arXiv:2603.04347v1 [math.ST] 4 Mar 2026

最小假设下的极端几何分位数及其与 Tukey 深度的联系

Sibsankar SINGHA
[a] *Marie KRATZ[b]
Sreekar VADLAMANI[c]

[a] T´el´巴黎电信学院，法国；电子邮箱：sibsankar.singha@telecom-paris.fr[b] 法国 ESSEC 商学院，IDO 部门与 CREAR；电子邮箱：kratz@essec.edu[c] 印度班加罗尔 TIFR-CAM；电子邮箱：sreekar@tifrbng.res.in

2026年3月5日

摘要

Chaudhury 提出的几何分位数（亦称空间分位数）是定义多元分位数的三种主要方法之一，已在文献中得到广泛研究。本文聚焦于这些分位数的极端行为。我们确立了新的极端性质，即在不要求任何矩条件的情况下，给出极端几何分位数范数的一般下界和上界。我们探讨了这些结果对刻画分布行为的影响。重要的是，该下界可直接与单变量分位数以及半空间（Tukey）深度中心区域联系起来，揭示了这两种多元分位数基本概念之间的新颖联系。

*本研究基于第一作者博士论文的部分内容，该论文是在他作为学生于 TIFR–CAM 就读期间，在另外两位作者的指导下完成的。研究期间，作者曾相互访问法国巴黎的 ESSEC 商学院和印度班加罗尔的 TIFR–CAM，受益良多。作者感谢接待机构提供的支持，并感谢 数学建模科学基金会´ (ANR-11-LABX-0023-01) 和 SERB–MATRICS 项目资助 MTR/2020/000629 提供的经费。

1 引言

高维统计学的最新进展重新激发了人们对理解数据集几何结构的工具的兴趣。人们提出了许多多元分位数和统计深度函数，用于确定多元数据的秩并识别离群值。

这些几何工具，包括深度函数和分位数，为多维空间中的数据集提供了非参数描述，因此在分类和回归等统计推断问题中尤为有用（例如，参见 [2, 5, 10, 9, 16, 19, 21]），也可用于异常值或异常检测（例如， [20]）以及几何学中的应用（例如， [14]）。 [8], 对几何分位数与中心向外的基于传输的分位数进行了近期比较讨论，并分析了这两种方法所诱导的不同等高线结构。

一个关键问题是，能否利用几何分位数刻画概率测度的尾部行为，也就是说，它们能否捕捉底层分布的本质特征。 [11] 证明，几何分位数能够唯一确定底层概率测度； [3] 随后利用这一性质构造了检验统计量。另请参见 [12]，其中介绍了从几何分位数恢复概率测度的解析反演方法。

因此，自然需要探究几何分位数的极端行为与其底层概率测度之间更清晰的联系。这也促成了本研究。关于几何分位数的总体层面分析已有若干结果，尤其涉及其渐近行为和极端行为（见 [7,17]），这些研究启发了我们的工作。

本研究确立了几何分位数的新极端性质，具体而言，在最少的模型假设下且无需任何矩条件，我们给出了其范数增长率的下界和上界。上界通过经典概率工具推导得出，而下界则采用几何方法获得。值得注意的是，下界与 Tukey 深度直接相关，因此我们能够进一步得到以单变量分位数表示的下界。这一点尤其重要，因为我们的结果不仅给出了几何分位数的一般速率，还揭示了两种不同分位数概念之间的联系。尽管这两种概念有很大差异，理解它们如何相互作用仍然很有意义。

这些一般界限对于分析具有极重尾、可能缺乏一阶或二阶矩的分布尤为重要。在标准矩假设下，可以得到更紧的结果。例如， [7] 指出，具有相同协方差矩阵的分布具有相同的一阶展开，

因此难以区分。这促使我们利用高阶项来改进尾部识别。遵循 [7], 提出的展开方法，我们扩展了他们的结果，并沿用其符号，以确保清晰明了并与他们的工作保持连贯。

最后，我们通过两个数值例子展示下界结果。所有几何分位数的计算均使用 [1, 3] 开发的算法在 Python 中实现。

记号。 除非另有说明，本文中的所有分析均在 $\mathbb{R}^d$中进行。 $\mathbb{R}^d$ 中的中心单位开球和单位球面分别记为 B^d和 S^{d-1}，而 $\langle \cdot, \cdot \rangle$ 和 $\|\cdot\|$ 分别表示 $\mathbb{R}^d$中的欧几里得内积和 ℓ^2-范数。

论文结构。 第2节回顾几何分位数和深度函数的定义及性质。第 3节给出我们的主要结果：几何分位数范数增长率的一般下界和上界，并着重阐述其与单变量分位数及 Tukey 深度的联系。第 4 节在可积性条件下进一步探讨几何和渐近性质。最后，附录A 通过数值例子展示几何分位数与半空间（Tukey）深度之间的比较，而附录B则在 [7]的基础上补充一项关于几何分位数方向的结果，其中包括一个收敛速率。

2 预备知识

定义 2.1 （几何分位数 [1]）。设 X 是一个取值于 $\mathbb{R}^d$ 的随机向量，满足$d > 1$，使得诱导测度 $\mathbb{P}$ 不支撑于任何一条直线，并设 $u \in \mathbb{S}^{d-1}$和 $0 \le \alpha < 1$。几何分位数 $q_X(\alpha u)$ 定义为以下优化问题的解：

$$q_X(\alpha u) = \arg\min_{q \in \mathbb{R}^d} \left(\mathbb{E}\big[\|X - q\| - \|X\|\big] - \langle \alpha u, q \rangle \right). \tag{2.1}$$

方程（2.1）的解的存在性和唯一性将在 [1, 第 2 节1]中讨论。在文献中， αu 被称为指标向量，它控制着分位数的中心性：当 α 接近 0 时，相应的分位数接近分布的中心（q_X（0）即中位数），因此称为中心分位数；当 α 接近 1 时，相应的分位数称为极端分位数。文献 [1] 指出，当 $d \ge 2$ 的分布 X 支撑在一条直线上时，唯一性论证便不再成立。因此，本文始终假设每个底层概率测度的支撑不包含在一条直线上。

由于方程（2.1）没有闭式解，通常需要寻找极小化点的刻画性质。特别地，在诱导测度 $\mathbb{P}$不支撑在一条直线上且为非原子的假设下，已证明对于任意 $0 < \alpha < 1$ 和 $u \in \mathbb{S}^{d-1}$，存在唯一的 $q \in \mathbb{R}^d$ 满足（2.1）当且仅当数对（αu， q） $\in \mathbb{R}^d \times \mathbb{R}^d$ 满足

$$\alpha u = -\mathbb{E}\left[\frac{X - q}{\|X - q\|}\right]. \tag{2.2}$$

这种等价关系已在经验情形中得到确立，见 [1, 定理 2.1。2], 其中充分性证明见 [7], 必要性和充分性证明见 [13, 定理 5.1，其中 $\rho(t) = t$]。这种对几何分位数的刻画通常更便于分析。

此外， [7]证明了，只要几何分位数有明确定义，

$$\|q_X(\alpha u)\| \underset{\alpha \to 1}{\to} \infty, \quad \text{and} \quad \frac{q_X(\alpha v)}{\|q_X(\alpha v)\|} \underset{(\alpha, v) \to (1, u)}{\longrightarrow} u, \tag{2.3}$$

对于 $0 < \alpha < 1$ 和 u， $v \in \mathbb{S}^{d-1}$。我们还需要几何分位数的仿射等变性这一性质（见 [1]），即对于任意正交矩阵Σ 和任意向量 $\theta \in \mathbb{R}^d$,

$$q_{\Sigma X + \theta}(\alpha u) = \Sigma\, q_X(\alpha \Sigma^\top u) + \theta. \tag{2.4}$$

我们来回顾半空间深度及其一些性质，这些性质在利用半空间深度定义极端几何分位数的下界时会用到。

定义 2.2 （半空间深度； [22]）。对于定义在 $\mathbb{R}^d$ 上的随机向量 $X \sim \mathbb{P}$，点 x 的半空间深度定义为： $HD(x, X) = \inf\{\mathbb{P}(X \in H) : H \in \mathcal{H}_x\}$，其中 $\mathcal{H}_x$ 表示 $\mathbb{R}^d$ 中包含 $x \in \mathbb{R}^d$ 的所有半空间构成的集合。

显然， HD 的取值范围为 $[0, 1]$。直观地说，深度较高的点更接近中心，而深度较低的点则相对更为极端。达到最高深度的点称为半空间中位数。如果达到最高深度的点不止一个，则将所有此类点的算术平均数定义为中位数。半空间深度函数 $x \mapsto HD(x, X)$ 满足以下性质：

- 仿射等变性：对于任意非奇异矩阵 Σ 和向量 $\beta \in \mathbb{R}^d$， $HD(\Sigma x + \beta, \Sigma X + \beta) = HD(x, X)$;
- 沿射线单调性：从中位数出发的任意射线上，半空间深度均为非增；

- 在无穷远处消失：$\lim_{t\to\infty} HD(tx, X) = 0$;
- 上半连续性：函数 $x \mapsto HD(x, X)$ 关于 $\mathbb{R}^d$ 是上半连续的。

给定 $\alpha \in (0,1)$ 的半空间深度中心区域定义为

$$\mathcal{D}(\alpha \mid X) = \{x \in \mathbb{R}^d : HD(x, X) \geq \alpha\}.$$

当上下文清楚时，我们省略对 X 的依赖，并记 $\mathcal{D}(\alpha)$ 而非 $\mathcal{D}(\alpha|X)$。在 [22, 15] 中证明了 $\mathcal{D}(\alpha)$ 可以表示为所有概率至少为 α 的半空间的交集，即

$$\mathcal{D}(\alpha) = \bigcap \{H : \mathbb{P}(X \in H) \geq \alpha\}.$$

因此， $\mathcal{D}(\alpha)$ 是一个凸集。
对于连续分布，半空间深度函数是连续的。在这种情况下，中心区域的边界可以写为

$$\partial\mathcal{D}(\alpha) = \{x \in \mathbb{R}^d : HD(x, X) = \alpha\}.$$

因此，对于任意 $x \in \mathcal{D}(\alpha)$ 以及任意包含 x 的半空间 H，我们有

$$\mathbb{P}(X \in H) \geq \alpha.$$

等价地，对于任意单位向量 $u \in \mathbb{S}^{d-1}$,

$$\mathbb{P}\big(u^\top (X - x) \geq 0\big) \geq \alpha. \tag{2.5}$$

我们将在定理的证明中用到这一性质3.3。

3 极端几何分位数范数的一般上下界

我们推导出几何分位数范数的无矩上界和下界。上界可由一个简单的论证得到，而较为精妙的下界则具有重要意义，其中包括与 Tukey 深度的联系，下一节将对此进行讨论。

定理 3.1 （上界）。设 $u \in S^{d-1}$ 且 $0 < \alpha < 1$。

1. 设 $k_\alpha > 0$ 满足 $\mathbb{P}(\|X\| > k_\alpha) < (1-\alpha)/2$。那么，我们有

$$\|q_X(\alpha u)\| \leq \frac{2\,k_\alpha\,\mathbb{P}(\|X\| \leq k_\alpha)}{1-\alpha-2\mathbb{P}(\|X\| > k_\alpha)}. \tag{3.1}$$

2. 若 $\mathbb{E}\|X\| < \infty$，则上界可改进为：

$$\|q_X(\alpha u)\| \leq \frac{2\mathbb{E}\|X\|}{1-\alpha}. \tag{3.2}$$

该定理通过去除二阶矩有限和多元正则变 variation 的假设，推广了 [6, 7] 的结果，具体如（3.1）所示。在假设一阶矩有限时，易于计算的上界（3.2）渐近更紧于（3.1）。此外，对于除一阶矩外不存在更高阶矩的正则变尾部分布，该上界在渐近意义下是锐的，因为它与 [6, 定理 1(ii)]中得到的速率一致。

定理 *3.1*的证明。
回忆 $q(\alpha u) = \arg\min f_{\alpha u}(q)$，其中

$$f_{\alpha u}(q) = \mathbb{E}\left[\|X-q\| - \|X\| - \langle \alpha u, q\rangle\right].$$

注意到，由于 $f_{\alpha u}(0) = 0$，我们必有

$$f_{\alpha u}(q(\alpha u)) \leq f(0) = 0. \tag{3.3}$$

对于任意 q 和 $k_\alpha > 0$，应用三角不等式可得以下下界，

$$\begin{aligned} f_{\alpha u}(q) &\geq \mathbb{E}\left[\{\|q\| - 2\|X\|\}\mathbb{1}_{\|X\|\leq k_\alpha}\right] + \mathbb{E}\left[-\|q\|\mathbb{1}_{\|X\|>k_\alpha}\right] - \alpha\|q\| \\ &\geq \|q\|\left\{1-\alpha-2\mathbb{P}(\|X\| > k_\alpha)\right\} - 2\,k_\alpha\mathbb{P}(\|X\| \leq k_\alpha). \end{aligned} \tag{3.4}$$

将 q 替换为 $q(\alpha u)$ 代入（3.4），并将其与（3.3）结合，我们得出

$$\|q(\alpha u)\|\left\{1-\alpha-2\mathbb{P}(\|X\| > k_\alpha)\right\} - 2\,k_\alpha\mathbb{P}(\|X\| \leq k_\alpha) \leq 0,$$

这就证明了定理的断言，只要 $\mathbb{P}(\|X\| > k_\alpha) < (1-\alpha)/2$。当 $E\|X\| < \infty$ 时，我们可以直接将 $f_{\alpha u}(q)$ 界定为：

$$f_{\alpha u}(q) \geq \|q\| - 2\mathbb{E}\|X\| - \alpha\|q\|. \tag{3.5}$$

结合（3.3）和（3.5）得到（3.2）。 □

定理3.1即使对于除一阶矩之外没有更高阶矩的分布也仍具有信息量。当假设 $\|X\|$ 的生存分布存在某种衰减时，我们可以推出 $\|q_X(\alpha u)\|$ 随 $\alpha \to 1$的衰减。即，

推论 3.2。

1. W当一阶矩存在时，上界(3.2) 给出 $O(1/(1-\alpha))$。
2. 假设对于某个 $\beta > 0$，所有充分大的 c 都满足 $\mathbb{P}(\|X\| > c) \leq A\, c^{-\beta}$。
   若$(1-\alpha)$ 取得足够小，使得$(1-\alpha)^{-1} > A$，则

$$\|q_X(\alpha u)\| = O\left((1-\alpha)^{-\left(1+\frac{1}{\beta}\right)\frac{\beta}{\beta-\epsilon}}\right), \ \forall\ 0 < \epsilon < \beta.$$

当高于一阶的矩不存在时，这是一个紧确的下界。当一阶矩存在时，下界 (3.2) 会变得更紧确，给出 $O(1/(1-\alpha))$ 速率。回顾一下，在假设存在至少二阶的矩时，精确速率为~~O~~$(1/\sqrt{(1-\alpha)})$，如 [7] 所示。我们将在备注 3.8 中回到这一点，届时将讨论多元正则变异（MRV）情形下的上界和下界。不过，值得先指出的是，除了 MRV 框架之外，这一下界在重尾情形下仍然成立（例如，允许边缘分布具有不同的尾部厚重程度），如第 1b 节的图 A 所示。

正如定理3.1所确立的，并在重尾情形（推论3.2）中明确指出的，上界通常并不精确，除非在命题3.9所讨论的特定情形中。

推论*3.2*的证明。设 $\mathbb{P}(\|X\| > c) \leq A\, c^{-\beta}$，对某个 $\beta > 0$ 以及足够大的 c。令 $(1-\alpha)/2 = Ac^{-\beta+\epsilon}$。则 $k_\alpha = (2A)^{1/(\beta-\epsilon)}(1-\alpha)^{-1/(\beta-\epsilon)}$。采用这一记号，我们有

$$\begin{aligned}
\|q(\alpha u)\| &\leq \frac{(2A)^{1/(\beta-\epsilon)}(1-\alpha)^{-1/(\beta-\epsilon)}\mathbb{P}(\|X\| \leq k_\alpha)}{\frac{(1-\alpha)}{2} - \mathbb{P}(\|X\| > k_\alpha)} \\
&\leq \frac{(2A)^{1/(\beta-\epsilon)}(1-\alpha)^{-1/(\beta-\epsilon)}}{\frac{(1-\alpha)}{2} - \frac{A^{\epsilon/(\beta-\epsilon)}(1-\alpha)^{\beta/(\beta-\epsilon)}}{2^{\beta/(\beta-\epsilon)}}} = \frac{(2A)^{1/(\beta-\epsilon)}(1-\alpha)^{-1/(\beta-\epsilon)}}{(1-\alpha)^{\beta/(\beta-\epsilon)}\left[\frac{(1-\alpha)^{-\epsilon/(\beta-\epsilon)}}{2} - \frac{A^{\epsilon/(\beta-\epsilon)}}{2^{\beta/(\beta-\epsilon)}}\right]}.
\end{aligned}$$

取足够小的 $(1-\alpha)$，使得 $(1-\alpha)^{-1} > A$，我们有

$$\|q_X(\alpha u)\| \leq c(1-\alpha)^{-\frac{\beta+1}{\beta-\epsilon}} = c(1-\alpha)^{-\left(1+\frac{1}{\beta}\right)\frac{\beta}{(\beta-\epsilon)}}, \ \forall \epsilon > 0. \qquad \square$$

我们现在讨论下界。由于分位数范数 $\|q_X(\alpha u)\|$ 发散至无穷，推导这样的下界同样是有意义的。我们通过基于几何的方法得到它。首先，我们证明极端几何分位数区域包含给定水平的半空间深度区域。然后，我们推导出用单变量分位数表示的下界。

定理 3.3. 设 X 为连续且完全支撑的随机向量（因此，其分布不集中在任何直线上）。设 $q_X(\alpha u)$ 为 $0 < \alpha < 1$ 和 $u \in \mathbb{S}^{d-1}$的几何分位数。则我们有

$$\left\{ x \in \mathbb{R}^d : HD\left(x|X\right) \geq \frac{1-\alpha^2}{M_\gamma} \right\} \subseteq \left\{ q_X(\beta u) : 0 \leq \beta \leq \alpha, u \in \mathbb{S}^{d-1} \right\}, \tag{3.6}$$

其中，参数 α 满足 $\alpha^2 \in \left(1 - \frac{M}{d+1}, \ 1\right)$ 且 $0 < M_\gamma < 1$ 对任意 $0 < \gamma < 1$定义为

$$M_\gamma = (1-\gamma) \inf_{u \in \mathbb{S}^{d-1}} \mathbb{P}(u^\top X \geq \|X\| \sqrt{1-\gamma^2}). \tag{3.7}$$

注3.4（关于几何常数 M_γ）。常数 M_γ 定义于（3.7），是连接几何分位数与 *Tukey* 深度的关键几何量。它在（3.6）给出的包含关系中，对深度水平起缩放惩罚作用。较小的 M_γ会提高深度阈值 $\left(1-\alpha^2\right)/M_\gamma$，从而缩小有保证的 *Tukey*区域并削弱界。

从几何上看，对于固定的 $\gamma \in$ (0,1)，事件 $\{u^\top X \geq \overline{\|X\|}\sqrt{1-\gamma^2}\}$ 对应于 X 落在一个以方向 u 为中心、半角为 $\theta =$反正弦(γ)的圆锥内。因此，M_γ 衡量了在所有方向 u上，此类圆锥所具有的最小角概率质量。它量化了每个开口角为 θ 的极冠中必然包含的概率质量。若分布在各个方向上均匀扩散，则 M_γ相对较大。若存在某个概率质量很小的方向（各向异性或偏度较强），则下确界会变小，使 M_γ减小。对于旋转不变分布（例如多元正态分布或多元柯西分布），角分布在球面上均匀分布。此时，M_γ 有显式形式

$$M_\gamma = \frac{1-\gamma}{2} I_{\gamma^2}\left(\frac{d-1}{2}, \frac{1}{2}\right),$$

其中，$I_x(a,b)$ 表示正则化不完全贝塔函数。
重要的是，该表达式仅取决于维数 d，而与径向分布无关。因此，高斯分布和柯西分布具有相同的 M_γ 值。因此，缩放因子 M_γ 完全取决于方向几何，而非尾部行为。

对于非对称分布，(3.7) 中的下确界在角向质量最小的方向上取得。在这种情况下 M_γ 会减小，使得该界更加保守。此外，贝塔函数表示表明，当 $d \to \infty$时，

$$I_{\gamma^2}\left(\frac{d-1}{2}, \frac{1}{2}\right) \to 0,$$

这意味着当 $d \to \infty$时， $M_\gamma \to 0$ ，体现了维数灾难的一种形式。

注释 3.5 (关于 α的可取范围。 参数 α 必须满足

$$\frac{1-\alpha^2}{M_\gamma} < \frac{1}{d+1}.$$

根据中心点定理（参见 *[18]*），这一条件保证半空间深度区域

$$\left\{ x \in \mathbb{R}^d : HD(x \mid X) \geq \frac{1-\alpha^2}{M_\gamma} \right\}$$

非空，因此在 (3.6)中给出了非平凡的包含关系。这一要求较为保守，因为它仅依赖于最大深度的普适下界 $1/(d+1)$。对于特定的分布类别，最大深度可能严格更大，从而扩大 α的可取范围。

例如，如果 X 的分布是半空间对称的，那么其最大深度等于 $1/2$。在这种情况下，非空性只要求

$$\frac{1-\alpha^2}{M_\gamma} < \frac{1}{2} \qquad \textit{or, equivalently,} \qquad 1 - \frac{M_\gamma}{2} < \alpha^2 < 1.$$

总之， α 的可取范围由几何常数M_γ 和该分布可达到的最大深度共同决定。

定理 *3.3*的证明。 我们采用基于几何的方法。证明的核心论证由以下命题给出。

命题 3.6。设 $q \in \mathbb{R}^d$ 并定义

$$U_q = \frac{X-q}{\|X-q\|},$$

其中 X 是一个取值于 $\mathbb{R}^d$ 且无原子的随机向量。设 $\bar{U}_q$ 是U_q 的一个独立副本。那么，

$$\|\mathbb{E}(U_q)\|^2 = 1 - \frac{1}{2}\,\mathbb{E}\|U_q - \bar{U}_q\|^2.$$

该命题通过得到对所有 q 均一致成立的解析界，对指标向量 $\mathbb{E}(U_q)$ 进行几何控制。

¯命题证明 3.6。 由于 $\|U_q\| = 1$ 几乎必然成立，对于一个独立副本 U_q，我们可以写出

$$\mathbb{E}\|U_q - \bar{U}_q\|^2 = 2 - 2\,\mathbb{E}\left(\langle U_q, \bar{U}_q\rangle\right) = 2 - 2\,\langle \mathbb{E}(U_q), \mathbb{E}(\bar{U}_q)\rangle = 2 - 2\,\|\mathbb{E}(U_q)\|^2,$$

由此得出该命题。 □

命题3.6 表明，任何关于 $\mathbb{E}\|U_q - \bar{U}_q\|^2$ 的下界都会给出关于 $\|\mathbb{E}(U_q)\|^2$的上界。

若 A， $B \subset S^{d-1}$ 是两个可测子集，且 $\mathbb{P}(U_q \in A) > 0, \mathbb{P}(U_q \in B) > 0$，并且$\delta = \min_{u\in A,\ v\in B} \|u - v\| > 0$，则

$$\mathbb{E}\|U_q - \bar{U}_q\|^2 \geq \mathbb{E}\big[\|U_q - \bar{U}_q\|^2 \mathbf{1}_{\{U_q\in A, \bar{U}_q\in B\}}\big] \geq \delta^2\,\mathbb{P}(U_q \in A)\,\mathbb{P}(U_q \in B),$$

因此，

$$\|\mathbb{E}(U_q)\|^2 \leq 1 - \frac{1}{2}\,\delta^2\,\mathbb{P}(U_q \in A)\,\mathbb{P}(U_q \in B) < 1. \tag{3.8}$$

我们将明确构造集合 A 和 B。

记 $d_g(u,\ v) = \arccos\langle u,\ v\rangle$为两点 u， $v \in \mathbb{S}^{d-1}$之间的大圆距离。函数 d_g 在空间 $\mathbb{S}^{d-1}$上定义了一个度量，满足 $0 \leq d_g(u,\ v) \leq \pi$，且与欧几里得距离的关系为 $\|u - v\| = 2\sin(d_g(u,\ v)/2)$。对任意 $q \neq 0$，定义

$$A_q^\gamma = \left\{ w \in \mathbb{S}^{d-1} : d_g\Big(-\frac{q}{\|q\|}, w\Big) \leq \arccos(\sqrt{1-\gamma^2}) \right\} \quad \text{with } 0 < \gamma < 1,$$

and

$$B_q = \left\{ w \in \mathbb{S}^{d-1} : d_g\Big(\frac{q}{\|q\|}, w\Big) \leq \frac{\pi}{2} \right\}.$$

显然，集合 A_q^γ 和 B_q 不相交。事实上，假设它们相交，并令 $u \in A_q^\gamma \cap B_q$。那么根据三角不等式，我们得到

$$\pi = d_g\left(-\frac{q}{|q|}, \frac{q}{|q|}\right) \leq d_g\left(-\frac{q}{|q|}, u\right) + d_g\left(u, \frac{q}{|q|}\right) < \pi,$$

这就产生了矛盾。

现在，让我们估计 δ 在（3.8）中的值，其中 $A = A_q^\gamma$ 且 $B = B_q$。对于 $u \in A_q^\gamma$ 和 $v \in B_q$，对 d_g应用三角不等式，可得

$$\begin{aligned}\pi = d_g\Big(-\frac{q}{\|q\|}, \frac{q}{\|q\|}\Big) &\leq d_g\Big(-\frac{q}{\|q\|}, u\Big) + d_g(u,v) + d_g\Big(v, \frac{q}{\|q\|}\Big) \\ &\leq \arccos(\sqrt{1-\gamma^2}) + d_g(u,v) + \frac{\pi}{2},\end{aligned}$$

这意味着 $d_g(u,\ v) \geq \frac{\pi}{2} - $ 反余弦$(\sqrt{1-\gamma^2}) =$ 反正弦$(\sqrt{1-\gamma^2})$。因此，

$$\delta = \min_{u\in A_q^\gamma, v\in B_q} \|u-v\| = \sqrt{2\Big(1-\cos(d_g(u,v))\Big)} \geq \sqrt{2(1-\gamma)}.$$

将这个下界代入 δ 不等式（3.8）中，可得

$$\|\mathbb{E}(U_q)\|^2 \leq 1 - \frac{1}{2}\,\delta^2\,\mathbb{P}(U_q \in A_q^\gamma)\,\mathbb{P}(U_q \in B_q) = 1 - (1-\gamma)\,\mathbb{P}(U_q \in A_q^\gamma)\,\mathbb{P}(U_q \in B_q). \tag{3.9}$$

接下来，我们证明存在一个常数 $M_\gamma > 0$，使得

$$0 < M_\gamma \leq (1-\gamma)\mathbb{P}(U_q \in A_q^\gamma), \quad \text{for all } q \neq 0.$$

首先，注意对于 $\frac{x}{\|x\|} \in A_q^\gamma$，

$$d_g\Big(\frac{x}{\|x\|}, -\frac{q}{\|q\|}\Big) \leq \arccos\Big(\sqrt{1-\gamma^2}\Big) \;\Rightarrow\; d_g\Big(\frac{x-q}{\|x-q\|}, -\frac{q}{\|q\|}\Big) \leq \arccos\Big(\sqrt{1-\gamma^2}\Big),$$

因此

$$\mathbb{P}(U_0 \in A_q^\gamma) \leq \mathbb{P}(U_q \in A_q^\gamma).$$

定义

$$M_\gamma = (1-\gamma) \inf_{q\in\mathbb{S}^{d-1}} \mathbb{P}(U_0 \in A_q^\gamma), \tag{3.10}$$

于是 $M_\gamma > 0$，因为 X 在 $\mathbb{R}^d$ 上完全支撑。最后，注意根据 U_q 的定义，

$$U_q \in B_q \quad \text{if and only if} \quad \Big\langle \frac{X-q}{\|X-q\|}, \frac{q}{\|q\|} \Big\rangle \geq 0,$$

这等价于 $\Big\langle X-q, \frac{q}{\|q\|} \Big\rangle \geq 0$。因此，利用（3.9）和（3.10），我们可以写成

$$\|\mathbb{E}(U_q)\|^2 < 1 - M_\gamma \cdot \mathbb{P}\left(\Big\langle X-q, \frac{q}{\|q\|}\Big\rangle \geq 0\right).$$

现在，令 $q \in \mathcal{D}(\frac{1-\alpha^2}{M_\gamma})$，其中 $\alpha^2 \in (1-\frac{M_\gamma}{d+1},\ 1)$。然后，由（2.5），可得 $\mathbb{P}\left(\Big\langle X-q, \frac{q}{\|q\|}\Big\rangle \geq 0\right) \geq \frac{1-\alpha^2}{M_\gamma}$，从而推出

$$\sup_{q\in D(\frac{1-\alpha^2}{M_\gamma}|X)} \|\mathbb{E}(U_q)\|^2 < 1 - M_\gamma \frac{1-\alpha^2}{M_\gamma} = \alpha^2.$$

换言之，

$$\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma}|X\right) \subseteq \left\{q \in \mathbb{R}^d : \|\mathbb{E}(U_q)\| \le \alpha\right\}. \tag{3.11}$$

此外，利用几何分位数的刻画（2.2），对任意 $0 < \alpha < 1$，有

$$\left\{q \in \mathbb{R}^d : \|\mathbb{E}(U_q)\| \le \alpha\right\} = \left\{q_X(\beta u) : 0 \le \beta \le \alpha,\, u \in \mathbb{S}^{d-1}\right\},$$

其中 $q_X(\beta u)$ 表示在（2.1）中定义的几何分位数。结合（3.11），定理 3.3 的证明即告完成。□

定理 3.3中的集合包含关系自然给出了几何分位数范数的一个下界，该下界以单变量分位数表示，如下述定理所述。

定理 3.7。设 X 是完全支撑于 $\mathbb{R}^d$ 上的连续随机向量。设 $\theta \in \arg\max HD(x|X)$ 为深度最高的点（即半空间中位数）。则

$$\|q_X(\alpha u) - \theta\| \;\ge\; \min_{\|u\|=1} \left| Q_{u^\top X}\left(1 - \frac{1-\alpha^2}{M_\gamma}\right) - u^\top \theta \right|, \tag{3.12}$$

其中 $Q_{u^\top X}(\beta) = \inf\{t \in \mathbb{R} : \mathbb{P}(u^\top X > t) < 1-\beta\}$（其中 $0 \le \beta < 1$）表示投影 $u^\top X$ 的单变量分位数函数，参数 α 满足 $\alpha^2 \in \left(1 - \frac{M_\gamma}{d+1}, 1\right)$，其中 $0 < M_\gamma < 1$ 定义于 (3.7)。

式（3.12）中的下界将多变量问题转化为一组涉及投影分布的一维约束。对于任意方向 $u \in \mathbb{S}^{d-1}$，几何分位数与半空间中位数 θ 之间的位移以下列分布的单变量分位数为下界：$u^\top X$；对 u 取最小值即可得到所述下界。因此，该下界是显式的，且在计算上易处理。当投影分布以最小位移达到给定深度水平时，该下界取等号；换言之，这发生在相关水平上尾部最轻的边缘分布中。

因此，根据（3.6），几何分位数在任何方向上的增长都不会慢于相应单变量深度约束所规定的速度。当二阶矩存在时，应用马尔可夫不等式可得

$$Q_{u^\top X}\left(1 - \frac{1-\alpha^2}{M_\gamma}\right) = O\big((1-\alpha)^{-1/2}\big), \quad \text{uniformly in } u,$$

而在 [7] 中确立的精确渐近速率为

$$(1-\alpha)^{-1/2}\,\frac{1}{2}\big(\operatorname{tr}\Sigma - \langle \Sigma u, u\rangle\big).$$

因此，相较于精确速率，我们的下界通常较为保守。它的主要优势在于对所有方向都具有一致性，而且不依赖矩假设或多元正则变差（MRV）等结构条件，也不要求密度存在。

此外，在满足 MRV 假设时，该界能够恢复精确速率，详情见[6],中的后续注释和命题。

注 3.8. MRV 条件下界的精确性。
如前所述，推论 *3.2*给出的上界的阶为 $(1-\alpha)^{-(1+\frac{1}{\beta})\frac{\beta}{\beta-\epsilon}}$，这可能远大于 $(1-\alpha)^{-1/\beta}$ 中推导出的 *MRV* 速率 *[6]*。由于 $\frac{\beta+1}{\beta-\epsilon} > \frac{1}{\beta}$，我们的上界随 $\alpha \uparrow 1$增长得更快，这凸显了这一一般估计的保守性。

相比之下，定理 *3.3* 所确立的一般下界则非常精确：对于具有密度但缺乏二阶以上矩的分布，它与精确的多项式增长速率 $(1-\lambda)^{-1/\beta}$ 相符，该速率由 *[6]* 中得到。在我们的设定中，有效深度参数对应于 $\lambda = \alpha^2$，因此起决定作用的量变为 $(1-\alpha^2)$。这一对应关系也与 *[7]*中的精细化结果一致。

我们在以下命题中总结了多元正则变差条件下的这些结果。

Prop命题 3.9. 在多元正则变差假设下，以下结论成立：

1。在多元正则变差*(1)*条件下，下界与上界一致，并给出速率

$$O\big((1-\alpha)^{-1}\big),$$

这与 *[6]* 中确立的精确速率一致。

2. 在多元正则变差*(β)*且 $\beta > 0$ 的条件下，下界给出速率

$$O\big((1-\alpha)^{-\beta}\big),$$

这与 *[6]* 在 $\beta \le 2$ 下的精确速率一致。然而，对于 $\beta \neq 1$，上界会变得过于保守，正如推论 *3.2* 中所讨论的那样。

定理证明3.7根据定理 3.3,
我们有

$$\theta \in \left\{ x \in \mathbb{R}^d : HD\left(x|X\right) > \frac{1-\alpha^2}{M_\gamma} \right\} \;\Rightarrow\; \theta \in \left\{ q_X(\beta u) : 0 \le \beta \le \alpha, u \in \mathbb{S}^{d-1} \right\}.$$

因此，连接 θ 和 $q_X(\alpha u)$ 的任意线段都必须与闭曲线$\left\{ x \in \mathbb{R}^d : HD(x|X) = \frac{1-\alpha^2}{M_\gamma} \right\}$相交。因此，

$$\|q_X(\alpha u) - \theta\| \;\ge\; \min_{x \in \partial\mathcal{D}(\frac{1-\alpha^2}{M_\gamma}|X)} \|x - \theta\|.$$

因此，只需证明

$$\min_{x \in \partial\mathcal{D}(\frac{1-\alpha^2}{M_\gamma}|X)} \|x - \theta\| = \min_{\|u\|=1} \left| Q_{u^\top X}\left(1 - \frac{1-\alpha^2}{M_\gamma}\right) - u^\top \theta \right|. \tag{3.13}$$

注意到，只需证明3.13对于 $\theta = 0$成立即可，这是由于几何分位数（2.4）和平移分位数具有平移不变性。事实上，对于 $\theta \neq 0$，令 $Y = X - \theta$。此时，结论可由以下蕴含关系得出

$$x \in \partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma} \mid Y\right) \;\Rightarrow\; x + \theta \in \partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma} \mid X\right),$$

结合方向分位数的平移不变性:

$$Q_{u^\top Y}(\cdot) = Q_{u^\top X}(\cdot) - u^\top \theta.$$

根据假设，点 $\theta = 0$ 达到最大深度，且该深度至少为 $\frac{1}{d+1}$（见[4]）。因此，任何包含 0 的半空间的概率都至少为 $\frac{1}{d+1}$。由于假设 $\frac{1-\alpha^2}{M_\gamma} < \frac{1}{d+1}$，可知对于任意 $u \in \mathbb{S}^{d-1}$，

$$\mathbb{P}(u^\top X \ge 0) \ge \frac{1}{d+1} > \frac{1-\alpha^2}{M_\gamma}.$$

因此，

$$0 \le\; Q_{u^\top X}\left(1 - \frac{1-\alpha^2}{M_\gamma}\right) \quad \text{for all } u \in \mathbb{S}^{d-1}.$$

此外，$Q_{u^\top X}(\cdot)$ 在 u 上连续（见 [14]），因此右侧（3.13）的最小值在某个 $u_* \in \mathbb{S}^{d-1}$ 处取得。因此，对于任意 $v \neq u_*$，有

$$0 \le Q_{u_*^\top X}\left(1 - \frac{1-\alpha^2}{M_\gamma}\right) \le Q_{v^\top X}\left(1 - \frac{1-\alpha^2}{M_\gamma}\right).$$

S由于 X 连续，根据单变量分位数的定义，我们有

$$\frac{1-\alpha^2}{M_\gamma} = \mathbb{P}\left(u_*^\top X \geq Q_{u_*^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right)\right) = \mathbb{P}\left(v^\top X \geq Q_{v^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right)\right) \tag{3.14}$$

$$\leq \mathbb{P}\left(v^\top X \geq Q_{u_*^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right)\right) \quad (\text{since } Q_{u_*^\top X} \leq Q_{v^\top X} \text{ by definition of } u_*)$$

$$\leq \mathbb{P}\left(v^\top X \geq (v^\top u_*)\, Q_{u_*^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right)\right), \tag{3.15}$$

其中这些不等式对所有 $v \in \mathbb{S}^{d-1}$ 成立。为证明最后一个不等式，将X 分解为其平行于和正交于 u_*，$X = (u_*^\top X)\, u_* + X_\perp$，其中$X_\perp := X - (u_*^\top X)\, u_*$，从而 $u_*^\top X_\perp = 0$。于是 $v^\top X = (v^\top u_*)(u_*^\top X) + v^\top X_\perp$，由此可得（3.15）成立。

定义

$$y = Q_{u_*^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right) u_*.$$

于是（3.15）与（3.14）相结合，可改写为

$$\mathbb{P}\big(u_*^\top (X-y) \geq 0\big) = \frac{1-\alpha^2}{M_\gamma} \leq \mathbb{P}\big(v^\top (X-y) \geq 0\big), \quad \text{for all } v \in \mathbb{S}^{d-1}.$$

因此，y 的半空间深度的最优方向在 u_*处达到。特别地，

$$HD(y \mid X) = \mathbb{P}\big(u_*^\top (X-y) \geq 0\big) = \frac{1-\alpha^2}{M_\gamma}.$$

这表明 $y \in \partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma} \mid X\right)$。因此，

$$\min_{x \in \partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma} \mid X\right)} \|x\| \;\leq\; \|y\| = \min_{\|u\|=1} \left| Q_{u^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right)\right|. \tag{3.16}$$

接下来，我们证明不等式 (3.16) 实际上是等式。为反证，假设该不等式为严格不等式。那么存在 $z \in \partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma}\right)$ 使得$\|z\| < \|y\|$，其中 y 对应于方~~向~~ u_*上的方向分位数。令$w \in \mathbb{S}^{d-1}$ 满足

$$HD(z \mid X) = \Pr\big(w^\top (X-z) \geq 0\big) = \frac{1-\alpha^2}{M_\gamma}.$$

Then

$$Q_{w^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right)=w^\top z\le\|z\|<\|y\|=Q_{u_*^\top X}\left(1-\frac{1-\alpha^2}{M_\gamma}\right),$$

w这与 u_* 在 $\|u\|=$上使 $\left|Q_{u^\top X}(1-\frac{1-\alpha^2}{M_\gamma})\right|$ 最小这一事实相矛盾 1.
因此，任何 $z\in\partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma}\right)$ 都满足 $\|z\|\ge\|y\|$。特别地，

$$\min_{x\in\partial\mathcal{D}\left(\frac{1-\alpha^2}{M_\gamma}\right)}\|x\|\ge\|y\|. \tag{3.17}$$

因此，结合式（3.16）和（3.17），即可得到该定理。 □

更一般地，我们的基于几何的方法通过给出不依赖于矩或光滑性假设的下界，推广了这些结果，因此适用于更广泛的分布类别。尽管前述结果刻画了方向异质性截至二阶的影响，但并未反映偏度或高阶尾部行为的影响；我们将在下一节的适当矩条件下对此进行研究。

4 欠可积条件

几何分位数的一阶渐近行为已在 [7] 中确立。当 $\mathbb{E}\|X\|^2<\infty$ 时，有

$$\lim_{\alpha\to1}\|q_X(\alpha u)\|^2(1-\alpha)=\frac{1}{2}\big(\operatorname{tr}\Sigma-\langle\Sigma u,u\rangle\big),$$

其中， Σ 是 X的协方差矩阵。这个主导项通过协方差结构反映了各向异性，但未能捕捉偏度或尾部不对称性等高阶特征。特别地，具有相同协方差矩阵的分布具有相同的一阶行为。

为了细化这一分析，我们将 [7] 的渐近展开方法推广到一阶之外。在更强的矩假设下，高阶项揭示了偏度与尾部行为的影响。特别地，假设 $\mathbb{E}\|X\|^3<\infty$，我们在定理4.1 中得到如下三阶展开。该方法自然可进一步推广，从而在更强的矩假设下得到更高阶的展开。

定理 4.1. 设 $u \in S^{d-1}$ 与 $\mathbb{P}$ 为 $\mathbb{R}^d$ 上的概率测度，不含原子，且其支撑不集中于任何直线。若 $\mathbb{E}\|X\|^3 < \infty$，则有

$$\|q_X(\alpha u)\|\Big[\|q_X(\alpha u)\|^2(1-\alpha) - \frac{1}{2}\left(tr\,\Sigma - \langle \Sigma u, u\rangle\right)\Big] \underset{\alpha\to 1}{\longrightarrow} \mathbb{E}\Big(\langle X, u\rangle \|X - \langle X, u\rangle\|^2\Big) - \Big\langle (2\Sigma + uu^T)u, (I - uu^T)\mathbb{E}X \Big\rangle. \tag{4.1}$$

注 4.2。 若 X 均值为零，则方程 (4.1) 可简化为：

$$\|q_X(\alpha u)\|\left(\|q_X(\alpha u)\|^2(1-\alpha) - \frac{1}{2}\left(tr\,\Sigma - \langle \Sigma u, u\rangle\right)\right) \underset{\alpha\to 1}{\longrightarrow} \mathbb{E}\left(\langle X, u\rangle \|X - \langle X, u\rangle\|^2\right).$$

此外，若假设 X 在 u 上的投影与其在 u的正交补上的投影相互独立，则有：

$$\|q_X(\alpha u)\|\left(\|q_X(\alpha u)\|^2(1-\alpha) - \frac{1}{2}\left(tr\,\Sigma - \langle \Sigma u, u\rangle\right)\right) \underset{\alpha\to 1}{\longrightarrow} 0.$$

定理4.1 补充了 [7], 中的引理 6.2–6.5，区分了具有极重尾（不存在二阶矩）和具有中等重尾（存在三阶矩）的分布。

请注意，附录给出了一个关于几何分位数方向的补充结果，纳入其中是为了完整起见，并更全面地呈现 [7], 中的发现；我们通过建立收敛速率对这些发现进行了扩展。

定理 *4.1* 的证明
记 $\{u, w_1, ..., w_{d-1}\}$ 为 $\mathbb{R}^d$的正交归一基。设 $b(\alpha)$和 $\{\beta_k(\alpha)\}_{k=1}^{d-1}$ 为实数，其定义如下

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} = b(\alpha)u + \sum_{k=1}^{d-1}\beta_k(\alpha)w_k. \tag{4.2}$$

利用 [7], ，我们可以写成

$$\begin{aligned} & \|q(\alpha u)\|\Big[\|q(\alpha u)\|^2\beta_k^2(\alpha) - |\mathbb{E}\langle X, w_k\rangle|^2\Big] \\ = \; & \Big[\|q(\alpha u)\|\Big(\|q(\alpha u)\|\beta_k(\alpha) - |\mathbb{E}\langle X, w_k\rangle|\Big)\Big]\Big[\|q(\alpha u)\|\beta_k(\alpha) + \mathbb{E}\langle X, w_k\rangle\Big] \\ \underset{\alpha\to 1}{\longrightarrow} \; & \operatorname{cov}(\langle X, u\rangle, \langle X, w_k\rangle)\; 2\mathbb{E}\langle X, w_k\rangle. \end{aligned} \tag{4.3}$$

利用分解（4.2）和 [7, 命题 6.3] 在假设 $\mathbb{E}\|X\|^3 < \infty$ 下，我们有

$$\|q(\alpha u)\|\left(\|q(\alpha u)^2(1-\alpha b(\alpha)) - \frac{1}{2}\,\mathbb{E}\|X - \langle X,u\rangle u\|^2\right) \tag{4.4}$$

$$\underset{\alpha\to 1}{\longrightarrow}\ \mathbb{E}\left(\langle X,u\rangle\left[\|X-\langle X,u\rangle u\|^2 - \langle X, \mathbb{E}(X-\langle X,u\rangle u)\rangle\right]\right) \triangleq f_X(u).$$

因此，根据等式（4.3）和（4.4），我们可以得出

$$\|q(\alpha u)\|\left(\|q(\alpha u)\|^2\Big(1-\alpha b(\alpha)\Big) - \frac{1}{2}\mathbb{E}\|X-\langle X,u\rangle u\|^2 - \frac{1}{2}\sum\left[\|q(\alpha u)\|^2\beta_k^2(\alpha) - |\mathbb{E}\langle X,w_k\rangle|^2\right]\right)$$

$$\underset{\alpha\to 1}{\longrightarrow} f_X(u) - \sum_{k=1}^{d-1}\mathbb{E}\langle X,w_k\rangle\ \mathrm{cov}(\langle X,u\rangle,\langle X,w_k\rangle. \tag{4.5}$$

但是，请注意，等式（4.5）的左端可以改写为

$$\|q(\alpha u)\|\left(\|q(\alpha u)\|^2\frac{1}{2}\left(1-\alpha^2+(b(\alpha)-\alpha)^2\right) - \frac{1}{2}\sum_{k=1}^{d-1}\mathrm{var}\langle X,w_k\rangle\right),$$

通过观察到（$1-b^2$（α））$=\sum_{k=1}^{d-1}\beta_k^2$（$\alpha$）。
此外，$\|q$（αu）$\|^3$（$\alpha - b$（α））$^2 \underset{\alpha\to 1}{\to} 0$，因为 $\|q$（αu）$\|$（$\alpha - b$（α））收敛于 0，而$\|q$（αu）$\|^2$（$\alpha - b$（α））收敛于一个有限极限，因为根据 $\alpha \to 1$，分别由 [7], 中的引理 6.2 和 6.4 可知。因此，我们得出

$$\|q(\alpha u)\|\left(\|q(\alpha u)\|^2\Big(1-\alpha\Big) - (\mathrm{tr}\Sigma - \langle \Sigma u,u\rangle)\right)$$

$$\underset{\alpha\to 1}{\longrightarrow} f_X(u) - \sum_{k=1}^{d-1}\mathbb{E}\Big(\langle X,w_k\rangle\Big)\ \mathrm{cov}\Big(\langle X,u\rangle,\langle X,w_k\rangle\Big).$$

汇总所有项后，我们得到

$$\|q_X(\alpha u)\|\left[\|q_X(\alpha u)\|^2(1-\alpha) - \frac{1}{2}\left(\mathrm{tr}\ \Sigma - \langle \Sigma u,u\rangle\right)\right] \underset{\alpha\to 1}{\longrightarrow}$$
$$\mathbb{E}\Big(\langle X,u\rangle\left[\|X-\langle X,u\rangle u\|^2 - \langle X,\mathbb{E}\left(X-\langle X,u\rangle u\right)\rangle\right]\Big)$$
$$-\sum_{k=1}^{d-1}\mathrm{cov}(\langle X,u\rangle,\langle X,w_k\rangle)\,\mathbb{E}(\langle X,w_k\rangle),$$

进一步简化后可得

$$\|q_X(\alpha u)\|\Big[\|q_X(\alpha u)\|^2(1-\alpha)-\frac{1}{2}\left(\operatorname{tr}\,\Sigma-\langle\Sigma u,u\rangle\right)\Big]\xrightarrow[\alpha\to 1]{}$$
$$\mathbb{E}\Big(\langle X,u\rangle\big[\|X-\langle X,u\rangle u\|^2\Big)-\langle(2\Sigma+\mu\mu^T)u,(I-uu^T)\mu\rangle.\quad\square$$

5 结论

本文研究了多元情形下几何分位数的极端行为。我们的主要贡献包括：在不作矩假设的情况下，推导几何分位数范数增长率的一般下界和上界；以及揭示多元几何分位数、单变量分位数与 Tukey 深度之间的联系。我们提供了几何学层面的见解，并分析了渐近性质，重点阐述其与半空间深度的关系，以及在可积性（矩）条件下的改进。更广泛地说，我们基于几何的方法扩展了以往的结果，给出了不依赖矩或光滑性假设的界，因此适用于广泛的分布类别，包括具有重尾的分布。除了无矩的一般界之外，我们还考虑了可积情形，并将 [7] 的渐近展开方法扩展至更高阶，以刻画偏度和更高阶的尾部效应。总的来说，我们的研究加深了对多元分位数几何学的理解。

参考文献

[1] P. Chaudhury. 多元数据分位数的一种几何概念。《美国统计学会会刊》，91:862–872，1996。[2] A. Cuevas、M. Febrero 和 R. Fraiman. 基于投影的深度概念对函数型数据进行稳健估计与分类。计算统计学，22(3):481–496，2007。[3] S. Dhar、B. Chakraborty 和 P. Chaudhuri. 使用分位数-分位数图及相关检验比较多元分布。《伯努利》，20(3):1484–1506，2014。[4] D. Donoho 和 M. Gasko. 基于半空间深度和投影离群度的位置估计的崩溃性质。统计年刊，20(4):1803–1827，1992。

[5] S. Dutta、S. Sarkar 和 A. K. Ghosh. 使用局部空间深度进行多尺度分类。《机器学习研究杂志》，17(1):7657–7686，2016。[6] S. Girard 和 G. Stupfler. 多元正则变 variation 框架下的极端几何分位数。《极值》，18:629–663，2015。[7] S. Girard 和 G. Stupfler. 极端几何分位数的有趣性质。REVSTAT，15(1):107–139，2017。[8] M. Hallin 和 D. Konen. 多元分位数：几何等高线与基于测度输运的等高线，第61–78页。施普林格自然瑞士，沙姆，2024。[9] M. Hallin、D. Paindaveine 和 M. ˇSiman. 多元分位数与多输出回归分位数：从 l_1 优化到半空间深度。《统计年刊》，38(2):635–669，2010。[10] M. Hubert 和 S. Van der Veeken. 偏斜数据的稳健分类。《数据分析与分类进展》，4(4):239–254，2010。[11] V. Koltchinskii. M–估计、凸性与分位数。《统计年刊》，25(2):435–477，1997。[12] D. Konen. 几何分布函数与分位数的偏微分方程刻画。《伯努利》，31(3):2077–2104，2025。[13] D. Konen 和 D. Paindaveine. 多元 ρ 分位数：一种空间方法。《伯努利》，28:1912–1934，2022。[14] L. Kong 和 I. Mizera. 分位数断层成像：将分位数用于多元数据。《中华统计学志》，第1589–1610页，2012。[15] R. Liu. 基于随机单纯形的数据深度概念。《统计学年鉴》，第405–414页，1990。[16] D. Paindaveine 和 M. ˇSiman. 计算多输出回归分位数区域。《计算统计学与数据分析》，56(4):840–853，2012。[17] D. Paindaveine 和 J. Virta. 最小假设下极端 d维空间分位数的行为。《当代统计学与计量经济学进展》。Christine Thomas-Agnan 荣休文集。施普林格，2021。

[18] R. Rado. 关于一般测度的定理。《伦敦数学学会杂志》，21:291–300，1946年。[19] P. Rousseeuw 和 A. Struyf. 刻画角对称性与回归对称性。《统计规划与推断杂志》，122(1-2):161–173，2004年。[20] G. Staerman、E. Adjakossa、P. Mozharovskyi、V. Hofer、J. Sen Gupta 和 S. Cl´emen¸con. 函数型异常检测：一项基准研究。《国际数据科学与分析杂志》，16:101–117，2023年。[21] A. Struyf 和 P. Rousseeuw. 半空间深度和回归深度刻画了经验分布。《多元分析杂志》，69(1):135–153，1999年。[22] J. Tukey. 数学与数据图示。载于《国际数学大会论文集》，第2卷，第523–531页，1975年。

附录

几何分位数与半空间深度的比较

本节将探讨定理3.3中所确立的集合包含关系的紧致性。更具体地说，给定水平为 α的几何分位数区域，我们要探究其中包含的最大半空间深度区域是什么。

从解析角度来看，通常很难确定哪条半空间深度等高线最适合包含在给定的几何分位数等高线内。因此，我们采用数值方法来探究这个问题。

令 $\partial G(\alpha)$ 表示几何分位数等高线，即水平为 α的几何分位数区域 $G(\alpha)$的边界。我们定义

$$\alpha_{\text{best}} = \max_{x \in \partial G(\alpha)} HD(x \mid X).$$

根据定义，半空间深度区域 $\mathcal{D}(\alpha_{\text{best}})$ 是包含在 $G(\alpha)$ 中的最大半空间深度区域。实际应用中，通过在 $\partial G(\alpha)$ 上选取一个稠密网格点集，计算各点的半空间深度并取最大值，以此估计 α_{best} 。

在图 1中，我们绘制了三条等高线：几何分位数等高线 $\partial G(\alpha)$、数值计算得到的半空间深度等高线 $\partial \mathcal{D}(\alpha_{\text{best}}$以及理论深度

等高线 $\partial\mathcal{D}\begin{pmatrix}1-\alpha^2\\ \underline{M_\gamma}\end{pmatrix}$，对应两种分布：高斯分布和重尾分布，即与高斯 copula 耦合的 Student t 分布。回顾（3.10），可知

$$M_\gamma = (1-\gamma)\inf_{q\in\mathbb{S}^{d-1}}\mathbb{P}(U_0 \in A_q^\gamma).$$

在数值实验中，我们固定 $\gamma = 0.01$，并通过最小化 $\mathbb{P}(U_0 \in A_q^\gamma)$ 来估计 M_γ，其中 $q \in \mathbb{S}^{d-1}$。然后，我们选择 α，使得

$$\alpha^2 \in \left(1 - \frac{M_\gamma}{d+1}, 1\right),$$

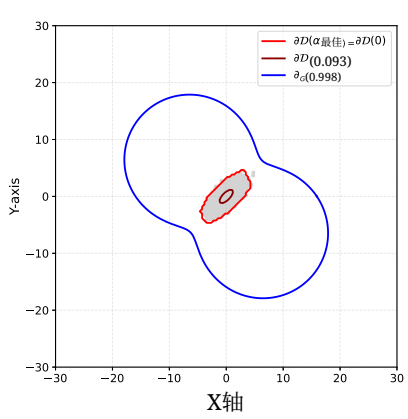
从而确保水平为 α 的几何分位数区域包含半空间深度中位数，符合定理 3.3 的要求。如果不满足这一条件，几何分位数区域可能无法包含任何非空的半空间深度区域。

注 A.1. 从数值角度来看，检查一个分位数区域是否包含于另一个分位数区域中十分简单。给定一个几何分位数区域，可以计算该区域内各点的半空间深度，并取其达到的最小值。此后，任何深度小于该最小值的半空间深度区域都必然包含于该几何分位数区域中。
然而，从解析角度来看，深度区域与分位数区域之间的一般包含关系在很大程度上仍未解决。定理 *3.3* 为几何分位数区域与半空间深度区域建立了部分解析联系，其中半空间深度通过其作为方向事件概率界的解释自然地出现；见公式 (3.9)。

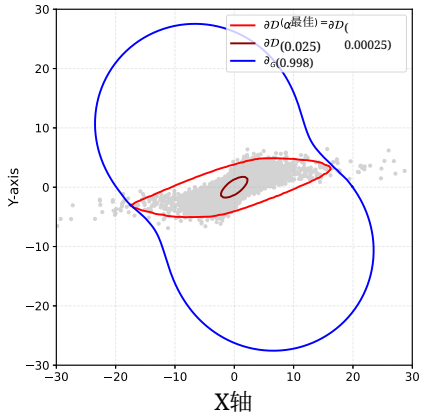
定理 3.3背后的几何包含关系及其在定理 3.7中陈述的推论，见图 1，其中我们考察两个例子：

1. 边缘为标准高斯分布、相关系数为 $\rho = 0.7$的二元高斯分布；见图 1a。2. 缺乏多元正则变异结构的重尾分布：自由度不同的 Student-t t边缘分布，$\nu_1 = 1.2$ 和 $\nu_2 = 2.2$通过高斯 Copula 耦合，参数为 $\rho = 0.7$；见图 1b。

除了比较这三条等高线外，这些图还凸显了几何分位数区域形状的差异。在高斯 copula 的情形下（第一行），几何分位数区域呈现出明显的细长形状；而在重尾示例中（第二行），等高线则显得更加圆润。



(a) 边缘分布为二元高斯分布，且边缘分布为标准高斯边缘分布，且 $\rho = 70\%$



(b) Student 的 t 边缘分布，参数为 $\nu_1 = 3$、$\nu_2 = 10$，通过高斯 copula 耦合（$\rho = 70\%$）

图1：不同二元分布的深度等高线

B 几何分位数方向的收敛速率

在本附录中，我们给出一项关于几何分位数方向（而非其大小）的补充结果。本文主要关注几何分位数的大小；为使内容更加完整并进一步拓展 [7]，我们证明了一个收敛速率。

引理A. 设 $u \in S^{d-1}$ 与 $\mathbb{P}$ 为 $\mathbb{R}^d$ 上无原子的概率测度，且其支撑不集中于任何直线。

若 $\mathbb{E}\|X\|^2 < \infty$，且 Σ 表示 X 的协方差矩阵，则有

$$\|q_X(\alpha u)\|\Big[q_X(\alpha u) - \big\{\|q_X(\alpha u)\|u + \mathbb{E}(X - \langle X, u\rangle u)\big\}\Big] \underset{\alpha\to 1}{\longrightarrow} -\frac{1}{2}\|\mathbb{E}(X - \langle X, u\rangle u)\|^2 u + (I - uu^T)\Sigma u.$$

此外，若假设 X 的均值为零，则极限可简化为

$$\|q_X(\alpha u)\|\left(q_X(\alpha u) - \|q_X(\alpha u)\|u\right) \underset{\alpha\to 1}{\longrightarrow} (I - uu^T)\Sigma u.$$

证明。 利用基表示式（4.2），可写为

$$
\begin{aligned}
&\|q(\alpha u)\|\left[q(\alpha u)-\|q(\alpha u)\|\,u-\mathbb{E}(X-\langle X,u\rangle u)\right]\\
&+\frac{1}{2}\,\|\mathbb{E}(X-\langle X,u\rangle u)\|^2 u-(I-uu^T)\Sigma u\\
=\;&\underbrace{\left[\|q(\alpha u)\|^2\,(b(\alpha)-1)+\frac{1}{2}\,\|\mathbb{E}(X-\langle X,u\rangle u)\|^2\right]u}_{I}\\
&+\underbrace{\sum_{k-1}^{d-1}\left(\|q(\alpha u)\|^2\,\beta_k(\alpha)-\|q(\alpha u)\|\,\mathbb{E}\langle X,w_k\rangle-\mathrm{cov}(\langle X,u\rangle,\langle X,w_k\rangle)\right)w_k}_{II}.
\end{aligned}
\tag{B.1}
$$

利用引理 6.4, [7], 我们观察到

$$
\|q(\alpha u)\|^2\beta_k(\alpha)-\|q(\alpha u)\|\mathbb{E}\langle X,w_k\rangle-\mathrm{cov}(\langle X,u\rangle,\langle X,w_k\rangle)\underset{\alpha\to 1}{\longrightarrow}0,
\tag{B.2}
$$

从而得出，(II B.1) 中的这一部分随着 $\alpha\to 1$趋于 0。
接下来，考虑 (I B.1) 中的这一部分。回忆一下，$\|q(\alpha u)\|^2(1-\alpha)\underset{\alpha\to 1}{\longrightarrow}\frac{1}{2}\,(\mathrm{tr}\Sigma-\langle\Sigma u,\ u\rangle)$，并且 $\|q(\alpha u)\|^2(\alpha b(\alpha)-1)\underset{\alpha\to 1}{\longrightarrow}-\frac{1}{2}\mathbb{E}\|X-\langle X,\ u\rangle u\|^2$分别由定理 2.2 和 [7,引理 6。3], 因此，我们可以得出

$$
\|q(\alpha u)\|^2(1-b(\alpha))\underset{\alpha\to 1}{\to}\frac{1}{2}\sum_{k=1}^{d-1}\left[\mathbb{E}\langle X-\langle X,u\rangle u,w_k\rangle\right]^2=\frac{1}{2}\|\mathbb{E}\,(X-\langle X,u\rangle u)\,\|^2,
\tag{B.3}
$$

这就证明了 (I B.1) 中的这一部分。
因此，我们得到：

$$
\begin{aligned}
&\|q_X(\alpha u)\|\left[q_X(\alpha u)-\left\{\|q_X(\alpha u)\|u+\mathbb{E}(X-\langle X,u\rangle u)\right\}\right]\\
&\qquad\underset{\alpha\to 1}{\longrightarrow}-\frac{1}{2}\,\|\mathbb{E}(X-\langle X,u\rangle u)\|^2 u+(I-uu^T)\Sigma u.
\end{aligned}
$$

将 $X-\mathbb{E}(X)$ 代入 X，即可直接得到该引理的简化形式，使其含义一目了然。

□