非单连通闭黎曼流形上的高斯核永远不是正定的

SIRAN LI

摘要。我们证明，对于任意非单连通闭黎曼流形 $(\mathcal{M}, g)$，其高斯核 $\exp\{-\lambda d_g^2(\bullet,\bullet)\}$ 都不是正定的，其中 d_g 为测地距离，且这一结论对任意 $\lambda > 0$均成立。我们的证明结合了 Da Costa–Mostajeran–Ortega 最近预印本 [9] 中的分析以及黎曼几何中的经典比较定理。

arXiv:2303.06558v1 [math.DG] 12 Mar 2023

0. 主结果

本文探讨通过黎曼流形上的测地距离定义的标准高斯核的数学性质。我们的主要结果如下：

主定理。设 $\mathcal{M}$ 为任意紧致、无边界的非单连通 $\mathcal{C}^\infty$-流形。对于 $\mathcal{M}$ 上的任意黎曼度量 g 以及任意参数 $\lambda > 0$，高斯核 $e^{-\lambda d_g^2(\bullet,\bullet)} : \mathcal{M} \times \mathcal{M} \to [0, \infty[$ 都不是正定的。这里，d_g 是 $(\mathcal{M}, g)$上的测地距离。

证明的关键依托于 Da Costa–Mostajeran–Ortega 最近的一项出色工作 [9]，该工作证明了上述定理对 $\mathcal{M} = $ 圆周 $\varrho\mathbb{S}^1$ 成立，适用于任意 $\varrho > 0$，因此也适用于球面、实射影空间、实格拉斯曼流形等。将 [9] 中的论证与流形上的经典变分法技巧相结合，我们便能得出该定理。

1. 背景与预备知识

1.1. 流形学习与核方法的交汇。近年来，机器学习领域中，专注于数据非欧几何特征的流形学习备受关注。另请参见 [1, 4, 27] 以及许多其他参考文献。我们主要关注它与核方法的联系；核方法在机器学习及其数学理论中同样发挥着重要作用。[10] Cucker–Smale 对核方法的数学基础作了详细综述。关于核方法在支持向量机、核主成分分析、受控随机微分方程和储备池计算等各类问题中的应用，可参见 [15,18,25,27,28] 及其中引用的大量文献。

现在考虑一个更一般的情形：(X, d) = 度量空间。对于应用于度量空间的核方法，其核心数学结构是一个对称函数 $k : X \times X \to \mathbb{R}$，称为核，它取决于观测之间的度量距离。在$(X, d)$上广泛使用的一类核是

$$k(x, y) := e^{-\lambda[d(x,y)]^q} \qquad \text{for } \lambda, q > 0.$$

参见 Chapelle–Haffner–Vapnik [7] 和 Feragen–Lauze–Hauberg [12]。

日期：2023年3月14日。2020年数学主题分类。46E22、46N30、68R12、62B11。关键词和短语。高斯核；核方法；等距嵌入；闭测地线；再生核希尔伯特空间（再生核希尔伯特空间）；非单连通流形。

当 $q=2$ 时，这种 k 被称为 (X,d) 上的 高斯核 。

任意完备黎曼流形（$\mathcal{M},g$）都具有度量空间结构，且其度量空间拓扑与流形拓扑相容。关于黎曼度量 g，$\mathcal{M}$ 上的度量——即距离函数——定义为

$$d_g(x,y) := \inf\left\{\gamma\colon [0,L]\to\mathcal{M}\colon \gamma\text{是一条分段 } \mathcal{C}^1 \text{ 曲线，满足 } \gamma(0)=x,\ \text{且 } \gamma(1)=y\right\}.$$

此处将“分段 $\mathcal{C}^1$”替换为“分段 $\mathcal{C}^\infty$”、“$\mathcal{C}^\infty$”或“$\mathcal{C}^1$”均无影响。度量空间（$\mathcal{M},d_g$）的完备性等价于黎曼流形（$\mathcal{M},g$）的测地完备性。参见例如 [5, 霍普夫–里诺定理，§7]。

1.2. **正定核**。在机器学习（ [10,12,15,18,25,27,28]）应用中，通常要求核函数将数据映射到再生核希尔伯特空间（RKHS）中，而这又意味着核的正定性。参见 Schölkopf–Smola [27]。

定义 1.1. $k\colon (X,d)\times(X,d)\to\mathbb{R}$ 若对任意 $N\in\mathbb{N}$ 和 $\{x_1,\dots,x_N\}\subset X$，格拉姆矩阵 $\mathcal{G}=\{\mathcal{G}_{ij}\}\in \mathrm{Mat}(N\times N;\mathbb{R})$（其中 $\mathcal{G}_{ij}:=k(x_i,x_j)$）是半正定的，则称其为正定。

这种核函数正定性的概念由 Mathias 和 Bochner 提出。参见 [2,26]。

正定核在泛函分析中也具有相当重要的理论意义。Schoenberg 的一条经典定理（ [26]）早在 1938年 就指出，可分度量空间（X,d）能够等距嵌入希尔伯特空间 $L^2(]0,1[)$，当且仅当对每个 $\lambda>0$，高斯核 $e^{-\lambda d^2(\bullet,\bullet)}$ 都是正定的。因此可以发现，$L^2(]0,1[)$包含 $\mathbb{R}^n$ 的副本，即配备弗罗贝尼乌斯范数／对数欧几里得度量的对称正定矩阵空间 Sym_n^{++}，以及配备射影度量的实格拉斯曼流形 $\mathrm{Gr}(k,n)$ 等。

定义 1.2。 设(X,d) 和(X',d') 为度量空间。若映射 $\iota:\ X\to X'$ 是一一映射，且对所有 $x,y\in X$，都有 $\iota(x)\ \iota(y)\big)=d(x,y)$，则称其为等距嵌入。

对于完备黎曼流形$(\mathcal{M},g)$ 和$(\mathcal{M}',g')$，若映射 $\iota:\ \mathcal{M}\to\mathcal{M}'$ 是 $\mathcal{C}^\infty$-流形范畴中的一个嵌入，并且$\iota\colon(\mathcal{M},d_g)\to\left(\mathcal{M},d_{g'}\right)$ 是度量空间之间的等距嵌入，则称其为度量空间等距嵌入。

注 1.3。 对于黎曼流形，按度量空间意义定义的等距嵌入与按微分几何意义定义的等距嵌入截然不同。参见附录 *A*。

另一方面，文献中也报道了各种负面结果。将高斯核在 (X,d) 上正定性的范围记为

$$\Lambda_+(X) := \left\{\lambda\in]0,\infty[\colon \text{the Gaussian kernel } e^{-\lambda d^2(\bullet,\bullet)} \text{ is positive definite}\right\}.$$

Feragen–Lauze–Hauberg [12] 证明，对于任意黎曼流形 $(\mathcal{M},g)$，范围 $\Lambda_+(\mathcal{M},d_g)$ 不可能是整个 $]0,\infty[$，除非 g 是欧几里得空间[1]。因此，根据 Schoenberg 定理 [26]，，只有平坦黎曼流形才能以等距方式嵌入度量空间意义下的 $L^2(]0,1[)$ 中。

然而，这仍留下一个重要问题：哪些黎曼流形具有较大的 $\Lambda_+\subset]0,\infty[$? 通过对 $\mathbb{S}^1$，上等分布点所对应的高斯核进行显式计算，Da Costa–Mostajeran–Ortega 在近期的预印本 [9]，中证明了以下结果

[1]事实上，[2，定理 1] 证明了一个更一般的结果：具有正定高斯核的测地度量空间只能属于 CAT(0) 类。 te 高斯核 属于 CAT(0) 类。

定理 1.4。 以下黎曼流形上的高斯核的正定范围为空（即 $\Lambda_+ = \emptyset$），且均配备典范度量：

- $\mathbf{S}^n$ *with* $n \geq 1$;
- *the real project spaces* $\mathbb{R}\mathbf{P}^n$ *with* $n > 1$;
- *the real Grassmannians* $\mathbf{Gr}(k, n)$ *with* $1 \leq k < n$;
- *the hyperbolic hyperboloid; and*
- *the torus* $\mathbf{T}(a, b) = \left\{ \begin{bmatrix} (a + b\cos\theta)\cos\phi \\ (a + b\cos\theta)\sin\phi \\ b\sin\theta \end{bmatrix} \in \mathbb{R}^3 : 0 \leq \theta, \phi < 2\pi \right\}$ *for any* $0 < a < b$.

实际上，我们在 §0 中的主定理为这样一类黎曼流形大大扩展了名单：对于任意 $\lambda > 0$, 其高斯核都不是正定的。若一个流形是紧致且无边界的，则称其为闭流形。将我们的结果与上述定理 1.4 结合，可推得

推论 1.5。 $\Lambda_+(\mathcal{M},\ g) = \emptyset$ 对于以下**闭**黎曼流形成立：

- 任意非单连通$(\mathcal{M},\ g)$;
- $\mathbf{S}^n$ 满足 $n \geq 2$;
- 实格拉斯曼流形$\mathbf{Gr}(k,\ n)$ 满足 $1 \leq k < n$。

鉴于 Istas 的以下结果 [17, 定理 2.12]，高斯核代表了一个临界或边界情形。另见 [12, 定理 3]。

命题 1.6. 对于任意完备黎曼流形 $(\mathcal{M}, g)$ 以及任意 $q > 2$，存在某个 $\lambda > 0$，使得核 $k(x, y) := e^{-\lambda[d_g(x,y)]^q}$ 不是正定的。

1.3. **进一步讨论**。 [9] 中指出，$\Lambda_+\left(\mathbf{S}^1\right) = \emptyset$ ——因此定理 1.4——可以通过 Wood [31] 和 Gneiting [14]的其他论证推导得出。尽管如此， [9] 中直接进行的傅里叶理论计算尤其适用于推广到更一般的黎曼流形类，本文将对此进行说明。

另一方面， [9] 表明，对于任意闭黎曼流形，都可以通过证明任何这样的 $(\mathcal{M}, g)$ 在度量空间意义下都包含圆周 $\mathbf{S}^1$ 的等距嵌入，来推导 $\Lambda_+(\mathcal{M}, g) = \emptyset$ 。事实上，这似乎对 $(\mathcal{M}, g)$ 施加了相当严格的限制。参见附录 A 中的定理。不过，目前我们既无法给出反例，也无法证实 [9]中的这一推测。

还需指出，在数学、统计学和机器学习的理论研究中，人们研究了定义在 $\mathbf{S}^n$ 及更一般的黎曼流形上的高斯核变体。参见 [3, 11, 12, 14] 及其他许多文献。

2. 当 $\pi_1(\mathcal{M}) \neq \{0\}$ 时，高斯核的非正定性

In t本节证明我们的主定理，该定理已在 §0 中陈述，现重述如下：

定理 2.1. 设 $\mathcal{M}$ 是任意闭（即，紧致且无边界）可微流形，满足 $\pi_1(\mathcal{M}) \neq \{0\}$。对于 $\mathcal{M}$ 上的任意黎曼度量 g 及任意参数 $\lambda > 0$，高斯核 $e^{-\lambda d_g^2(\bullet,\bullet)} : \mathcal{M} \times \mathcal{M} \to [0, \infty[$ 不可能是正定的。

也就是说，我们证明对于非单连通闭流形（$\mathcal{M}, g$），

$$\Lambda_+(\mathcal{M}, g) := \left\{\lambda \in [0, \infty[:\ e^{-\lambda d_g^2(\bullet,\bullet)} : \mathcal{M} \times \mathcal{M} \to [0, \infty[\text{ is positive definite}\right\} = \emptyset.$$

证明的关键是以下线性代数引理，该引理由 Da Costa–Mostajeran–Ortega 通过显式计算得出 [9,第 5 页–7]：

引理 2.2 (见 [9])。 给定任意 $\lambda > 0$。该循环矩阵

$$\mathcal{K} = \begin{bmatrix} 1 & \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) & \exp\left(-\lambda\left(\frac{4\pi}{N}\right)^2\right) & \cdots & \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) \\ \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) & 1 & \cdots & \cdots & \exp\left(-\lambda\left(\frac{4\pi}{N}\right)^2\right) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) & \exp\left(-\lambda\left(\frac{4\pi}{N}\right)^2\right) & \exp\left(-\lambda\left(\frac{6\pi}{N}\right)^2\right) & \cdots & 1 \end{bmatrix}$$

当某个足够大的自然数 N 可被整除时，具有一个负特征值 *4*.

文献 [9] 中利用引理 2.2 证明：对于任何允许圆周 S^1 等距嵌入的度量空间等距嵌入的流形，高斯核都不是正定的。我们的关键观察是，引理 2.2 与黎曼几何中的经典比较定理（例如见 例如 [5, 8]）结合起来，使我们能够推导出主定理 2.1。

本文中，圆周指配备典范测度的 S^1，即通过将 $\mathsf{S}^1 \subset \mathbb{R}^2$ 作为单位圆 $\{z \in \mathbb{R}^2 : |z| = 1\}$ 纳入而由二维勒贝格测度诱导的测度。

定理 *2.1* 的证明。 设 $\mathcal{M}$ 是一个闭的、非单连通的 $\mathcal{C}^\infty$-流形。取 $\mathcal{M}$ 上的黎曼度量 g 和一个正数 λ。我们的论证主要依赖于 É. Cartan 的经典定理（ [6]; 另见 do Carmo [5, 定理 12.2.2]）：

定理。 对于任意黎曼流形 $(\mathcal{M}, g)$，若其满足 $\pi_1(\mathcal{M}) \neq \{0\}$，则每个非平凡的 1-同伦类中都存在一条最短闭测地线。 _____

我们希望选出“所有这些最短闭测地线中最短的一条”。然而请注意，这样的测地线可能有无穷多条，因为一般而言，我们只知道 $\pi_1(\mathcal{M})$ 是有限生成的（但不是有限的）。为避免生成元不唯一带来的复杂情况，我们采用以下论证，从而能够选出 $(\mathcal{M}, g)$ 上“几乎”最短的非平凡闭测地线。

对于任意（分段可微）环路 γ，记

$$\Omega_{[\gamma]} := \text{ the free homotopy class of loops containing } \gamma.$$

于是定义

$$\kappa_0 := \inf\left\{\text{长度}[\gamma] : \gamma \text{ 是 } \Omega_{[\gamma]} \text{ 中的一条最短闭测地线，其中 } \Omega_{[\gamma]} \neq \{0\} \text{ 属于 } \pi_1(\mathcal{M})\right\}.$$

下确界先对所有闭测地线 γ 取值，然后再对所有自由同伦类 $\Omega_{[\gamma]}$ 取值。

首先，我们断言

$$\kappa_0 > 0.$$

为说明这一点，若 $\kappa_0 = 0$，则会存在一条闭测地线 γ'，其长度可以任意短，且满足 $\Omega_{[\gamma']} \neq \{0\}$。然而，由于 $\mathcal{M}$ 是一个紧致黎曼流形，存在一个正数 $\eta_0 = \eta_0(\mathcal{M})$，使得任何长度小于 η_0 的闭曲线都是零伦的。事实上，可以在 $\mathcal{M}$ 上选取一个由测地正规坐标图构成的有限开覆盖，并取该覆盖的勒贝格数作为 η_0。这就导致了矛盾。

对于稍后将要指定的 $\varepsilon \in \left]0, \frac{\kappa_0}{100}\right]$，我们选取 $\gamma_\star$，它是其自身自由同伦类（非平凡）中长度最短的闭测地线，使得

$$\operatorname{Length}[\gamma_\star] \leq \kappa_0 + \varepsilon.$$

然后，取某个稍后指定的大数 $N \in \mathbb{N}$，选取 N 个点 $\{x_1, \dots, x_N\} \subset \gamma_\star$，使其关于度量诱导距离 d_g 限制在 $\gamma_\star$ 上等分布。具体而言，对于任意 $1 \leq j < i \leq N$，我们要求

$$\left[d_g|_{\gamma_\star}\right](x_i, x_j) := \begin{cases} \frac{2\pi}{N}(i-j) & \text{if } i-j \leq \frac{N}{2}, \\ \frac{2\pi}{N}(j+N-i) & \text{if } i-j > \frac{N}{2}. \end{cases}$$

换言之，$\{x_1, \dots, x_N\}$ 在 $\gamma_\star$ 上满足与圆周 S^1 上的等分布点相同的距离关系——这里距离是相对于 $\gamma_\star$ 上的 d_g 而言的。根据构造，这是可行的，因为 $\gamma_\star$ 与 S^1 微分同胚；特别地，最短条件蕴含了 $\gamma_\star$ 的单纯性或素性。

接下来，我们断言，相对于整个流形 $\mathcal{M}$ 上的 d_g 来看，这些 N 个点的距离关系几乎相同。更准确地说：

$$0 \leq \left[d_g|_{\gamma_\star}\right](x_i, x_j) - \left[d_g|_{\mathcal{M}}\right](x_i, x_j) \leq \varepsilon \qquad \text{for any } 1 \leq j < i \leq N. \tag{♠}$$

断言的证明（♠）。î $d_g|_{\gamma_\star}$ ó $(x_i, x_j) - [d_g|_{\mathcal{M}}](x_i, x_j)$ 的非负性显然成立，因为距离是通过取下确界来定义的。

为证明另一个不等式，假设对于某些 i, j，差值大于 ε。那么就会存在一条曲线 σ (sigma，意为“short-cut”)，其端点为 x_i, x_j，与 $\gamma_\star^+$ 拼接后构成一条分段-$\mathcal{C}^\infty$ 环路，满足

$$\operatorname{Length}\left[\sigma \# \gamma_\star^+\right] < \operatorname{Length}[\gamma_\star] - \varepsilon.$$

此处及下文中，$\gamma_\star^+$ 表示连接 x_i 和 x_j 的 $\gamma_\star$ 上两条测地线段中较长的一条，$\gamma_\star^-$ 表示较短的一条。（若两条线段长度相等，则选择显而易见；在这种情况下，我们可以放心地省略细节。）

上述不等式意味着

$$\operatorname{Length}\left[\sigma \# \gamma_\star^+\right] < \kappa_0.$$

因此，根据 κ_0 的定义中的极小性条件，环路 $\sigma \# \gamma_\star^+$ 必须是零伦的。然而，在这种情况下，环路 $\sigma \# \gamma_\star^-$ 不可能是零伦的，否则 $\gamma_\star$ 必须是零伦的；这可通过依次收缩 $\sigma \# \gamma_\star^+$ 和 $\sigma \# \gamma_\star^-$ 得出。

因此，由于 σ（即，$\operatorname{Length}[\sigma] < \operatorname{Length}[\gamma_\star^-] - \varepsilon$）的短性、$\gamma_\star^\pm$ 的定义以及 $\gamma_\star$ 的选取，我们推得

$$\begin{aligned} \operatorname{Length}\left[\sigma \# \gamma_\star^-\right] &< 2\operatorname{Length}\left[\gamma_\star^-\right] - \varepsilon \\ &\leq \operatorname{Length}\left[\gamma_\star^+\right] + \operatorname{Length}\left[\gamma_\star^-\right] - \varepsilon \\ &= \operatorname{Length}[\gamma_\star] - \varepsilon \\ &\leq \kappa_0. \end{aligned}$$

这与 κ_0 的定义相矛盾。 □

现在让我们考虑格拉姆矩阵

$$\mathcal{G} = \mathcal{G}_N \in \mathrm{Mat}\,(N \times N; \mathbb{R}_+)$$

参数为 $\lambda > 0$ 的高斯核。其元素为

$$\mathcal{G}_{ij} = e^{-\lambda d_g^2(x_i, x_j)} \qquad \text{for } i, j \in \{1, 2, \dots, N\},$$

其中，$\{x_1, \dots, x_N\}$ 是上文选取的闭测地线 $\gamma_\star$ 上的点。这里，距离函数 d_g 必须理解为 $\left[d_g\big|_{\mathcal{M}}\right]$，而不是受限版本。

估计式 **(♠)** 以及对 $e^{-\lambda(A+\varepsilon)} - e^{-\lambda A}$ 在 $A > 0$ 处应用泰勒展开可得

$$\|\mathcal{G} - \mathcal{K}\|_\infty := \max_{i,j \in \{1,2,\dots,N\}} |\mathcal{G}_{ij} - \mathcal{K}_{ij}| \le C_0 \varepsilon,$$

其中 $C_0 = C_0(\lambda)$，而 $\mathcal{K}$ 是引理 2.2 中出现的循环矩阵：

$$\mathcal{K} = \begin{bmatrix} 1 & \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) & \exp\left(-\lambda\left(\frac{4\pi}{N}\right)^2\right) & \cdots & \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) \\ \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) & 1 & \cdots & \cdots & \exp\left(-\lambda\left(\frac{4\pi}{N}\right)^2\right) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \exp\left(-\lambda\left(\frac{2\pi}{N}\right)^2\right) & \exp\left(-\lambda\left(\frac{4\pi}{N}\right)^2\right) & \exp\left(-\lambda\left(\frac{6\pi}{N}\right)^2\right) & \cdots & 1 \end{bmatrix}.$$

最后，回顾引理 2.2（参见 Da Costa–Mostajeran–Ortega [9, 第 5–7] 页），当 $N = N$ 足够大且 λ 与 $\equiv 0$ 模 4 同余时，$\mathcal{K}$ 有一个负特征值。另一方面，具有负特征值是一个开条件，并且特征值随矩阵元素连续变化[2]。因此，取足够小的 $\varepsilon = \varepsilon$，当 λ 与 > 0 同时，格拉姆矩阵 $\mathcal{G}$ 也会在某个足够大的 $N = N$ 下具有负特征值，其中 λ 与 $\in \mathbb{N}$ 同时。

因此，根据核正定性的定义 1.1，我们得出结论：高斯核 $e^{-\lambda d_g^2(\bullet,\bullet)} : \mathcal{M} \times \mathcal{M} \to [0, \infty[$ 不可能是正定的。这里，$\lambda > 0$ 和 $\mathcal{M}$上的黎曼度量 g 均为任意选择。定理得证。□

Let 我们来评论一下，为什么上述证明不适用于单连通黎曼流形 s.

尽管在单连通的（$\mathcal{M}, g$）上仍然存在闭测地线 $\gamma_\star$ ——这是 Lyusternik–Fet 定理 [22]；另见 Klingenberg [20, 附录 § A.1] ——但我们无法以任何意义保证 $\gamma_\star$ 足够短。Synge（[30]；参见另见 [5, §9，推论 3.10(a)]）通过证明不存在最短闭测地线，得出了具有正截面曲率的闭的可定向偶数维黎曼流形是单连通的。直观地说，这是因为在 $\mathcal{M}$中任何环路都是零伦的，因此没有什么能阻止 $\gamma_\star$ 缩短其长度。因此，在上述定理 2.1 的证明中，我们没有（♠）的类似结论，因为总会存在缩短 $\gamma_\star$长度的“捷径”。这使基于引理 2.2 的特征值稳定性论证失效。

3. 备注

本文证明了，对于每个参数 $\lambda > 0$，任意非单连通闭黎曼流形 $(\mathcal{M}, g)$ 上的高斯核 $e^{-\lambda d_g^2(\bullet,\bullet)}$ 都不是正定的。事实上，我们——主要依赖 [9], 中的计算，尤其是引理 2.2——证明了

[2]这可由 Kato经典扰动定理得出 [19, 第126–127页，定理52]。这一结论曾在 [13], 中被提出，该文建立了著名的 Ger˙sgorin 圆盘定理，但其中的论证似乎并不完整。关于严格证明及相关讨论，参见近期工作 [21] 。

存在大量的 $N = N(\lambda)$，使得对于 $\mathcal{M}$上一条非平凡闭测地线上均匀分布的 N 个样本点，相应的高斯核 $N \times N$ 格拉姆矩阵 $\mathcal{G} \equiv \mathcal{G}_N$ 不是半正定的。

然而，对于任意固定的 N，格拉姆矩阵 $\mathcal{G}_N$ 在 L^∞-范数下收敛到单位矩阵，当 $\lambda \to \infty$ 时成立。因此，随着参数 $\lambda > 0$ 增大，我们需要选取更多$N \to \infty$样本点，才能检测出高斯核不具正定性的情况。参见 [9, §5]。由此我们提出以下

猜想/问题。在完备黎曼流形上，高斯核 $e^{-\lambda d_g^2(\bullet,\bullet)}$ 在较粗的尺度上是正定的；而对于维数为 ≥ 1 的紧流形，在较细的尺度上则会失去正定性。请对这一现象给出定量描述并加以证明。

还应指出（参见 [9, §3]），可以构造具有复杂 $\emptyset \subsetneq \Lambda_+\left(\mathcal{M}^0\right) \subsetneq]0, \infty[$ 的有限度量空间 $\mathcal{M}^0$（它们是零维流形）。此外，Sra [29]证明了，对于带有 S-散度的 $X = \mathrm{Sym}^{++}(n)$，有 $\Lambda_+(X) = \bigcup_{i=1}^{n-2} \left\{\frac{i}{2}\right\} \cup \left[\frac{n-1}{2}, \infty\right[$。

附录

A. 等距嵌入的两种概念。黎曼流形之间的度量空间意义下的等距嵌入定义见定义 1.2。随后，备注 1.3 指出，这与黎曼流形通常（微分几何意义下）的等距嵌入概念大不相同。本附录旨在进一步阐述这一问题。回顾一下，

定义。设(X, d) 和(X', d') 是度量空间。若映射 $\iota: X \to X'$ 是一一映射，且对所有 $x, y \in X$，都有 $\iota(x)\ \iota(y)\big) = d(x, y)$，则称其为等距嵌入。

定义。设$(\mathcal{M}, g)$ 和$(\mathcal{M}', g')$ 是完备黎曼流形。若 $\iota: (\mathcal{M}, g) \hookrightarrow (\mathcal{M}', g')$ 满足：ι 是 $\mathcal{C}^\infty$-流形之间的嵌入，且 $\iota: (\mathcal{M}, d_g) \hookrightarrow \left(\mathcal{M}', d_{g'}\right)$ 是度量空间等距嵌入，则称之为度量空间等距嵌入。

上述定义本质上是非局部的——不同于微分几何中等距嵌入概念要求向量场长度逐点一致，这里要求测地距离一致，而测地距离是通过曲线的长度定义的。

相比之下，黎曼流形之间的微分几何等距嵌入概念，如 Nash 的开创性著作 [23, 24]（参见 Han– Hong 的全面综述 [16]），定义如下 — $\iota: (\mathcal{M}, g) \hookrightarrow (\mathcal{M}', g')$ 是等距嵌入，当且仅当 ι是$\mathcal{C}^\infty$维流形之间的嵌入，且对于 $\mathcal{M}$上的任意向量场 X, Y，都有 $g'\big(\mathrm{d}\iota(X), \mathrm{d}\iota(Y)\big) = g(X, Y)$。这是一个局部条件：它仅要求由黎曼度量诱导的内积在每一对点 $\big(p, \iota(p)\big) \in \mathcal{M} \times \mathcal{M}'$ 处相同；或者等价地，要求每一对 ι-相关向量$X\big|_p \in T_p\mathcal{M}$ 和 $\mathrm{d}\iota(X)\big|_{\iota(p)} \in T_{\iota(p)}\mathcal{M}'$ 的长度相同，其中长度分别相对于 g 和 g'而言。

应当强调，这两种黎曼流形之间的等距嵌入概念本质上是不同的。考虑下面这个简单的例子。

例子。圆周 $\mathbf{S}^1$ 显然可以通过单位圆在 $\mathbb{R}^2$中的包含映射 ι，在微分几何的意义下等距嵌入到二维欧几里得空间中。然而，ι 并不是度量空间意义下的等距嵌入，因为 $\mathbf{S}^1$ 上的任意一对对跖点$\{p_+, p_-\}$ 之间的内蕴距离为 π，而它们的外在距离（即相对于环境空间上的欧几里得距离而言）在 $\mathbb{R}^2$中等于2 。

理解这一现象的一种方式是，在 $\mathbb{R}^2$中， $p_{\pm}$之间的直径线段相对于沿圆周测得的$\mathbf{S}^1$ 上的测地距离而言，起到了“捷径”的作用。

更一般地，根据纳什定理 [24], ，任何黎曼流形 $(\mathcal{M}, g)$ 都可以按照微分几何意义等距嵌入某个欧几里得空间。然而，除非流形 $(\mathcal{M}, g)$ 本身是欧几里得平坦的，否则这种等距嵌入绝不会是度量空间等距嵌入，正如 Feragen–Lauze–Hauberg [12] 所证明的那样。

能够将圆周 $\varrho\mathbf{S}^1$ 以某个缩放因子 $\varrho > 0$ 作度量空间等距嵌入的黎曼流形示例包括：对于 $n \geq 1$， $\mathbf{S}^n$ ；对于 $n > 1$， $\mathbb{RP}^n$ ；以及对于 $1 \leq k < n$，实格拉斯曼流形 $\mathbf{Gr}(k, n)$。事实上，定理 1.4 正是通过这种方式在 [9] 中得到证明的。

我们通过下面这个简单的观察来结束对等距嵌入概念的澄清，该观察进一步阐明了上述示例。

定理。 设$(\mathcal{M}, g)$ 和$(\mathcal{M}', g')$ 是完备黎曼流形。

(1) 若子流形 $(\mathcal{M}, g) \subset (\mathcal{M}', g')$（即包含映射 ι）在度量空间意义下是等距嵌入，则 ι 必为全测地的，因此也是微分几何意义下的等距嵌入。*(2)* 设 $\pi : (\mathcal{M}, g) \to (\mathcal{M}, g)$ 是度量空间 $(\mathcal{M}, d_g)$ 到自身的满射等距映射。那么 π 在微分几何意义下是等距的。特别地，若 $f : (\mathcal{M}, g) \to (\mathcal{M}, g)$ 是度量空间等距嵌入且为满射，则它是微分几何意义下的等距嵌入。

证明。 根据定义 1.2，作为度量空间等距嵌入的 $\iota : (\mathcal{M}, g) \hookrightarrow (\mathcal{M}', g')$ 保持测地距离。特别地，它将测地线映为测地线。这就证明了第一条结论。第二条结论是迈尔斯–斯廷罗德的一个定理 [8]。□

致谢。SL 的研究得到国家自然科学基金委员会（国家自然科学基金委员会）项目（编号：12201399）、上海交通大学–伦敦大学学院联合种子基金以及上海现代分析前沿科学中心的资助。

参考文献

[1] M. Belkin、P. Niyogi 和 V. Sindhwani，《流形正则化：从有标签和无标签样本中学习的几何框架》。《机器学习研究杂志》（JMLR），**7** (2006), 2399–2434.[2] S. Bochner，《傅里叶积分讲义》，莱比锡，1932年。[3] V. Borovitskiy、A. Terenin、P. Mostowsky 和 M. Deisenroth，黎曼流形上的 Matérn高斯过程，载于：H. Larochelle、M. Ranzato、R. Hadsell、M. Balcan 和 H. Lin（编），《神经信息处理系统进展》第33卷，第12426–12437页。Curran Associates, Inc.（2020）。[4] S. Calinon，黎曼流形上的高斯分布：机器人学习与自适应控制的应用，《IEEE机器人与自动化杂志》**27**(2020), 33–45。[5] M. P. do Carmo，《黎曼几何：数学理论与应用》，Birkhäuser Boston, Inc.，美国马萨诸塞州波士顿，1992年。[6] É. Cartan，《黎曼空间几何讲义》，（法文）第2版。*Gauthier-Villars*，巴黎，1946年。[7] O. Chapelle、P. Haffner 和 V. N. Vapnik。用于基于直方图的图像分类的支持向量机。《*IEEE*神经网络汇刊》，**10**(1999), 1055–1064。[8] J. Cheeger 和 D. G. Ebin，《黎曼几何中的比较定理》。北荷兰数学文库，第9卷。North-Holland Publishing Co.，阿姆斯特丹—牛津；American Elsevier Publishing Co., Inc.，纽约，1975年。[9] N. Da Costa、C. Mostajeran 和 J.-P. Ortega，圆上的高斯核及允许圆等距嵌入的空间。arXiv预印本：2302.10623v1 [cs.LG]。

[10] F. Cucker 和 S. Smale，论学习的数学基础。《美国数学会通报》39 (2001), 1–49。[11] A. Feragen 和 S. Hauberg，开放问题：流形和度量空间上的核方法。正定测地指数核的概率是多少？载于：V. Feldman、A. Rakhlin、O. Shamir（编），第29届学习理论年会。《机器学习研究论文集》第49卷，第1647–1650页。PMLR，哥伦比亚大学，美国纽约州纽约市（2016年6月23日至26日）。[12] A. Feragen、F. Lauze 和 S. Hauberg，测地指数核：曲率与线性之间的冲突。载于：2015年IEEE计算机视觉与模式识别会议（CVPR），第3032–3042页（2015年）。[13] S. Gerˇsgorin，论矩阵特征值的界。《苏联科学院数学辑刊》1 (1931), 749–754。[14] T. Gneiting，球面上的严格与非严格正定函数，《伯努利》19 (2013), 1327–1349。[15] L. Grigoryeva 和 J.-P. Ortega，通过规范化对循环网络进行降维。《几何与力学》13 (2021), 647–677。[16] Q. Han 和 J.-X. Hong，《黎曼流形在欧几里得空间中的等距嵌入》，美国数学会：普罗维登斯，2006年。[17] J. Istas，流形指标分数场，《ESAIM：概率与统计》16 (2012), 222–276。[18] S. Jayasumana、R. Hartley、M. Salzmann、H. Li 和 M. Harandi，采用高斯径向基函数核的黎曼流形核方法。《IEEE模式分析与机器智能汇刊》37 (2015), 2464–2477。[19] T. Kato，《线性算子扰动理论简明导论》，施普林格出版社，1982年。[20] W. Klingenberg，《闭测地线讲义》。《数学科学基础丛书》第230卷。施普林格出版社，柏林—纽约，1978年。x+227 页。ISBN：3-540-08393-6。[21] C.-K. Li 和 F. Zhang，特征值连续性与 Gerˇsgorin 定理，《电子线性代数杂志》35 (2019), 619–625。[22] L. A. Lyusternik 和 A. I. Fet，闭流形上的变分问题。《苏联科学院院刊》（新辑）81 (1951), 17–18 [（俄文）]。[23] J. Nash，C^1 等距嵌入，《数学年刊》60 (1954), 383–396。[24] J. Nash，黎曼流形的嵌入问题，《数学年刊》63 (1956), 20–63。[25] C. Salvi、T. Cass、J. Foster、T. Lyons 和 W. Yang，签名核是古尔萨偏微分方程的解，《SIAM数学与数据科学杂志》3 (2021), 873–899。[26] I. J. Schoenberg，度量空间与正定函数。《美国数学会汇刊》44 (1938), 522–536。[27] B. Schölkopf 和 A. Smola，《核方法学习：支持向量机、正则化、优化及其他》。自适应计算与机器学习。MIT出版社，2002年。[28] B. Schölkopf、A. Smola 和 K. R, Müller，作为核特征值问题的非线性成分分析，《神经计算》10 (1998), 1299–1319。[29] S. Sra，正定矩阵与 S-散度，《美国数学会会刊》144 (2016), 2787–2797。[30] J. L. Synge，论正曲率空间的连通性。《数学季刊》（牛津版）7 (1936), 316–320。[31] A. T. A. Wood，直线上的截断协方差函数何时是圆上的协方差函数？《统计与概率通讯》24 (1995), 157–164。Siran Li：上海交通大学数学科学学院与上海数学中心，中国上海市闵行区东川路800号科学楼6号（200240）

电子邮箱地址：siran.li@sjtu.edu.cn