附录 A 的 A1 节中预备引理的证明。

引理 *A1* 的证明。由于 U_1 的分量相互独立且已标准化，经过简单计算可得 $\mathbb{E}(\|U_1\|^2) = p$，并且

$$\mathbb{E}(\|\Gamma U_1\|^2) = \mathbb{E}(U_1^\top \Gamma^\top \Gamma U_1) = \mathrm{tr}\{\Gamma^\top \Gamma \mathbb{E}(U_1 U_1^\top)\} = \mathrm{tr}(\Omega).$$

在条件 C.1 下，$U_1 = (U_{1,1}, \ldots, U_{1,p})^\top$ 的分量是相互独立的次指数随机变量，且 $\max_{1\leqslant j\leqslant p} \|U_{1,j}\|_{\psi_\alpha} \leqslant c_0$。应用 Wang、Peng 和 Li（2015）中引理 S2.1 证明所用的集中不等式，对每个 $t \geqslant 0$,

$$\mathbb{P}\left(\left|\|U_1\|^2 - p\right| \geqslant t\right) \leqslant C_1 \exp\left\{-C_2\left(p^{-1}t^2\right)^{\alpha/(4\alpha+4)}\right\}. \tag{S.18}$$

and

$$\mathbb{P}\left\{\left|\|\Gamma U_1\|^2 - \mathrm{tr}(\Omega)\right| \geqslant t\right\} \leqslant C_1 \exp\left[-C_2\left\{\frac{t^2}{\mathrm{tr}(\Omega^2)}\right\}^{\alpha/(4\alpha+4)}\right]. \tag{S.19}$$

对任意固定的 $0 < \epsilon < 1$，令

$$\mathcal{A}_1 = \{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_1\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\}$$

and

$$\mathcal{A}_2 = \{(1-\epsilon)\mathrm{tr}(\Omega) \leqslant \|\Gamma U_1\|^2 \leqslant (1+\epsilon)\mathrm{tr}(\Omega)\}.$$

将 (S.18) 中的 $t = \epsilon p^{(1+\delta)/2}$ 和 (S.19) 中的 $t = \epsilon \mathrm{tr}(\Omega)$ 代入，可得

$$\mathbb{P}(\mathcal{A}_1) \geqslant 1 - C_1 \exp\left\{-C_2(\epsilon^2 p^\delta)^{\alpha/(4\alpha+4)}\right\}$$

and

$$\mathbb{P}(\mathcal{A}_2) \geqslant 1 - C_1 \exp\left[-C_2\left\{\frac{\epsilon^2 \mathrm{tr}^2(\Omega)}{\mathrm{tr}(\Omega^2)}\right\}^{\alpha/(4\alpha+4)}\right].$$

根据条件 C.3,

$$\mathrm{tr}(\Omega^2) = \sum_{j=1}^{p}\sum_{\ell=1}^{p}\omega_{j\ell}^2 \leqslant \bar{M}p \max_{1\leqslant\ell\leqslant p}\sum_{j=1}^{p}|\omega_{j\ell}| \leqslant \bar{M}p a_0(p)\,.$$

由于 $\mathrm{tr}(\Omega) = p$ 和 $a_0(p) \asymp p^{1-\delta}$，因此我们得出

$$\frac{\mathrm{tr}^2(\Omega)}{\mathrm{tr}(\Omega^2)} \geqslant \frac{p^2}{\bar{M}p a_0(p)} \asymp p^\delta\,.$$

因此，存在某些正数常数 c_1 和 c_2，使得

$$\mathbb{P}(\mathcal{A}_1) \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}\,.$$

and

$$\mathbb{P}(\mathcal{A}_2) \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

当 p 足够大时，结论成立。因此，我们完成了这个引理的证明。 □

引理 *A2* 的证明。 (i) 由于 $U_i = (U_{i,\ 1},\ \ldots\ldots,\ U_{i,\ p})^\top$ 的分量是独立同分布的标准化次指数随机变量，经简单代数运算可得

$$\begin{aligned}
\mathbb{E}(\|U_i\|^4) &= \mathbb{E}\left\{\left(\sum_{j=1}^{p} U_{i,j}^2\right)^2\right\} \\
&= \sum_{j=1}^{p} \mathbb{E}(U_{i,j}^4) + \sum_{1\leqslant j_1\neq j_2\leqslant p} \mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2) \\
&= p\mathbb{E}(U_{i,j}^4) + p(p-1)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(\|U_i\|^6) &= \sum_{j=1}^{p} \mathbb{E}(U_{i,j}^6) + 3\sum_{1\leqslant j_1\neq j_2\leqslant p} \mathbb{E}(U_{i,j_1}^4)\mathbb{E}(U_{i,j_2}^2) \\
&\quad + \sum_{1\leqslant j_1\neq j_2\neq j_3\leqslant p} \mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2) \\
&= p\mathbb{E}(U_{i,j}^6) + 3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)\,.
\end{aligned}$$

此外,

$$\begin{aligned}
\mathbb{E}(\|U_i\|^8) &= \sum_{j=1}^{p} \mathbb{E}(U_{i,j}^8) + 4\sum_{1\leqslant j_1\neq j_2\leqslant p} \mathbb{E}(U_{i,j_1}^6)\mathbb{E}(U_{i,j_2}^2) \\
&\quad +3\sum_{1\leqslant j_1\neq j_2\leqslant p} \mathbb{E}(U_{i,j_1}^4)\mathbb{E}(U_{i,j_2}^4) \\
&\quad +6\sum_{1\leqslant j_1\neq j_2\neq j_3\leqslant p} \mathbb{E}(U_{i,j_1}^4)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2) \\
&\quad + \sum_{1\leqslant j_1\neq j_2\neq j_3\neq j_4\leqslant p} \mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2)\mathbb{E}(U_{i,j_4}^2) \\
&= p\mathbb{E}(U_{i,j}^8) + 4p(p-1)\mathbb{E}(U_{i,j_1}^6) + 3p(p-1)\{\mathbb{E}(U_{i,j_1}^4)\}^2 \\
&\quad +3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)(p-3)\,.
\end{aligned}$$

The 任意正整数 k 的 $\mathbb{E}(\|U_i\|^{2k}) = p^k + O(p^{k-1})$ 结果可通过以下方式验证

$$\begin{aligned}\mathbb{E}(\|U_i\|^{2k}) &= \sum_{1\leqslant j_1\neq\cdots\neq j_k\leqslant p} \mathbb{E}(U_{i,j_1}^2)\times\cdots\times\mathbb{E}(U_{i,j_k}^2)\{1+O(p^{-1})\} \\ &= p^k + O(p^{k-1})\,.\end{aligned}$$

此外，由于对于所有 $u\geqslant 0$, $\{1+u-(u-1)^2\}/2\leqslant u^{1/2}\leqslant(1+u)/2$，我们可以得到，对于所有正整数 k, $\mathbb{E}(\|U\|^k) = p^{k/2}+O(p^{k/2-1})$。

(ii) 将 $\Lambda_{j\ell} = \sum_{j_1=1}^p \Gamma_{j_1 j}\Gamma_{j_1\ell}$ 记为 $\Gamma^\top\Gamma$中的第 ℓ) 个元素，（j,，则

$$\begin{aligned}\mathbb{E}(\|\Gamma U_i\|^4) &= \mathbb{E}\left\{\left(\sum_{j=1}^p\sum_{\ell=1}^p \Lambda_{j\ell}U_{i,j}U_{i,\ell}\right)^2\right\} \\ &= \sum_{j=1}^p \Lambda_{jj}^2\mathbb{E}(U_{i,j}^4) + 2\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2) \\ &\quad + \sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_1}\Lambda_{j_2j_2}\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2) \\ &= \mathbb{E}(U_{i,j}^4)\sum_{j=1}^p\Lambda_{jj}^2 + 2\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2 + \sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_1}\Lambda_{j_2j_2} \\ &= \left(\sum_{j=1}^p\Lambda_{jj}\right)^2 + \{\mathbb{E}(U_{i,j}^4)-1\}\sum_{j=1}^p\Lambda_{jj}^2 + 2\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2 \\ &= \{\mathrm{tr}(\Omega)\}^2 + O\{\mathrm{tr}(\Omega^2)\}\end{aligned}$$

根据条件 C.3, $\sum_{j=1}^p\Lambda_{jj}^2+\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2 = \sum_{j=1}^p\sum_{\ell=1}^p\Lambda_{j\ell}^2 = \mathrm{tr}(\Omega^2)$ 和 $\mathrm{tr}(\Omega^2)\lesssim p^{2-\delta}$。类似地，我们可以证明

$$\begin{aligned}\mathbb{E}(\|\Gamma U_i\|^6) &= \sum_{1\leqslant j_1\neq j_2\neq j_3\leqslant p}(\Lambda_{j_1j_1}\Lambda_{j_2j_2}\Lambda_{j_3j_3} + \Lambda_{j_1j_2}^2\Lambda_{j_3j_3} \\ &\qquad + \Lambda_{j_1j_2}\Lambda_{j_1j_3}\Lambda_{j_2j_3})\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2)\{1+O(p^{-1})\} \\ &= p^3 + O(p^{3-\delta})\end{aligned}$$

以及 $\mathbb{E}(\|\Gamma U_i\|^{12}) = p^6 + O(p^{6-\delta})$。

与第 (i) 部分的证明类似，结果 $\mathbb{E}(\|\Gamma U_i\|) = p^{1/2}+O(p^{1/2-\delta})$和 $\mathbb{E}(\|\Gamma U_i\|^3) = p^{3/2}+O(p^{3/2-\delta})$ 直接由 $\mathbb{E}(\|\Gamma U_i\|^2) = p$, $\mathbb{E}(\|\Gamma U_i\|^4) = p^2+O(p^{2-\delta})$, $\mathbb{E}(\|\Gamma U_i\|^6)$ =得出

$p^3 + O(p^{3-\delta})$, $\mathbb{E}(\|\Gamma U_i\|^{12}) = p^6 + O(p^{6-\delta})$ 和 $\{1 + u - (u-1)^2\}/2 \leqslant u^{1/2} \leqslant (1+u)/2$ 对所有 $u \geqslant 0$均成立。

(iii) 现在我们考虑 $\mathbb{E}\{\|\Gamma S(U_i)\|^2\}$。对于 $i = 1, \ldots, n$, 令

$$\mathcal{A}_{1i} = \{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_i\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\}$$

对于固定的 $0 < \epsilon < 1$。根据引理 A1 以及 $\|\Gamma U_i\|^2 \leqslant \mathrm{tr}(\Omega)\|U_i\|^2$ 这一事实,

$$\begin{aligned}
\mathbb{E}\{\|\Gamma S(U_i)\|^2\} &= \mathbb{E}(\|\Gamma U_i\|^2 \|U_i\|^{-2}) \\
&= p^{-1}\mathbb{E}\{\|\Gamma U_i\|^2\} + \mathbb{E}\left\{\|\Gamma U_i\|^2 \left(\|U_i\|^{-2} - p^{-1}\right)\right\} \\
&= 1 + \mathbb{E}\left\{\|\Gamma U_i\|^2 \left(\|U_i\|^{-2} - p^{-1}\right)\right\},
\end{aligned}$$

其中

$$\begin{aligned}
& \mathbb{E}\left\{\|\Gamma U_i\|^2 \left(\|U_i\|^{-2} - p^{-1}\right)\right\} \\
\leqslant\ & p^{-1}\mathbb{E}\left(\|\Gamma U_i\|^2 \|U_i\|^{-2} \left|\|U_i\|^2 - p\right|\right) \\
=\ & p^{-1}\mathbb{E}\left\{\|\Gamma U_i\|^2 \|U_i\|^{-2} \left|\|U_i\|^2 - p\right| \mathbb{I}(\mathcal{A}_{1i})\right\} \\
& + p^{-1}\mathbb{E}\left\{\|\Gamma U_i\|^2 \|U_i\|^{-2} \left|\|U_i\|^2 - p\right| \mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant\ & p^{-1}\{p - \epsilon p^{(1+\delta)/2}\}^{-1}\mathbb{E}\left(\|\Gamma U_i\|^2 \left|\|U_i\|^2 - p\right|\right) \\
& + p^{-1}\mathrm{tr}(\Omega)\mathbb{E}\left\{\left|\|U_i\|^2 - p\right| \mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant\ & p^{-1}\{p - \epsilon p^{(1+\delta)/2}\}^{-1}\left\{\mathbb{E}(\|\Gamma U_i\|^4)\right\}^{1/2}\left\{\mathbb{E}(\left|\|U_i\|^2 - p\right|^2)\right\}^{1/2} \\
& + \left\{\mathbb{E}(\left|\|U_i\|^2 - p\right|^2)\right\}^{1/2}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\
\leqslant\ & p^{-1}(p - \epsilon p^{1-\delta})^{-1}\{p^2 + O(p^{2-\delta})\}^{1/2} \times O(p^{1/2}) \\
& + O(p^{1/2}) \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
=\ & O(p^{-1/2}).
\end{aligned}$$

由此可得 $\mathbb{E}\{\|\Gamma S(U_i)\|^2\} = 1 + O(p^{-1/2})$。

同理，最后一个结果可由此得出

$$\begin{aligned}\mathbb{E}\{\|\Gamma S(U_i)\|^4\} &= p^{-2}\mathbb{E}\{\|\Gamma U_i\|^4\} + \mathbb{E}\left\{\|\Gamma U_i\|^4\left(\|U_i\|^{-4} - p^{-2}\right)\right\} \\ &= 1 + O(p^{-\delta}) + \mathbb{E}\left\{\|\Gamma U_i\|^4\left(\|U_i\|^{-4} - p^{-2}\right)\right\},\end{aligned}$$

其中

$$\begin{aligned}&\mathbb{E}\left\{\|\Gamma U_i\|^4\left(\|U_i\|^{-4} - p^{-2}\right)\right\} \\ \leqslant\ & p^{-2}\mathbb{E}\left(\|\Gamma U_i\|^4\|U_i\|^{-4}\left|\|U_i\|^4 - p^2\right|\right) \\ =\ & p^{-2}\mathbb{E}\left\{\|\Gamma U_i\|^4\|U_i\|^{-4}\left|\|U_i\|^4 - p^2\right|\mathbb{I}(\mathcal{A}_{1i})\right\} \\ & + p^{-2}\mathbb{E}\left\{\|\Gamma U_i\|^4\|U_i\|^{-4}\left|\|U_i\|^4 - p^2\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\ \leqslant\ & p^{-2}(p - \epsilon p^{1-\delta})^{-2}\mathbb{E}\left(\|\Gamma U_i\|^4\left|\|U_i\|^4 - p^2\right|\right) \\ & + p^{-2}\{\mathrm{tr}(\Omega)\}^2\mathbb{E}\left\{\left|\|U_i\|^4 - p^2\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\ \leqslant\ & p^{-2}(p - \epsilon p^{1-\delta})^{-2}\left\{\mathbb{E}(\|\Gamma U_i\|^6)\right\}^{2/3}\left\{\mathbb{E}(\left|\|U_i\|^4 - p^2\right|^3)\right\}^{1/3} \\ & + \left\{\mathbb{E}(\left|\|U_i\|^4 - p^2\right|^2)\right\}^{1/2}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\ \leqslant\ & p^{-2}(p - \epsilon p^{1-\delta})^{-2} \times O(p^2) \times O(p^{3/2}) \\ & + O(p^{3/2}) \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\ =\ & O(p^{-1/3}).\end{aligned}$$

(iv) 由于 ν_i 和 $S(U_i)$ 相互独立，

$$\begin{aligned}&\mathbb{E}(\nu_i^{-1})\mathbb{E}\{\|\Gamma S(U_i)\|^{-1}\} \\ =\ & \mathbb{E}(\nu_i^{-1}\|\Gamma U_i\|^{-1}\|U_i\|) \\ =\ & \mathbb{E}(R_i^{-1}\|U_i\|) \\ =\ & \mathbb{E}\{R_i^{-1}\|U_i\|\mathbb{I}(\mathcal{A}_{1i})\} + \mathbb{E}\{R_i^{-1}\|U_i\|\mathbb{I}(\mathcal{A}_{1i}^c)\} \\ \leqslant\ & \{p + \epsilon p^{(1+\delta)/2}\}^{1/2}\mathbb{E}\{R_i^{-1}\mathbb{I}(\mathcal{A}_{1i})\} + \{\mathbb{E}(R_i^{-4})\}^{1/4}\{\mathbb{E}\|U_i\|^4\}^{1/4}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\ \lesssim\ & \{p + \epsilon p^{(1+\delta)/2}\}^{1/2}\mathbb{E}(R_i^{-1}) + \zeta_4^{1/4} \times p^{1/2} \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\ \lesssim\ & \zeta_1 p^{1/2},\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E}(\nu_i^{-2})\mathbb{E}\{\|\Gamma S(U_i)\|^{-2}\} \\
= \;& \mathbb{E}\{R_i^{-2}\|U_i\|^2\mathbb{I}(\mathcal{A}_{1i})\} + \mathbb{E}\{R_i^{-2}\|U_i\|^2\mathbb{I}(\mathcal{A}_{1i}^c)\} \\
\leqslant \;& \{p+\epsilon p^{(1+\delta)/2}\}\mathbb{E}\{R_i^{-2}\mathbb{I}(\mathcal{A}_{1i})\} + \{\mathbb{E}(R_i^{-4})\}^{1/2}\{\mathbb{E}\|U_i\|^6\}^{1/3}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/6} \\
\lesssim \;& \{p+\epsilon p^{(1+\delta)/2}\}\mathbb{E}(R_i^{-2}) + \zeta_4^{1/2}\times p\times c_1^{1/6}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/6\} \\
\lesssim \;& \zeta_2 p\,,
\end{aligned}$$

此外，我们还有

$$\begin{aligned}
& \mathbb{E}(\nu_i^{-3})\mathbb{E}\{\|\Gamma S(U_i)\|^{-3}\} \\
= \;& \mathbb{E}\{R_i^{-3}\|U_i\|^3\mathbb{I}(\mathcal{A}_{1i})\} + \mathbb{E}\{R_i^{-3}\|U_i\|^3\mathbb{I}(\mathcal{A}_{1i}^c)\} \\
\leqslant \;& \{p+\epsilon p^{(1+\delta)/2}\}^{3/2}\mathbb{E}\{R_i^{-3}\mathbb{I}(\mathcal{A}_{1i})\} + \{\mathbb{E}(R_i^{-4})\}^{3/4}\{\mathbb{E}\|U_i\|^{18}\}^{1/6}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/12} \\
\lesssim \;& \{p+\epsilon p^{(1+\delta)/2}\}^{3/2}\mathbb{E}(R_i^{-3}) + \zeta_4^{3/4}\times p^{3/2}\times c_1^{1/12}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/12\} \\
\lesssim \;& \zeta_3 p^{3/2}\,.
\end{aligned}$$

由柯西-施瓦茨不等式和詹森不等式可得

$$[\mathbb{E}\{\|\Gamma S(U_i)\|^{-1}\}]^{-1} \leqslant \mathbb{E}\{\|\Gamma S(U_i)\|\} \leqslant [\mathbb{E}\{\|\Gamma S(U_i)\|^2\}]^{1/2} = 1+O(p^{-1/2})\,,$$
$$[\mathbb{E}\{\|\Gamma S(U_i)\|^{-2}\}]^{-1} \leqslant \mathbb{E}\{\|\Gamma S(U_i)\|^2\} = 1+O(p^{-1/2})\,,$$

and

$$[\mathbb{E}\{\|\Gamma S(U_i)\|^{-3}\}]^{-1} \leqslant \mathbb{E}\{\|\Gamma S(U_i)\|^3\} \leqslant [\mathbb{E}\{\|\Gamma S(U_i)\|^4\}]^{3/4} = 1+O(p^{-1/3})\,.$$

于是，本部分的结果立即成立。至此，我们完成了该引理的证明。 □

引理 *A3* 的证明。 (i) 对于 $i=1,\ldots,n$，令 $\mathcal{A}_{2i}=\{(1-\epsilon)\mathrm{tr}(\Omega)\leqslant\|\Gamma U_i\|^2\leqslant(1+\epsilon)\mathrm{tr}(\Omega)\}$，其中 $0<\epsilon<1$ 固定。回顾，Γ_j 是 Γ 的第 j 行，且 $W_{i,j}=\Gamma_j U_i/\|\Gamma U_i\|$，则

$$Q_{j\ell} \;=\; n^{-1}\sum_{i=1}^n R_i^{-1}W_{i,j}W_{i,\ell} = n^{-1}\sum_{i=1}^n \nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i)\|\Gamma U_i\|^{-3}$$

$$
\begin{aligned}
&= n^{-1}p^{-3/2}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i) \\
&\quad + n^{-1}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i)\left(\|\Gamma U_i\|^{-3} - p^{-2/3}\right),
\end{aligned}
$$

其中最后一项满足

$$
\begin{aligned}
&\left|\mathbb{E}\left\{n^{-1}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i)\left(\|\Gamma U_i\|^{-3} - p^{-2/3}\right)\right\}\right| \\
\leqslant\ & p^{-3/2}\mathbb{E}\left\{\nu_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-3}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\right\} \\
=\ & p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-2}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\right\} \\
=\ & p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-2}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i})\right\} \\
&\quad + p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-2}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i}^c)\right\} \\
\leqslant\ & (1-\epsilon)^{-1}p^{-5/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i})\right\} \\
&\quad + p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i}^c)\right\} \\
\lesssim\ & p^{-5/2}\{\mathbb{E}(R_i^{-4})\}^{1/4}\left[\mathbb{E}\left\{\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|^4\right\}\right]^{1/4}\left\{\mathbb{E}\left(\left|\|\Gamma U_i\|^3 - p^{3/2}\right|^2\right)\right\}^{1/2} \\
&\quad + p^{-3/2}\{\mathbb{E}(R_i^{-4})\}^{1/4}\left\{\mathbb{E}\left(\left|\|\Gamma U_i\|^3 - p^{3/2}\right|^2\right)\right\}^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/4} \\
\lesssim\ & \zeta_1 p^{-1-\delta/2}.
\end{aligned}
$$

由此可得

$$
Q_{j\ell} = n^{-1}p^{-3/2}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i) + O_p(\zeta_1 p^{-1-\delta/2}).
$$

对于 $i = 1, \ldots, n$, 令 $\mathcal{A}_{1i} = \{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_i\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\}$, 其中 $0 < \epsilon < 1$ 固定。根据引理 A1,

$$
\begin{aligned}
&\mathbb{E}\left[\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}^2\right] \\
=\ & \mathbb{E}\left\{\|U_i\|^{-4}(\Gamma_j U_i U_i^\top\Gamma_\ell^\top)^2\right\} \\
=\ & \mathbb{E}\left\{\|U_i\|^{-4}(\Gamma_j U_i U_i^\top\Gamma_\ell^\top)^2\mathbb{I}(\mathcal{A}_{1i})\right\} + \mathbb{E}\left\{\|U_i\|^{-4}(\Gamma_j U_i U_i^\top\Gamma_\ell^\top)^2\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\lesssim\ & \{p - \epsilon p^{(1+\delta)/2}\}^{-2}\mathbb{E}\left\{(\Gamma_j U_i U_i^\top\Gamma_\ell^\top)^2\right\} + p^2\mathbb{P}(\mathcal{A}_{1i}^c)
\end{aligned}
$$

$$\begin{aligned} &\lesssim \{p-\epsilon p^{(1+\delta)/2}\}^{-2} + p^2 \times c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\} \\ &\lesssim p^{-2}\,. \end{aligned}$$

接下来，我们可以证明

$$\begin{aligned} & n^{-1}p^{-3/2}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i) \\ = \;& n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top + O_p(\zeta_1 p^{-7/6})\,, \end{aligned}$$

其中最后一个相等性由

$$\begin{aligned} & \mathbb{E}|p^{-3/2}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top(\|U_i\|^2-p)| \\ \lesssim \;& p^{-3/2}\{\mathbb{E}(\nu_i^{-3})\}^{1/3}\left(\mathbb{E}\left[\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}^2\right]\right)^{1/2}\left[\mathbb{E}\left\{(\|U_i\|^2-p)^6\right\}\right]^{1/6} \\ \lesssim \;& \zeta_1 p^{-7/6}\,. \end{aligned}$$

因此，我们得到

$$Q_{j\ell} = n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top + O_p(\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2})\,.$$

由于 ν_i 和 $S(U_i)$ 相互独立，因此有

$$\begin{aligned} & \mathbb{E}\left\{n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\} \\ = \;& p^{-1/2}\mathbb{E}(\nu_i^{-1})\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\,, \end{aligned}$$

其中 $\mathbb{E}(\nu_i^{-1}) \lesssim p^{1/2}\zeta_1$ 由引理 A2 得出。

根据引理 A1，并注意到 $\Gamma_j\Gamma_\ell^\top = \omega_{j\ell}$,

$$\begin{aligned} & \mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\} \\ = \;& \mathbb{E}\left(\Gamma_j U_i U_i^\top\Gamma_\ell^\top\|U_i\|^{-2}\right) \\ = \;& p^{-1}\mathbb{E}\left(\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right) + \mathbb{E}\left\{\Gamma_j U_i U_i^\top\Gamma_\ell^\top\left(\|U_i\|^{-2}-p^{-1}\right)\right\} \\ = \;& p^{-1}\omega_{j\ell} + \mathbb{E}\left\{\Gamma_j U_i U_i^\top\Gamma_\ell^\top\left(\|U_i\|^{-2}-p^{-1}\right)\right\} \end{aligned}$$

$$
\begin{aligned}
&\leqslant p^{-1}|\omega_{j\ell}| + \mathbb{E}\left(\left|\Gamma_j U_i U_i^\top \Gamma_\ell^\top\right| \left|\|U_i\|^{-2} - p^{-1}\right|\right) \\
&= p^{-1}|\omega_{j\ell}| + p^{-1}\mathbb{E}\left(\left|\Gamma_j U_i U_i^\top \Gamma_\ell^\top\right| \|U_i\|^{-2} \left|\|U_i\|^2 - p\right|\right) \\
&= p^{-1}|\omega_{j\ell}| + p^{-1}\mathbb{E}\left\{\left|\Gamma_j U_i U_i^\top \Gamma_\ell^\top\right| \|U_i\|^{-2} \left|\|U_i\|^2 - p\right| \mathbb{I}(\mathcal{A}_{1i})\right\} \\
&\quad + p^{-1}\mathbb{E}\left\{\left|\Gamma_j U_i U_i^\top \Gamma_\ell^\top\right| \|U_i\|^{-2} \left|\|U_i\|^2 - p\right| \mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
&\lesssim p^{-1}|\omega_{j\ell}| + \{p^2 - \epsilon p^{(3+\delta)/2}\}^{-1}\mathbb{E}\left(\left|\Gamma_j U_i U_i^\top \Gamma_\ell^\top\right| \left|\|U_i\|^2 - p\right|\right) \\
&\quad + \mathbb{E}\left\{\left|\|U_i\|^2 - p\right| \mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
&\leqslant p^{-1}|\omega_{j\ell}| + \{p^2 - \epsilon p^{(3+\delta)/2}\}^{-1}\left[\mathbb{E}\left\{\left(\Gamma_j U_i U_i^\top \Gamma_\ell^\top\right)^2\right\}\right]^{1/2}\left[\mathbb{E}\left\{\left(\|U_i\|^2 - p\right)^2\right\}\right]^{1/2} \\
&\quad + \left[\mathbb{E}\left\{\left(\|U_i\|^2 - p\right)^2\right\}\right]^{1/2}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\
&\leqslant p^{-1}|\omega_{j\ell}| + O(p^{-3/2}) + O(p^{1/2}) \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
&\lesssim p^{-1}|\omega_{j\ell}| + O(p^{-3/2}),
\end{aligned}
$$

其中倒数第二个不等式是由.....得出的

$$
\begin{aligned}
\mathbb{E}\left\{\left(\|U_i\|^2 - p\right)^2\right\} &= \mathbb{E}(\|U_i\|^4 - 2p\|U_i\|^2 + p^2) \\
&= p\mathbb{E}(U_{i,j}^4) + p(p-1) - 2p^2 + p^2 \\
&= O(p).
\end{aligned}
$$

因此，可得

$$
\mathbb{E}\left\{n^{-1}p^{-1/2}\sum_{i=1}^n \nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top\right\} \lesssim \zeta_1 p^{-1}|\omega_{jj}| + O(\zeta_1 p^{-3/2}).
$$

此外，当 $\mathbb{E}(\nu_i^{-2}) \lesssim p\zeta_2$ 时，我们可以得出结论:

$$
\begin{aligned}
&\mathrm{Var}\left\{n^{-1}p^{-1/2}\sum_{i=1}^n \nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top\right\} \\
= \; & n^{-1}p^{-1}\mathbb{E}(\nu_i^{-2})\mathbb{E}\left[\left\{\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top\right\}^2\right] \\
& - n^{-1}p^{-1}\{\mathbb{E}(\nu_i^{-1})\}^2\left[\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top\right\}\right]^2 \\
\lesssim \; & \zeta_1^2 n^{-1}p^{-2}.
\end{aligned}
$$

由切比雪夫不等式可得

$$\left| n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top \right| \lesssim \zeta_1 p^{-1}|\omega_{j\ell}| + O_p(\zeta_1 n^{-1/2}p^{-1} + \zeta_1 p^{-3/2})\,.$$

最后，我们得到 $|Q_{j,\ \ell}| \lesssim \zeta_1 p^{-1}|\omega_{j\ell}| + O_p(\zeta_1 n^{-1/2}p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$。

(ii) 根据第 **(i)** 部分的证明，我们知道 $Q_{j\ell} = Q_{0,\ j\ell} + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$，其中$Q_{0,\ j\ell}$ 是随机矩阵 $Q_0 = n^{-1}p^{-1/2}\sum_{i=1}^n \nu_i^{-1}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top$ 的第 $(j,\ \ell)$ 个分量。此外，
$\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\} \lesssim p^{-1}|\omega_{j\ell}| + O(p^{-3/2})$。由此可得

$$\begin{aligned}
& \mathrm{tr}\left\{\left(\mathbb{E}\left[\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \\
= & \sum_{j=1}^{p}\sum_{\ell=1}^{p}\left[\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\right]^2 \\
\lesssim & \ p^{-2}\sum_{j=1}^{p}\sum_{\ell=1}^{p}|\omega_{j\ell}|^2 + p^{-5/2}\sum_{j=1}^{p}\sum_{\ell=1}^{p}|\omega_{j\ell}| + p^{-1} \\
\lesssim & \ p^{-1}a_0(p) + p^{-3/2}a_0(p) + p^{-1} \\
\lesssim & \ p^{-\delta}\,.
\end{aligned}$$

这意味着

$$\begin{aligned}
\mathrm{tr}[\{\mathbb{E}(Q_0)\}^2] &= p^{-1}\{\mathbb{E}(\nu_i^{-1})\}^2\mathrm{tr}\left\{\left(\mathbb{E}\left[\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \\
&\lesssim p^{-1-\delta}
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E}\{\mathrm{tr}(Q_0^2)\} \\
= & \ n^{-1}p^{-1}\mathrm{tr}\left(\mathbb{E}\left[\nu_i^{-2}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right) \\
& +(1-n^{-1})p^{-1}\mathrm{tr}\left\{\left(\mathbb{E}\left[\nu_i^{-1}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \\
= & \ n^{-1}p^{-1}\mathbb{E}(\nu_i^{-2})\mathbb{E}\left\{\|\Gamma S(U_i)\|^4\right\} \\
& +(1-n^{-1})p^{-1}\{\mathbb{E}(\nu_i^{-1})\}^2\mathrm{tr}\left\{\left(\mathbb{E}\left[\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \\
= & \ O(n^{-1}p^{-1}) + \mathrm{tr}[\{\mathbb{E}(Q_0)\}^2](1-n^{-1})\,.
\end{aligned}$$

因此，我们有

$$\mathrm{tr}[\mathbb{E}(Q_0^2) - \{\mathbb{E}(Q_0)\}^2] = O(n^{-1}p^{-1})\,.$$

我们完成该引理的证明。 □

引理 *A4* 的证明。 回顾， Γ_j 是 Γ的第 j行，并记 $\Gamma_{j\ell}$ 为 Γ的第 $(j,\ \ell)$个元素，则

$$\Gamma_j U_i = \sum_{\ell=1}^{p} \Gamma_{j\ell} U_{i,\ell}.$$

注意到$\omega_{j\ell} = \sum_{j_1=1}^{p} \Gamma_{jj_1}\Gamma_{\ell j_1}$，则

$$\mathrm{Var}(\Gamma_j U_i) = \sum_{\ell=1}^{p} \Gamma_{j\ell}^2 = \omega_{jj}$$

and

$$\begin{aligned}
& \mathbb{E}\{(\Gamma_j U_i)^4\} \\
= \ & \mathbb{E}\left\{\left(\sum_{\ell=1}^{p} \Gamma_{j\ell} U_{i,\ell}\right)^4\right\} \\
= \ & \sum_{\ell=1}^{p} \Gamma_{j\ell}^4 \mathbb{E}(U_{i,\ell}^4) + 6 \sum_{1\leqslant \ell_1 \neq \ell_2 \leqslant p} \Gamma_{j\ell_1}^2 \Gamma_{j\ell_2}^2 \mathbb{E}(U_{i,\ell_1}^2)\mathbb{E}(U_{i,\ell_2}^2) \\
\lesssim \ & \omega_{jj}^2\,.
\end{aligned}$$

(i) 对于 $i = 1,\ .....,\ n$，令 $\mathcal{A}_{2i} = \{(1-\epsilon)\mathbf{tr}(\Omega) \leqslant \|\Gamma U_i\|^2 \leqslant (1+\epsilon)\mathbf{tr}(\Omega)\}$，其中 $0 < \epsilon < 1$为固定值，则

$$\mathbb{P}(\mathcal{A}_{2i}) \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

根据引理 A1 的证明，可得

$$\begin{aligned}
\mathbb{E}(W_{i,j}^4) &= \mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4\} \\
&= \mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4 \mathbb{I}(\mathcal{A}_{2i})\} + \mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4 \mathbb{I}(\mathcal{A}_{2i}^c)\}
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-2}\mathbb{E}\{(\Gamma_j U_i)^4\} + \mathbb{P}(\mathcal{A}_{2i}^c) \\
&\lesssim \omega_{jj}^2\{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-2} + c_1\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\} \\
&\lesssim \omega_{jj}^2\{\mathrm{tr}(\Omega)\}^{-2}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}(W_{i,j}^2) &\geqslant \mathbb{E}\{\|\Gamma U_i\|^{-2}(\Gamma_j U_i)^2\mathbb{I}(\mathcal{A}_{2i})\} \\
&\geqslant \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\mathbb{I}(\mathcal{A}_{2i})\} \\
&= \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} - \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
&\geqslant \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} - \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}[\mathbb{E}\{(\Gamma_j U_i)^4\}]^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/2} \\
&\gtrsim \omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} - \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} \times \omega_{jj} \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
&\gtrsim \omega_{jj}\{\mathrm{tr}(\Omega)\}^{-1},
\end{aligned}
$$

由此我们得出

$$
\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^4\} \lesssim \zeta_1^{-4}p^{-2}\omega_{jj}^2 \lesssim \bar{M}^2
$$

and

$$
\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \gtrsim \zeta_1^{-2}p^{-1}\omega_{jj} \gtrsim \underline{m}.
$$

（ii）与部分（i）的证明类似，对于任意 $\varrho \geqslant 1$,

$$
\begin{aligned}
\mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\right\} &= \mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\mathbb{I}(\mathcal{A}_{1i})\right\} + \mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
&\lesssim \zeta_1^{-\varrho}\{\mathrm{tr}(\Omega)\}^{-\varrho/2}\mathbb{E}\{|\Gamma_j U_i|^{\varrho}\} + \zeta_1^{-\varrho}\mathbb{P}(\mathcal{A}_{1i}^c) \\
&\lesssim \mathbb{E}\{|\Gamma_j U_i|^{\varrho}\} + p^{\varrho/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}.
\end{aligned}
$$

由于对于某个常数 c_0，$\max_{1\leqslant j\leqslant p}\|U_{i,j}\|_{\psi_\alpha} \leqslant c_0$，根据 Koike (2021) 中的引理 B.4，我们有 $\|\Gamma_j U_i\|_{\psi_\alpha} \lesssim c_0$。于是，我们知道，对于任意 $\varrho \geqslant 1$，根据等价的

次指数性质（Koike 2021）。因此,

$$\mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\right\} \lesssim \varrho^{\varrho/\alpha}$$

对任意 $\varrho \geqslant 1$ ，当 p 足够大时，这表明 $\zeta_1^{-1}W_{i,\ j}$ 服从次指数分布，因此$\|\zeta_1^{-1}W_{i,\ j}\|_{\psi_\alpha} \lesssim \bar{B}$。

（iii）通过简单的代数运算,

$$\begin{aligned}\mathbb{E}(W_{i,j}^2) &= p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} + \mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2} - p^{-1})\} \\ &= p^{-1}\omega_{jj} + \mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2} - p^{-1})\}\,,\end{aligned}$$

其中 $\mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2} - p^{-1})\}$ 满足

$$\begin{aligned}&\left|\mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2} - p^{-1})\}\right| \\ \leqslant\ & p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\} \\ =\ & p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i})\} \\ & + p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\ \leqslant\ & p^{-1}\{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2|\|\Gamma_j U_i\|^2 - p|\} + p^{-1}\mathbb{E}\{\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\ \leqslant\ & p^{-2}(1-\epsilon)^{-1}[\mathbb{E}\{(\Gamma_j U_i)^4\}]^{1/2}\{\mathbb{E}(|\|\Gamma_j U_i\|^2 - p|^2)\}^{1/2} \\ & + p^{-1}\{\mathbb{E}(|\|\Gamma_j U_i\|^2 - p|^2)\}^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/2} \\ \lesssim\ & p^{-2}\times p^{1-\delta/2} + p^{-1}\times p^{1-\delta/2}\times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\ \lesssim\ & p^{-1-\delta/2}.\end{aligned}$$

此外，对于 $1 \leqslant j \neq \ell \leqslant p$，我们有

$$\begin{aligned}\mathbb{E}(W_{i,j}W_{i,\ell}) &= p^{-1}\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)\} + \mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2} - p^{-1})\} \\ &= p^{-1}\omega_{j\ell} + \mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2} - p^{-1})\}\,,\end{aligned}$$

其中 $\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2} - p^{-1})\}$ 满足

$$
\begin{aligned}
& \left|\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2} - p^{-1})\}\right| \\
\leqslant\ & p^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\} \\
=\ & p^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i})\} \\
& + p^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
\leqslant\ & p^{-1}\{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma_j U_i\|^2 - p|\} + p^{-1}\mathbb{E}\{\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
\leqslant\ & p^{-2}(1-\epsilon)^{-1}[\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|^2\}]^{1/2}\{\mathbb{E}(|\|\Gamma_j U_i\|^2 - p|^2)\}^{1/2} \\
& + p^{-1}\{\mathbb{E}(|\|\Gamma_j U_i\|^2 - p|^2)\}^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/2} \\
\lesssim\ & p^{-2} \times p^{1-\delta/2} + p^{-1} \times p^{1-\delta/2} \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
\lesssim\ & p^{-1-\delta/2}.
\end{aligned}
$$

(iv) 根据 (ii) 部分，$\zeta_1^{-1}W_1, \ldots, \zeta_1^{-1}W_n$ 是满足 $\|\zeta_1^{-1}W_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$ 的 p 维独立同分布随机向量，对所有 $i = 1, \ldots, n$ 和 $j = 1, \ldots, p$ 均成立。根据 van der Vaart与Wellner（1996）的引理 2.2.2,

$$
\left\| \max_{1\leqslant i\leqslant n} \max_{1\leqslant j\leqslant p} |\zeta_1^{-1}W_{i,j}| \right\|_{\psi_\alpha} \lesssim \log^{1/\alpha}(np)\,.
$$

与第 (i) 部分的证明类似，我们可以证明

$$
\begin{aligned}
\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} &= \zeta_1^{-2}\mathbb{E}\{\|\Gamma U_i\|^{-2}(\Gamma_j U_i)^2\mathbb{I}(\mathcal{A}_{1i})\} \\
&\quad + \zeta_1^{-2}\mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4\mathbb{I}(\mathcal{A}_{1i}^c)\} \\
&\leqslant \zeta_1^{-2}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} + \zeta_1^{-2}\mathbb{E}\{\mathbb{I}(\mathcal{A}_{1i}^c)\} \\
&\leqslant \zeta_1^{-2}\omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} + \zeta_1^{-2}c_1\exp\{-c_2 p^{\delta/(4+4\alpha)}\} \\
&= \zeta_1^{-2}\omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\{1+o(1)\}\,.
\end{aligned}
$$

由此可得

$$
\max_{1\leqslant j\leqslant p}\sum_{i=1}^n \mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \lesssim \max_{1\leqslant j\leqslant p}\sum_{i=1}^n \zeta_1^{-2}\omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} \lesssim n\max_{1\leqslant j\leqslant p}\omega_{jj} \leqslant \bar{M}n\,,
$$

应用 Chernozhukov、Chetverikov 和 Kato（2017）的引理 E.1，可得，对于 $\alpha \geqslant 1$and $n^{-1/2}\log^{3/2}(np) \lesssim 1$,

$$\begin{aligned}\mathbb{E}\left(\left|n^{-1/2}\sum_{i=1}^{n}\zeta_1^{-1}W_i\right|_\infty\right) &\lesssim n^{-1/2}\{n^{1/2}\log^{1/2}(p)+\log^{1/\alpha}(np)\log(p)\}\\ &\lesssim \log^{1/2}(np)\,.\end{aligned}$$

根据 ψ_α 范数的性质，有

$$\left\|\max_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}|\zeta_1^{-1}W_{i,j}|^2\right\|_{\psi_{\alpha/2}} \lesssim \log^2(np).$$

根 据 Chernozhukov、Chetverikov 和 Kato（2017）的引理 E.3，我们有 t

$$\mathbb{E}\left(\left|n^{-1}\sum_{i=1}^{n}(\zeta_1^{-1}W_i)^2\right|_\infty\right) \lesssim n^{-1}\{\bar{M}n+\log^2(np)\log(p)\} \lesssim \bar{M}\,.$$

至此，我们完成了该引理的证明。 □

˜ ˆ ˜ ˜引理 *A5* 的证明。令 $X_i = X_i - \boldsymbol{\theta}_n$和 $R_i = \|X_i\|$，其中 $i = 1,\ldots,n$。根据引理 1 的证明，$\|\hat{}\boldsymbol{\theta}_n\| = O_p(\zeta_1^{-1}n^{-1/2})$ 且 $\max_{1\leqslant i\leqslant n} R_i^{-1} = O_p(\zeta_1 n^{1/4})$。于是 $R_i^{-1}\|\hat{}\boldsymbol{\theta}_n\|$ 满足

$$R_i^{-1}\|\hat{\boldsymbol{\theta}}_n\| = O_p(n^{-1/2}) \text{ and } \max_{1\leqslant i\leqslant n} R_i^{-1}\|\hat{\boldsymbol{\theta}}_n\| = O_p(n^{-1/4})\,.$$

由于根据泰勒展开，$\tilde{R}_i^{-1} = R_i^{-1}\|W_i - R_i^{-1}\hat{}\boldsymbol{\theta}_n\|^{-1} = R_i^{-1}\left(1-2R_i^{-1}W_i^\top\hat{}\boldsymbol{\theta}_n + R_i^{-2}\|\hat{}\boldsymbol{\theta}_n\|^2\right)^{-1/2}$,

$$\tilde{R}_i^{-1} = R_i^{-1}\left(1+R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \tilde{\delta}_{1i}\right),$$

其中 $\tilde{\delta}_{1i}$ 满足 $\tilde{\delta}_{1i} = O_p(n^{-1})$，且 $\max_{1\leqslant i\leqslant n}\tilde{\delta}_{1i} = O_p(n^{-1/2})$。由此可得

$$\tilde{R}_i^{-1} = R_i^{-1}(1+\tilde{\delta}_{2i})\,,$$

其中 $\tilde{\delta}_{2i} = R_i^{-1}W_i^\top\hat{}\boldsymbol{\theta}_n - 2^{-1}R_i^{-2}\|\hat{}\boldsymbol{\theta}_n\|^2 + \tilde{\delta}_{1i}$ 满足 $\tilde{\delta}_{2i} = O_p(n^{-1/2})$，且 $\max_{1\leqslant i\leqslant n}\tilde{\delta}_{2i} =$

$O_p(n^{-1/4})$。因此，

$$\tilde{R}_i^{-1} = O_p(\zeta_1) \quad \text{and} \quad \max_{1 \leqslant i \leqslant n} \tilde{R}_i^{-1} = O_p(\zeta_1 n^{1/4}) .$$

~ ~ ~记 $W_i = X_i/\|X_i\|$为 $i = 1, \ldots, n$。则，

$$\begin{aligned} \tilde{W}_i &= \tilde{R}_i^{-1}(X_i - \hat{\boldsymbol{\theta}}_n) \\ &= R_i^{-1}(X_i - \hat{\boldsymbol{\theta}}_n)(1 + \tilde{\delta}_{2i}) \\ &= (W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1 + \tilde{\delta}_{2i}) . \end{aligned}$$

我们首先证明 $\|\tilde{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1} n^{-1/2})$。注意到， $\tilde{\boldsymbol{\theta}}_n$ 使其最小化

$$L_n^*(\boldsymbol{\beta}) = \sum_{i=1}^{n} \|Z_i \tilde{X}_i - \boldsymbol{\beta}\| ,$$

这是关于 $\boldsymbol{\beta}$的严格凸函数。因此，若能证明 $L_n^*(\boldsymbol{\beta})$ 存在一个 $\zeta_1 n^{1/2}$-相合局部极小解，则该局部极小解必定是 $L_n^*(\boldsymbol{\beta})$的一个 $\zeta_1 n^{1/2}$-相合全局极小解。存在 $\zeta_1 n^{1/2}$-相合局部极小解这一事实，可由如下结论推出：对于任意小的 $\varepsilon > 0$，存在一个常数 C_0，它不依赖于 n 和 p，使得

$$\liminf_n \mathbb{P}\left\{ \inf_{q \in \mathbb{R}^p,\ \|q\| = C_0} L_n^*(\zeta_1^{-1} n^{-1/2} q) > L_n^*(0) \right\} > 1 - \varepsilon, \tag{S.20}$$

由于 $|Z_i| = 1$，我们将 $\|Z_i \tilde{X}_i - \zeta_1^{-1} n^{-1/2} q\|$ 改写为

$$\begin{aligned} & \|Z_i \tilde{X}_i - \zeta_1^{-1} n^{-1/2} q\| \\ = \ & \tilde{R}_i \left(1 - 2\zeta_1^{-1} n^{-1/2} \tilde{R}_i^{-1} Z_i q^\top \tilde{W}_i + \zeta_1^{-2} n^{-1} \tilde{R}_i^{-2} \|q\|^2\right)^{1/2} . \end{aligned}$$

由于 $|\zeta_1^{-1} n^{-1/2} \tilde{R}_i^{-1} Z_i q^T \tilde{W}_i| = O_p(n^{-1/2})$ 和 $\zeta_1^{-2} n^{-1} \tilde{R}_{1i}^{-2} \|q\|^2 = O_p(n^{-1})$，通过泰勒展开，我们得到

$$\begin{aligned} & \|Z_i \tilde{X}_i - \zeta_1^{-1} n^{-1/2} q\| \\ = \ & \tilde{R}_i - \zeta_1^{-1} n^{-1/2} Z_i q^\top \tilde{W}_i + 2^{-1} \zeta_1^{-2} n^{-1} \tilde{R}_i^{-1} \|q\|^2 \end{aligned}$$

$$-2^{-1}\zeta_1^{-2}n^{-1}\tilde{R}_i^{-1}q^\top\tilde{W}_i\tilde{W}_i^\top q + O_p(\zeta_1^{-1}n^{-3/2})\,.$$

然后，

$$\begin{aligned}
&\zeta_1\left\{L_n^*(\zeta_1^{-1}n^{-1/2}q) - L_n^*(0)\right\}\\
=\;& \zeta_1\sum_{i=1}^n\left(\|Z_i\tilde{X}_i - \zeta_1^{-1}n^{-1/2}q\| - \|\tilde{X}_i\|\right)\\
=\;& -n^{-1/2}q^\top\left(\sum_{i=1}^n Z_i\tilde{W}_i\right) + 2^{-1}\zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n\tilde{R}_i^{-1}\\
&-2^{-1}\zeta_1^{-1}n^{-1}q^\top\left(\sum_{i=1}^n\tilde{R}_i\tilde{W}_i\tilde{W}_i^\top\right)q + O_p(n^{-1/2})\,.
\end{aligned}\tag{S.21}$$

由于 $\mathbb{E}^*\left(n^{-1/2}\sum_{i=1}^n Z_i\tilde{W}_i\right) = 0$ 且

$$\mathbb{E}^*\left(\left\|n^{-1/2}\sum_{i=1}^n Z_i\tilde{W}_i\right\|^2\right) = n^{-1}\sum_{i=1}^n\tilde{W}_i^\top\tilde{W}_i = 1,$$

我们得到

$$\left|n^{-1/2}q^\top\sum_{i=1}^n Z_i\tilde{W}_i\right| \leqslant \|q\|\left\|n^{-1/2}\sum_{i=1}^n Z_i\tilde{W}_i\right\| = O_p(\|q\|)\,.$$

与此同时，由于 $\zeta_1^{-1}n^{-1}\sum_{i=1}^n R_i^{-1} = 1 + O_p(n^{-1/2}),$，我们有

$$\begin{aligned}
\zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n\tilde{R}_i^{-1} &= \zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n R_i^{-1}(1+\tilde{\delta}_{2i})\\
&= \|q\|^2\{1 + O_p(n^{-1/4})\}\,.
\end{aligned}$$

经简单代数运算可得

$$\begin{aligned}
&n^{-1}\sum_{i=1}^n\tilde{R}_i^{-1}\tilde{W}_i\tilde{W}_i^\top\\
=\;& n^{-1}\sum_{i=1}^n R_i^{-1}(W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)^\top(1+\tilde{\delta}_{2i})
\end{aligned}$$

$$
\begin{aligned}
= \quad & n^{-1}\sum_{i=1}^{n} R_i W_i W_i^{\top}(1+\tilde{\delta}_{2i}) - 2n^{-1}\sum_{i=1}^{n} R_i^{-2} W_i \hat{\boldsymbol{\theta}}_n^{\top}(1+\tilde{\delta}_{2i}) \\
& + n^{-1}\sum_{i=1}^{n} R_i^{-3}\hat{\boldsymbol{\theta}}_n\hat{\boldsymbol{\theta}}_n^{\top}(1+\tilde{\delta}_{2i})\,.
\end{aligned}
$$

与 Cheng 等人（2019）的证明类似，并利用引理 A3 中关于 $Q = n^{-1}\sum_{i=1}^{n} R_i^{-1}W_iW_i^{-1}$ 的结果，我们可以证明 $n^{-1}q^{\top}\sum_{i=1}^{n} R_i W_i W_i^{\top} q(1+\tilde{\delta}_{2i}) = O_p(\zeta_1 n^{-1/2} + \zeta_1 p^{-(1/6\wedge\delta/2)})$。此外，由于

$$
n^{-1}\sum_{i=1}^{n} R_i^{-2} q^{\top} W_i \leqslant n^{-1}\sum_{i=1}^{n} R_i^{-2}\|q\|\|W_i\| = \|q\| n^{-1}\sum_{i=1}^{n} R_i^{-2} = O_p(\zeta_1^2)
$$

以及 $n^{-1}\sum_{i=1}^{n} R_i^{-3} = \zeta_3\{1+o_p(1)\}$，我们有

$$
\begin{aligned}
& n^{-1}q^{\top}\sum_{i=1}^{n} R_i^{-2} W_i \hat{\boldsymbol{\theta}}_n^{\top} q(1+\tilde{\delta}_{2i}) \\
= \quad & n^{-1}\sum_{i=1}^{n} R_i^{-2} q^{\top} W_i (1+\tilde{\delta}_{2i})(\hat{\boldsymbol{\theta}}_n^{\top} q) = O_p(\zeta_1 n^{-1/2})\,.
\end{aligned}
$$

and

$$
n^{-1}q^{\top}\sum_{i=1}^{n} R_i^{-3}\hat{\boldsymbol{\theta}}_n\hat{\boldsymbol{\theta}}_n^{\top} q(1+\tilde{\delta}_{2i}) = n^{-1}\sum_{i=1}^{n} R_i^{-3}(1+\tilde{\delta}_{2i})\|q^{\top}\hat{\boldsymbol{\theta}}_n\|^2 = O_p(\zeta_1 n^{-1})\,.
$$

因此，我们得到

$$
\begin{aligned}
& 2^{-1}\zeta_1^{-1} n^{-1}\|q\|^2 \sum_{i=1}^{n}\tilde{R}_i^{-1} + 2^{-1}\zeta_1^{-1} n^{-1} q^{\top}\left(\sum_{i=1}^{n}\tilde{R}_i\tilde{W}_i\tilde{W}_i^{\top}\right) q \\
= \quad & 2^{-1}\|q\|^2 + O_p(n^{-1/4} + p^{-\delta})\,.
\end{aligned}
$$

选取一个足够大的常数 C_0，则 (S.21) 中的第二项占主导，因此 $\zeta_1\left\{L_n^*(\zeta_1^{-1}n^{-1/2}q) - L_n^*(0)\right\} > 0$。于是，我们有 $\|\tilde{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1}n^{-1/2})$。

记 $\Theta_i = Z_i\boldsymbol{\theta}_n + \boldsymbol{\theta}_n$ 为 $i = 1, \ldots, n$。则

$$
\max_{1\leqslant i\leqslant n}\|\Theta_i\| \leqslant \|\hat{\boldsymbol{\theta}}_n\| + \|\tilde{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1}n^{-1/2}).
$$

˜回顾一下， $\boldsymbol{\theta}_n$ 满足

$$\sum_{i=1}^{n} \frac{Z_i\tilde{X}_i - \tilde{\boldsymbol{\theta}}_n}{\|Z_i\tilde{X}_i - \tilde{\boldsymbol{\theta}}_n\|} = \sum_{i=1}^{n} \frac{Z_iW_i - R_i^{-1}\Theta_i}{\|Z_iW_i - R_i^{-1}\Theta_i\|} = 0\,,$$

这等价于

$$n^{-1}\sum_{i=1}^{n}(Z_iW_i - R_i^{-1}\Theta_i)\left(1 - 2Z_iR_i^{-1}W_i^{\top}\Theta_i + R_i^{-2}\|\Theta_i\|^2\right)^{-1/2} = 0\,,$$

其中 $|R_i^{-1}W_i^{\top}\Theta_i| = O_p(n^{-1/2})$, $R_i^{-2}\|\Theta_i\|^2 = O_p(n^{-1})$,

$$\max_{1\leqslant i\leqslant n}|R_i^{-1}W_i^{\top}\Theta_i| = O_p(n^{-1/4}) \quad \text{and} \quad \max_{1\leqslant i\leqslant n} R_i^{-2}\|\Theta_i\|^2 = O_p(n^{-1/2})\,.$$

泰勒展开得到

$$n^{-1}\sum_{i=1}^{n}(Z_iW_i - R_i^{-1}\Theta_i)(1 + Z_iR_i^{-1}W_i^{\top}\Theta_i - 2R_i^{-2}\|\Theta_i\|^2 + \tilde{\delta}_{3i}) = 0$$

其中 $\delta_{3i} = O_p\{(Z_iR_i^{-1}W_i^{\top}\Theta_i - R_i^{-2}\|\Theta_i\|^2)^2\} = O_p(n^{-1})$, 且 $\max_{1\leqslant i\leqslant n}\delta_{3i} = O_p(n^{-1/2})$。于是,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} Z_iW_i(1 - 2R_i^{-2}\|\Theta_i\|^2 + \tilde{\delta}_{3i}) + n^{-1}\sum_{i=1}^{n} R_i^{-1}(W_i^{\top}\Theta_i)W_i \\
= \;& n^{-1}\sum_{i=1}^{n} Z_iW_i(1 - 2R_i^{-2}\|\Theta_i\|^2 + \tilde{\delta}_{3i}) + n^{-1}\sum_{i=1}^{n} Z_iR_i^{-1}W_iW_i^{\top}\hat{\boldsymbol{\theta}}_n \\
& + n^{-1}\sum_{i=1}^{n} R_i^{-1}W_iW_i^{\top}\tilde{\boldsymbol{\theta}}_n \\
= \;& n^{-1}\sum_{i=1}^{n} R_i^{-1}\Theta_i(1 + \tilde{\delta}_{3i} + \tilde{\delta}_{4i}) \\
= \;& n^{-1}\sum_{i=1}^{n} R_i^{-1}\tilde{\boldsymbol{\theta}}_n(1 + \tilde{\delta}_{3i} + \tilde{\delta}_{4i}) + n^{-1}\sum_{i=1}^{n} Z_iR_i^{-1}\hat{\boldsymbol{\theta}}_n(1 + \tilde{\delta}_{3i} + \tilde{\delta}_{4i})\,,
\end{aligned}$$

其中 $\tilde{\delta}_{4i} = Z_iR_i^{-1}W_i^{\top}\Theta_i - 2R_i^{-2}\|\Theta_i\|^2 = O_p(\tilde{\delta}_{3i}^{1/2})$ 满足 $\max_{1\leqslant i\leqslant n}\tilde{\delta}_{4i} = O_p(n^{-1/4})$。

引理 1 的证明蕴含 $\lceil\boldsymbol{\theta}|_{\infty} = O_p\{n^{-1/2}\log^{1/2}(np)\}$。由于 $\mathbb{E}^*\left(n^{-1}\sum_{i=1}^{n} Z_iR_i^{-1}\right) = 0$ 且 $\mathbb{E}^*\left\{\left(n^{-1}\sum_{i=1}^{n} Z_iR_i^{-1}\right)^2\right\} = n^{-2}\sum_{i=1}^{n} R_i^{-2} = O_p(n^{-1}\zeta_2)$, 我们有 $n^{-1}\sum_{i=1}^{n} Z_iR_i^{-1} = O_p(\zeta_1 n^{-1/2})$。

由于 Z_i 有界，按照引理A4（iii）的证明，可以直接证明 $|n^{-1/2}\sum_{i=1}^n Z_i W_i|_\infty = O_p\{p^{-1/2}\log^{1/2}(np)\}$。因此，类似于引理 1 的证明，我们得到

$$|\tilde{\boldsymbol{\theta}}|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}$$

and

$$\begin{aligned}&\left|n^{-1}\sum_{i=1}^n R_i^{-1}W_i W_i^\top \tilde{\boldsymbol{\theta}}_n\right|_\infty \\ = \;& O_p\{\zeta_1 n^{-1/2}p^{-(1/6\wedge\delta/2)}\log^{1/2}(np) + \zeta_1 n^{-1}\log^{1/2}(np)\}\,.\end{aligned}$$

与此同时，有 $|n^{-1}\sum_{i=1}^n R_i^{-1}| = \zeta_1 + O_p(\zeta_1 n^{-1/2})$。最后，

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n Z_i W_i + \tilde{C}_n\,, \tag{S.22}$$

˜和 C_n 是满足条件的余项

$$|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-\delta-(1/6\wedge\delta/2)}\log^{1/2}(np)\}\,.$$

我们完成了这个引理的证明。 □

附录 C：补充模拟结果

本节报告补充模拟结果。C1 节展示了关于 $\rho = 0.2$ 和 0.5 的同时置信区间模拟结果。C2 节报告了关于高维位置参数全局检验的模拟结果。

C.1 关于同时置信区间的补充模拟结果

ˆ 表 A4 报告了基于 $\boldsymbol{\theta}_n$ 的 $\rho = 0.2$¯和 0.5 时同时置信区间的覆盖概率和中位数长度，基于样本均值 X_n 的同时置信区间结果则列于括号中。ˆ 我们发现，基于 $\boldsymbol{\theta}_n$ 、 $\rho = 0.2$ 和 0.5 的同时置信区间，其表现与主文中 $\boldsymbol{\theta}_n$ 和 0.

8 时的表现相似。同时置信区间取得了令人满意的覆盖概率，¯并且其长度远短于基于 X_n 的同时置信区间；后者基于多元 t分布，而该分布具有重尾特征。

ˆ 表 A4：基于 $\boldsymbol{\theta}_n$ 的同时置信区间的覆盖概率（%）和中位数长度，基于 X_n 的同时置信区间的结果¯列于括号内。

				$\boldsymbol{\theta}=\boldsymbol{\theta}_1$				$\boldsymbol{\theta}=\boldsymbol{\theta}_2$			
				覆盖概率		中位数长度		覆盖概率		中位数长度	
模型	ρ	n	p	90%	95%	90%	95%	90%	95%	90%	95%
I	0.2	100	100	89.8 (89.9)	94.5 (94.5)	0.65 (0.65)	0.69 (0.69)	88.8 (88.7)	94.4 (94.4)	0.65 (0.65)	0.69 (0.69)
			1000	88.7 (88.7)	94.5 (94.3)	0.77 (0.77)	0.80 (0.80)	90.0（89.6）	94.7（94.8）	0.77（0.77）	0.80（0.80）
		200	100	89.0 (88.9)	94.3 (94.1)	0.46 (0.46)	0.49 (0.49)	88.8 (88.8)	94.0 (94.2)	0.46 (0.46)	0.49 (0.49)
			1000	89.8 (89.8)	94.4 (94.4)	0.55 (0.55)	0.57 (0.57)	88.7 (89.2)	94.6 (94.3)	0.55 (0.55)	0.57 (0.57)
	0.5	100	100	89.6 (89.8)	94.5 (94.4)	0.65 (0.65)	0.69 (0.69)	88.4 (88.8)	94.0 (94.1)	0.65 (0.65)	0.69 (0.69)
			1000	88.4 (88.4)	94.3 (94.3)	0.77 (0.77)	0.80 (0.80)	87.4 (87.4)	94.1 (94.2)	0.77 (0.77)	0.80 (0.80)
		200	100	90.9 (90.9)	95.1 (95.2)	0.46 (0.46)	0.49 (0.49)	89.7（90.0）	95.3（95.0）	0.46（0.46）	0.49（0.49）
			1000	89.0 (89.0)	94.2 (94.3)	0.55 (0.55)	0.57 (0.57)	88.8 (88.6)	94.3 (94.0)	0.55 (0.55)	0.57 (0.57)
II	0.2	100	100	89.2 (88.8)	94.8 (94.2)	0.71 (1.05)	0.75 (1.12)	88.4（88.8）	93.7（94.3）	0.71（1.05）	0.75（1.11）
			1000	89.0 (89.4)	94.1 (94.8)	0.84 (1.24)	0.88 (1.30)	89.0（88.9）	94.4（94.6）	0.84（1.24）	0.88（1.30）
		200	100	90.7 (89.8)	95.3 (94.7)	0.50 (0.76)	0.53 (0.80)	89.2 (89.7)	94.0 (94.4)	0.50 (0.76)	0.53 (0.80)
			1000	88.6 (89.5)	94.2 (94.6)	0.59 (0.90)	0.62 (0.93)	89.0 (90.6)	95.0 (95.1)	0.59 (0.90)	0.62 (0.94)
	0.5	100	100	89.2 (87.9)	93.6 (93.9)	0.71 (1.05)	0.75 (1.12)	89.4 (88.6)	94.6 (94.1)	0.71 (1.05)	0.75 (1.11)
			1000	89.2 (88.9)	94.4 (94.2)	0.84 (1.24)	0.88 (1.30)	90.0（89.4）	94.7（94.6）	0.84（1.25）	0.88（1.30）
		200	100	89.4 (90.0)	94.1 (94.6)	0.50 (0.76)	0.53 (0.80)	89.7 (88.6)	95.0 (93.6)	0.50 (0.76)	0.53 (0.80)
			1000	90.0 (89.9)	95.6 (94.8)	0.59 (0.90)	0.62 (0.94)	88.8 (89.5)	93.8 (94.4)	0.59 (0.89)	0.62 (0.93)
III	0.2	100	100	89.6 (89.5)	95.0 (95.1)	0.65 (0.66)	0.69 (0.70)	89.4 (89.4)	94.6 (94.6)	0.65 (0.66)	0.69 (0.70)
			1000	89.3 (88.8)	94.5 (94.5)	0.78 (0.78)	0.82 (0.82)	90.3 (90.7)	95.0 (94.9)	0.78 (0.78)	0.82 (0.82)
		200	100	89.2 (89.0)	94.4 (94.4)	0.46 (0.46)	0.49 (0.49)	90.0 (89.6)	95.1 (95.2)	0.46 (0.46)	0.49 (0.49)
			1000	89.7 (89.7)	94.6 (94.8)	0.55 (0.55)	0.57 (0.58)	90.4（90.6）	95.0（95.0）	0.55（0.55）	0.57（0.57）
	0.5	100	100	88.9 (89.3)	94.0 (94.6)	0.65 (0.65)	0.69 (0.69)	88.0（88.5）	94.2（94.0）	0.65（0.65）	0.69（0.69）
			1000	89.1 (89.2)	94.3 (94.2)	0.78 (0.78)	0.81 (0.81)	89.2 (88.9)	94.1 (94.0)	0.78 (0.78)	0.81 (0.81)
		200	100	89.6 (89.7)	95.0 (94.4)	0.46 (0.46)	0.49 (0.49)	89.6 (89.7)	94.9 (94.4)	0.46 (0.46)	0.49 (0.49)
			1000	89.0 (89.1)	94.3 (94.4)	0.55 (0.55)	0.57 (0.57)	89.3 (89.6)	95.4 (95.0)	0.55 (0.55)	0.57 (0.57)

C.2 高维位置参数全局检验的模拟

本节报告基于 T_n （中位数检验）的一样本高维位置参数检验的表现，并将其与三种替代方法进行比较：Chen 和 Qin（2010）的检验（CQ 检验）；基于 T_{Mean} （均值检验）并采用自助法¯近似的 X_n检验；以及 Wang、Peng 和 Li（2015）基于 T_{WPL}提出的检验（WPL 检验）。我们考虑与第 5.1 节相同的数据生成模型（I、II 和 III）。对于 $\boldsymbol{\theta}$，我们令其前$\lfloor c_0 \log p \rfloor$ 个分量非零，其余元素均为零。 c_0 取 0.5 或 1。 $\boldsymbol{\theta}$ 中非零项的大小为 $\kappa(\log p/n)^{1/2}$, ，其中 κ 取 0 到 5 之间的值。注意， $\kappa = 0$ 表示零假设。对于每个样本量，我们考虑 $n = 50$ 或 100，以及 $p = 100$ 和 1000。

图 A3–A10 绘出了模型 I 和模型 II 下四种检验方法（CQ、均值、中位数和 WPL 检验）在 5% 显著性水平下的经验尺寸（$\kappa = 0$）和检验功效（$\kappa \neq 0$）。$\kappa = 0$ 的结果表明，在不同情形下，这四种检验的经验尺寸均接近名义显著性水平。当 $\kappa \neq 0$时，随着 κ 增大，这些检验的功效也随之提高，即信号越强，检验功效越高。对于高斯数据，基于 T_{Mean} 的均值检验与基于 T_n 的中位数检验具有相近的检验功效，且二者均优于 CQ 检验和 WPL 检验，这两种检验属于 L_2-范数类型。此外，当数据来自多元 t分布时，中位数检验优于均值检验，这表明在重尾性下，基于样本空间中位数的程序优于基于样本均值的程序。总之，当备择假设稀疏且底层分布为重尾分布时，基于 T_n 的中位数检验在四种检验方法中更具优势。

其次，图 A11 展示了在 $\rho = 0$下，模型 III 中四种检验（CQ、均值、中位数、WPL）的经验尺寸和检验功效。可以看出，即使模型 III 不属于椭圆分布族，中位数检验的尺寸仍可控制在名义水平 $\alpha = 0.05$，WPL 检验也同样如此。我们还可以看到，中位数检验和均值检验的检验功效优于 CQ 检验和 WPL 检验，尤其是在 $c_0 = 0.5$ 中 $\boldsymbol{\theta}$ 的非零元素数量相对较少时。

参考文献

Belloni, A., Chernozhukov, V., Chetverikov, D., Hansen, C., 和 Kato, K. (2018年) 高维计量经济学与广义 GMM。*arXiv* 预印本 *arXiv:1806.01888*。

Chen, S. X. 和 Qin, Y. (2010) 一种适用于高维数据的双样本检验及其在基因集检验中的应用，$Ann. Statist.$ **38** (**2**)，808–835。

Cheng,G. Liu, B. Peng, L; Zhang, B and Zheng, S. (2019).检验两个高维空间符号协方差矩阵的相等性。《斯堪的纳维亚统计学杂志》**46**, 257–271.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2017).高维中的中心极限定理与自助法。《概率年刊》**45**(**4**), 2309–2352.

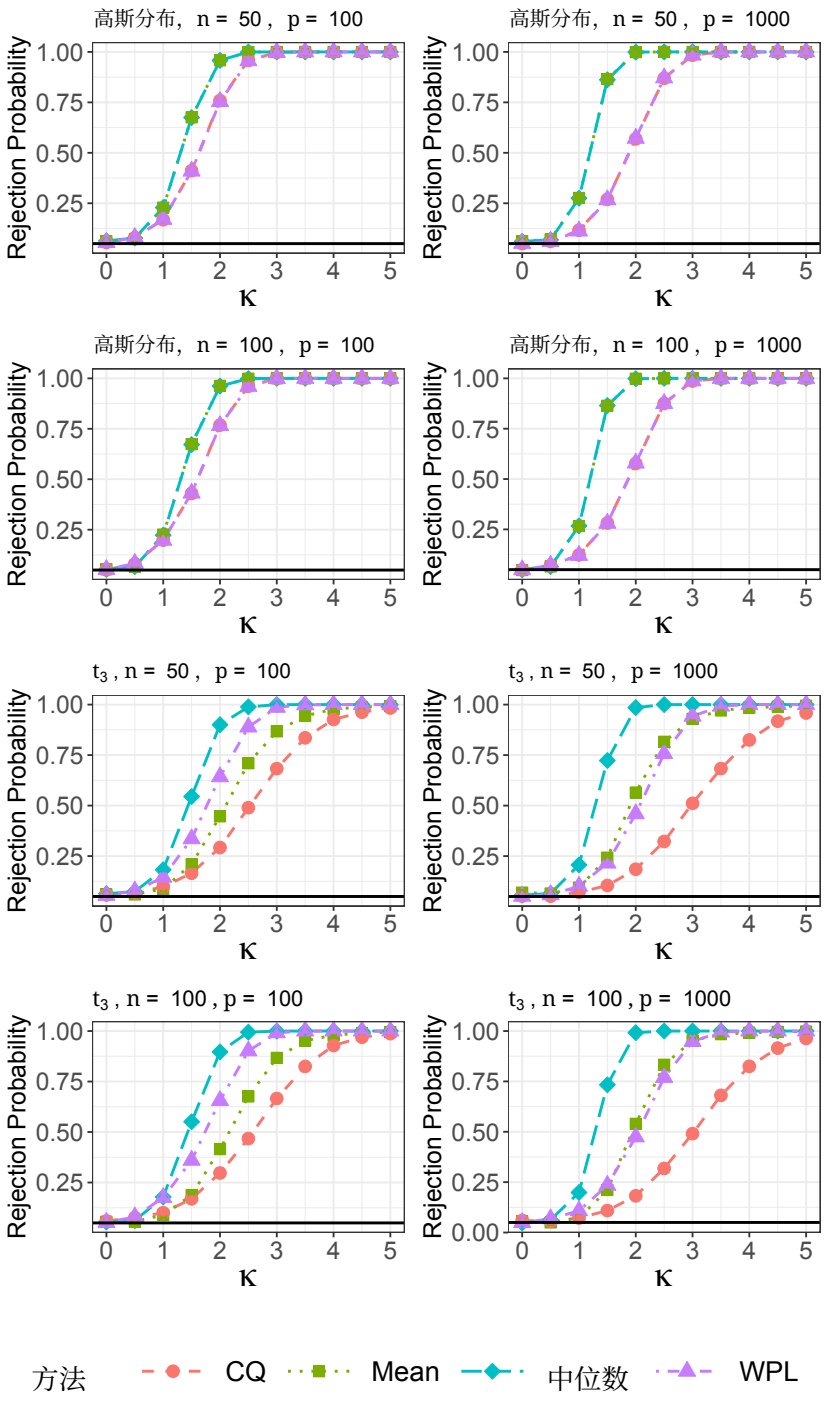


图 A3：模型 I 和模型 II 中，$c_0 = 0.5$ 和 $\rho = 0$条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，t_3 表示自由度为 3 的多元 t分布。

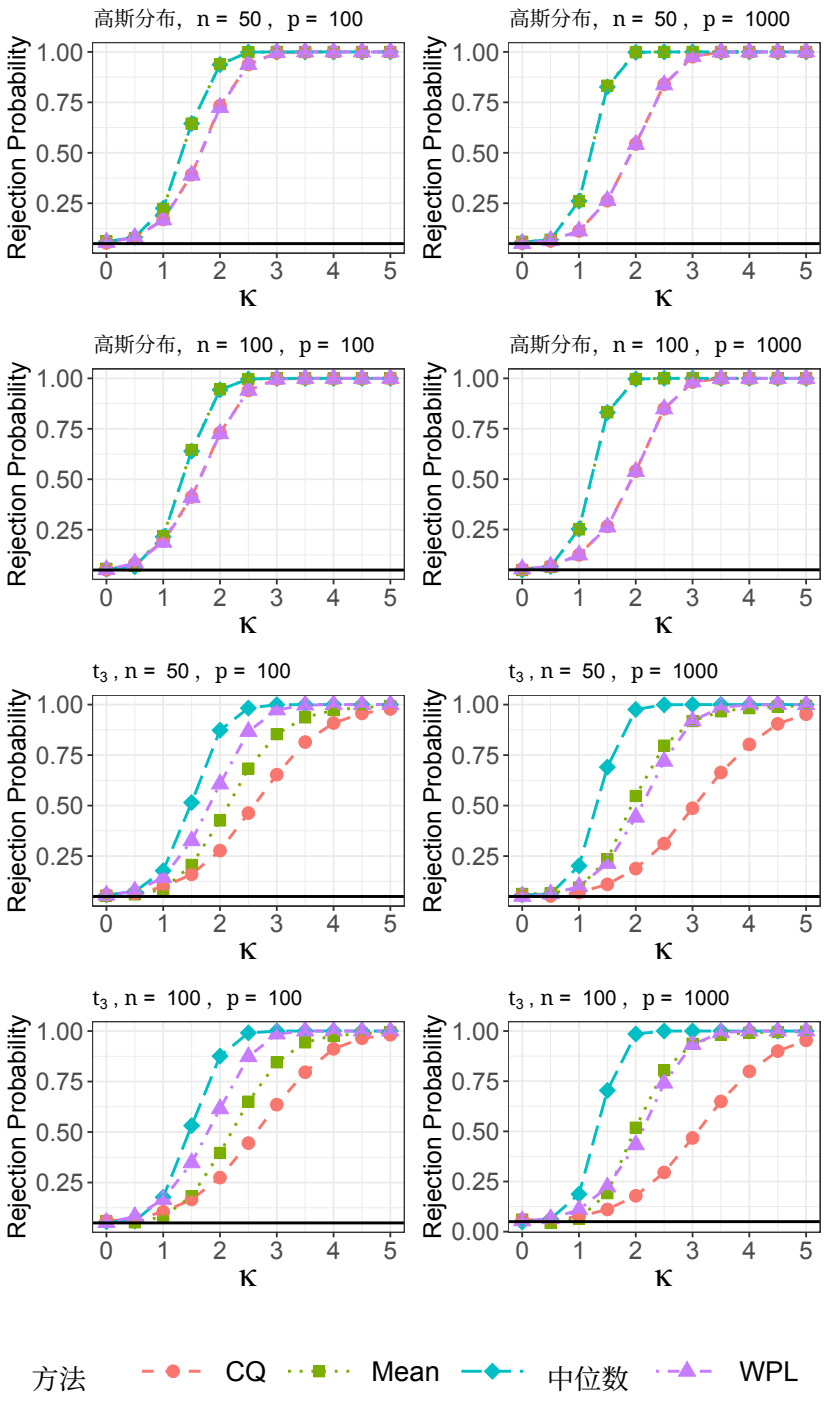


图 A4：模型 I 和模型 II 在 $c_0 = 0.5$ 和 $\rho = 0.2$条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，而 t_3 表示具有 3 个自由度的多元 t分布。

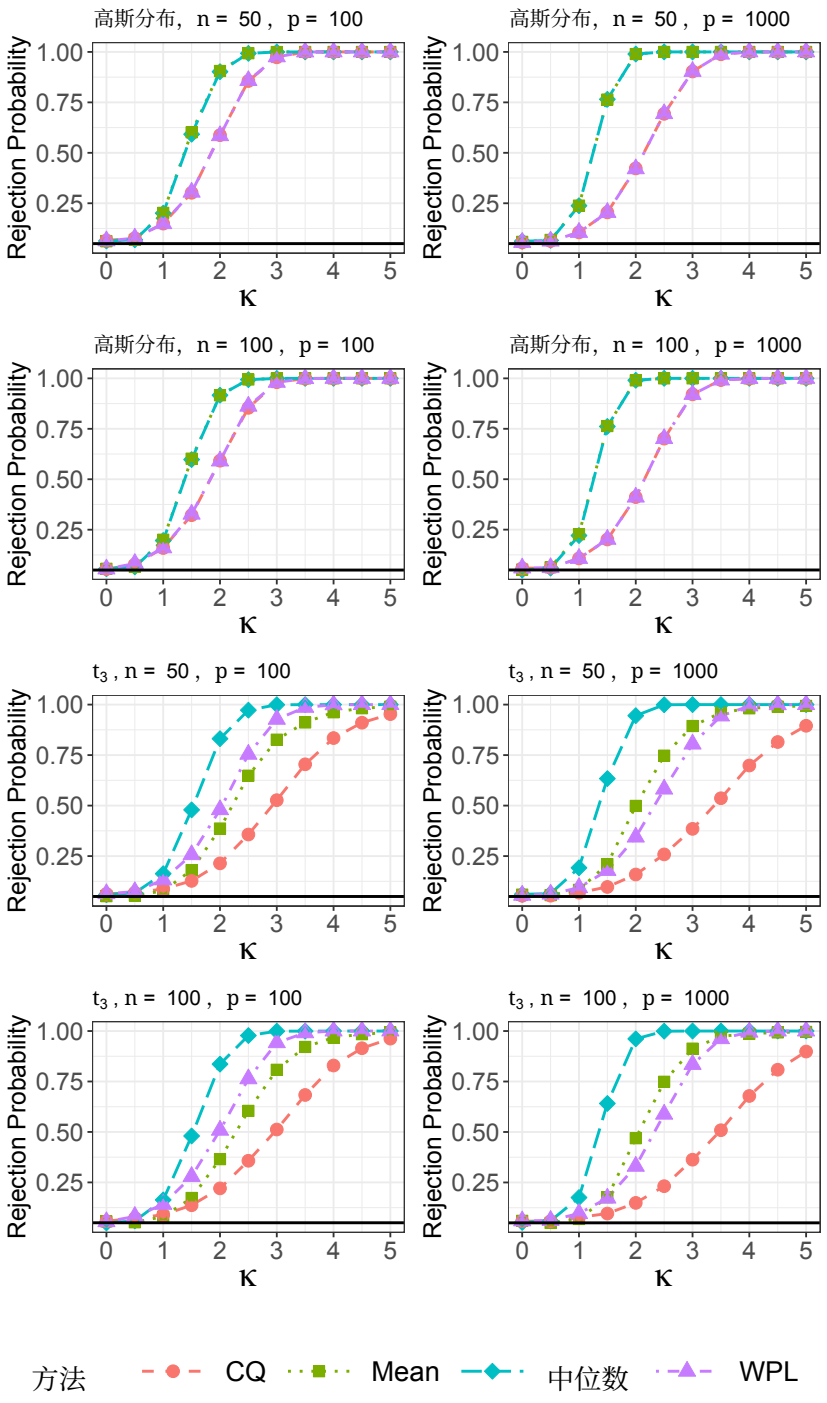


图 A5：模型 I 和模型 II 在 $c_0 = 0.5$ 和 $\rho = 0.5$条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，t_3 表示自由度为 3 的多元 t分布。

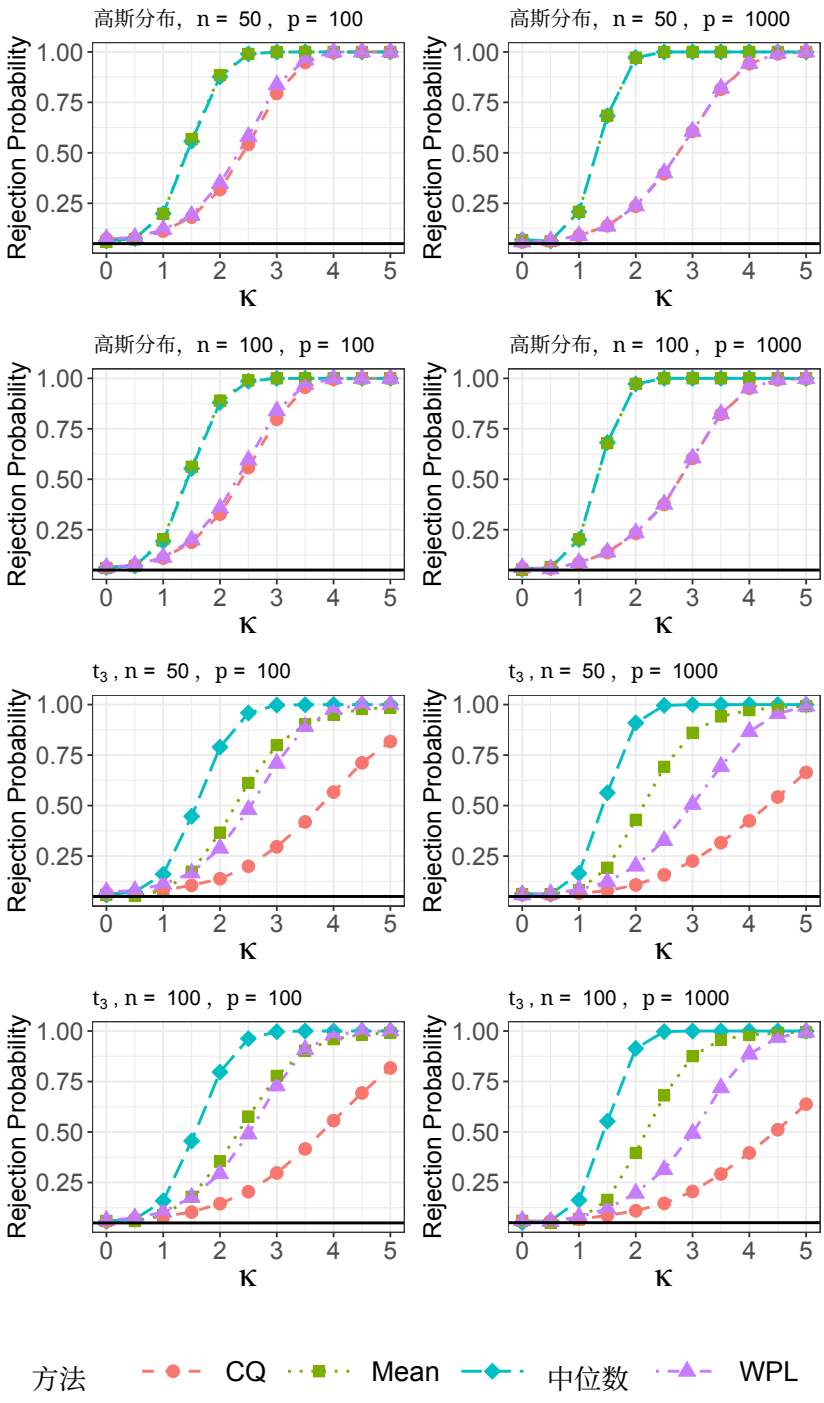


图 A6：模型 I 和模型 II 中，$c_0 = 0.5$ 和 $\rho = 0.8$条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，t_3 表示具有 3 个自由度的多元 t分布。

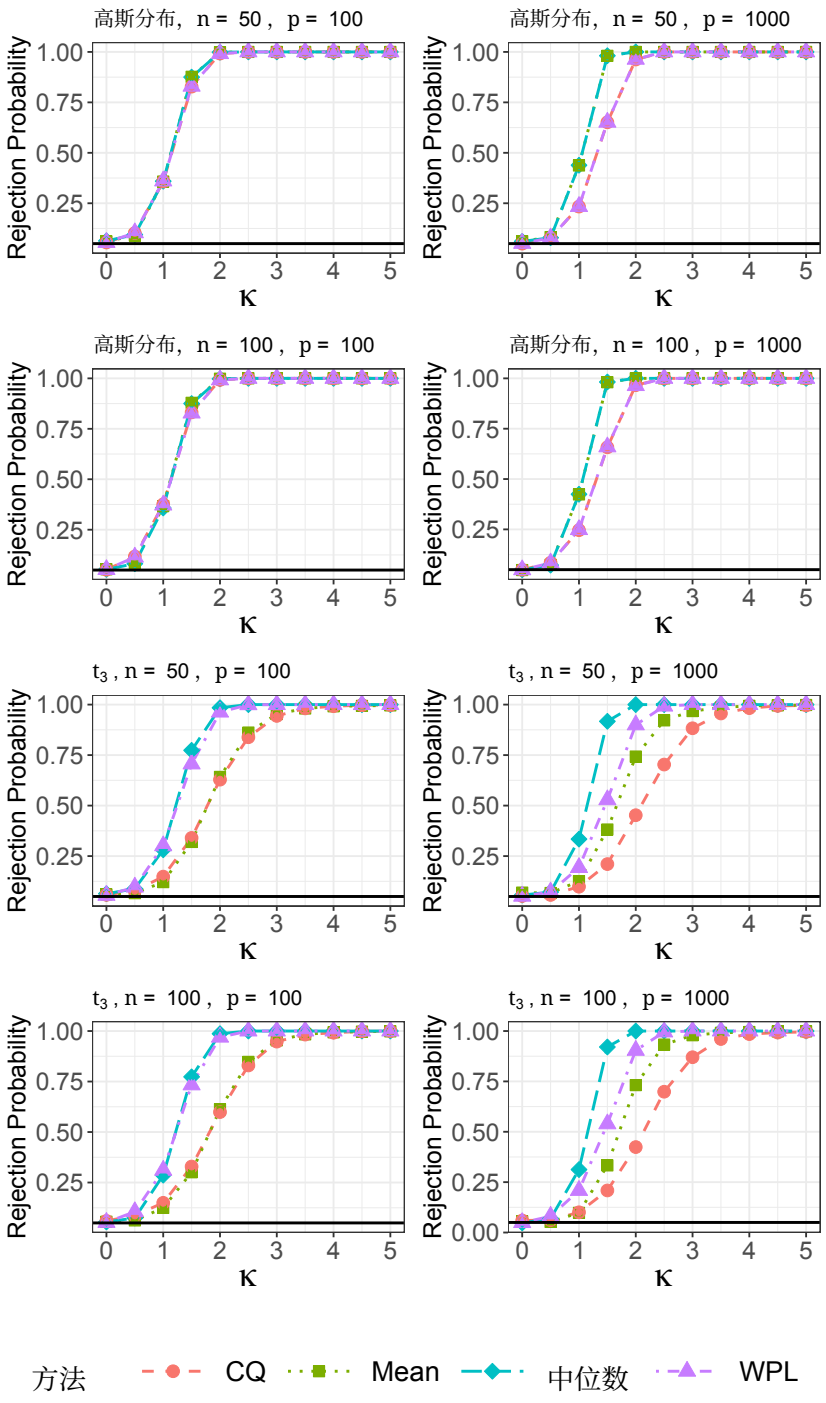


图 A7：模型 I 和模型 II 在 $c_0 = 1$ 和 $\rho = 0$条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。黑色水平线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，而 t_3 表示具有 3 个自由度的多元t分布。

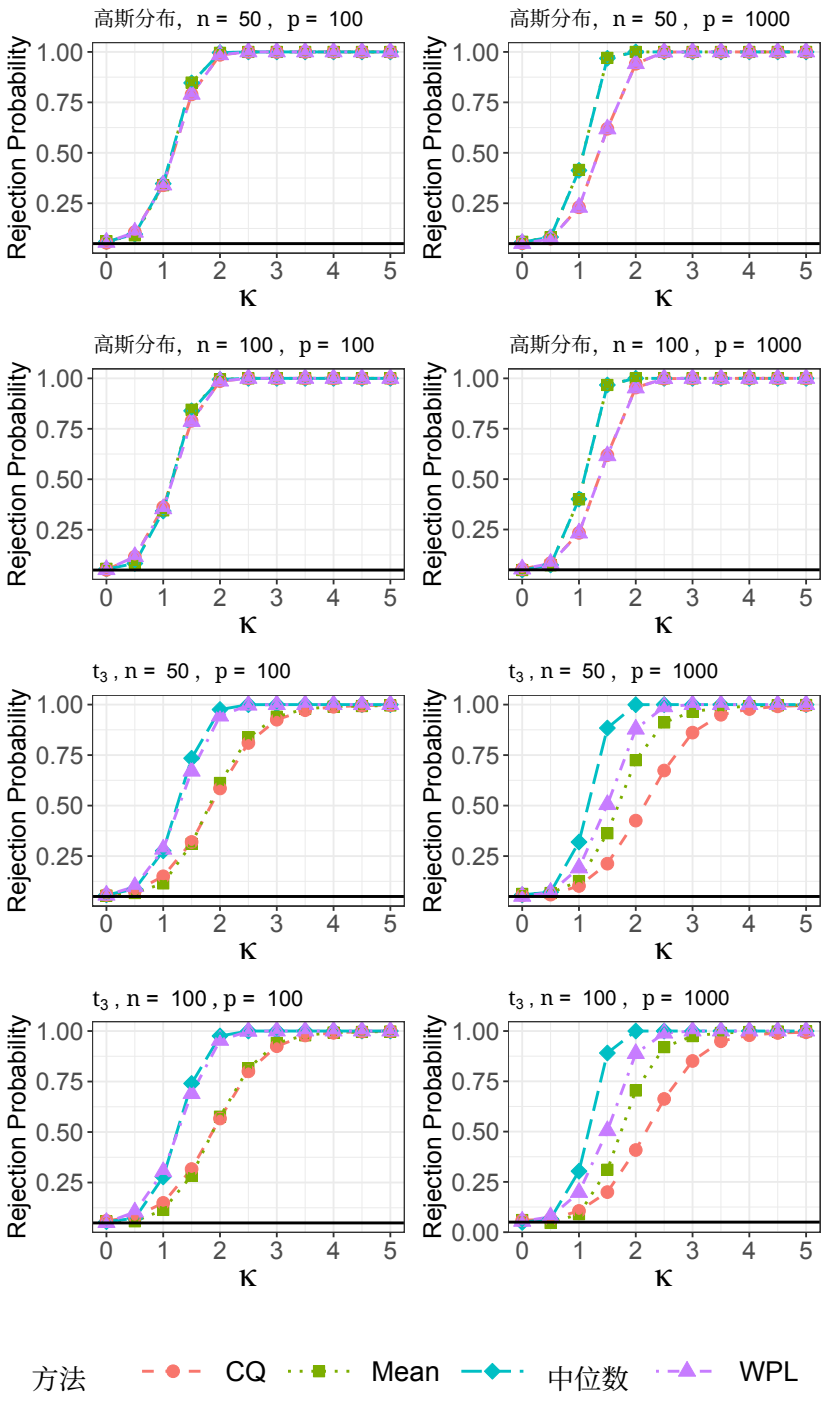


图 A8：模型 I 和模型 II 在 $c_0 = 1$ and $\rho = 0.2$ 下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，而 t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

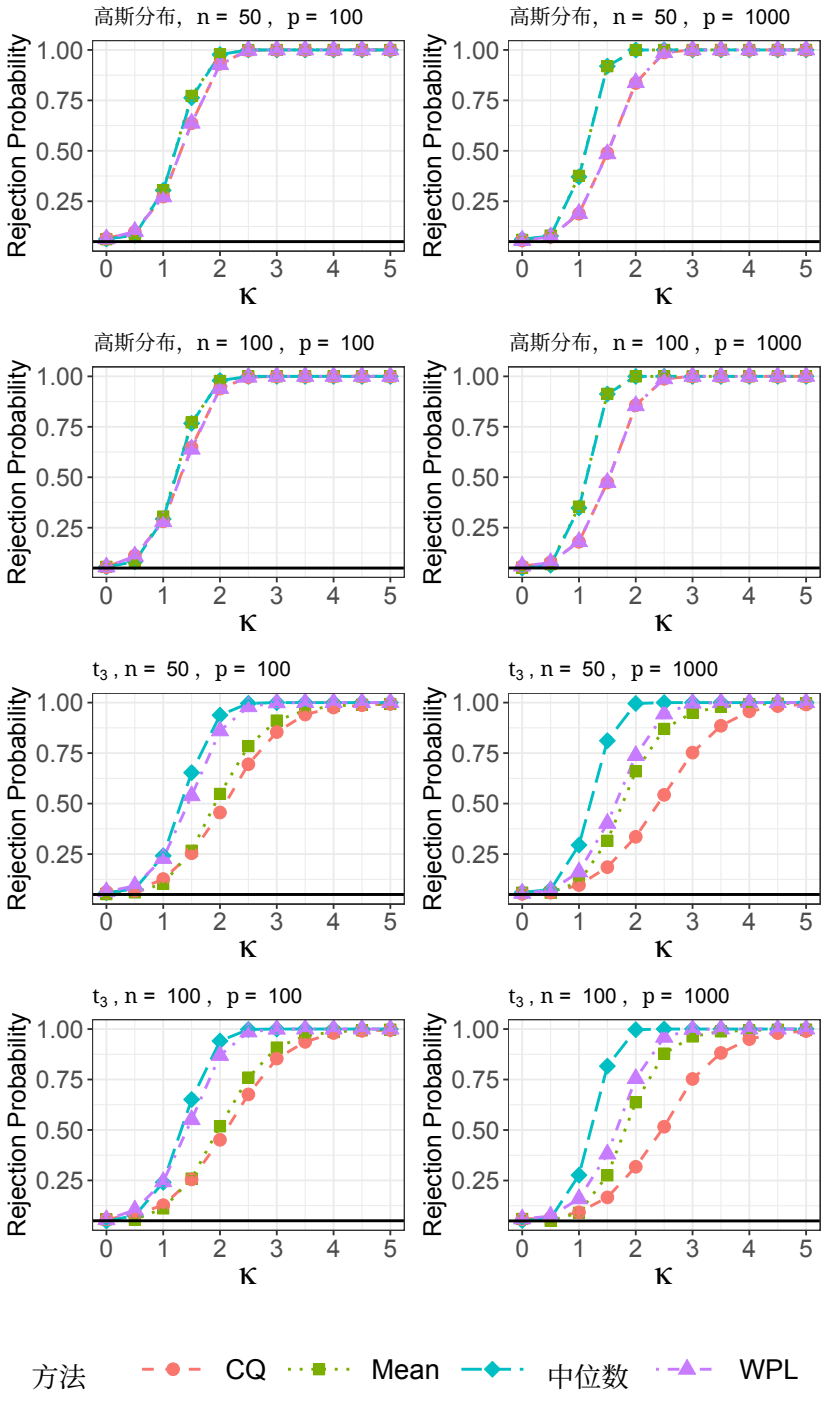


图 A9：模型 I 和模型 II 在 $c_0 = 1$ 和 $\rho = 0.5$条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布，t_3 表示具有 3 个自由度的多元 t分布。

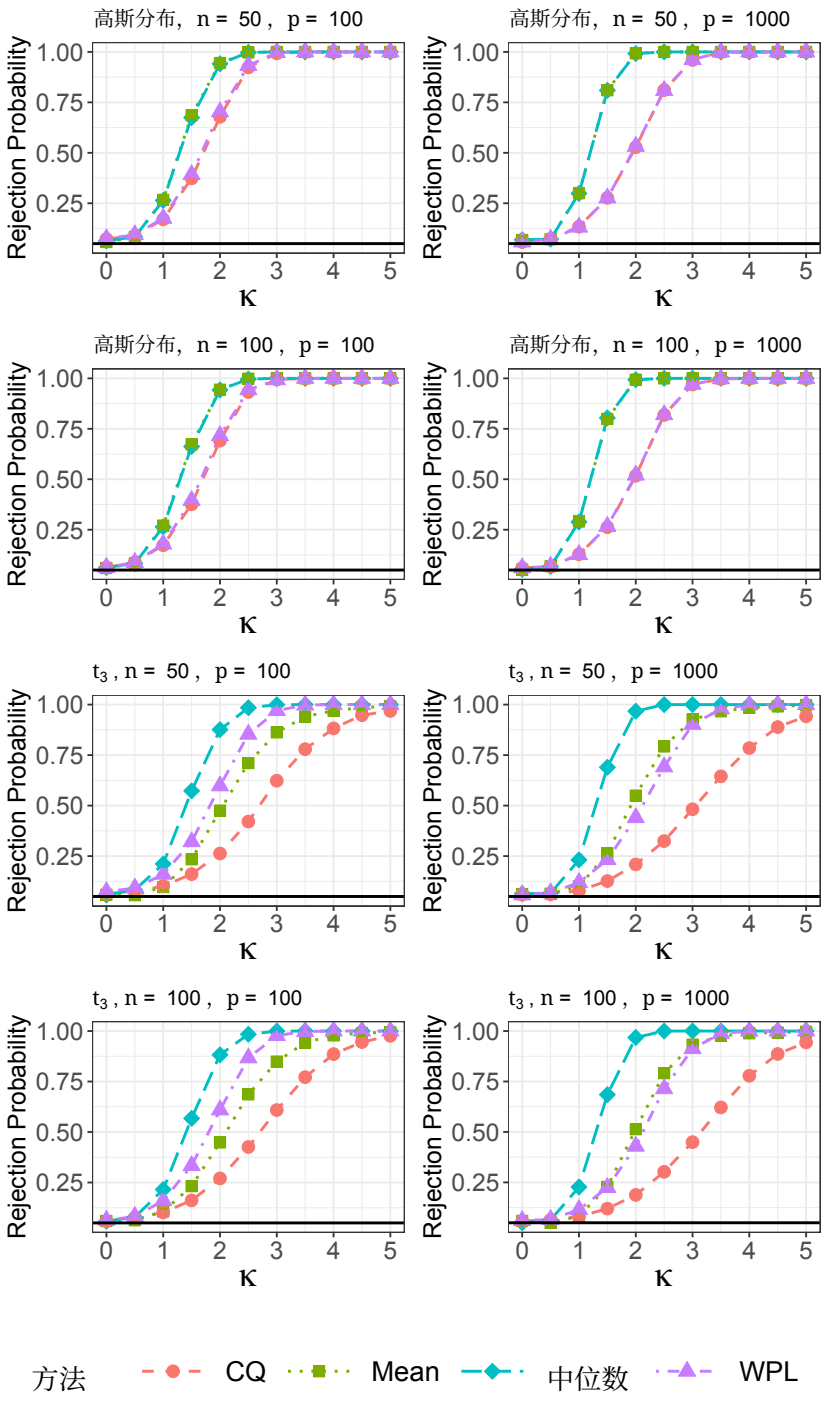


图 A10：模型 I 和模型 II 在 $c_0 = 1$ and $\rho = 0.8$ 条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。黑色水平线表示名义 5% 显著性水平。“高斯分布”表示多元正态分布， t_3 denotes 多元t分布，其自由度为 3。

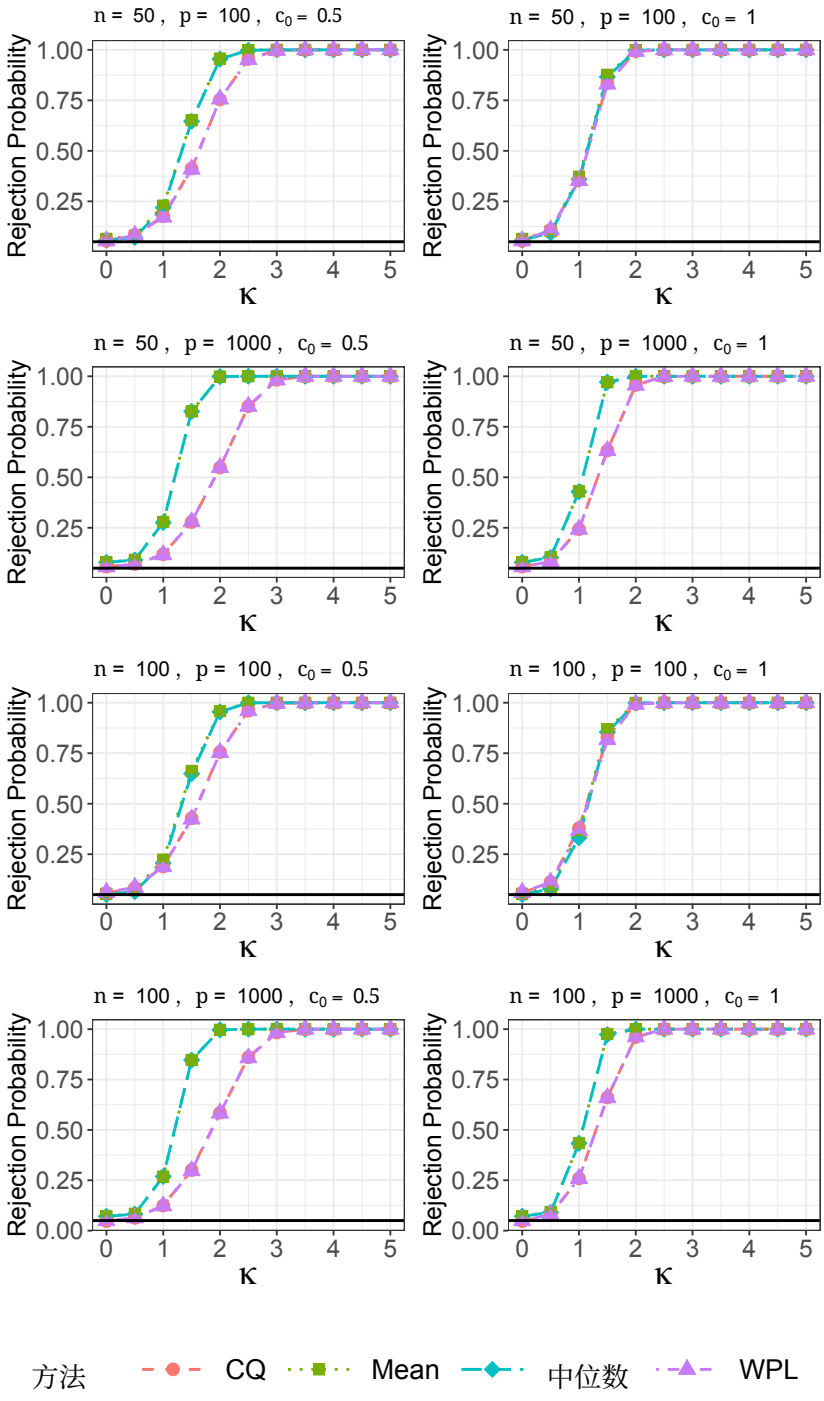


图 A11：模型 III 在 $\rho = 0$ 条件下四种检验方法（CQ、均值、中位数、WPL）的经验尺寸和检验功效。水平黑色实线表示名义 5% 显著性水平。

Koike, Y.（2021）。关于超矩形高维中心极限定理中维数依赖性的注记。日本统计与数据科学杂志。**1**，257–297。

Liu, W. 和 Shao, Q.-M.（2014）。具有错误发现率控制的大规模 t 检验中的相变与正则化自助法。统计学年鉴。**42**，2003–2025。

Rudelson, M. 和 Vershynin, R.（2013）。Hanson–Wright 不等式与次高斯集中性。概率电子通讯。**18**，1–9。

Vershynin, R.（2018）。高维概率。剑桥大学出版社，2018年。

Wang, L.、Peng, B. 和 Li, R.（2015）。一种用于均值向量的高维非参数多元检验。美国统计学会杂志。**110**，1658–1669。