

# 高维两样本位置参数秩和检验\*

邱涛<sup>1</sup> 吴文琦<sup>1</sup> 许王莉<sup>1,2</sup>

(1. 中国人民大学统计学院, 北京 100872; 2. 中国人民大学应用统计科学研究中心, 北京 100872)

**摘要** 对于一维数据, 很容易对数据进行排序. 在高维情况下, 文章定义了“最小向量”并通过比较每个样本与“最小向量”之间的欧式距离对高维数据进行排序赋秩. 对高维两样本位置参数是否相同的假设检验问题, 根据所提出的数据排序方法将一维 Wilcoxon-Mann-Whitney 秩和检验推广到高维情形. 与其他位置检验方法相比, 方法具有不依赖于分布, 平移, 缩放不变的特性, 同时能良好应用于高维小样本的位置检验问题. 理论上, 文章证明该检验统计量在原假设下的渐近分布是正态分布及在一些弱相依条件下证明了该检验具有良好的功效. 数值分析结果表明, 对于不同总体分布的位置参数是否相等的检验问题, 所提出的检验具有更好的有效性.

— 创新方法  
— 解决的问题  
— 优点  
— 具体工作

**关键词** 高维两样本, 位置参数, Wilcoxon-Mann-Whitney 秩和检验.

MR(2000) 主题分类号 62G10, 62G20

## Rank-Sum Tests for Two-Sample Location Parameter with High Dimensional Data

QIU Tao<sup>1</sup> WU Wenqi<sup>1</sup> XU Wangli<sup>1,2</sup>

(1. School of Statistics, Renmin University of China, Beijing 100872;  
2. Center for Applied Statistics, Renmin University of China, Beijing 100872)

**Abstract** It is simple to sort the sample for univariate data. In this paper, we define the “minimum vector” and rank the high-dimensional data by comparing the Euclidean distance between each sample and the “minimum vector”. For checking the adequacy whether the two-sample location parameter with high dimension is equal, the one-dimensional Wilcoxon Mann Whitney rank sum test is extended to the high-dimensional case according to the proposed data sorting method. Compared with other existing methods, the proposed test is distribution free and can be applied to

\* 国家自然科学基金项目 (11971478), 教育部人文社会科学重点研究基地重大项目 (16JJD910002), “中央高校建设世界一流大学 (学科) 和特色发展引导专项资金” (KYGJD2019003), 北京市自然科学基金 (Z190002), 中国人民大学公共健康与疾病预防控制交叉学科重大创新平台建设成果资助课题.

收稿日期: 2020-01-07, 收到修改稿日期: 2020-03-30.

通信作者: 许王莉, Email: wxu@ruc.edu.cn.

编委: 朱利平.

high-dimension, low-sample-size situations. What's more, the new approach has the property of translation and scale invariance. In theory, we prove that the proposed statistic is normal distribution asymptotically under null hypothesis. The behavior of the test under alternative hypothesis is also investigated under some weakly conditions. The numerical analysis show that the proposed test is superior to the compared statistics in almost all scenarios.

**Keywords** High dimensional two-sample, location parameter, Wilcoxon-Mann-Whitney rank-sum test.

## 1 引言

假设  $p$  维随机向量  $\boldsymbol{x} = (X_1, X_2, \dots, X_p)^T$  和  $\boldsymbol{y} = (Y_1, Y_2, \dots, Y_p)^T$  分别服从  $F$  与  $G$  分布, 其中  $F$  与  $G$  是未知的连续函数, 且满足  $F(\boldsymbol{x}) = G(\boldsymbol{x} - \boldsymbol{u})$ . 本文感兴趣的问题是在总体  $\boldsymbol{x}$  和  $\boldsymbol{y}$  尺度参数相同的情况下, 检验它们的位置参数是否有差异, 即检验

$$H_0: \boldsymbol{u} = \mathbf{0}_{p \times 1}, \quad H_1: \boldsymbol{u} \neq \mathbf{0}_{p \times 1}, \quad (1.1)$$

其中  $\boldsymbol{u} = (u_1, u_2, \dots, u_p)^T$ ,  $\mathbf{0}_{p \times 1}$  是  $p$  维零向量. 两总体位置参数检验的问题是统计学中最基本的研究问题之一 (Lehmann 和 Romano<sup>[1]</sup>), 并在生物信息学、地质成像、信号处理、天体测量、金融等领域有着广泛的应. 例如, 在现代遗传学研究中, Barry 等<sup>[2]</sup> 通过比较来自相同组织的微阵列数据的均值, 以此来识别两组患者之间基因表达的差异.

当  $p = 1$  时, 若  $\boldsymbol{x}$  和  $\boldsymbol{y}$  服从正态分布, 上述检验转化为方差相等的情况下, 均值是否相等的检验, 经典的方法有  $U$  检验和  $t$  检验. 在总体正态性假设不满足或者总体分布未知时, 众多非参数方法已经被提出来, Wilcoxon<sup>[3]</sup> 提出两样本秩和检验; Mann 和 Whitney<sup>[4]</sup> 将该结果推广到两组样本容量不等的情况; Brown 和 Mood<sup>[5]</sup> 利用中位数位置的检验来判别两总体位置参数的差异; Van der Warden<sup>[6]</sup> 提出了正态得分检验, 通过将样本的秩换成正态分位点, 以此降低异常值和尾部数据对总体位置的影响.

当  $p \geq 2$  时, Wald 和 Wolfowitz<sup>[7]</sup>, Chatterjee 和 Sen<sup>[8]</sup>, Puri 和 Sen<sup>[10]</sup>, Möttönen 和 Oja<sup>[11]</sup> 以及 Choi 和 Marden<sup>[12]</sup> 定义了样本在每个维度的符号或秩, 通过使用分量的符号或秩将单变量秩过程扩展到多元问题, Chakraborty 和 Chaudhuri<sup>[13, 14]</sup> 基于空间符号与空间秩提出了多元与高维两样本 Mann-Whitney-Wilcoxon 秩和检验. 但这些过程不是仿射不变的, 同时 Bickel<sup>[15]</sup> 指出, 如果变量高度相关, 这些检验的功效将大幅度下降. Brown 和 Hettmansperger<sup>[16]</sup>, Hettmansperger 和 Oja<sup>[17]</sup> 和 Hettmansperger 等<sup>[18]</sup> 基于  $L_1$  准则对符号及秩检验进行推广, 提出了多元两样本或多样本秩检验与符号检验, Randles<sup>[19]</sup> 提出了基于交叉点的多元两样本符号检验. 这些检验满足仿射不变性, 但这些检验统计量并不是严格意义上与总体分布无关, 它们至少是有条件的, 或者是渐近分布无关的. Liu 和 Singh<sup>[20]</sup> 根据数据深度定义了衡量两个分布偏离程度的指标, 并以此指标对 Wilcoxon 秩和检验做高维推广, 提出精确的高维秩和检验. Rosenbaum<sup>[21]</sup> 提出基于主成分分析的位置检验和基于数据深度的尺度检验. 尽管这两种方法具有仿射不变性, 同时又有不依赖于总体分布的特性, 但这两种方法都需要大量的样本来支撑, 在小样本情形它们的检验效果不佳. 另一方面, Friedman 和 Rafsky<sup>[22]</sup> 基于最小生成树将 Wald-Wolfowitz 游程检验、Kolmogorov-Smirnov 最

开门见山: 具体问题.

5/41  $\approx$  12%

研究意义: 应用广并举例

3/41  $\approx$  7% 说明.

既有的研究方法

(优缺点, 埋下伏笔, 暗示自己的方法继承发展)

24/41  $\approx$  59%

大偏差检验推广到高维情形; Mondal 等<sup>[23]</sup> 提出了最邻近的高维两样本同分布检验; Biswas 等<sup>[24]</sup> 利用最短哈密顿路径将游程检验进行高维推广. 这三类检验都是基于图方法且具有不依赖于总体分布的特性. 但是, 以上三种检验的功效受到诸如近邻数或权重函数等调整参数的影响, 在高维情形很难去选择最优的调整参数以获得最佳的检验功效. 本文提出了一种新的高维数据排序方法, 定义为“最小向量”的排序方法, 并基于此排序方法对样本排序赋秩, 将 Wilcoxon-Mann-Whitney 秩和检验推广到高维情形. 基于“最小向量”的排序方法是通过比较每个观测点与“最小向量”点之间的欧式距离进行排序, 因此对数据整体进行一致的平移、放缩都不会改变排列的次序, 保证了该过程具有平移、放缩不变性. 同时, 基于 Wilcoxon-Mann-Whitney 秩和检验的特性, 所提出的检验方法与总体分布无关, 而且当样本量与维数都趋向无穷时无论二者数量关系如何, 检验统计量在原假设下的渐近分布都是标准正态. 因此, 该方法具有很好的统计性质, 分析结果表明该统计量适用于样本量远远小于维数的情形.

20%

自己提出的方法:

① 怎么做

② 优点

③ 模拟效果

9/41 ≈ 22%

## 2 基于“最小向量”排序的检验统计量

### 2.1 基于“最小向量”排序赋秩的定义

假设  $\{x_i = (X_{i1}, X_{i2}, \dots, X_{ip})^T, i = 1, 2, \dots, n_1\}$  和  $\{y_j = (Y_{j1}, Y_{j2}, \dots, Y_{jp})^T, j = 1, 2, \dots, n_2\}$  分别取自总体  $x$  和  $y$  的  $p$  维独立样本. 首先定义“最小向量”  $z_0$ ,

$$z_0 = (Z_{01}, Z_{02}, \dots, Z_{0p})^T, \tag{2.1}$$

其中  $Z_{0r} = \min\{X_{1r}, X_{2r}, \dots, X_{n_1r}, Y_{1r}, Y_{2r}, \dots, Y_{n_2r}\}, r = 1, 2, \dots, p$ . “最小向量”排序赋秩方法是通过比较每个样本与“最小向量”  $z_0$  之间的欧式距离进行小到大的排序, 进而赋予每个样本对应的秩. 在这里选择“最小向量”  $z_0$  作为参照点是基于“最小向量”排序方法是将一维数据按照其数值大小进行排序方法的高维推广, 使得来自不同分布的样本整体上位于以  $z_0$  为“原点”的笛卡尔坐标系的同一侧, 从而减少来自不同位置参数的观测点与参照点的欧式距离相等的情况, 进而保证更多来自不同分布的观测点所赋予的秩有差异.

### 2.2 检验统计量

定义

$$z_i = x_i, \quad i = 1, 2, \dots, n_1, \quad z_{j+n_1} = y_j, \quad j = 1, 2, \dots, n_2. \tag{2.2}$$

首先, 计算每个样本  $\{z_i, i = 1, 2, \dots, n_1 + n_2\}$  同参照点  $z_0$  的欧氏距离, 记作  $\|z_i - z_0\|$ . 然后, 按照每个样本对应的欧氏距离由小到大进行排序, 得到每一个样本  $z_i$  的秩, 记为  $\{R_i, i = 1, 2, \dots, n_1 + n_2\}$ . 接下来, 计算所有来自  $F$  分布样本点的秩和, 即  $T = W_x = \sum_{i=1}^{n_1} R_i$ , 同时通过将  $T$  标准化来构造检验统计量  $\tilde{T} = \{n_1 n_2 (n_1 + n_2 + 1) / 12\}^{-\frac{1}{2}} \{T - n_1 (n_1 + n_2 + 1) / 2\}$ . 直观上看, 如果两总体分布存在位置参数的差异, 那么检验统计量  $T$  的值出现偏大或者偏小情形. 反之, 如果原假设成立, 即两总体分布不存在位置参数差异, 那么检验统计量  $T$  的值会处于中间大小. 因此, 当  $T$  的值出现过于大或者过于小时, 认为数据倾向于拒绝原假设.

### 3 检验统计量的渐近性质

本节将探讨检验统计量  $\tilde{T}$  在原假设成立的渐近分布, 以及该检验的功效. 在此之前, 首先给出推导原假设下渐近分布的条件.

**条件 1** 当  $n_1, n_2 \rightarrow \infty$ , 有  $n_1/(n_1 + n_2) \rightarrow \lambda, 0 < \lambda < 1$ .

条件 1 是两样本检验的常用条件, 表示两个总体的样本数量比较均衡. 在原假设成立的条件下, 如下定理 1 给出了检验统计量的渐近性质.

**定理 1** 若条件 1 成立, 则当  $n_1 \rightarrow \infty, n_2 \rightarrow \infty$  且原假设  $H_0$  成立的条件下, 有  $\tilde{T}$  依分布收敛到标准正态分布.

证 若原假设  $H_0$  成立, 则  $\mathbf{x}$  与  $\mathbf{y}$  独立同分布, 由此可知样本  $\{\mathbf{x}_i = (X_{i1}, X_{i2}, \dots, X_{ip})^T, i = 1, 2, \dots, n_1\}$  和  $\{\mathbf{y}_j = (Y_{j1}, Y_{j2}, \dots, Y_{jp})^T, j = 1, 2, \dots, n_2\}$  所对应的秩来自同一分布, 且满足

$$\begin{aligned} P(R_i = k) &= 1/(n_1 + n_2), \quad k = 1, 2, \dots, n_1 + n_2, \\ P(R_i = k, R_j = l) &= 1/\{(n_1 + n_2)(n_1 + n_2 - 1)\}, \quad 1 \leq k \neq l \leq n_1 + n_2. \end{aligned} \quad (3.1)$$

由此直接计算得到,

$$\begin{aligned} E(T) &= n_1(n_1 + n_2 + 1)/2, \quad \text{var}(T) = n_1 n_2 (n_1 + n_2 + 1)/12, \\ \text{cov}(R_i, R_j) &= -(n_1 + n_2 + 1)/12, \quad 1 \leq i \neq j \leq n_1 + n_2. \end{aligned} \quad (3.2)$$

结合 (3.1)-(3.2) 式, 类似于 Mann 和 Whitney<sup>[4]</sup> 第 4 节对秩和检验统计量依分布收敛到标准正态分布的证明, 可以得到定理 1 成立, 在这里我们省略定理 1 的详细证明. 值得注意的是, 对于定理 1, 对样本量与样本维数没有做任何的数量关系的假设, 换一句话说, 在这里维数  $p$  可以是固定的, 也可以是发散的. 这说明所提出的检验统计量  $\tilde{T}$  适用范围较为广泛且具备较强适用性.

接下来将在高维数据情形, 即维度  $p \rightarrow \infty$  时, 讨论该检验功效的渐近性质. 在分析检验功效之前, 首先给出一些定义, 记号及一些适当的条件. 记  $X_i$  和  $Y_i$  分别为  $p$  维随机向量  $\mathbf{x}$ 、 $\mathbf{y}$  的第  $i$  维分量. 令

$$\begin{aligned} E\mathbf{x} &= \mathbf{u}_x, \quad E\mathbf{y} = \mathbf{u}_y, \quad E\mathbf{z}_0 = \mathbf{u}_{z_0}, \\ \Sigma_x &= \text{cov}(\mathbf{x}), \quad \Sigma_y = \text{cov}(\mathbf{y}), \quad \Sigma_{z_0} = \text{cov}(\mathbf{z}_0). \end{aligned} \quad (3.3)$$

**定义 1** 称随机向量  $\mathbf{x} = (X_1, X_2, \dots, X_p)^T$  为  $p$  维  $m$ -相依随机向量, 若对给定的正整数  $m$ , 及任意的  $0 < i, j \leq p$ , 如果  $|i - j| > m$ , 则  $X_i$  与  $X_j$  相互独立.

**条件 2**  $1 \leq i \leq p$  时,  $X_i$  和  $Y_i$  的四阶矩一致有界.

**条件 3**  $\mathbf{x}$  与  $\mathbf{y}$  是  $p$  维  $m$ -相依随机向量.

**条件 4**  $\text{tr}(\Sigma_x)/p \rightarrow \sigma_x^2, \text{tr}(\Sigma_y)/p \rightarrow \sigma_y^2, \|\mathbf{u}_x - \mathbf{u}_y\|^2/p \rightarrow \nu_{xy}^2$ , 其中  $\text{tr}(\cdot)$  表示矩阵的迹.

**条件 5**  $\{\|\mathbf{u}\|^2 + 2(\mathbf{u}_y - \mathbf{u}_{z_0})^T \mathbf{u}\}/p \rightarrow c$ , 其中  $\mathbf{u}$  为 1.1 式定义,  $c$  为不等于 0 的常数.

条件 2-4 是类似于 Biswas 等<sup>[24]</sup> 给出两样本检验的常用条件. 条件 2 强调的是  $\mathbf{x}$  与  $\mathbf{y}$  每一个维度分量四阶矩有界; 条件 3 说明随机向量  $\mathbf{x}$  与  $\mathbf{y}$  各自分量之间是弱相关的; 条件 4 要求的是随机向量  $\mathbf{x}$  与  $\mathbf{y}$  协方差的迹与  $p$  同阶, 且条件 2-3 满足, 可以得证条件 4 成立. 此外, 条件 2-3 成立能够保证序列  $\{(X_i, Y_i), i = 1, 2, \dots, p\}$  具有弱收敛性质, 即当  $p \rightarrow \infty$  时,

$$\|\mathbf{x} - \mathbf{y}\|^2/p \xrightarrow{\text{Pr}} \sigma_x^2 + \sigma_y^2 + \nu_{xy}^2 + \rho_{xy}, \quad (3.4)$$

其中,  $\rho_{xy} = \lim_{p \rightarrow \infty} E\{(x - \mathbf{u}_x)^\top (y - \mathbf{u}_y)/p\}$ ,  $\text{pr}$  表示依概率收敛. 此外, 在 (3.4) 式中, 由于  $x$  与  $y$  独立, 故  $\rho_{xy} = 0$ .

条件 5 主要是为了保证 (3.7) 式中  $\nu_{xz_0}^2 \neq \nu_{yz_0}^2$ . 根据  $\mathbf{u}_{z_0}$  的定义可知,  $\mathbf{u}_y - \mathbf{u}_{z_0} \geq 0$ . 如果  $\mathbf{u} \geq 0$  或  $\mathbf{u} \leq 0$ , 且  $\sum_{i=1}^p u_i = O(p)$ , 条件 5 成立. 由此可知, 条件 5 也是一个比较容易满足的条件. 记  $Z_{1-\frac{\alpha}{2}}$  为标准正态分布  $(1 - \frac{\alpha}{2}) \times 100\%$  的分位数, 在备择假设成立的条件下, 如下定理 2 给出了检验功效的渐近性质.

**定理 2** 若条件 1-5 满足, 备择假设  $H_1$  成立, 则当  $n_1 \rightarrow \infty, n_2 \rightarrow \infty, p \rightarrow \infty$ , 有  $|\tilde{T}| \xrightarrow{\text{pr}} \infty$  及  $P(|\tilde{T}| > Z_{1-\alpha/2}) \rightarrow 1$ .

证 回顾 (2.1) 式对  $z_0$  的定义, 若  $x$  与  $y$  是  $p$  维  $m$ -相依随机向量, 则  $z_0$  也是  $p$  维  $m$ -相依随机向量. 由此可知, 若条件 2-3 成立且  $p \rightarrow \infty$  时, 有

$$\begin{aligned} \|\mathbf{u}_x - \mathbf{u}_{z_0}\|^2/p &\rightarrow \nu_{xz_0}^2, \quad \|\mathbf{u}_y - \mathbf{u}_{z_0}\|^2/p \rightarrow \nu_{yz_0}^2, \quad \text{tr}(\boldsymbol{\Sigma}_{z_0})/p \rightarrow \sigma_{z_0}^2, \\ E\{(\mathbf{x} - \mathbf{u}_x)^\top (z_0 - \mathbf{u}_{z_0})/p\} &\rightarrow \rho_{xz_0}, \quad E\{(\mathbf{y} - \mathbf{u}_y)^\top (z_0 - \mathbf{u}_{z_0})/p\} \rightarrow \rho_{yz_0}. \end{aligned} \quad (3.5)$$

类似 (3.4) 式的推导, 可知当  $p \rightarrow \infty$  时, 有

$$\begin{aligned} \|\mathbf{x} - z_0\|^2/p &\xrightarrow{\text{pr}} \sigma_x^2 + \sigma_{z_0}^2 + \nu_{xz_0}^2 + \rho_{xz_0}, \\ \|\mathbf{y} - z_0\|^2/p &\xrightarrow{\text{pr}} \sigma_y^2 + \sigma_{z_0}^2 + \nu_{yz_0}^2 + \rho_{yz_0}. \end{aligned} \quad (3.6)$$

备择假设  $H_1$  成立, 即  $\mathbf{u}_x - \mathbf{u}_y = \mathbf{u} \neq 0$  且  $\sigma_x^2 = \sigma_y^2$ , 则  $\rho_{xz_0} = \rho_{yz_0}$ . 结合 (3.5)-(3.6) 式可知

$$\|\mathbf{x} - z_0\|^2/p - \|\mathbf{y} - z_0\|^2/p \xrightarrow{\text{pr}} \nu_{xz_0}^2 - \nu_{yz_0}^2. \quad (3.7)$$

另一方面, 由于  $\mathbf{u}_x = \mathbf{u}_y + \mathbf{u}$ , 直接计算可得

$$\begin{aligned} \nu_{xz_0}^2 - \nu_{yz_0}^2 &= \lim_{p \rightarrow \infty} (\|\mathbf{u}_x - \mathbf{u}_{z_0}\|^2/p - \|\mathbf{u}_y - \mathbf{u}_{z_0}\|^2/p) \\ &= \lim_{p \rightarrow \infty} \{(\mathbf{u}_x - \mathbf{u}_y)^\top (\mathbf{u}_x + \mathbf{u}_y - 2\mathbf{u}_{z_0})/p\} \\ &= \lim_{p \rightarrow \infty} [\|\mathbf{u}\|^2 + 2(\mathbf{u}_y - \mathbf{u}_{z_0})^\top \mathbf{u}]/p. \end{aligned} \quad (3.8)$$

由条件 5 可知  $\nu_{xz_0}^2 - \nu_{yz_0}^2$  不等于 0, 由此在概率测度下保证来自总体  $x$  的样本与  $z_0$  的欧式距离与来自总体  $y$  的样本与  $z_0$  的欧式距离不同, 进而保证了来自不同分布的观测点所赋予的秩整体有差异. 在这里不妨假设  $\nu_{xz_0}^2 > \nu_{yz_0}^2$ , 则来自总体  $x$  的样本与  $z_0$  的欧式距离比来自总体  $y$  的样本与  $z_0$  的欧式距离都要大. 由此可知, 当  $p \rightarrow \infty$  时, 来自总体  $x$  的样本所对应的秩的取值范围为  $n_2 + 1 \leq R_i \leq n_1 + n_2$ . 所以, 当  $p \rightarrow \infty$  时,

$$T/\{n_1(n_1 + 2n_2 + 1)/2\} \xrightarrow{\text{pr}} 1. \quad (3.9)$$

若  $\nu_{xz_0}^2 < \nu_{yz_0}^2$ , 则来自总体  $x$  的样本与  $z_0$  的欧式距离与来自总体  $y$  的样本与  $z_0$  的欧式距离都要小. 类似 (3.9) 式的证明, 当  $p \rightarrow \infty$  时, 有

$$T/\{n_1(n_1 + 1)/2\} \xrightarrow{\text{pr}} 1. \quad (3.10)$$

结合 (3.9)-(3.10) 式, 可知当  $p \rightarrow \infty$  时, 有

$$|\tilde{T}|/\{3n_1n_2/(n_1 + n_2 + 1)\}^{\frac{1}{2}} \xrightarrow{\text{pr}} 1. \quad (3.11)$$

所以, 当  $n_1 \rightarrow \infty, n_2 \rightarrow \infty, p \rightarrow \infty$ , 有

$$|\tilde{T}| \xrightarrow{\text{pr}} \infty. \quad (3.12)$$

则, 由 (3.11) 式可知该检验功效为,

$$\lim_{p, n_1, n_2 \rightarrow \infty} P\left(|\tilde{T}| > Z_{1-\alpha/2}\right) = \lim_{p, n_1, n_2 \rightarrow \infty} P\left(\left\{3n_1n_2/(n_1+n_2+1)\right\}^{\frac{1}{2}} > Z_{1-\alpha/2}\right). \quad (3.13)$$

所以,  $\lim_{p, n_1, n_2 \rightarrow \infty} P\left(|\tilde{T}| > Z_{1-\alpha/2}\right) = 1$ , 即当样本量与样本维度同时趋向无穷时该检验功效趋向于 1. 此即得证定理 2 成立. 由此可知, 若  $\left\{3n_1n_2/(n_1+n_2+1)\right\}^{\frac{1}{2}} > Z_{1-\alpha/2}$ , 则拒绝原假设. 特别地, 取显著性水平  $\alpha = 0.05$ , 则只需要  $\min\{n_1, n_2\} \geq 3$  就保证在备择假设下  $\left\{3n_1n_2/(n_1+n_2+1)\right\}^{\frac{1}{2}} > Z_{0.975}$  成立. 另一方面, 定理 1 可知当样本量趋向无穷时, 检验统计量渐近服从标准正态分布. 模拟显示, 只需要  $\min\{n_1, n_2\} \geq 5$ , 既满足了该检验犯第一类错误的概率接近 0.05, 同时又保证了在备择假设下  $\left\{3n_1n_2/(n_1+n_2+1)\right\}^{\frac{1}{2}} > Z_{0.975}$  成立. 由此可知该检验方法对数据量的要求小, 对维度高, 样本稀缺的问题也适用.

#### 4 数值模拟

上一节主要探讨了检验统计量  $\tilde{T}$  的渐近性质, 本节将考察统计量  $\tilde{T}$  的有限样本性质并在有限样本情形下比较该统计量与其他已有统计量的优劣, 已有统计量包括 Puri 和 Sen<sup>[10]</sup>, Möttönen 和 Oja<sup>[11]</sup>, Chakraborty 和 Chaudhuri<sup>[14]</sup>, Rosenbaum<sup>[21]</sup> 和 Biswas 等<sup>[24]</sup>. 为了方便书写, 记所提出的检验方法为 QWX, 上述其他 5 种检验方法分别记为 PS, MO, CC, R 及 BMG. 考虑如下 4 种模型: 数据来自重尾, 轻尾, 弱相关及不对称分布模型.

**例 1** 设  $\{x_i, i = 1, 2, \dots, n_1\}$  和  $\{y_i, i = 1, 2, \dots, n_2\}$  分别取自  $p$  维自由度为 3 的  $t$  分布  $t_3(\mathbf{0}_{p \times 1}, \mathbf{I}_p)$  和  $t_3(\delta \mathbf{1}_{p \times 1}, \mathbf{I}_p)$ .  $\mathbf{0}_{p \times 1}$  和  $\mathbf{1}_{p \times 1}$  分别表示  $x$  和  $y$  的位置向量,  $\mathbf{I}_p$  是  $p \times p$  单位矩阵表示  $x$  和  $y$  的尺度参数矩阵.

**例 2** 设  $\{x_i, i = 1, 2, \dots, n_1\}$  和  $\{y_i, i = 1, 2, \dots, n_2\}$  分别取自  $p$  维正态分布  $N(\mathbf{0}_{p \times 1}, \mathbf{I}_p)$  和  $N(\delta \mathbf{1}_{p \times 1}, \mathbf{I}_p)$ .  $\mathbf{0}_{p \times 1}$  和  $\mathbf{1}_{p \times 1}$  分别表示  $x$  和  $y$  的位置向量,  $\mathbf{I}_p$  是  $p \times p$  单位矩阵表示  $x$  和  $y$  的尺度参数矩阵.

**例 3** 设  $\{x_i, i = 1, 2, \dots, n_1\}$  和  $\{y_i, i = 1, 2, \dots, n_2\}$  分别取自  $p$  维正态分布  $N(\mathbf{0}_{p \times 1}, \Sigma)$  和  $N(\delta \mathbf{1}_{p \times 1}, \Sigma)$ , 其中  $\Sigma = (0.5^{|i-j|})_{p \times p}$  是  $p \times p$  矩阵.

**例 4** 设  $X_{ik}, Y_{ik}, k = 1, 2, \dots, p$  分别独立取自  $Laplace_p(0, 0.5)$  和  $Laplace_p(\delta, 0.5)$  分布, 其中  $Laplace_p(\mu, b)$  表示拉普拉斯分布函数,  $\mu$  是位置参数,  $b$  是尺度参数.

在这 4 组模拟实验中, 考虑两总体位置参数  $\delta = 0, 0.15, 0.2$ ; 样本量分别取为  $n_1 = n_2 = 5, 10, 30, 90$ ; 维度  $p = 10, 30, 90, 300, 900, 2000$ . 在这里设定显著性水平  $\alpha = 0.05$ , 实验重复次数为 1000. 由于 PS 和 MO 检验不适用于计算样本量小于样本维度的情况, 因此只给出样本量大于样本维度情况下的结果, 样本量小于样本维度情况下的结果记为“-”.

例 1 和例 2 考虑的是样本数据来自重尾分布及轻尾分布情形. 表 1 展示了 CC 检验在  $n_1 = 5$  时不能很好地控制第一类错误, 同时表 1 和表 2 也展示了 PS 和 MO 检验在  $p = n_1$  时检验水平接近 0, 其他所有的检验水平都比较接近 0.05. 这主要是因为 CC 检验在重尾情形及 PS 和 MO 检验在重尾与轻尾情形对样本量要求相对较大. 表 1 和表 2 还显示所有检验的检验功效都随着样本量及样本维度的增加而更有效, 同时这五种检验的检验功效在样本数据来自轻尾分布比来自重尾分布整体要好. 具体来说, 表 1 展示了当  $n_1 = 5, 10$  时 CC 检验表现最好紧跟着 QWX, R, BMG 检验, PS 与 MO 几乎失去功效. 当  $n_1 = 30, 90$  时, QWX 表现最好其次是 CC, R, BMG 检验. 表 2 展示了 QWX 检验在例 2 所有设置下的检验功效

都比其他五种检验要好,特别是小样本高维情形,其他4种检验的检验功效相对QWX检验要差很多,甚至有些情形失去功效.

表1 例1模拟检验水平与检验功效(%)

(Table 1 Test level and power of simulation for Example 1 (%))

$\delta$	$p$	$n_1 = n_2 = 5$						$n_1 = n_2 = 10$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	-	-	3.1	3.0	5.1	7.2	0.2	0.0	4.3	5.9	4.7	4.4
	30	-	-	4.2	4.9	5.9	12.0	-	-	4.1	5.5	5.5	6.2
	90	-	-	3.2	4.0	5.8	9.4	-	-	4.3	6.2	5.2	6.4
	300	-	-	3.4	3.9	5.0	10.8	-	-	4.8	5.2	4.5	7.6
	900	-	-	3.5	4.4	5.6	9.1	-	-	5.0	4.7	4.5	6.8
	2000	-	-	3.6	3.5	5.4	9.8	-	-	3.9	4.2	5.5	7.7
0.15	10	-	-	3.7	4.7	10.1	10.6	0.5	0.2	6.7	5.3	10.4	7.4
	30	-	-	2.6	4.2	14.1	11.4	-	-	5.4	6.7	20.6	9.3
	90	-	-	6.0	5.2	22.1	12.0	-	-	9.5	8.2	34.8	11.8
	300	-	-	6.2	3.9	23.9	14.3	-	-	10.4	9.5	45.9	20.6
	900	-	-	7.8	5.4	25.0	19.0	-	-	17.1	11.5	54.2	37.2
	2000	-	-	13.4	5.4	27.8	27.9	-	-	26.2	10.1	55.8	64.2
0.2	10	-	-	4.2	4.8	9.6	9.0	0.1	0.1	6.9	7.3	15.0	10.0
	30	-	-	5.9	5.7	17.8	10.7	-	-	9.2	6.8	31.5	11.8
	90	-	-	6.0	4.6	28.0	14.8	-	-	9.1	8.6	52.2	17.3
	300	-	-	8.5	6.5	38.0	18.9	-	-	15.7	10.8	64.8	42.1
	900	-	-	15.2	8.1	37.0	36.0	-	-	35.2	14.2	69.4	77.4
	2000	-	-	30.1	7.6	37.3	58.0	-	-	53.7	20.9	69.5	94.0
$\delta$	$p$	$n_1 = n_2 = 30$						$n_1 = n_2 = 90$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	2.9	3.8	5.4	3.7	4.8	3.9	5.9	6.8	5.6	4.2	6.1	4.6
	30	0.2	0.9	5.2	5.0	5.9	4.3	3.8	3.7	5.1	5.2	5.6	3.0
	90	-	-	4.4	4.9	4.9	6.2	1.0	0.6	4.8	3.9	4.6	5.8
	300	-	-	4.2	4.4	5.5	6.0	-	-	5.3	3.9	4.8	4.0
	900	-	-	4.7	3.8	5.4	4.8	-	-	4.4	5.6	5.5	4.8
	2000	-	-	4.9	5.0	4.6	5.8	-	-	5.7	3.6	6.3	6.0
0.15	10	7.7	6.5	8.6	8.2	23.1	10.9	32.2	24.0	12.8	10.9	58.6	31.2
	30	2.8	2.0	8.0	10.0	54.7	18.6	51.6	32.8	16.5	16.5	95.6	64.0
	90	-	-	10.8	10.9	85.3	34.6	46.5	19.0	19.5	26.7	100.0	96.5
	300	-	-	19.1	17.1	97.0	74.4	-	-	35.9	51.5	100.0	100.0
	900	-	-	34.4	24.0	98.2	99.4	-	-	71.4	84.0	100.0	100.0
	2000	-	-	58.0	37.2	99.1	99.9	-	-	95.6	96.2	100.0	100.0

续表 1 例 1 模拟检验水平与检验功效 (%)

(Table 1 Test level and power of simulation for Example 1 (%)(Continued))

$\delta$	$p$	$n_1 = n_2 = 30$						$n_1 = n_2 = 90$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0.2	10	13.8	13.3	9.7	7.6	41.9	17.6	57.3	42.3	19.2	16.3	86.4	57.7
	30	4.0	2.4	13.3	15.1	77.7	36.5	86.3	62.4	27.9	30.9	100.0	93.0
	90	-	-	18.9	21.4	96.6	68.5	90.4	52.8	39.2	57.9	100.0	99.9
	300	-	-	34.7	34.1	99.7	98.9	-	-	70.7	91.9	100.0	100.0
	900	-	-	69.7	54.1	99.9	100.0	-	-	99.2	99.9	100.0	100.0
	2000	-	-	95.6	73.4	99.9	100.0	-	-	100.0	100.0	100.0	100.0

例 3 考虑的是样本数据来自总体协方差不为单位阵, 即随机变量维度之间存在相关性的情形. 类似于表 1 和表 2 所展示的结果, 表 3 同样说明 PS 和 MO 检验在  $p = n_1$  时检验水平接近 0. 表 3 展示了 QWX 检验在例 3 所有设置下检验功效要优于其他检验, 特别是在样本量小于等于 30 的情况下优势明显. 其他检验整体表现如下,  $CC \geq R \geq BMG$ , PS 与 MO 在样本量相对维度过小时几乎失去功效. 另一方面, 比较表 2 和表 3 结果, 发现在 QWX 检验在维度  $p$  较大时几乎不受随机变量维度之间弱相关性的影响, QWX 检验在例 2 和例 3 的功效没有太大区别, 而其他四种检验对随机变量维度之间存在相关性比较敏感, 它们在例 3 中的表现整体要差于在例 2 中的表现.

表 2 例 2 模拟检验水平与检验功效 (%)

(Table 2 Test level and power of simulation for Example 2 (%))

$\delta$	$p$	$n_1 = n_2 = 5$						$n_1 = n_2 = 10$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	-	-	3.4	4.9	6.1	5.3	0.1	0.2	4.0	4.3	5.8	4.3
	30	-	-	3.2	4.3	6.6	6.2	-	-	3.7	4.3	5.8	5.1
	90	-	-	2.7	3.4	4.8	7.2	-	-	5.5	5.3	5.2	5.1
	300	-	-	2.9	4.4	5.7	7.5	-	-	4.6	4.1	4.5	5.1
	900	-	-	2.1	3.0	6.5	7.7	-	-	4.4	3.7	5.1	5.3
	2000	-	-	2.8	4.0	4.9	6.5	-	-	3.6	4.7	6.7	5.6
0.15	10	-	-	4.1	4.3	9.1	7.1	0.1	0.2	7.4	6.7	13.5	7.6
	30	-	-	4.4	6.4	17.7	7.2	-	-	7.8	7.1	37.1	7.8
	90	-	-	5.6	7.4	50.2	8.7	-	-	11.9	10.1	81.3	14.2
	300	-	-	8.6	8.1	89.7	13.1	-	-	17.4	14.7	99.6	29.5
	900	-	-	12.9	12.4	99.9	23.0	-	-	38.5	29.2	100.0	63.4
	2000	-	-	26.0	19.6	100.0	43.8	-	-	65.3	46.5	100.0	90.4
0.2	10	-	-	4.5	5.2	12.5	7.0	0.1	0.3	8.1	7.5	23.2	10.0
	30	-	-	6.3	5.6	30.4	10.1	-	-	10.9	8.9	55.1	17.5
	90	-	-	8.8	6.8	66.7	12.6	-	-	19.7	14.8	95.0	27.2
	300	-	-	16.3	14.8	99.3	25.6	-	-	39.5	28.8	100.0	61.6
	900	-	-	38.8	21.7	100.0	51.5	-	-	77.5	59.2	100.0	96.2
	2000	-	-	61.5	43.1	100.0	82.1	-	-	97.8	86.5	100.0	100.0



续表 2 例 2 模拟检验水平与检验功效 (%)

(Table 2 Test level and power of simulation for Example 2 (%) (Continued))

$\delta$	$p$	$n_1 = n_2 = 30$						$n_1 = n_2 = 90$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	4.1	4.4	4.5	4.4	6.0	3.7	4.5	4.5	4.3	3.1	5.2	3.4
	30	0.6	0.3	5.5	4.3	5.5	5.6	3.8	3.4	4.1	4.7	4.9	3.1
	90	-	-	4.8	5.1	6.2	5.7	0.5	0.7	5.4	4.3	4.6	6.9
	300	-	-	4.4	3.7	4.7	4.6	-	-	4.6	3.6	4.7	5.7
	900	-	-	5.5	3.4	6.1	5.6	-	-	6.7	4.4	4.2	4.62
	2000	-	-	5.4	4.3	6.0	4.3	-	-	6.0	5.0	5.1	5.3
0.15	10	11.4	11.7	10.1	7.7	36.9	15.7	48.3	50.9	22.8	11.5	85.1	49.5
	30	4.2	3.7	16.8	9.3	81.7	27.9	76.8	79.0	34.2	20.8	99.9	89.4
	90	-	-	25.6	16.1	99.9	55.4	81.4	85.7	58.0	44.3	100.0	99.9
	300	-	-	51.0	37.7	100.0	95.5	-	-	94.3	83.6	100.0	100.0
	900	-	-	90.8	71.3	100.0	100.0	-	-	100.0	99.8	100.0	100.0
	2000	-	-	99.9	96.1	100.0	100.0	-	-	100.0	100.0	100.0	100.0
0.2	10	22.3	23.0	16.6	9.1	59.6	28.9	79.6	81.6	41.0	18.6	97.0	82.5
	30	8.9	10.2	30.7	17.6	95.7	56.0	98.2	98.8	62.9	41.4	100.0	99.6
	90	-	-	49.4	34.4	100.0	92.4	99.7	99.9	92.8	82.5	100.0	99.9
	300	-	-	89.8	74.8	100.0	100.0	-	-	100.0	99.9	100.0	100.0
	900	-	-	100.0	99.0	100.0	100.0	-	-	100.0	100.0	100.0	100.0
	2000	-	-	100.0	100.0	100.0	100.0	-	-	100.0	100.0	100.0	100.0

表 3 例 3 模拟检验水平与检验功效 (%)

(Table 3 Test level and power of simulation for Example 3 (%))

$\delta$	$p$	$n_1 = n_2 = 5$						$n_1 = n_2 = 10$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	-	-	3.7	2.9	6.6	5.9	0.2	0.2	3.1	4.4	4.8	4.7
	30	-	-	3.0	3.8	5.0	6.4	-	-	4.4	5.3	4.8	2.5
	90	-	-	1.8	3.9	5.5	6.1	-	-	4.3	5.1	5.3	4.6
	300	-	-	3.3	4.0	5.1	6.2	-	-	2.7	4.4	3.5	5.0
	900	-	-	2.9	4.6	5.4	4.2	-	-	5.2	4.5	5.5	4.5
	2000	-	-	2.5	2.6	4.8	6.4	-	-	4.6	6.1	5.8	5.8
0.15	10	-	-	4.8	2.8	5.4	5.4	0.1	0.3	7.9	4.3	9.0	7.8
	30	-	-	4.8	5.8	10.6	7.2	-	-	10.2	5.7	16.2	7.8
	90	-	-	4.8	5.7	18.6	7.7	-	-	9.4	8.1	37.0	11.2
	300	-	-	9.4	7.5	54.6	10.8	-	-	16.2	12.0	85.4	22.5
	900	-	-	12.9	10.0	91.4	20.6	-	-	29.8	21.0	99.9	42.9
	2000	-	-	17.9	14.1	99.5	27.0	-	-	51.8	32.9	100.0	74.6

续表3 例3 模拟检验水平与检验功效 (%)

(Table 3 Test level and power of simulation for Example 3 (%) (Continued))

$\delta$	$p$	$n_1 = n_2 = 5$						$n_1 = n_2 = 10$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	-	-	4.9	5.7	10.3	7.2	0.2	0.3	9.4	5.5	13.5	7.4
	30	-	-	6.6	6.8	14.7	9.9	-	-	11.2	6.7	27.5	14.3
	90	-	-	7.8	6.8	34.5	11.2	-	-	21.0	13.6	61.3	20.7
	300	-	-	12.8	8.6	74.5	19.6	-	-	32.6	22.1	97.8	44.2
	900	-	-	26.9	16.1	99.1	36.5	-	-	60.1	39.5	100.0	84.2
	2000	-	-	43.2	28.8	100.0	61.0	-	-	90.5	66.1	100.0	98.4
$\delta$	$p$	$n_1 = n_2 = 30$						$n_1 = n_2 = 90$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	4.0	4.3	3.1	4.1	4.6	2.6	4.6	4.5	4.8	3.9	4.3	3.2
	30	1.2	0.3	5.2	4.0	5.1	4.5	3.3	3.5	4.8	3.9	5.1	3.7
	90	-	-	5.2	3.4	5.3	3.3	0.4	0.7	5.0	3.1	5.0	3.3
	300	-	-	4.4	5.0	4.4	3.9	-	-	5.5	3.9	4.7	4.6
	900	-	-	3.9	2.7	4.2	5.6	-	-	4.4	4.9	5.2	5.2
	2000	-	-	5.7	4.9	4.0	4.7	-	-	4.2	4.5	4.3	3.7
0.15	10	6.6	6.9	15.5	8.1	19.0	12.6	19.1	20.1	42.0	7.5	47.6	31.5
	30	1.3	1.5	22.8	8.4	39.0	22.3	23.3	22.7	58.7	11.6	85.5	62.9
	90	-	-	29.6	11.8	86.2	38.2	14.2	14.0	70.0	26.4	100.0	95.3
	300	-	-	51.3	25.2	100.0	77.8	-	-	91.5	61.1	100.0	100.0
	900	-	-	80.3	54.9	100.0	98.9	-	-	100.0	95.7	100.0	100.0
	2000	-	-	97.5	83.6	100.0	100.0	-	-	100.0	100.0	100.0	100.0
0.2	10	9.7	9.3	25.3	6.3	29.5	17.8	33.8	33.7	66.4	8.9	70.0	57.1
	30	2.7	2.4	37.9	12.3	66.2	41.6	50.3	52.8	81.5	21.8	98.2	90.2
	90	-	-	52.2	23.3	98.2	70.8	42.2	44.6	90.7	50.2	100.0	99.9
	300	-	-	83.0	50.1	100.0	98.6	-	-	99.8	94.1	100.0	100.0
	900	-	-	99.3	89.1	100.0	100.0	-	-	100.0	100.0	100.0	100.0
	2000	-	-	100.0	99.3	100.0	100.0	-	-	100.0	100.0	100.0	100.0

例4 考虑的是样本数据来自不对称分布. 类似于表3的结果, 表4表示QWX检验的检验水平接近0.05且在例4所有设置下, 检验功效都要优于其他检验, 特别是在样本量相对样本维度要小的情形下.

总体而言, 通过比较4组模拟结果, 相较于其他位置检验的方法, QWX检验的优势都十分明显, 充分的展示该方法具有不依赖于分布, 平移, 缩放不变的特性, 同时能良好应用于高维小样本的位置检验问题.

表4 例4 模拟检验水平与检验功效 (%)  
(Table 4 Test level and power of simulation for Example 4 (%))

$\delta$	$p$	$n_1 = n_2 = 5$						$n_1 = n_2 = 10$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	-	-	3.2	3.7	4.8	5.9	0.1	0.0	6.0	4.5	6.1	4.0
	30	-	-	3.8	4.8	7.0	6.9	-	-	4.8	4.4	5.1	5.2
	90	-	-	3.4	4.0	5.5	5.4	-	-	4.7	4.3	5.4	5.9
	300	-	-	4.3	3.9	6.4	5.8	-	-	4.3	5.8	5.4	5.2
	900	-	-	2.8	3.8	5.6	6.0	-	-	4.8	3.3	4.9	4.6
	2000	-	-	2.2	4.4	5.2	7.7	-	-	4.7	4.7	5.1	6.4
0.15	10	-	-	4.3	4.2	11.1	6.5	0.1	0.1	6.8	5.4	15.5	8.9
	30	-	-	5.0	4.6	15.9	8.2	-	-	8.1	7.1	31.9	9.1
	90	-	-	6.9	6.0	33.5	8.9	-	-	11.5	9.7	71.2	13.2
	300	-	-	10.1	7.9	78.2	11.1	-	-	22.6	13.1	99.6	28.8
	900	-	-	20.6	11.3	99.6	24.1	-	-	47.5	24.8	100.0	60.7
	2000	-	-	41.1	16.8	100.0	38.7	-	-	75.8	41.0	100.0	91.8
0.2	10	-	-	5.4	7.3	12.7	7.0	0.4	0.0	8.8	7.1	21.6	10.8
	30	-	-	6.0	5.5	24.9	10.2	-	-	13.7	10.0	49.7	17.8
	90	-	-	9.9	6.6	55.6	14.2	-	-	24.3	13.4	90.2	29.5
	300	-	-	22.8	11.3	95.3	24.1	-	-	44.8	23.7	100.0	63.5
	900	-	-	47.8	21.3	100.0	51.2	-	-	85.8	52.2	100.0	96.7
	2000	-	-	82.7	35.9	100.0	81.7	-	-	99.5	79.5	100.0	100.0
$\delta$	$p$	$n_1 = n_2 = 30$						$n_1 = n_2 = 90$					
		PS	MO	R	BMG	QWX	CC	PS	MO	R	BMG	QWX	CC
0	10	4.0	4.2	5.8	4.7	5.1	4.2	4.4	4.2	5.2	4.2	4.5	3.3
	30	0.7	0.9	4.8	4.3	4.1	4.0	3.6	3.2	5.7	5.7	4.1	3.3
	90	-	-	4.5	4.6	6.0	4.5	0.8	1.0	4.2	3.4	4.7	4.8
	300	-	-	4.4	4.0	5.1	5.0	-	-	5.0	4.5	4.8	5.9
	900	-	-	5.2	5.0	5.3	4.9	-	-	4.6	4.2	4.6	5.2
	2000	-	-	4.3	3.9	6.0	3.9	-	-	6.2	3.6	5.8	4.4
0.15	10	18.3	13.0	13.6	7.6	37.3	16.5	71.5	54.0	27.1	13.3	85.8	54.9
	30	7.1	4.8	16.9	10.5	77.9	24.7	96.8	83.3	36.2	20.7	100.0	87.4
	90	-	-	28.7	17.8	99.7	58.3	98.6	85.3	59.8	39.5	100.0	99.8
	300	-	-	53.4	35.8	100.0	95.5	-	-	94.5	80.0	100.0	100.0
	900	-	-	93.1	70.6	100.0	100.0	-	-	100.0	99.9	100.0	100.0
	2000	-	-	100.0	92.8	100.0	100.0	-	-	100.0	100.0	100.0	100.0
0.2	10	33.5	24.6	18.4	12.1	58.2	30.5	95.4	88.1	43.1	22.7	97.9	85.9
	30	20.0	10.2	30.0	18.2	97.7	55.0	100.0	99.6	64.7	40.2	100.0	99.9
	90	-	-	53.9	34.8	100.0	92.8	100.0	99.7	91.6	80.9	100.0	100.0
	300	-	-	93.2	71.2	100.0	100.0	-	-	100.0	99.7	100.0	100.0
	900	-	-	100.0	98.2	100.0	100.0	-	-	100.0	100.0	100.0	100.0
	2000	-	-	100.0	100.0	100.0	100.0	-	-	100.0	100.0	100.0	100.0

## 5 实例分析

本文选取来自加利福尼亚大学的机器学习数据仓库 (the UC Irvine machine learning repository) 中的语法面部表情分类数据集 (grammatical facial expression)<sup>1</sup>. 本数据集包含 WH 问题 (WH-question), 肯, 否定回答问题 (yes/no question), 疑问问题 (doubt question), 主题 (topic), 否认 (negation), 断言 (assertion), 条件句 (conditional clause), 强调 (focus), 关系从句 (relative clause) 九个子数据集. 每个子数据集中数据维度都为 300 维, 由 100 个三维数据点坐标构成. 依次表示左眼, 右眼, 左眉毛, 右眉毛, 鼻子, 嘴, 脸型, 左虹膜, 右虹膜, 鼻尖, 左眉上面的线, 右眉上面的线周围取点的三维坐标, 具体如图 1 所示.

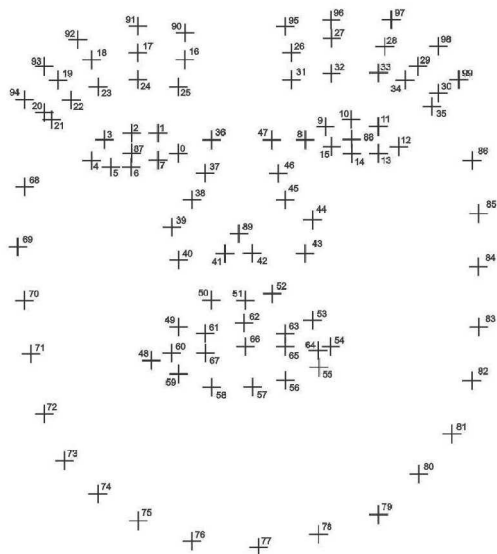


图 1 人脸面部数据采集点位置

(Figure 1 Attribute points locations on user face)

样本标签分为语法面部表情和非语法面部表情两类. 本文抽取数据集中的肯、否定回答子数据集进行实验. 为了考察不同检验方法的检验功效随样本量变化, 每次从数据集随机抽取  $n_1 = n_2 = 5, 10, 15, 20, 25, 30, 35, 40, 45, 50$  个样本. 在这里设定显著性水平  $\alpha = 0.05$ , 实验重复次数为 500. 由于本文探究的假设检验问题仅涉及位置问题的检验, 假设分布尺度相同, 因此先对数据的尺度做了标准化. 另外 PS 和 MO 检验不适用于样本量小于样本维度的情况, 因此这两个检验在这部分不予讨论.

检验结果如图 2 所示, 由折线图可以看出, 在样本量小于 20 的时候, R 检验表现最好, 其次为本文提出的 QWX 检验. 当样本量在 20 到 30 之间的时候, QWX 检验表现最佳, 其次为 CC 与 BMG 检验, 当样本量继续增大时, QWX, CC 和 BMG 检验功效均逼近 1, 表现最优, R 检验的功效则一直在 90% 上下波动.

1. <https://archive.ics.uci.edu/ml/datasets>.

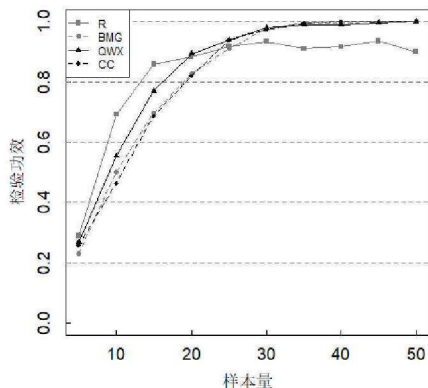


图 2 基于语法面部表情数据集的测试功效

(Figure 2 The powers performance in the analysis of grammatical facial expression dataset)

## 6 结论

研究方法+问题(+手段)

本文利用“最小向量”的高维排序方法将一维 Wilcoxon-Mann-Whitney 秩和两样本位置参数检验推广到高维情形。QWX 检验继承了 Wilcoxon-Mann-Whitney 秩和检验不依赖于整体分布的性质。另一方面, QWX 检验具有平移, 缩放不变性, 且在原假设下检验统计量的渐近分布是标准正态, 在一些弱相依条件下证明了 QWX 检验在高维小样本情形具有良好的功效。模拟实验展示 QWX 检验对于不同分布的位置参数检验表现出更好的一致性, 尤其是能更好应用于高维小样本数据场景。

——方法理论优点

——总结模拟效果

## 参 考 文 献

- [1] Lehmann E L, Romano J P. Testing Statistical Hypotheses. New York: Springer Science and Business Media, 2006.
- [2] Barry W T, Nobel A B, Wright F A. Significance analysis of functional categories in gene expression studies: A structured permutation approach. *Bioinformatics*, 2005, **21**(9): 1943–1949.
- [3] Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1945, **1**(6): 80–83.
- [4] Mann H B, Whitney D R. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 1947, **18**(1): 50–63.
- [5] Brown G, Mood A. Homogeneity of several samples. *The American Statistician*, 1948, **2**: 22.
- [6] Van der Warden B L. *Mathematische Statistik*. Berlin: Springer, 1957.
- [7] Wald A, Wolfowitz J. Statistical tests based on permutations of the observations. *The Annals of Mathematical Statistics*, 1944, **15**(4): 358–372.
- [8] Chatterjee S K, Sen P K. Non-parametric tests for the bivariate two-sample location problem. *Calcutta Statistical Association Bulletin*, 1964, **13**(12): 18–58.
- [10]

- [10] Puri M L, Sen P K. On a class of multivariate multisample rank-order tests. *The Indian Journal of Statistics, Series A*, 1966, **28**: 353–376.
- [11] Möttönen J, Oja H. Multivariate spatial sign and rank methods. *Journaltitle of Nonparametric Statistics*, 1995, **5**(2): 201–213.
- [12] Choi K, Marden J. An approach to multivariate rank tests in multivariate analysis of variance. *Journal of the American Statistical Association*, 1997, **92**(440): 1581–1590.
- [13] Chakraborty A, Chaudhuri P. A Wilcoxon-Mann-Whitney-type test for infinite-dimensional data. *Biometrika*, 2015, **101**(102): 239–246.
- [14] Chakraborty A, Chaudhuri P. Tests for high-dimensional data based on means, spatial signs and spatial ranks. *The Annals of Statistics*, 2017, **101**(45): 771–779.
- [15] Bickel P J. On some asymptotically nonparametric competitors of hotelling's  $T^2$ . *The Annals of Mathematical Statistics*, 1965, **36**(1): 160–173.
- [16] Brown B, Hettmansperger T. Affine invariant rank methods in the bivariate location model. *Journal of the Royal Statistical Society: Series B*, 1987, **49**(3): 301–310.
- [17] Hettmansperger T P, Oja H. Affine invariant multivariate multisample sign tests. *Journal of the Royal Statistical Society: Series B*, 1994, **56**(1): 235–249.
- [18] Hettmansperger T P, Möttönen J, Oja H. Affine invariant multivariate rank tests for several samples. *Statistica Sinica*, 1998, **8**(3): 785–800.
- [19] Randles R, Methods R. A two sample extension of the multivariate interdirection sign test. *L1-Statistical Analysis*, 1992, **49**: 295–302.
- [20] Liu R Y, Singh K. A quality index based on data depth and multivariate rank tests. *Journal of the American Statistical Association*, 1993, **88**(421): 252–260.
- [21] Rosenbaum P R. An exact distribution-free test comparing two multivariate distributions based on adjacency. *Journal of the Royal Statistical Society: Series B*, 2005, **67**(4): 515–530.
- [22] Friedman J H, Rafsky L C. Multivariate generalizations of the wald-wolfowitz and smirnov two-sample tests. *The Annals of Statistics*, 1979, **7**(4): 697–717.
- [23] Mondal P K, Biswas M, Ghosh A K. On high dimensional two-sample tests based on nearest neighbors. *Journal of Multivariate Analysis*, 2015, **141**: 168–178.
- [24] Biswas M, Mukhopadhyay M, Ghosh A K. A distribution-free two-sample run test applicable to high-dimensional data. *Biometrika*, 2014, **101**(4): 913–926.