

专题综述

双曲核网络

东南大学 方鹏飞

本文是东南大学团队针对双曲核网络方面的一系列研究成果，发表于ICCV 2021^[1]、IJCV2023^[2]。该系列工作研究的问题是围绕双曲空间（hyperbolic space），一种具有负曲率的曲面空间，定义有效的核函数并验证其有效性。近期研究表明，将数据嵌入至双曲空间作为其表征空间能够有效提升诸如小样本学习、零样本学习、度量学习等多种机器学习应用的性能。然而，由于双曲空间的曲面性质导致在该空间中的操作变得困难，例如，求解一系列双曲空间中的点的Fréchet均值需要迭代的算法。在欧式空间中，核方法不仅具备丰富的理论性质，同时具有强大的表征能力。本文研究双曲空间中的核方法，包含双曲正定核以及双曲不定核。这为表征学习带来两个好处：1-核化将使得表征兼具核机的强大表示能力以及双曲空间的编码能力；2-核空间（希尔伯特空间或克莱空间）丰富的结构可以简化多种双曲空间内的操作。具体而言，系列工作以嵌入方程思路展开，设计了多个双曲嵌入方程，并定义一系列双曲核函数。通过在一系列学习任务中验证提出的双曲核函数的有效性。

一、研究背景

在机器学习社区，欧式空间一直作为特征空间，用以编码诸如图像或文本等数据。其主要原因是高维欧式空间是我们所熟悉三维空间的自然概括，且在欧式空间中计算距离与相似度等算子较为简易。然而，将数据编码至欧式空间可能会损害或者扭曲数据的结构信息，从而丢失数据中固有的复杂空间几何信息。例如，欧式空间难以编码图结构数据中的层次信息^[3]。近期，系列研究表明，相较于将数据嵌入至欧式空间，双曲空间作为表征空间能够刻画数据的层次结构，同时，该结构信息益于大量机器学习应用，包括文本蕴含^[4]、机器翻译^[5]、

语言视觉推理^[5]、图像分类和检索^[6]、图数据分类^[7]等任务。

双曲空间是一种具有恒定负曲率的曲面空间。该负曲率特性赋予双曲空间具有编码数据层次结构的性质。这是由于在双曲空间中，其体积会随着半径呈指数级增长^[8]，从而提升了数据在该空间中的表示能力。尽管有一系列工作已经成功利用双曲空间进行推理^[4-7]，但在非线性空间中进行数学运算等困难阻碍了其更广泛的使用。例如，在欧式空间中计算一系列数据的均值很简单，但双曲空间中的平均需要用Fréchet均值来近似。然而，计算Fréchet均值需要迭代算法，这使得算法或者神经网络的计算成本变得更加高昂^[9,10]。这促使我们研究双曲空间中的核方法，以便能够无缝地利用核机以分析双曲数据。

核方法可以认为是一种衡量数据间相似度的度量。在欧式空间中，许多常见的核函数被定义为欧式距离（亦称为欧式空间的测地线距离）的函数。以常用的径向基核（RBF kernel），记为 $k(x, y) = \exp(-\xi d^2(x, y))$ ，为例。有一种猜想是一旦知道测地线距离，就可以在曲面空间（双曲空间就是其中之一）中构造有效的正定核。不幸的是，情况并非如此，正如Jayasumana等人^[11]和Feragen等人^[12]所示，由于曲面空间与平坦的欧式空间不存在等距关系，故利用测地线距离定义的RBF核是非正定的。有趣的是，在曲面空间上定义正定核的难度现在被认为是机器学习领域中的一个悬而未决的问题^[13]。

在本文中，我们使用庞加莱模型解决了在双曲空间中定义正定核的难题。在这里，我们提出了几个有效的正定双曲核，包括功能强大的通用核（universal kernel）。为此，我们首先利用引理构造一个有效的庞

加莱线性核。利用该引理，我们进一步为双曲几何定义了有效的庞加莱RBF核和庞加莱拉普拉斯核。同时，我们也提出了庞加莱二项式核。同时，针对上述核函数在神经网络优化中因调参等原因产生的繁琐问题，本文紧接着定义了庞加莱径向核，通过在庞加莱球中设计了一个通用公式，即对多核核函数进行加权和，获得双曲多核学习方法。本文在大量实验中验证上述双曲核函数的表达能力。

二、双曲核网络

2.1 前言：双曲空间

一个 n 维的双曲空间 \mathbb{H}^n 是一种具有恒定负曲率的黎曼流形^[14]。庞加莱球是 n 维的双曲空间的一种建模形状，在庞加莱球中，所有的点（样本）均嵌入至一个 n 维的球体内。正式情况下，曲率为 c 的庞加莱模型定义为 $\mathbb{D}_c^n = \{z \in \mathbb{R}^n: c\|z\| < 1\}$ 。它的黎曼度量定义为 $g_c^{\mathbb{D}}(z) = \lambda_c^2(z) \cdot g^E$ ，其中 $\lambda_c(z)$ 是适形因子，定义为 $\lambda_c(z) = \frac{2}{1-c\|z\|^2}$ ，以及 $g^E = I_n$ 是欧式度量。此外，我们借助莫比乌斯陀螺矢量空间（Möbius gyrovector space）便于矢量运算。具体而言，在庞加莱球中的莫比乌斯加法可定义为：

$$z_i \oplus_c z_j = \frac{(1 + 2c\langle z_i, z_j \rangle + c\|z_j\|^2)z_i + (1 - c\|z_i\|^2)z_j}{1 + 2c\langle z_i, z_j \rangle + c^2\|z_i\|^2\|z_j\|^2}$$

其中， $z_i, z_j \in \mathbb{D}_c^n$ 表示庞加莱球中的两个点。针对两个样本的测地线距离定义为：

$$d_c(z_i, z_j) = \frac{2}{\sqrt{c}} \tanh^{-1}(\sqrt{c}\| -z_i \oplus_c z_j \|).$$

在黎曼集合中，一个点的切空间是一个内积空间，其包含了所有方向上与该点相切的向量。对于庞加莱球中的一个点， $z \in \mathbb{D}_c^n$ ， z 点的切空间记为 $T_z \mathbb{D}_c^n$ 。其中，指数映射定义为：

$$\mathcal{T}_z(p) = z \oplus_c \left(\tanh\left(\sqrt{c} \frac{\lambda_c(z)\|p\|}{2}\right) \frac{p}{\|p\|} \right)$$

其中 $p \in T_z \mathbb{D}_c^n$ 。指数映射将在 z 的切平面中的一个向量 p 映射至庞加莱球 \mathbb{D}_c^n 。其逆过程称为对数映射，将庞加莱球中的一个向量 $q \in \mathbb{D}_c^n$ 映射至 z 点的切空间，记为：

$$\mathcal{J}_z(p) = \frac{2}{\sqrt{c}} \lambda_c(z) \tanh^{-1}(\sqrt{c}\| -z \oplus_c p \|) \frac{-z \oplus_c p}{\| -z \oplus_c p \|}$$

值得注意的是，指数映射与对数映射互为逆函数，满足： $\mathcal{T}_z(\mathcal{J}_z(p)) = p \in T_z \mathbb{D}_c^n$ 。在本文中，我们巧妙利用对数映射定义嵌入方程，以构建相应的双曲核函数。

2.2 双曲正定核

在本文中，我们首先利用双曲几何的切空间来定义一组有效的正定核。具体而言，我们首先提出一个映射， $f_{\mathbb{D}}(z): \mathbb{D}_c^n \rightarrow \mathbb{R}^n$ ，即将庞加莱球中一个向量映射至欧式空间，定义为：

$$f_{\mathbb{D}}(z) := \tanh^{-1}(\sqrt{c}\|z\|) \frac{z}{\sqrt{c}\|z\|}$$

上式的物理意义是将庞加莱球中的向量 z 映射至原点的切空间。在原点的切空间，每一个向量 $f_{\mathbb{D}}(z)$ 不仅能估计其在原庞加莱球中的原始向量 z ，同时，我们ICCV21^[1]的论文中通过证明曲线长度等效定理揭示在两个空间中优化度量的等效性。基于该映射，本文提出了一系列正定核，具体如下：

庞加莱正切核：在欧式空间中，最简单的正定核为线性核。我们在双曲空间中也定义相应的线性核，记为： $k^{\tan}(z_i, z_j) = \langle f_{\mathbb{D}}(z_i), f_{\mathbb{D}}(z_j) \rangle$ 。该核可以理解在庞加莱球切空间中定义的线性核，故将其命名为庞加莱正切核。该核没有参数，适合对原型的快速验证。

庞加莱径向基核：本文通过利用定义的映射方程 $f_{\mathbb{D}}(z)$ 以及在该空间的度量 $d_c(\cdot, \cdot)$ ，提出庞加莱径向基核，表示为： $k^{\text{rbf}}(z_i, z_j) = \exp(-\xi \|f_{\mathbb{D}}(z_i) - f_{\mathbb{D}}(z_j)\|^2)$ 。在ICCV21^[1]文中，通过证明 $\|f_{\mathbb{D}}(z_i) - f_{\mathbb{D}}(z_j)\|^2$ 是负定性质进而证明庞加莱径向基核 $k^{\text{rbf}}(z_i, z_j)$ 为正定核。庞加莱径向基核具有强大的通用估计（universal approximation）能力。

庞加莱拉普拉斯核：另一个常用且具备通用估计能力的核称之为拉普拉斯核。在双曲空间中，本文定义庞加莱拉普拉斯核，记为： $k^{\text{lap}}(z_i, z_j) = \exp(-\xi \|f_{\mathbb{D}}(z_i) - f_{\mathbb{D}}(z_j)\|)$ 。其正定性的证明过程与庞加莱径向核证明过程类似。

庞加莱二项式核：除了上述指数型核函数，本文继续构造庞加莱二项式核函数，定义为 $k^{\text{bin}}(z_i, z_j) = (1 - \langle f_{\mathbb{D}}(z_i), f_{\mathbb{D}}(z_j) \rangle)^{-\alpha}$ 。该核函数通过泰勒分解获得非负性

完全泰勒级数证明其正定性，并且具备通用估计能力。

庞加莱径向核：上述核函数为单一的核函数，其在优化的过程中常常受数据、模型等影响，需要调试其参数。该问题启发我们在双曲空间中研究多核学习 (multiple-kernel learning, MKL) 方法，使得核函数能够根据数据或者模型本身学习相应的核函数。具体而言，基于 $f_{\mathbb{D}}(z)$ ，我们提出一个新的映射函数 $g_{\mathbb{D}}(z) := \frac{f_{\mathbb{D}}(z)}{\|f_{\mathbb{D}}(z)\|}$ ，并提出庞加莱径向核： $k^{\text{rad}}(z_i, z_j) = \sum_{m=-2}^{\infty} a_m k_m^{\text{cos}}(z_i, z_j)$ ，其中 $k_m^{\text{cos}}(z_i, z_j)$ 表示为余弦核。

三、实验结果

3.1 小样本学习

小样本学习 (Few-shot Learning) 是一种新的学习范式，其通过少量样本学习数据的表征空间，并使得该表征空间能够有效适用于未见过的测试数据^[15-17]。通常而言，小样本学习通过元学习的方式训练神经网络。具体而言，每次迭代均会采样一个情节的数据用以训练神经网络，该训练方式称之为 N-way K-shot，并通过该训练方式实现在每个情节中对 N 个类别的识别任务。小样本学习主要基于度量学习思想展开，用以学习同时适用于可见样本与未见样本的度量空间。在匹配网络 (Matching Network)^[15]中，模型通过学习与样本相关度量以确定查询样本的类别。受匹配网络启发，原型网络 (Prototypical Network, ProtoNet)^[16]设计了与类别相关的度量方法，即每个类别的所有样本均被认为是对该类别的描述。在关系网络 (Relation Network)^[17]中，

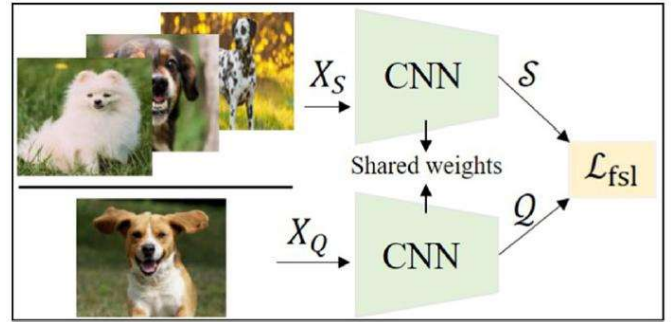


图 1 原型网络示意图

通过研究显式地建模支持样本与查询样本间的非线性相似度关系，使得所学习的隐度量空间能够自适数据本身的分布情况。

在我们的实验中，采用 5-way 1-shot 或者 5-way 5-shot 以评估模型。该实验采用原型网络模型 (ProtoNet)^[15]用以训练神经网络特征提取器。图 1 为原型网络的示意图。为验证所提出的双曲核函数的优越性，我们采用具有双曲表征能力原型网络为基线，记为 Hyper ProtoNet^[6]。同时，我们采用 Conv-4 以及 ResNet-18 作为模型的骨干网络，用以提取数据的特征。同时，我们采用 tiered-ImageNet 以及 FC100 数据集用以验证模型提出的双曲核函数的效果。实验结果如表 1 所示。

如表 1 所示，本文所提出的双曲核函数较基线模型具有显著提升。在 tiered-ImageNet 数据集上，最简单的庞加莱正切核在 Conv-4 和 ResNet-18 骨干网络分别提升 (0.29%&2.46%) / (1.03%&1.56%)。在 FC100

Model	Backbone	tiered-ImageNet		FC100	
		5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot
Hyper ProtoNet [†] (Khruikov et al., 2020)	Conv-4	54.44 ± 0.23	71.96 ± 0.20	37.59 ± 0.19	51.76 ± 0.19
Poincaré tangent kernel	Conv-4	54.73 ± 0.22	74.37 ± 0.18	37.66 ± 0.17	52.29 ± 0.18
Poincaré RBF kernel	Conv-4	<u>57.78 ± 0.23</u>	76.11 ± 0.18	<u>38.93 ± 0.18</u>	<u>54.40 ± 0.18</u>
Poincaré Laplace kernel	Conv-4	57.33 ± 0.22	<u>76.48 ± 0.18</u>	37.99 ± 0.17	53.54 ± 0.18
Poincaré binomial kernel	Conv-4	56.72 ± 0.22	75.87 ± 0.18	38.32 ± 0.18	53.50 ± 0.18
Poincaré radial kernel	Conv-4	57.96 ± 0.22	76.87 ± 0.18	39.24 ± 0.17	54.82 ± 0.18
Hyper ProtoNet [†] (Khruikov et al., 2020)	ResNet-18	62.28 ± 0.23	74.50 ± 0.21	40.64 ± 0.20	52.50 ± 0.30
Poincaré tangent kernel	ResNet-18	63.31 ± 0.23	76.06 ± 0.23	42.18 ± 0.26	54.32 ± 0.32
Poincaré RBF kernel	ResNet-18	<u>64.52 ± 0.22</u>	76.82 ± 0.21	43.84 ± 0.23	56.01 ± 0.30
Poincaré Laplace kernel	ResNet-18	64.38 ± 0.22	<u>77.16 ± 0.21</u>	<u>43.22 ± 0.23</u>	<u>55.47 ± 0.30</u>
Poincaré binomial kernel	ResNet-18	64.12 ± 0.23	76.44 ± 0.23	42.60 ± 0.24	55.08 ± 0.32
Poincaré radial kernel	ResNet-18	65.33 ± 0.21	77.48 ± 0.20	44.12 ± 0.20	56.28 ± 0.26

[†] indicates the network was self-implemented. 1st / 2nd best in “bold” / “(underline)”

表 1 双曲核函数在小样本学习中的实验结果

数据集上，最简单的庞加莱正切核在 Conv-4 和 ResNet-18 骨干网络分别提升 (0.07%&0.53%) / (1.54%&1.82%)。其余的双曲单核函数均有进一步的提升。其中，庞加莱径向基核在 Conv-4 为骨干网络时表现整体最优，而庞加莱拉普拉斯核在 ResNet-18 为骨干网络时表现整体最优。相较于双曲单核函数，多核学习方法也在小样本学习中展示其优越性，庞加莱径向核在 tiered-ImageNet 以及 FC100 中均获得最优性能。其在 tiered-ImageNet 数据集对 Conv-4 和 ResNet-18 骨干网络分别提升 (3.52%&4.91%) / (3.05%&2.98%)；在 FC100 数据集对 Conv-4 和 ResNet-18 骨干网络分别提升 (1.65%&3.06%) / (3.48%&3.78%)，进一步揭示多核学习方法对提升双曲核函数表征能力的优越性。

双曲核函数的优越性在零样本学习、度量学习、行人重识别等机器学习应用中均有验证，感兴趣的读者请参考文献[1,2]。

四、总结

本文在双曲空间提出了一系列正定核，用以将双曲空间中的表征映射至希尔伯特空间。为了定义此类核函数，我们利用庞加莱球的恒等切空间（即庞加莱球原点的切空间），并进一步在恒等切线空间中定义有效的正定核函数。所提出的核函数包括功能强大的通用核函数（即庞加莱径向基核函数、庞加莱拉普拉斯核函数、庞加莱二项式核函数和庞加莱径向核函数）。我们在小样本学习等任务中评估了所提出双曲核函数的有效性。实验结果表明，这些核函数对提升双曲空间中的表征取得积极效果。未来的工作包括探索双曲负定核函数在相关机器学习应用中的有效性。

五、前期相关工作

核方法作为机器学习领域的重要理论框架，其系统性研究已在支持向量机 (SVM)、主成分分析 (PCA) 及聚类算法等[1]经典模型中展现出显著优势。其理论的核心在于通过将原始数据映射至高维（甚至无限维）的

再生核希尔伯特空间 (Reproducing Kernel Hilbert Space, RKHS)，在该空间中用线性模型解决原线性不可分问题。为规避显式求解高维映射的计算复杂度，核技巧利用核函数直接计算再生核希尔伯特空间内的样本相似度。基于此理论，研究者相继提出了多项式核、径向基核及拉普拉斯核等经典核函数形式^[18]。为改善单核方法的适用性，多核学习 (Multiple Kernel Learning, MKL)^[19,20]框架应运而生。在多核学习中，核函数为基核的凸组合，权重从数据中习得，使得所学习到的核机器能够最大限度匹配数据的内在结构^[19,21]。

近年来，为增强结构化数据表征能力，将核方法扩展至非线性几何空间的研究方向备受关注。在非欧几里得几何空间定义有效正定核 (Positive Definite Kernel, PD Kernel) 的通用策略是采用适配的距离度量。Jayasumana 等人^[22]开创性地建立了对称正定矩阵空间的高斯核理论体系，随后该成果被进一步拓展至 Grassmann 流形^[11]。Harandi 等人^[23]系统探索了 Grassmann 流形上的核函数构造方法，提出基于等效嵌入映射的正定 Grassmannian 核。而 Le 与 Yamada^[24]将使用 Fisher 信息度量的核函数用于持续图。Jayasumana 等人^[25]则发展了一系列紧致流形（包括 n 维球面、Grassmann 流形及形状流形）的径向基核统一框架。在现有研究中，Cho 等人^[26]提出的双曲空间中的支持向量机方法与本文工作最具相关性，其虽尝试在双曲几何中引入核方法以构建非线性决策边界，但所提出的核函数因缺乏正定性导致两个本质缺陷：其一，不定核 (Indefinite Kernel) 不是一致逼近核，违背了通用逼近性质^[27]；其二，不定核因其需要稳定损失值，导致训练困难^[28]。

本文针对上述理论的局限，构建了双曲几何空间的正定核理论框架。作为对不定核的理论补充，本文通过双曲空间的核化过程，将双曲数据嵌入高维（可能无限维）希尔伯特空间，使得所得表征能够充分受益于核机器的理论优势。本文后续章节将详细阐述理论推导过程，并在多组具有挑战性的实际应用中评估算法的有效性。

责任编辑 魏秀参

参考文献

- [1] P. Fang, M. Harandi, and L. Petersson. Kernel methods in hyperbolic spaces. In ICCV 2021.
- [2] P. Fang, M. Harandi, Z. Lan and L. Poincaré kernels for hyperbolic representations. IJCV, 2023.
- [3] Q. Liu, M. Nickel, and D. Kiela. Hyperbolic graph neural networks. In NeurIPS 2019.
- [4] O. Ganea, G. Bécigneul, and T. Hofmann. Hyperbolic neural networks. In NeurIPS 2018.
- [5] C. Gulcehre, M. Denil, M. Malinowski, A. Razavi, R. Pascanu, K. Hermann, P. Battaglia, V. Bapst, D. Raposo, A. Santoro, and N. Freitas. Hyperbolic attention networks. In ICLR 2019.
- [6] V. Khrulkov, L. Mirvakhabova, E. Ustinova, I. Oseledets, and V. Lempitsky. Hyperbolic image embeddings. In CVPR 2020.
- [7] Q. Liu, M. Nickel, and D. Kiela. Hyperbolic graph neural networks. In NeurIPS 2019.
- [8] M. Hamann. On the tree-likeness of hyperbolic spaces. Mathematical Proceedings of the Cambridge Philosophical Society, 2017.
- [9] H. Karcher. Riemannian center of mass and mollifier smoothing. In Communications on Pure and Applied Mathematics, 1977.
- [10] A. Lou, I. Katsman, Q. Jiang, S. Belongie, S. Lim, and C. Sa. Differentiating through the fréchet mean. In ICML 2020.
- [11] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi. Kernel methods on Riemannian manifolds with gaussian RBF kernels. TPAMI, 2015.
- [12] A. Feragen, F. Lauze, and S. Hauberg. Geodesic exponential kernels: When curvature and linearity conflict. In CVPR 2015.
- [13] A. Feragen, and S. Hauberg. Open problem: Kernel methods on manifolds and metric spaces. What is the probability of a positive definite geodesic exponential kernel. In CoLT 2016.
- [14] P. Absil, R. Mahony, and R. Sepulchre. Optimization algorithms on matrix manifolds. Princeton University Press. 2007.
- [15] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and H. Wierstra. Matching networks for one shot learning. In NeurIPS 2016.
- [16] J. Snell, K. Swersky, and R. Zemel. Prototypical networks for few-shot learning. In NeurIPS 2017.
- [17] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. Torr, and T. Hospedales. Learning to compare: Relation network for few-shot learning. In CVPR 2018.
- [18] T. Hofmann, B. Schölkopf, and A. J. Smola. Kernel methods in machine learning. The Annals of Statistics, 2008.
- [19] A. Rakotomamonjy, F. R. Bach, S. Canu, and Y. Grandvalet. SimpleMKL. JMLR, 2008.
- [20] G. R. G. Lanckriet, N. Cristianini, P. Bartlett, L. E. Ghaoui, and M. I. Jordan. Learning the kernel matrix with semidefinite programming. JMLR, 2004.
- [21] T. Wang, L. Zhang, and W. Hu. Bridging deep and multiple kernel learning: A review. Information Fusion, 2021.
- [22] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi. Kernel methods on the riemannian manifold of symmetric positive definite matrices. In CVPR 2013.
- [23] M. T. Harandi, M. Salzmann, S. Jayasumana, R. Hartley, and H. Li. Expanding the family of grassmannian kernels: An embedding perspective. In ECCV 2014.
- [24] T. Le, and M. Yamada. Persistence fisher kernel: A Riemannian manifold kernel for persistence diagrams. In NeurIPS 2018.
- [25] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi. Optimizing over radial kernels on compact manifolds. In CVPR 2014.
- [26] H. Cho, B. DeMeo, J. Peng, and B. Berger. Large-margin classification in hyperbolic space. In ICML 2019.

[27] C. A. Micchelli, Y. Xu, and H. Zhang. Universal kernels. JMLR, 2006.

[28] C. S. Ong, X. Mary, S. Canu, and A. J. Smola. Learning with non-positive kernels. In ICML 2004.



方鹏飞

东南大学计算机科学与工程学院教授，博士生/硕士生导师。分布于 2017 年与 2022 年在澳大利亚国立大学获得硕士与博士学位，随后于澳大利亚蒙纳士大学担任博士后研究员。2023 年加入东南大学，任副教授，教授。长期从事机器学习、计算机视觉领域研究工作。

Email: fangpengfei@seu.edu.cn