arXiv:2602.05225v1 [math.ST] 5 Feb 2026

度量空间值 Fr´echet 回归

L´aszl´ó Gy¨örfi[1*]，Pierre Humbert[2†]，Batiste LeBars[3]

[1] 计算机科学与信息论系，布达佩斯技术与经济大学[2]巴黎萨克雷大学、法国国家科学研究中心、埃夫里大学，埃夫里数学与建模实验室，法国埃夫里-库尔库罗讷，邮编 91037[3]里尔大学、法国国家信息与自动化研究所、法国国家科学研究中心、里尔中央理工学院、UMR 9189、CRIStAL，法国里尔 F-59000

摘要

我们研究如何根据取值于可分度量空间的数据，估计 Fr´echet 均值和条件 Fr´echet 均值。与已有成熟方法的欧几里得空间不同，目前尚无适用于所有度量空间的实用估计量。因此，我们提出一种基于随机量化技术、可计算的 Fr´echet 均值估计量，并证明它在任意可分度量空间中均具有普适一致性。此外，我们还提出另一种条件 Fr´echet 均值估计量，利用数据驱动划分和量化，并证明当输出空间为任意巴拿赫空间时，该估计量具有普适一致性。

1 引言

设 $(\mathcal{X}, \rho_{\mathcal{X}})$ 和 $(\mathcal{Y}, \rho_{\mathcal{Y}})$ 是两个配备了其 Borel σ-代数的可分度量空间，且 (X, Y) 是一对随机变量，分别取值于 $\mathcal{X}$ 和 $\mathcal{Y}$。我们用 μ_X 和 μ_Y 表示它们各自的边缘概率分布。在本文中，我们研究 Fr´echet 均值和条件 Fr´echet 均值的估计问题，其中两个变量 X 和 Y 均取度量空间中的值。

给定损失函数 $\ell : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_+$, i f

$$R(y) = \mathbb{E}[\ell(Y, y)]$$

表示在 y 处的风险，则 Fr ´echet 均值定义为

$$\bar{m}_{Fr} = \arg\min_{y \in \mathcal{Y}} R(y) \,, \tag{1}$$

也称为重心或质心 [Fr´echet, 1948]。本文假设 Fr´echet 均值 $\bar{m}_{Fr}$ 定义良好，这意味着 $\arg\min_{y \in \mathcal{Y}}$ 存在。然而，通常无法保证极小值点的唯一性，参见 Hein [2009]。相反，我们假设在极小值点集合中存在一个（可能未知的）可测选择规则。

接下来，我们回顾一些可以应用我们的 $\bar{m}_{Fr}$ 估计量的例子。

*通信联系人：gyorfi@cs.bme.hu。

†本研究的部分工作是在 P.H. 任职于索邦大学 LPSM 期间完成的。

例 1 设 $\mathcal{Y}$ 是一个范数为 $\|\cdot\|$ 的可分希尔伯特空间，并且

$$\ell(y, y') = \|y - y'\|^2 .$$

如果 $\mathcal{Y} = \mathbb{R}$，那么这个代价函数称为平方损失。由于

$$\mathbb{E}[\ell(Y, y)] = \mathbb{E}\big[\|Y - y\|^2\big] = \mathbb{E}\big[\|Y - \mathbb{E}[Y]\|^2\big] + \|\mathbb{E}[Y] - y\|^2 ,$$

可得

$$\bar{m}_{Fr} = \mathbb{E}[Y],$$

也就是说，Fr ´echet 均值等于期望。关于可分巴拿赫空间中期望的定义，请参阅 Beck *[1963]*。期望的估计问题$\mathbb{E}[Y]$ 很简单，因为样本均值是强一致的。

例 2 相对于例 1，我们将代价函数修改为:

$$\ell(y, y') = \|y - y'\| .$$

如果 $\mathcal{Y} = \mathbb{R}$，那么这个代价函数称为绝对损失，而 $\bar{m}_{Fr}$ 是中位数，它可能不唯一，参见 Milasevic 和 Ducharme *[1987]*。在这种情形下，中位数的估计很简单，因为样本中位数是一致的。然而， $\mathcal{Y} = \mathbb{R}^d$ 与 $d \geq 2$ 的设定要困难得多，因为目前尚无对应 Fr ´echet 均值的已知样本估计量。

例 3 令 $\mathcal{Y} \subset L_1(\mathbb{R}^d)$ 为定义在 $\mathbb{R}^d$上的概率密度集合，也就是说， Y 是一个随机密度。给定两个密度 f 和 g，令

$$\ell(f, g) = \|f - g\|_1 = \int_{\mathbb{R}^d} |f(z) - g(z)|\lambda(dz) ,$$

其中， λ 表示勒贝格测度。于是， $\bar{m}_{Fr}$ 是随机密度的质心。在一种等价的设定中， $\mathcal{Y}$ 是 $\mathbb{R}^d$上的绝对连续概率测度集合。若ν_f 和ν_g 分别是具有密度 f 和 g的概率测度，则定义

$$\ell(\nu_f, \nu_g) = TV(\nu_f, \nu_g)$$

其全变差距离为 TV。根据 Scheff ´e 定理,

$$TV(\nu_f, \nu_g) = \frac{1}{2}\|f - g\|_1 .$$

这意味着，若 f^* 是密度的质心，而 ν^* 是相应概率测度的质心，则

$$\nu^* = \nu_{f^*} .$$

例 4 在网络分析中， $\mathcal{Y}$ 可以是图空间或图拉普拉斯矩阵空间（例如，参见 Ginestet *[2012],* Ginestet 等人、 *[2017],* Kolaczyk 等人、 *[2020],* Lunag ´omez 等人、*[2021],* Severn 等人 *[2022]*）。

例 5 Bhattacharya 和 Patrangenaru *[2003, 2005]* 研究了 $\mathcal{Y}$ 为黎曼流形的情形。Ahidar-Coutrix 等人 *[2020]*在一般条件下给出了紧测地度量空间中的收敛速率。最近，Lugosi 和 Matabuena *[2024]* 展示了一些临床数据的实际示例，如磁共振成像图像或生物传感器曲线。

对于取值于度量空间的回归问题， X 称为预测变量，而 Y 是相应的响应变量。我们要寻找一个函数 $m:\ \mathcal{X}\to\mathcal{Y}$，使得 $m(X)\approx Y$。这种近似由损失函数 ℓ 来衡量，因此 m 的风险定义为

$$R(m)=\mathbb{E}[\ell(Y,m(X))]\ . \tag{2}$$

在标准回归设定中， $\mathcal{X}$ 和 $\mathcal{Y}$ 都是有限维欧几里得空间，且 $\ell(y,y')=\|y-y'\|^2$，其中 $\|\cdot\|$ 表示欧几里得范数。本文考虑一种更一般的设定，其中 X 和 Y 可以取值于更复杂的空间，例如曲线空间、词空间，或函数空间等其他无限维空间。此时，假设最优回归函数是使方程 (2) 中的风险最小化的函数，其解由以下回归函数给出：

$$m_{Fr}(x)=\operatorname*{arg\,min}_{y\in\mathcal{Y}}\mathbb{E}[\ell(Y,y)\mid X=x]\ , \tag{3}$$

其中 $x\in\mathcal{X}$。函数 m_{Fr} 通常称为给定 X 时 Y 的条件 Fr´échet 均值，或 Fr´échet 回归函数。

例6 在例1的设定中，可得

$$\begin{aligned}\mathbb{E}[\ell(Y,y)\mid X=x]&=\mathbb{E}\big[\|Y-y\|^2\mid X=x\big]\\&=\mathbb{E}\big[\|Y-\mathbb{E}[Y\mid X=x]\|^2\mid X=x\big]+\|\mathbb{E}[Y\mid X=x]-y\|^2\ .\end{aligned}$$

因此，

$$m_{Fr}(x)=\mathbb{E}[Y\mid X=x]\ ,$$

也就是说，条件 Fr´échet 均值 $m_{Fr}(x)$ 等于条件期望 $\mathbb{E}[Y\mid X=x]$。若 $\mathcal{X}$ 是任意可分度量空间，且 $\mathcal{Y}=\mathbb{R}$，则 Hanneke 等人 *[2021]* 以及 Gy¨orfi 和 Weiss *[2021]* 提出了 $m_{Fr}(\cdot)$ 的普遍一致估计。

例7 对于例2中的损失函数，条件 Fr´échet 均值是分位数回归的一种扩展。对于特定情形 $\mathcal{X}=\mathbb{R}^d$ 和 $\mathcal{Y}=\mathbb{R}$，请参见 Koenker 和 Hallock*[2001]*。

例8 最近，Nguyen 等人 *[2026]* 提出了一种有趣的密度-密度回归方案，其中 $\mathcal{X}$ 和 $\mathcal{Y}$ 均为多元连续分布函数的集合，所采用的度量是 Wasserstein 度量。

例9 如果 $\mathcal{X}=\mathbb{R}^d$ 和 $\mathcal{Y}$ 是图拉普拉斯矩阵的空间，那么 Zhou 和 M ¨uller *[2022]*研究了相应条件 Fr´échet 均值的估计。

在现有文献中，Fr´échet 均值或条件均值估计量是通过在 $\mathcal{Y}$ 上最小化损失函数得到的。然而，在实践中，在度量空间上计算这一最小值并不总是简单明了。在 Fr´échet 均值的情形下，一种可能的方法是使用所谓的受限 Fréchet 均值，即最小值不是在 $\mathcal{Y}$ 上取得，而仅在一个样本上取得（例如参见 Sverdrup-Thygeson [1981], Ginestet [2012]）。在条件 Fr´échet 均值的情形下，Petersen 和 M¨uller [2019] 提出了一种针对特定度量空间 $\mathcal{Y}$ 的计算方法，但没有给出系统性的计算方法。尽管这一问题凸显了针对一般 $\mathcal{Y}$ 或一般损失函数进行实际实现的复杂性，近期的研究进展已开始着手解决这一问题。特别是，Cohen and Kontorovich [2022]开发了一种普遍一致的 Fr´échet 回归估计量及相应算法。

他们的方法首先针对 δ 的多个取值，构造 $(X_1,\dots,X_n)$ 的 δ-网，然后通过一个特定的最小化问题，输出每个所得单元的中心点。然而，他们的方法存在一个重大局限：当 $\mathcal{Y}$ 不是可数度量空间时，他们假设可以访问一个预言机过程，该过程能够提供 $\mathcal{Y}$ 的 ε-网，而这是一项相当强的假设。总体而言，在一般情况下，最小化通常需要根据所考虑度量空间的具体结构逐一处理。因此，我们论文的主要目标是提出一种具有较小计算复杂度的普遍一致 Fr´échet 回归估计量。我们的贡献概括如下。

贡献。

1. 在第 2 节中，我们介绍了一种新颖且实用的 Fr´échet 均值估计，其构造方式与 Sverdrup-Thygeson[1981] 的受限 Fr´échet 均值略有不同。在损失函数满足有界性条件的情况下，我们证明，只要 Fr´échet 均值属于 Y 的分布支撑集，该估计量就是强一致的，见定理 1。

2. 关于条件 Fr´échet 均值，在第 3 节中，当 $\mathcal{X}$ 和 $\mathcal{Y}$ 都是可分度量空间时，我们引入了 Gy¨orfi 和 Weiss [2021], 的 PROTO-NN 估计的扩展。我们证明，在温和的假设下，我们的估计具有强普遍一致性，即当条件 Fr´échet 均值属于 Y 的分布支撑集时，在该估计处评估的风险几乎处处收敛于条件 Fr´échet 均值的风险，见定理2。

2 经验Fr´échet均值

给定来自分布 μ_Y 的一些数据 Y_1、.....、Y_n，Fr´échet均值（公式（1））的一个自然估计量是

$$\bar{m}_{Fr,n} = \operatorname*{arg\,min}_{y\in\mathcal{Y}} \sum_{i=1}^{n} \ell(Y_i, y)\ . \tag{4}$$

经验Fr´échet均值的渐近理论如今已经十分成熟。例如，当空间 $\mathcal{Y}$ 是可分有限拟度量空间时，Ziezold [1977] 分析了其大样本性质。除渐近理论之外，各种情形下也已经建立了非渐近结果。Le Gouic等 [2022] 和 Sch¨otz [2019] 证明了Fr´échet均值与其经验对应量之间距离的上界。关于近期结果，我们还可参阅Sch¨otz [2022], 、Aveni和Mukherjee [2024], 、Evans和Jaffe [2024], 、Hotz等 [2024],、Jaffe [2024] 及其中的参考文献。最近，Brunel 和Serres[2024] 针对同一量给出了有限样本集中不等式。最后，在稳健Fr´échet均值的研究方面也取得了重大进展，其中具有代表性的贡献包括Hsu和Sabato [2016], 、Yun和Park [2023], 、Kim等 [2025] 以及Bartl等[2025]。需要注意的是，这些论文没有考虑如何在实践中求解对整个空间 $\mathcal{Y}$ 进行最小化这一困难的算法问题。

据我们所知，唯一针对一般空间 $\mathcal{Y}$ 并同时配有算法的估计由 Sverdrup-Thygeson [1981] 定义如下：

$$\widetilde{m}_n = \operatorname*{arg\,min}_{y\in\{Y_j\}_{j=1}^n} \sum_{i=1}^{n} \ell(Y_i, y)\ . \tag{5}$$

该估计称为受限 Fr´échet 均值，可解释为一种重新代入估计。在 Fr´échet 均值属于 Y的支撑集这一条件下，可证明其强一致性，见下文条件（7）。

我们的估计如下:

算法 1。如上所述，基本思路是对空间 $\mathcal{Y}$进行量化，从而使最小化成为可能。为此，我们假设可以获得标签 Y的一组独立同分布副本 Y_1'，...，Y_n'。设 $\mathcal{Y}_n$ 为原型 Y_1'，...，Y_n'构成的集合，则我们的 Fr´echet 均值估计定义为

$$\bar{m}_n = \underset{y \in \mathcal{Y}_n}{\arg\min} \sum_{i=1}^{n} \ell(Y_i, y) \; . \tag{6}$$

我们的估计可以解释为一种数据拆分估计，其中将 $2n$ 样本拆分为学习样本 Y_1，...，Y_n和测试数据 Y_1'，...，Y_n'。请注意，无论原型的数量是多少， $\mathcal{Y}_n$， m_n 都取值于 μ_Y 的支撑集 $\mathcal{S}$ 中——其存在性由 $\mathcal{Y}$ 可分这一事实保证（Parthasarathy[2005] 中的定理 2.1，第 27 页）。在理论方面，两种算法在相同的假设条件下都是普遍一致的，但考虑算法 1 时，证明更简单。

假设 1 如果 $\mathcal{S} \subsetneq \mathcal{Y}$，这可能会引发一致性问题，因为 $\bar{m}_{Fr}$ 可以取值于 $\mathcal{Y} \backslash \mathcal{S}$，而 $\bar{m}_n$ 则永远不会。因此，以下我们假设:

$$\bar{m}_{Fr} \in \mathcal{S} \; . \tag{7}$$

该假设意味着

$$\min_{y \in \mathcal{Y}} R(y) = \min_{y \in \mathcal{S}} R(y) \; .$$

现在我们可以陈述一致性结果。

定理 1 假设 (7) 成立。此外，假设对于所有 y、y'、$y'' \in \mathcal{Y}$，

$$\sup_{y, y' \in \mathcal{Y}} \ell(y, y') \leq L < \infty \; ,$$

$$|\ell(y, y') - \ell(y, y'')| \leq c \cdot \ell(y', y'') \tag{8}$$

对于某个有限常数 c，且

$$\ell(y, y') \leq \rho_{\mathcal{Y}}(y, y')^{\alpha} \tag{9}$$

其中 $\alpha > 0$。那么

$$\lim_n R(\bar{m}_n) = R(\bar{m}_{Fr}) \quad a.s. \tag{10}$$

3 Fr´échet 回归估计

为了估计 Fr´échet 回归函数 m_{Fr} （公式 3），我们给定一组n带标签训练数据:

$$\mathcal{D}_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\} \tag{11}$$

这些数据均独立同分布地抽取自同一分布。在本节中，我们关注于构造一个估计量 $\text{Est}_n = \text{Est}_n(\mathcal{D}_n)$: $\mathcal{X} \to \mathcal{Y}$，使其具有普遍强相合性。换言之，我们寻求满足以下条件的 Est_n：对于 (X,Y) 的每一个分布，都有 $\mathbb{P}\left\{ \lim_{n\to\infty} R(\text{Est}_n) = R\left(m_{Fr}\right) \right\} = 1$。

过去，Fr´échet 回归估计问题的研究主要集中在预测变量空间 $\mathcal{X}$ 或响应空间 $\mathcal{Y}$ 为非欧几里得空间的情形。当 $\mathcal{X}$ 是可分度量空间且 $\mathcal{Y} = \mathbb{R}$时，最常用的估计量是 k 近邻（k-NN）估计量。遗憾的是，k-NN 估计量并非普遍一致，参见 C´érou 和 Guyader [2006], 、Collins 等人 [2020] 以及 Kumari 和 Pestov [2024]。Hanneke 等人 [2021] 中的估计量 OptiNet 和 Gy¨örfi 与 Weiss[2021] 中的估计量 Proto-NN 都是普遍一致的回归估计量，参见下一节对 Proto-NN的扩展。

一般而言，对于预测变量空间 $\mathcal{X}$ 和响应空间 $\mathcal{Y}$ 均为度量空间这一更一般的情形，相关研究相对较少。例如，Hein [2009] 以及 Steinke 等人 [2010], 提出了一种基于 Nadaraya-Watson 方法的估计量，但只有在 $\mathcal{X}$ 和 $\mathcal{Y}$ 均为黎曼流形这一严格假设下，才证明了该估计量的一致性。Petersen 和 M¨üller [2019], 将 Fan 和 Gijbels [1996] 提出的局部线性回归扩展到考虑 $\mathcal{Y}$ 为度量空间的情形，但其框架假设 $\mathcal{X}$ 是 $\mathbb{R}^d$。此外，他们对逐点相合性的证明依赖于特定的分布假设，因此无法保证普适一致性。随后，Chen 和 M¨üller[2022] 通过给出一致收敛速率改进了这些结果，但其结论仍依赖于分布假设。最近，Capitaine 等人 [2024]提出了 Fr´échet 树和 Fr´échet 随机森林，Qiu 等人 [2024] 则针对预测变量空间 $\mathcal{X}$和响应空间 $\mathcal{Y}$均为度量空间的情形，提出了一种随机森林加权的局部 Fr´échet 回归方法。然而，他们的理论结果同样依赖于对分布的假设。

我们的回归估计量的构造：现在，我们给出 Fr´échet 回归函数 m_{Fr} 的普遍一致估计量。该估计量自然扩展了 Gy¨orfi 和 Weiss [2021] 提出的 Proto-NN 估计量。我们的方法与 Gy¨orfi 和 Weiss [2021] 的方法之间的主要区别在于，前一篇论文中的输出空间 $\mathcal{Y}$ 是实数值集合，而本文考虑的 $\mathcal{Y}$ 可以是任意可分巴拿赫空间。

估计量构造如下。对于整数 $k \geq 1$，我们假设除了带标签样本 $\mathcal{D}_n$ 外，还可以访问一组独立的无标签样本，称为原型，记作 $\mathcal{X}_k = \{X_1', \ldots, X_k'\}$，其中 X_i' 是 X 的独立同分布副本。设 $\mathcal{X}$ 的数据驱动划分 $\mathcal{P}_k$ 满足: $\mathcal{P}_k$ 是以核点集 $\mathcal{X}_k$ 为核点的 Voronoi 划分，即

$$\mathcal{P}_k = \{A_{k,1}, A_{k,2}, \ldots, A_{k,k}\} \tag{12}$$

使得 $A_{k,\ j}$ 是围绕核 X_j' 的 Voronoi 胞腔，

$$A_{k,j} = \left\{ x \in \mathcal{X} : j = \underset{1\leq i\leq k}{\arg\min}\, \rho_{\mathcal{X}}(x, X_i') \right\},$$

其中平局处理依据索引进行，即若 X_i' 和 X_j' 到 x的距离相等，则当 $i < j$时，判定 X_i' “更近”。本文假设并列发生的概率为零。

据此，我们可以定义条件期望$\mathbb{E}[\ell(Y,y) \mid X=x]$的一个估计量，并希望对 $\mathcal{Y}$进行最小化（见公式 3）。尽管一个直观的估计量可以是经验估计量

$$\frac{\sum_{i=1}^{n} \ell(Y_i,y)\mathbb{I}_{\{X_i=x\}}}{\sum_{i=1}^{n} \mathbb{I}_{\{X_i=x\}}},$$

需要注意的是，对于许多数据分布（例如任何连续随机变量 X），事件 $\{X=x\}$ 的概率测度可能为零，从而导致估计效果不佳，甚至无法计算。相反，我们建议用 $\mathbb{E}[\ell(Y,y) \mid X \in A_{k,\ j}]$来近似条件期望 $\mathbb{E}[\ell(Y,y) \mid X=x]$，适用于所有$x \in A_{k,\ j}$。最后，对于$x \in A_{k,\ j}$，我们取:

$$\frac{\sum_{i=1}^{n} \ell(Y_i,y)\mathbb{I}_{\{X_i\in A_{k,j}\}}}{\sum_{i=1}^{n} \mathbb{I}_{\{X_i\in A_{k,j}\}}},$$

将其作为方程 (3) 中条件期望的最终估计。估计 $m_{Fr}(x)$（其中$x \in A_{k,\ j}$）的一种自然方法是，寻找使上述经验估计最小化的 $y \in \mathcal{Y}$。一般来说，不可能在整个空间 $y \in \mathcal{Y}$ 上进行最小化，因此我们提出一种算法（称为算法 2），它在该空间的适当量化上进行最小化，与算法 1 的方法类似。

算法 2. 与上一节一样，设 $\mathcal{Y}_n$ 为原型集合 $Y_1', \ldots, Y_n'$。那么，我们的 Fr´echet 回归估计定义为

$$m_n(x) = \operatorname*{arg\,min}_{y\in\mathcal{Y}_n} \sum_{i=1}^{n} \ell(Y_i,y)\mathbb{I}_{\{X_i\in A_{k,j}\}}, \quad \text{if } x \in A_{k,j}.$$

因此，估计量 $m_n(\cdot)$ 是一个分段常值函数，定义在 Voronoi 胞腔$A_{k,\ 1}, \ldots, A_{k,\ k}$上，且取值于 $\mathcal{Y}_n$。注意，无论 $\mathcal{Y}_n$ 中原型的数量是多少， m_n 的取值都属于 μ_Y 的支撑集 $\mathcal{S}$——其存在性由 $\mathcal{Y}$ 可分（Parthasarathy [2005] 第 27 页定理 2.1）保证。

假设 2 与 Fr´echet 均值的情况一样，如果 $\mathcal{S} \subsetneq \mathcal{Y}$，这可能会引发一致性问题，因为 $m_{Fr}(x)$ 的取值可能属于 $\mathcal{Y}\backslash\mathcal{S}$，而 $m_n(x)$ 则绝不会如此。因此，以下我们假设对所有 $x \in \mathcal{X}$，

$$m_{Fr}(x) \in \mathcal{S}\ . \tag{13}$$

这是一个较弱的条件。例如，如果 $\mathcal{Y}$ 是一个向量空间且 $\mathcal{S}$ 是凸的，那么该条件便得到满足。条件 (13) 蕴含

$$\min_{y\in\mathcal{Y}} \mathbb{E}[\ell(Y,y) \mid X=x] = \min_{y\in\mathcal{S}} \mathbb{E}[\ell(Y,y) \mid X=x]\ .$$

现在可以陈述我们的主要一致性结果。

定理2 假设 (13) 成立， $\mathcal{Y}$ 是可分巴拿赫空间，并且对于 $\mathcal{X}_k$，并列发生的概率为零。假设满足定理 1 中关于代价函数 ℓ 的条件。若 $k=k_n$ 满足 $k_n \to \infty$ 且 $k_n \ln n/n \to 0$，则

$$\lim_n R(m_n) = R(m_{Fr}) \quad a.s. \tag{14}$$

定理2中的极限关系（14）提供了我们所期待的强一致性保证。这一结论具有普适性，因为它对数据的分布和 Fr´echet 回归函数 $m_{Fr}(\cdot)$ 均未作任何假设。有趣的是，即使在

风险 $R(m)$ 存在多个极小值这一困难情形下，一致性依然成立。唯一的假设是针对损失函数 ℓ 作出的。关于 L 的有界性假设相当标准，例如 Cohen and Kontorovich[2022] 将其称为有界直径假设。事实上，当 ℓ 是度量时，方程（8）和（9）中的另外两个假设并非必要。具体而言，根据反三角不等式，（8）可直接由 $c = 1$ 满足；（9）也同理，因为它在这里仅用于确保 ℓ 由一个连续函数一致有界。

4 证明

4.1 定理 1 的证明

注意， $\mathcal{Y}_n$ 是一个大小为 n 的有限集合，且 $\bar{m}_n$ 取值于 $\mathcal{Y}_n$，也就是说，如果我们将经验风险记为

$$R_n(y) := \frac{1}{n}\sum_{i=1}^{n} \ell(y, Y_i)\,,$$

那么，

$$\bar{m}_n = \arg\min_{y\in\mathcal{Y}_n} R_n(y)\,. \tag{15}$$

现在我们将超额风险分解为两项，即估计误差和近似误差：

$$R(\bar{m}_n) - R(\bar{m}_{Fr}) = R(\bar{m}_n) - \min_{y\in\mathcal{Y}_n} R(y) + \min_{y\in\mathcal{Y}_n} R(y) - R(\bar{m}_{Fr})\,.$$

控制估计误差 $R(\bar{m}_n) - \min_{y\in\mathcal{Y}_n} R(y)$

估计误差满足以下界：

$$\begin{aligned}
R(\bar{m}_n) - \min_{y\in\mathcal{Y}_n} R(y) &= \max_{y\in\mathcal{Y}_n}(R(\bar{m}_n) - R(y)) \\
&= \max_{y\in\mathcal{Y}_n}(R(\bar{m}_n) - R_n(\bar{m}_n) + R_n(\bar{m}_n) - R_n(y) + R_n(y) - R(y)) \\
&\le \max_{y\in\mathcal{Y}_n}(R(\bar{m}_n) - R_n(\bar{m}_n) + R_n(y) - R(y)) \\
&\le 2\max_{y\in\mathcal{Y}_n}|R_n(y) - R(y)|\,,
\end{aligned}$$

其中我们利用了：对于任意 $y \in \mathcal{Y}_n$，根据定义 (15)， $R_n(\bar{m}_n) - R_n(y) \le 0$。最后，结合并集界、霍夫丁不等式以及假设 $\ell(y, y') \le L$，可得对所有 $\varepsilon > 0$:

$$\begin{aligned}
\mathbb{P}\left(\max_{y\in\mathcal{Y}_n}|R_n(y) - R(y)| > \varepsilon \mid \mathcal{Y}_n\right) &\le n\max_{y\in\mathcal{Y}_n}\mathbb{P}(|R_n(y) - R(y)| > \varepsilon) \\
&\le 2ne^{-2n\varepsilon^2/L^2} \\
&= 2e^{-2n\varepsilon^2/L^2 + \ln n}\,.
\end{aligned}$$

因此，博雷尔-坎特利引理推出

$$R(\bar{m}_n) - \min_{y\in\mathcal{Y}_n} R(y) \to 0 \quad a.s.$$

并证明了估计误差的一致性。

控制近似误差 $\min_{y\in\mathcal{Y}_n} R(y) - R(\bar{m}_{Fr})$

根据条件（假设 1），近似误差 I_n 具有如下形式

$$I_n = R(\bar{m}_n) - R(\bar{m}_{Fr}) = \min_{y\in\mathcal{Y}_n} R(y) - \min_{y\in\mathcal{Y}} R(y) = \min_{y\in\mathcal{Y}_n} R(y) - \min_{y\in\mathcal{S}} R(y).$$

任务是证明

$$I_n = \min_{y\in\mathcal{Y}_n} R(y) - \min_{y\in\mathcal{S}} R(y) \to 0 \quad \text{a.s.}$$

对于任意 $0 < s$，可得

$$\begin{aligned}
\mathbb{P}\{I_n > s\} &= \mathbb{P}\left\{\min_{y\in\mathcal{Y}_n} R(y) - \min_{y\in\mathcal{S}} R(y) > s\right\} \\
&= \mathbb{P}\left\{\min_{i=1,\dots,n} R(Y_i') - \min_{y\in\mathcal{S}} R(y) > s\right\} \\
&= \mathbb{P}\left\{R(Y_1') - \min_{y\in\mathcal{S}} R(y) > s\right\}^n.
\end{aligned}$$

因此,

$$\sum_{n=1}^{\infty} \mathbb{P}\{I_n > s\} < \frac{1}{\mathbb{P}\{R(Y_1') - \min_{y\in\mathcal{S}} R(y) \le s\}} < \infty,$$

if

$$\mathbb{P}\left\{R(Y_1') - \min_{y\in\mathcal{S}} R(y) \le s\right\} > 0.$$

令 $\tilde{y} = \arg\min_{y\in\mathcal{S}} R(y)$，则

$$\begin{aligned}
\mathbb{P}\left\{R(Y_1') - \min_{y\in\mathcal{S}} R(y) \le s\right\} &= \mathbb{P}\left\{\mathbb{E}[\ell(Y, Y_1') \mid Y_1'] - \mathbb{E}[\ell(Y, \tilde{y})] \le s\right\} \\
&\ge \mathbb{P}\left\{c \cdot \ell(Y_1', \tilde{y}) \le s\right\} \\
&\ge \mathbb{P}\left\{\rho_{\mathcal{Y}}(Y_1', \tilde{y}) \le (s/c)^{1/\alpha}\right\} \\
&> 0
\end{aligned}$$

其中，前两个不等式用到了 (8) 和 (9)，最后一个不等式成立是因为支撑集 $\mathcal{S}$的性质。证明至此结束。

□

4.2 定理2的证明

我们首先引入一类函数集，它们的取值属于 $\mathcal{Y}_n$ ，并且在沃罗诺伊单元$\{A_{k,j}\}_{j=1}^k$上为常数:

$$\mathcal{F}_{n,k} := \mathcal{F}(\mathcal{X}_k, \mathcal{Y}_n) := \left\{ f : f(x) = \sum_{j=1}^{k} y_j \mathbb{I}\{x \in A_{k,j}\}, y_j \in \mathcal{Y}_n \right\}.$$

注意，由于我们假设不存在平局，因此有 k 个不同的 Voronoi 胞腔。因此，$\mathcal{F}_{n,\,k}$ 是一个大小为 n^k 的有限集合，并且 m_n 属于 $\mathcal{F}_{n,\,k}$，也就是说，如果我们将经验风险记为

$$R_n(f) := \frac{1}{n}\sum_{i=1}^{n} \ell(f(X_i), Y_i),$$

那么，

$$m_n = \arg\min_{f \in \mathcal{F}_{n,k}} R_n(f) \ . \tag{16}$$

现在我们将超额风险分解为两项，即估计误差和近似误差：

$$R(m_n) - R(m_{Fr}) = R(m_n) - \min_{f \in \mathcal{F}_{n,k}} R(f) + \min_{f \in \mathcal{F}_{n,k}} R(f) - R(m_{Fr}) \ .$$

证明分为两部分。一部分证明估计误差$R(m_n) - \min_{f \in \mathcal{F}_{n,\,k}} R(f)$几乎必然趋于 0；另一部分证明近似误差 $\min_{f \in \mathcal{F}_{n,\,k}} R(f) - R(m_{Fr})$也几乎必然趋于 0。近似误差的控制也分为两项，因为如后文所述，会出现两类近似误差：一类源于使用 $\mathcal{Y}_n$ 中的原型对 $\mathcal{Y}$ 空间进行离散化，另一类源于使用 Voronoi 划分对 $\mathcal{X}$ 空间进行近似。

控制估计误差 $R(m_n) - \min_{f \in \mathcal{F}_{n,\,k}} R(f)$

估计误差的上界如下：

$$\begin{aligned}
& R(m_n) - \min_{f \in \mathcal{F}_{n,k}} R(f) \\
&= \max_{f \in \mathcal{F}_{n,k}} \left(R(m_n) - R(f) \right) \\
&= \max_{f \in \mathcal{F}_{n,k}} \left(R(m_n) - R_n(m_n) + R_n(m_n) - R_n(f) + R_n(f) - R(f) \right) \\
&\leq \max_{f \in \mathcal{F}_{n,k}} \left(R(m_n) - R_n(m_n) + R_n(f) - R(f) \right) \\
&\leq 2 \max_{f \in \mathcal{F}_{n,k}} \left| R_n(f) - R(f) \right| ,
\end{aligned}$$

其中我们利用了如下事实：对于任意 $f \in \mathcal{F}_{n,\,k}$，根据定义 (16)，$R_n(m_n) - R_n(f) \leq 0$。最后，由并集界、霍夫丁不等式以及假设 $\ell(y, y') \leq L$可知，对于所有 $\varepsilon > 0$:

$$\begin{aligned}
\mathbb{P}\left(\max_{f \in \mathcal{F}_{n,k}} |R_n(f) - R(f)| > \varepsilon \mid \mathcal{X}_k, \mathcal{Y}_n \right) &\leq |\mathcal{F}_{n,k}| \cdot \max_{f \in \mathcal{F}_{n,k}} \mathbb{P}(|R_n(f) - R(f)| > \varepsilon) \\
&\leq 2n^k e^{-2n\varepsilon^2/L^2} \\
&= 2e^{-2n\varepsilon^2/L^2 + k \ln n} \ .
\end{aligned}$$

因此，条件 $k_n \ln n/n \to 0$ 连同博雷尔-坎特利引理意味着

$$R(m_n) - \min_{f \in \mathcal{F}_{n,k}} R(f) \to 0 \quad a.s.$$

并证明了估计误差的一致性。

控制近似误差 $\min_{f \in \mathcal{F}_{n,\,k}} R(f) - R(m_{Fr})$。

首先请注意，给定一个沃罗诺伊单元 $A_{k_n,\ j}$，当观测到一个新点 X'_{k_n+1} 时，该单元$A_{k_n,\ j}$ 要么保持不变，要么被划分为若干个新的沃罗诺伊单元。这意味着 $\mathcal{F}_{n,\ k_n} \subseteq \mathcal{F}_{n+1,\ k_n}$、$\mathcal{F}_{n,\ k_n} \subseteq \mathcal{F}_{n,\ k_{n+1}}$以及 $\mathcal{F}_{n,\ k_n} \subseteq \mathcal{F}_{n+1,\ k_{n+1}}$。因此，对于任意固定的 k,

$$\min_{f \in \mathcal{F}_{n,k}} R(f) \downarrow$$

作为 $n \to \infty$，对于任意固定的 n，

$$\min_{f \in \mathcal{F}_{n,k}} R(f) \downarrow$$

当 $k \to \infty$ 时。这意味着对于任意 $k_n \uparrow$，

$$\min_{f \in \mathcal{F}_{n,k_n}} R(f) \downarrow \tag{17}$$

当 $n \to \infty$ 时。

我们需要证明

$$\lim_n \left(\min_{f \in \mathcal{F}_{n,k_n}} R(f) - R(m_{Fr}) \right) = 0 \qquad a.s. \tag{18}$$

由 (17) 可知，$\min_{f \in \mathcal{F}_{n,k_n}} R(f) - R(m_{Fr})$ 几乎必然收敛到一个非负随机变量。由于期望为 0 的非负随机变量几乎必然为 0，要证明 (18)，只需证明该极限的期望等于 0。根据控制收敛定理，我们有：

$$\mathbb{E}\left\{ \lim_n \left(\min_{f \in \mathcal{F}_{n,k_n}} R(f) - R(m_{Fr}) \right) \right\} = \lim_n \left(\mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k_n}} R(f) \right\} - R(m_{Fr}) \right) .$$

因此，任务 (18) 可归结为证明

$$\lim_n \left(\mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k_n}} R(f) \right\} - R(m_{Fr}) \right) = 0 . \tag{19}$$

可得

$$\lim_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k_n}} R(f) \right\} \geq \inf_{k \in \mathbb{N}} \inf_{n \in \mathbb{N}} \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k}} R(f) \right\}.$$

对于任意固定的 k，单调性意味着

$$\lim_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k_n}} R(f) \right\} \leq \lim_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k}} R(f) \right\} = \inf_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k}} R(f) \right\}.$$

因此，

$$\lim_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k_n}} R(f) \right\} \leq \inf_k \inf_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k}} R(f) \right\}.$$

由此，

$$\lim_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k_n}} R(f) \right\} = \inf_k \inf_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k}} R(f) \right\}. \tag{20}$$

结合此等式与 (19)，只需证明

$$\lim_k \lim_n \mathbb{E}\left\{ \min_{f \in \mathcal{F}_{n,k}} R(f) \right\} - R(m_{Fr}) = 0. \tag{21}$$

现在，我们来介绍

$$\mathcal{F}_{\infty,k} := \left\{ f : f(x) = \sum_{j=1}^{k} y_j \mathbb{I}\{x \in A_{k,j}\}, y_j \in \mathcal{S} \right\} .$$

由此可得

$$\mathbb{E}\left\{\min_{f\in\mathcal{F}_{n,k}} R(f)\right\} - R(m_{Fr}) = I_{n,k} + J_k$$

其中

$$I_{n,k} = \mathbb{E}\left\{\min_{f\in\mathcal{F}_{n,k}} R(f)\right\} - \mathbb{E}\left\{\min_{f\in\mathcal{F}_{\infty,k}} R(f)\right\}$$

and

$$J_k = \mathbb{E}\left\{\min_{f\in\mathcal{F}_{\infty,k}} R(f)\right\} - R(m_{Fr})\ .$$

控制 $I_{n,k}$:

回顾一下， $\mathcal{Y}_n := \{Y'_1, \ldots, Y'_n\}$ 并记作

$$y^*_{k,j} = \arg\min_{y\in\mathcal{S}} \mathbb{E}[\ell(Y,y)\mathbb{I}\{X\in A_{k,j}\}].$$

Put

$$G(y,x) = \mathbb{E}\{\ell(Y,y) \mid X = x\}\ .$$

对于函数 $f(x) = \sum_{j=1}^k y_j \mathbb{I}\{x\in A_{k,j}\}$,

$$\begin{aligned} R(f) &= \int_{\mathcal{X}} \mathbb{E}\{\ell(Y,f(x)) \mid X = x\}\mu(dx) \\ &= \int_{\mathcal{X}} G(f(x),x)\mu(dx) \\ &= \sum_{j=1}^k \int_{A_{k,j}} G(y_j,x)\mu(dx)\ . \end{aligned}$$

Then

$$\begin{aligned} &\min_{f\in\mathcal{F}_{n,k}} R(f) - \min_{f\in\mathcal{F}_{\infty,k}} R(f) \\ &= \min_{f\in\mathcal{F}_{n,k}} \mathbb{E}[\ell(Y,f(X))] - \min_{f\in\mathcal{F}_{\infty,k}} \mathbb{E}[\ell(Y,f(X))] \\ &= \sum_{j=1}^k \min_{y\in\mathcal{Y}_n} \int_{A_{k,j}} G(y,x)\mu(dx) - \sum_{j=1}^k \min_{y\in\mathcal{S}} \int_{A_{k,j}} G(y,x)\mu(dx) \\ &= \sum_{j=1}^k \left(\min_{y\in\mathcal{Y}_n} \int_{A_{k,j}} G(y,x)\mu(dx) - \int_{A_{k,j}} G(y^*_{k,j},x)\mu(dx)\right)\ , \end{aligned}$$

结合定理的条件可得

$$\begin{aligned} I_{n,k} &\le c\sum_{j=1}^k \mathbb{E}\left[\min_{y\in\mathcal{Y}_n} \ell(y,y^*_{k,j})\mu(A_{k,j})\right] \\ &= ck\mathbb{E}\left[\min_{y\in\mathcal{Y}_n} \ell(y,y^*_{k,1})\mu(A_{k,1})\right] \end{aligned}$$

$$\leq ck\mathbb{E}\left[\min\{L, \min_{y\in\mathcal{Y}_n} \rho_{\mathcal{Y}}(y, y^*_{k,1})^\alpha\}\mu(A_{k,1})\right]$$
$$= ck\mathbb{E}\left[\min\left\{L, \rho_{\mathcal{Y}}(Y'_{(1,n)}(y^*_{k,1}), y^*_{k,1})^\alpha\right\}\mu(A_{k,1})\right],$$

其中，$Y'_{(1,n)}(y^*_{k,\ 1})$ 是 $y^*_{k,\ 1}$ 在 $\mathcal{Y}_n$ 中的最近邻。由于 $\mathcal{Y}_n$ 是可分度量空间，且 $y^*_{k,\ 1} \in \mathcal{S}$，根据 Cover-Hart 引理 [Cover 和 Hart, 1967], 我们知道 $\rho_{\mathcal{Y}}(Y'_{(1,\ n)}(y^*_{k,\ 1})\, y^*_{k,\ 1}) \to 0$ 几乎必然成立。根据控制收敛定理，这意味着对于任意固定的 k

$$\lim_n I_{n,k} = 0\ ,$$

因此 $\lim_k \lim_n I_{n,k} = 0$ 几乎必然。

J_k的控制:

根据条件，$\mathcal{Y}$ 是一个可分巴拿赫空间。因此，结合标准的 Lusin 定理 [Dudley, 2018, Theorem 7.5.2] 与 Dugundji 延拓定理 [Dugundji,1951], 我们可以将 Lusin 定理表述为以下一般形式：对于任意 $\delta > 0$，存在连续函数 $\widehat{m}:\ \mathcal{X} \to \mathcal{Y}$ 以及闭集 $K \subset \mathcal{X}$，使得 $\mu(K^c) < \delta$，其中$K^c:= \mathcal{X}\backslash K$，并且函数 $\widehat{m}$与 m_{Fr}在 K上一致。事实上，Lusin 定理给出了从 K 到 $\mathcal{Y}$的连续函数 $\widehat{m}$ 。若 $\mathcal{Y}$ 是局部凸拓扑向量空间（巴拿赫空间即属此类），则 Dugundji 定理允许将 $\widehat{m}$延拓为一个从 $\mathcal{X}$ 到 $\mathcal{Y}$ 的连续函数。

对于记号

$$f^*_k(x) = \widehat{m}(X'_{(1,k)}(x)) = \widehat{m}(X'_j), \quad \text{if} \quad x \in A_{k,j},$$

可得

$$\begin{aligned}
J_k &= \mathbb{E}\left\{\min_{f\in\mathcal{F}_{\infty,k}} R(f)\right\} - R(m_{Fr}) \\
&\leq \mathbb{E}\{R(f^*_k)\} - R(m_{Fr}) \\
&= \mathbb{E}\left\{\int G(\widehat{m}(X'_{(1,k)}(x)), x)\mu(dx)\right\} - \int G(m_{Fr}(x), x)\mu(dx) \\
&\leq \left|\mathbb{E}\left\{\int G(\widehat{m}(X'_{(1,k)}(x)), x)\mu(dx)\right\} - \int G(\widehat{m}(x), x)\mu(dx)\right| \\
&\quad + \left|\int G(\widehat{m}(x), x)\mu(dx) - \int G(m_{Fr}(x), x)\mu(dx)\right|.
\end{aligned}$$

第二项很容易控制:

$$\begin{aligned}
\left|\int G(\widehat{m}(x), x)\mu(dx) - \int G(m_{Fr}(x), x)\mu(dx)\right| &\leq c\int \ell(\widehat{m}(x), m_{Fr}(x))\mu(dx) \\
&= c\int_{K^c} \ell(\widehat{m}(x), m_{Fr}(x))\mu(dx) \\
&\leq cL\mu(K^c) \\
&\leq cL\delta.
\end{aligned}$$

对于第一项，Cover-Hart 引理、$\widehat{m}$ 的连续性以及控制收敛定理意味着

$$\left|\mathbb{E}\left\{\int G(\widehat{m}(X'_{(1,k)}(x)), x)\mu(dx)\right\} - \int G(\widehat{m}(x), x)\mu(dx)\right|$$

$$\leq c \int \mathbb{E}\Big\{\min\Big\{L, \rho_{\mathcal{Y}}(\widehat{m}(X'_{(1,k)}(x)), \widehat{m}(x))^{\alpha}\Big\}\Big\}\mu(dx)$$
$$\to 0 \, ,$$

当 $k \to \infty$ 时。因此，

$$\limsup_k J_k \leq cL\delta \, .$$

证明至此结束。 □

致谢

P.H. 谨向索邦大学的 Emergence 项目 MARS 致谢。

参考文献

Adil Ahidar-Coutrix、Thibaut Le Gouic 和 Quentin Paris。《度量空间中经验重心的收敛速率：曲率、凸性与可延拓测地线》。《概率论及相关领域》，177(1):323–368，2020年。

Andrea Aveni 和 Sayan Mukherjee。《广义 Fréchet 均值的一致性》 arXiv 预印本，arXiv:2408.07534，2024。

Daniel Bartl、Gabor Lugosi、Roberto Imbuzeiro Oliveira 和 Zoraida F Rico。度量空间中的稳健次高斯均值估计量。arXiv 预印本 arXiv:2509.13606，2025。

Anatole Beck。关于大数强律。载于《遍历理论：国际研讨会论文集》，第 21–53 页。学术出版社，纽约，1963年。

Rabi Bhattacharya 和 Vic Patrangenaru。流形上内在与外在样本均值的大样本理论。《统计年刊》，31(1)：1–29，2003年。

Rabi Bhattacharya 和 Vic Patrangenaru。流形上内在与外在样本均值的大样本理论——II。《统计年刊》，33，2005。

Victor-Emmanuel Brunel 和 Jordan Serres。度量空间中经验重心的集中性。载于《国际算法学习理论会议》，第 337–361 页。PMLR，2024。

Louis Capitaine、J´er´emie Bigot、Rodolphe Thi´ebaut 和 Robin Genuer。针对非欧氏预测变量的度量空间值回归的 Fr´echet 随机森林。《机器学习研究期刊》，25(355):1–41，2024。

Fr´ed´eric C´erou 和 Arnaud Guyader。无限维最近邻分类。ESAIM：概率与统计，10:340–355，2006。

Yaqing Chen 和 Hans-Georg M¨uller。局部 Fr´echet 回归的一致收敛及其在度量空间值轨迹的极值定位与时间扭曲中的应用。统计年刊，50(3):1573–1592，2022。

Dan Tsir Cohen 和 Aryeh Kontorovich。度量损失学习。载于《学习理论会议》，第 662–700 页。PMLR，2022。

Benoˆıt Collins、Sushma Kumari 和 Vladimir G Pestov。度量空间中 k-近邻规则的普适一致性与 Nagata维数。ESAIM：概率与统计，24：914–934，2020。Thomas Cover 和 Peter Hart。最近邻模式分类。IEEE信息论汇刊，13(1)：21–27，1967。Richard M. Dudley。实分析与概率论。Chapman and Hall/CRC，2018。James Dugundji。Tietze 定理的一种扩张。太平洋数学杂志，1951。Steven N Evans 和 Adam Q Jaffe。Fr´echet 均值集合的极限定理。Bernoulli，30(1)：419–447，2024。范剑青和 Irene Gijbels。《局部多项式建模及其应用：统计学与应用概率专著 66》，第 66 卷。CRC出版社，1996。Maurice Fr´echet。度量空间中任意性质的随机元。载于《亨利·庞加莱研究所年刊》，第 10 卷，第 215–310 页，1948。Cedric E Ginestet。图值随机变量的 Fr´echet 样本均值集合的强一致性。arXiv 预印本 arXiv:1204.3183，2012。Cedric E Ginestet、Jun Li、Prakash Balachandran、Steven Rosenberg 和 Eric D Kolaczyk。功能性神经影像中的网络数据假设检验。应用统计学年刊，第 725–750 页，2017。L´aszl´o Gy¨orfi 和 Roi Weiss。度量空间中多类原型算法的普适一致性与收敛速率。《机器学习研究杂志》，22(151)：1–25，2021。Steve Hanneke、Aryeh Kontorovich、Sivan Sabato 和 Roi Weiss。度量空间中的普适贝叶斯一致性。统计年刊，49：2129–2150，2021。Matthias Hein。针对度量空间值输出的稳健非参数回归。神经信息处理系统进展，22，2009。Thomas Hotz、Huiling Le 和 Andrew TA Wood。光滑紧黎曼流形上内禀 Fréchet 均值的中心极限定理。《概率论及相关领域》，189(3)：1219–1246，2024。Daniel Hsu 和 Sivan Sabato。重尾下的损失最小化与参数估计。《机器学习研究杂志》，17(18)：1–40，2016。Adam Quinn Jaffe。无限维中的 Fr´echet 均值。arXiv 预印本 arXiv:2410.17214，2024。Jakwang Kim、Jiyoung Park 和 Anirban Bhattacharya。度量空间中的稳健估计：借助 Fr\’echet 中位数实现指数集中。arXiv 预印本 arXiv:2504.14161，2025。Roger Koenker 和 Kevin F Hallock。分位数回归。《经济展望杂志》，15(4)：143–156，2001。Eric D Kolaczyk、Lizhen Lin、Steven Rosenberg、Jackson Walters 和 Jie Xu。无标签网络的均值：几何刻画与渐近行为。统计学年刊，2020。

Sushma Kumari 和 Vladimir G Pestov。度量空间中 k近邻规则的普适一致性与 Nagata 维数。II。ESAIM：概率与统计，28:132–160，2024。Thibaut Le Gouic、Quentin Paris、Philippe Rigollet 和 Austin J Stromme。Alexandrov空间与 Wasserstein空间中经验重心的快速收敛。欧洲数学会期刊，25(6):2229–2250，2022。G´abor Lugosi 和 Marcos Matabuena。度量空间中的不确定性量化。arXiv 预印本 arXiv:2405.05110，2024。Sim´on Lunag´omez、Sofia C Olhede 和 Patrick J Wolfe。通过图距离对网络总体建模。美国统计协会期刊，116(536):2023–2040，2021。P Milasevic 和 GR Ducharme。空间中位数的唯一性。统计学年刊，15(1): 1332–1333，1987。Khai Nguyen、Yang Ni 和 Peter Mueller。贝叶斯多元密度-密度回归。arXiv 预印本 arXiv:2601.02640，2026。Kalyanapuram Rangachari Parthasarathy。《度量空间上的概率测度》，第352卷。美国数学会，2005。Alexander Petersen 和 Hans-Georg M¨uller。具有欧氏预测变量的随机对象 Fr´echet 回归。统计年刊，47:691–719，2019。Rui Qiu、Zhou Yu 和 Ruoqing Zhu。针对随机对象的随机森林加权局部 Fr´echet 回归。《机器学习研究杂志》，25(107):1–69，2024。Christof Sch¨otz。通过四元组不等式得到广义 Fr´echet 均值的收敛速率。电子统计学期刊，13:4280–4345，2019。Christof Sch¨otz。Fr´echet 均值集推广形式的强大数定律。统计学，56(1):34–52，2022。Katie E Severn、Ian L Dryden 和 Simon P Preston。网络样本的流形值数据分析及其在语料库语言学中的应用。应用统计学年刊，16(1):368–390，2022。Florian Steinke、Matthias Hein 和 Bernhard Sch¨olkopf。一般黎曼流形之间的非参数回归。SIAM成像科学期刊，3(3):527–563，2010。Harald Sverdrup-Thygeson。紧度量空间中随机变量集中趋势与离散度测度的强大数定律。统计年刊，页码 141–145，1981。Ho Yun 和 Byeong U Park。非正曲率空间中均值的几何中位数的指数集中。Bernoulli，29(4):2927–2960，2023。Yidong Zhou 和 Hans-Georg M¨uller。利用图拉普拉斯矩阵进行网络回归。《机器学习研究杂志》，23(320):1–41，2022。Herbert Ziezold。关于拟度量空间中随机元素的期望图形与强大数定律。载于《第七届布拉格信息论、统计决策函数、随机过程会议暨1974年欧洲统计学家会议论文集：1974年8月18日至23日在布拉格举行》，页码 591–602。Springer，1977。