arXiv:2305.01789v1 [stat.ME] 2 May 2023

研究论文

等距黎曼流形上的多元内在局部多项式回归：正定数据的应用。

Ronaldo García Reyes[1,2], Ying Wang[2], MinLi[2], Marlis Ontivero-Ortega 的[1], Deirel Paz-Linares[1,2], LídiceGalánGarcía和 Pedro Antonio Valdes Sosa[1,2]

[1]成都脑科学研究院临床医院，中国电子科技大学，中国成都，610054。电子邮箱：ronaldo@neuroinformatics-collaboratory.org,pedro.valdes@neuroinformatics-collaboratory.org。[2]神经信息学，古巴神经科学中心，哈瓦那，11300，古巴哈瓦那。。

关键词：多元内在局部多项式回归，黎曼流形，等距黎曼流形，欧几里得空间拉回度量，正定流形，多元协变量，对数乔列斯基度量，Rie-SNE

MSC 分类号：主要分类– 53B12, 62R30, 62G05;次要分类– 6208, 6211

摘要

本文介绍了一种利用等距黎曼流形（IRMs）的新型非参数黎曼回归方法。所提出的技术——等距黎曼流形上的内禀局部多项式回归（ILPR-IRMs）——能够在黎曼流形之间实现全局数据映射，同时保留其底层几何结构。本文将内在局部多项式回归（ILPR）推广至任意黎曼流形和等距映射上的多变量协变量。具体而言，对于配备欧几里得拉回度量（EPMs）的流形，本文推导出多元 ILPR（ILPR-EPM）的闭式解析公式。本文还确立了多元局部线性情形下 ILPR-EPM 的渐近统计性质，包括渐近偏差公式，从而证明了估计量的一致性。本文聚焦于对称正定（SPD）流形上的一组黎曼度量，展示了该方法的潜在应用；这一流形在机器学习和神经科学领域中均有应用。研究表明，SPD 流形上的若干度量属于 EPM，因此可以得到 SPD 流形上多元 ILPR 估计量的闭式解析表达式。本文在 SPD 流形的模拟数据上，针对两种特定的 EPM——对数-乔列斯基度量和对数欧几里得度量——评估了 ILPR 估计量的性能，并在缩放流形和协变量维度时，将其与采用仿射不变度量的外禀 LPR 进行比较。结果表明，采用对数-乔列斯基度量的 ILPR 计算速度更快，并且在误差与时间之间实现了优于其他度量的权衡。最后，本文采用 SPD 流形上的对数-乔列斯基度量，实现了 Rie-SNE 的高效内禀版本，用于可视化高维 SPD 数据。实现 ILPR-EPM 及其他相关计算的代码可在 GitHub 页面获取，网址为：https://github.com/ronald1129/Multivariate-Intrinsic-Local-Polynomial-Regression-on-Isometric-Riemannian-Manifolds.git。

目录

1 引言 2 一般记号 3 背景3.1 黎曼流形.3.2 对称正定流形（SPD）.3.3 等距黎曼流形（IRM）与欧几里得拉回度量（EPM）.3.4 多元回归的泰勒展开.**4 等距黎曼流形上的多元内在局部多项式回归（ILPR-IRM）**4.1 欧几里得空间上的多元局部多项式回归（LPR）. 2 5 6 6 6 7 8 9 9

4.2 任意黎曼矩阵流形上的多元内在局部多项式回归（ILPR）.10411 43 4 13. 等距黎曼流形上 ILPR 问题之间的关系.ILPR-EPM 的精确解.

5 对称正定流形上的模拟。 **15**

5 1 模拟 I . DTI 类型数据 .
175.2 模拟 II. 高维数据.18

6 讨论与结论 20 A 关于向量的偏导数定义 22 B 引理（3.4）的证明 22 C 引理（4.4）的证明 23 D 定理（4.5）的证明 23 E 定理（4.6）的证明 26 E.1 ILP 方程（4.9）的表达式.26 E.2 ILE 方程（4.11）的精确解析表达式.26 F 正则性条件 28 G 定理（4.7）的证明 28 G.1 $\widehat{\beta}_0(x;1,\mathrm{h})$ 的渐近偏差.28 G.2 估计量 $\widehat{\alpha}_0(x;1,\mathrm{h})$ 的一致性.30

1. 引言

回归刻画因变量 $Y_i \in \mathrm{Y}$ 与自变量 $X_i \in \mathrm{X}$ 之间的关系。这种关系可根据数据集 $\mathbf{\Sigma} = \{(X_i,\ Y_i) \in \mathrm{X} \times \mathrm{Y} : i = 1,\ 2,\ \cdots N\}$ 推断得到。其目标是找出最匹配的曲线或超曲面，以描述变量之间的关系。非参数回归对变量之间关系的函数形式作出尽可能少的假设。在欧几里得空间中，若 $\mathrm{X} \times \mathrm{Y} = \mathbb{R}^p \times \mathbb{R}^n$，且 $\mathbf{\Sigma}$ 对应于样本总体（X，Y）的一组独立同分布实现，则非参数回归旨在确定一个光滑的回归函数 $m(\cdot)$，使其满足

$$m(x) = \mathbb{E}[Y|X = x], \tag{1.1}$$

而不对 $m(\cdot)$ 的形状作参数化规定。与假定特定数学方程（例如，$m(X) = \alpha_0 + \alpha_1^T X$）的参数回归不同，非参数回归提供了一种更灵活的方法，其形状可根据数据 $\mathbf{\Sigma}$ [1, 2, 3] 推断出 $m(\cdot)$。

欧几里得空间中的一种特定非参数回归是局部多项式回归（LPR）。该方法针对每个数据点，在其周围的局部邻域内将局部多项式（LP）回归模型拟合到数据，而不是用单个全局多项式拟合整个数据集。LPR 能够采用更具适应性的模型来处理数据中的复杂非线性关系。严格来说，k 阶 LPR 是一种非参数方法，

它由对一个 $k+1$ 次可微回归函数 $m(x)$ 及其导数 $\partial_x^j m(X)$ 的估计组成，$0 < j \leq k$。与全局多项式回归相比，LPR 能有效处理变量之间的关系在数据集各处存在异质性的情况。欧几里得空间中的其他非参数回归方法包括核回归和样条回归 [1, 2, 3, 4]。

黎曼值数据 $\mathrm{Y} = \mathrm{M}_n \neq \mathbb{R}^n$ 出现在医学影像、计算解剖学、机器学习、计算机视觉、金融等领域 [5, 6, 7, 8]。与欧几里得空间不同，黎曼流形 $(\mathrm{M}n, g)$ 是一种非线性数学结构，在局部与欧几里得空间相似。具体而言，在每个点处，都可以用一个切空间来近似它，而切空间本质上是一个欧几里得空间 [9, 10]。非线性意味着该结构并非平坦，而且 $\mathrm{M}n$ 中元素的线性组合不一定仍属于该流形。因此，对于黎曼流形 $(\mathrm{M}n, g)$ 上的数据，像欧几里得均值 $(\widehat{Y} = \sum Y_i/N)$ 这样简单的集中趋势估计量通常与数据的几何结构不一致 $(\widehat{Y} \notin \mathrm{M}n)$。计算黎曼值数据均值的恰当方法需要使用 Karcher 均值 [11]。对黎曼值数据进行回归时也会遇到相同情况。流形值数据集 $\mathrm{X} \times \mathrm{Y} = \mathbb{R}^p \times \mathrm{M}n$ 中数据点之间的关系可能复杂且非线性，并且取决于底层流形的几何结构，因此传统的基于欧几里得空间的回归技术（如 LPR）并不适用。在这种情况下，需要采用黎曼流形回归方法，准确地建模数据中的关系 [12, 13, 14]。

黎曼回归方法的主要特点是将经典回归技术改写为黎曼运算的形式，例如指数映射、对数映射、平行移动和内在距离（这些概念在3.1节中定义）。这种方法使我们能够避免回归估计量与数据几何结构不一致的问题 [11]。在黎曼流形上的回归中，主要有两种方法：外在方法和内在方法。内在方法是指未知回归函数 $m(X)$ 属于同一流形，或属于能够保留所有原始几何信息的等价黎曼流形，即 $m(X) \in \mathcal{M}$。另一方面，对于外在方法， $m(X)$ 属于更高维的欧几里得空间，该空间无法保留 $\mathcal{M}$ 的底层几何结构。

外在方法包含两个步骤：首先，将黎曼流形上的取值数据映射到更高维的欧几里得空间；其次，使用经典回归方法对映射后的数据进行回归。虽然这种方法简单直接，但映射后的数据可能无法准确反映数据中真实存在的关系 [6, 15, 16, 17, 18, 19, 10]。

为流形 M 配备度量的方式有无穷多种 g。对于每一种度量，选择映射到更高维欧几里得空间的嵌入映射也有无穷多种方式。一些作者（如 Lin）采用等变嵌入，以尽可能尊重黎曼流形的底层几何结构 [15]。在实践中，文献中的许多作者通过在流形中一个方便的点处使用对数映射来定义这些嵌入 [6, 15]。如果某个点包含了尽可能多的有关数据变异性的信息，那么它就是一个方便的点。通常选择数据的 Karcher 均值作为这个点。

将数据映射到高维空间后，我们可以选择任何已知的经典回归技术来进行回归。例如，我们可以采用线性回归、核回归、局部多项式回归、样条回归或加性回归来处理多个描述变量。高维欧几里得空间中的回归方法取决于我们要解决的问题。最后，我们使用嵌入函数 [15, 10] 的“局部逆映射”，将回归结果映射回原始黎曼流形。在对称正定（SPD）流形的背景下，仿射不变度量是首选，因为它具有合同不变性， [20, 7, 6, 21, 22] 已对此作出证明。文献中一种广泛使用的方法是在仿射不变度量下应用外蕴局部多项式回归，从而实现非参数局部多项式回归 [15, 6, 10]。

在黎曼流形上进行回归的内在方法，是直接运用基于黎曼几何的方法执行回归。这种方法能够准确捕捉数据中的复杂关系。在各种内在方法中，多元广义线性模型（MGLM）将欧几里得空间中的经典参数线性回归方法推广到黎曼流形上 [23]。切空间上的加性回归模型（ARM）将加性回归中的经典结果推广到具有李群结构的黎曼流形上 [2, 24]。Fréchet 回归（FR）将非参数回归方法从欧几里得空间推广到度量空间。这类回归仅利用元素之间的距离，因此适用于一般度量空间，即使在这些空间中泰勒展开可能无法进行 [13]。最后，内在局部多项式回归（ILPR）是一种非参数回归方法，它将欧几里得空间中的经典 LPR 推广到黎曼流形上 [12]。内在方法面临的主要难题在于由此产生的回归优化问题十分复杂。通常难以求得精确解，因此往往采用数值技术，而这类技术的计算成本可能很高，也会降低内在方法的精度 [13, 12]。因此，我们亟需新的理论工具来简化其中复杂的数学问题，并使内在优化问题能够更直接地求解。

本文旨在为 Yuan 及其同事提出的多元内在局部多项式回归（ILPR）发展新的理论结果 [12]。目前，ILPR 的最新研究成果缺少一个正式定义，无法处理任意黎曼流形上的多元协变量；同时，对于对称正定（SPD）流形上的新度量（如对数-乔列斯基度量 [25, 14]），也缺少计算 ILPR 的通用表达式。此外，ILPR 尚无闭式解析表达式，也缺乏关于其在任意黎曼流形上的渐近性质的一般性结果。

我们通过在任意黎曼流形上正式定义多元内在局部多项式回归，并研究等距映射对估计的影响，提出了解决这些问题的方法。黎曼等距映射是保持黎曼流形几何结构的数学函数，它能够将一个黎曼流形变换为另一个，同时保持点与点之间的距离和角度不变 [26]。等距黎曼流形（IRMs）是黎曼流形的一个子集，它们通过等距映射具有相似的几何性质。因此，研究几何结构更简单的等距黎曼流形，可以简化对复杂黎曼流形的分析。

核心思想是，可以通过在等距黎曼流形上表述内在黎曼回归方法（如 ILPR），使其得到简化。这种简化使得黎曼流形上的回归问题更容易求解。因此，本文旨在拓展 ILPR 的分析工具，使其不再局限于 SPD 流形，并将其推广到任意黎曼流形上的多变量情形。为此，我们提出一种在等距黎曼流形上计算 ILPR 的新方法（ILPR-IRMs），能够在保持底层几何结构的同时，实现从一个黎曼流形到另一个黎曼流形的全局数据映射。ILPR-IRMs 使我们能够在最便于处理的黎曼流形上表达回归问题。找到化简后问题的解后，我们便可通过内在黎曼流形等距映射恢复原始黎曼流形上的解。因此，我们将探讨以下问题：

(a) 我们将多元内在局部多项式回归的定义推广到任意黎曼流形上的响应，
这些响应依赖于多变量协变量。这个新定义阐明了 IRM 上的等距映射如何影响 ILP
定理（4.5）中。(b) 定理（4.6）推导出了配备欧几里得拉回度量（EPM）的任意黎曼流形上多元内在局部多项式回归的精确解析表达式。该估计量具有闭式表达式，与欧几里得空间中的 LPR 相仿。$\widehat{\alpha}_0(x;\, k,\, h) = f^{-1}\left(f(\mathbf{Y})\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^{-}e_0)\otimes I_n\right)\right)$。(c) 已知对数欧几里得度量、幂欧几里得度量和乔列斯基度量均为 EPM。在引理（3.4）中，我们首次证明对数-乔列斯基度量也是 EPM。(d) 在定理（4.6）中，我们给出了配备某种度量的 SPD 流形上 ILPR 的精确解析表达式

EPM。我们使用 SPD 流形上的模拟数据，将该方法与采用对数-乔列斯基度量和对数欧几里得度量的方法，以及采用仿射不变距离的成熟外在方法进行比较。这些模拟在两种不同的设置下进行，针对 SPD(3) 和 SPD(n) 上的数据，协变量维度各不相同
协变量。基于模拟数据，使用对数-乔列斯基度量进行 ILPR 的计算
速度更快，并且在误差—时间与其他考量指标之间实现了最佳权衡。然而，尽管外在的仿射不变度量略为准确，但所需的计算时间显著更长
(e) 我们开发了 Log-Cholesky Rie t-SNE，这是一种基于 Rie-SNE 方法 [27], 的降维技术，用于可视化 ILPR 在高维模拟回归数据上的性能。(f) 用于实现 ILPR-EPM、Log-Cholesky Rie t-SNE 及其他相关计算的代码可

GitHub 页面链接：https://github.com/ronald1129/Multivariate-Intrinsic-Local-Polynomial-Regression-on-Isometric-Riemannian-Manifolds.git。

2. 一般记号

表 *1.* 基本记号。

流形	描述	Bib
$\mathbb{R}^p$	p 维欧几里得空间	–
$\mathrm{M}_n(\mathbb{R})$	实分量 $n \times n$ 矩阵空间	–
$\mathrm{GL}_n(\mathbb{R})$	$\mathrm{GL}_n(\mathbb{R}) = \{Y \in \mathrm{M}_n(\mathbb{R}) \colon \det(Y) \neq 0\}$	–
$\mathcal{M}$	任意流形	(3.1)
$\mathcal{T}_Y\mathcal{M}$	流形 M 在 Y处的切空间（$Y \in \mathrm{M}$）	(3.1)
$\mathcal{M}_n$	$\mathrm{M}_n(\mathbb{R})$ 的子流形	[28]
Sym_n	对称矩阵空间$_n = \{Y \in \mathrm{M}_n(\mathbb{R}) \colon Y = Y^T\}$	[22]
LT_n	$\mathrm{LT}_n = \{Y \in \mathrm{M}_n(\mathbb{R}) \colon \forall i < j, Y_{ij} = 0\}$	[25]
Diag_n	$\mathrm{Diag}_n = \{Y \in \mathrm{M}_n(\mathbb{R}) \colon \forall i \neq j, Y_{ij} = 0\}$	[25]
SPD_n	对称正定$_n = \{Y \in \mathrm{M}_n(\mathbb{R}) \colon \forall X \in \mathbb{R}^n, \ X^T YX > 0\}$	[22]
矩阵运算	$Y, Y_1, Y_2 \in \mathrm{M}_n(\mathbb{R})$ & $x, X \in \mathbb{R}^p$	–
O_n, I_n	$\mathrm{M}_n(\mathbb{R})$ 的零矩阵和单位矩阵	–
Y^T, Y^{-1}	转置矩阵和非奇异矩阵 Y的逆矩阵	–
列$_i(Y)$, 行$_i(Y)$	第 i 列，共 Y列；第 i 行，共 Y行	–
$\mathrm{colb}_i(Y)$, $\mathrm{rowb}_i(Y)$	分块矩阵的列或行 i Y	[29]
$\mathrm{tr}(Y)$	迹算子，$\mathrm{tr}(Y_{n\times n}) = \sum_{1\le i\le n} Y_{ii}$	–
$\exp(Y)$	矩阵指数，$\exp(Y) = \sum_{i=0}^{\infty} Y^n/n!$	[22]
$\log(Y)$	矩阵对数，$\log(Y) = \exp^{-1}(Y)$	[22]
$Y_1 \otimes Y_2$	Y_1 与 Y_2的克罗内克积	[30]
$Y_1 \boxtimes Y_2$	未分块时 Y_1 和 Y_2 的 Tracy-Singh 分块克罗内克积 Y_1	[29]
$Y_1 : Y_2$	弗罗贝尼乌斯内积，$Y_1 : Y_2 = \mathrm{tr}(Y_1 Y_2^T)$	[31]
$\|\|Y\|\|_2$	Frobenius范数，$\|\|Y\|\|_2 = \sqrt{Y : Y}$	[31]
$Y_{n\times m}^{\otimes j}$	Kronecker幂，$Y_{n\times m}^{\otimes j} = Y_{n\times m} \otimes Y^{\otimes}{}_{(j-1)\,n\times m}$, $Y_{n\times m}^{\otimes 0} = I_m$	[32]
$\partial_x^j Y(X)$	阶数为 j 的向量值矩阵函数的偏导数	[32]
$\mathcal{L}(Y)$	对称正定矩阵的 Cholesky分解 Y	[25]
$\mathbb{D}(Y)$, $\lfloor Y \rfloor$	Y的对角部分，Y的严格下三角部分	[25]
$\mathrm{pow}_\tau(Y)$	$\mathrm{pow}_\tau(Y) = \exp(\tau \log Y)$	[14]
$(Y)1/2$	$(Y)_{1/2} = \lfloor Y \rfloor + \mathbb{D}(Y)/2$	[25]
缩写	–	–
IRM	等距黎曼流形	(3.3)
EPM	欧几里得拉回度量	(3.3)
LPR	欧几里得空间中的多元局部多项式回归	(4.1)
ILPR	多元内在局部多项式回归	(4.2)
ILE	内在局部估计量	(4.3)

3. 背景

3.1. 黎曼流形

流形 $\mathcal{M}$ 是一个局部类似于欧几里得空间的拓扑空间。具体而言，对于每个元素 $Y \in \mathcal{M}$，都可以关联一个称为切空间的欧几里得空间，记作 $T_Y\mathcal{M}$。假设我们在切空间上定义一个映射 $g_S\colon \ T_Y\mathcal{M} \times T_Y\mathcal{M} \to \mathbb{R}$，使得 g_Y 是一个关于 $Y \in \mathcal{M}$ 平滑变化的内积。在这种情况下，二元组 $(\mathcal{M},\ g)$ 构成一个黎曼流形，其中 g 是黎曼度量 [9, 10]。

在流形上，曲线是一个可微映射 $\gamma\colon [t_0, t_1] \to \mathrm{M}$，使得对所有 $t \in [t_0, t_1]$, $\gamma'(t) \in T_{\gamma(t)}\mathrm{M}$。黎曼流形 $(\mathrm{M},\ g)$ 上定义的几何结构，结合曲线的概念，使我们能够定义两个元素 Y_1 和 Y_2 在 M 中的测地距离，记作 $\mathrm{dist}_g(Y_1,\ Y_2)$，即 Y_1 和 Y_2 之间最短路径的长度

$$\mathrm{dist}_g(Y_1, Y_2) = \min_{\gamma \in \tilde{\gamma}_{t_0,t_1}(Y_1,Y_2)} \int_{t_0}^{t_1} dt \sqrt{g_{\gamma(t)}(\gamma'(t), \gamma'(t))}, \tag{3.1}$$

其中 $\tilde{\gamma}_{t_0,\ t_1}(Y_1,\ Y_2)$ 是所有满足 $\gamma(t_0) = S$ 且 $\gamma(t_1) = K$ 的曲线的集合。测地线是 M 上的一条光滑曲线，其切向量 γ' 沿曲线移动时长度和方向保持不变。对于 $V \in T_Y\mathrm{M}$，存在唯一一条测地线，记为 $\gamma_{Y,\ V}(t)$，其定义域包含 $[0, 1]$，满足 $\gamma_{Y,\ V}(0) = S$ 且 $\gamma'_{Y,\ V}(0) = V$。

一般而言，黎曼流形是一个非线性空间。与欧几里得空间不同，线性组合可能不再属于该流形。不过，我们可以定义三种基本运算来操作和研究流形：指数映射、对数映射和平行移动。指数映射记为 $\exp_Y\colon \mathrm{T}_Y\mathrm{M} \to \mathrm{M}$，定义为 $\exp_Y(V) = \gamma_{Y,\ V}(1)$，其中 $\gamma_{Y,\ V}$ 是满足初始条件 $\gamma_{Y,\ V}(0) = Y$ 且 $\gamma'_{Y,\ V}(0) = V$ 的唯一测地线。对数映射记为 $\log_Y\colon \mathrm{M} \to T_Y\mathrm{M}$，定义为指数映射的逆映射，即 $\log_Y = \exp_Y^{-1}$。最后，平行移动记为 $\pi_{Y_1 \to Y_2}\colon T_{Y_1}\mathrm{M} \to T_{Y_2}\mathrm{M}$，是黎曼流形上的线性映射，可沿曲线移动切向量 $V \in T_{Y_1}\mathrm{M}$，同时保持其方向不变。这些运算对于研究和操作黎曼流形至关重要，并在机器学习、计算机视觉和医学影像等领域具有实际应用 [9, 10, 11, 12]。

3.2. 对称正定流形（*SPD*）

对称正定（SPD）流形是一个数学对象，广泛应用于计算机视觉、信号处理和机器学习等领域。本文将介绍在该流形上使用的分析工具。它由一组对称正定矩阵构成，这些矩阵通常用于表示数据集中的协方差或相似性信息，更准确地说 [22, 11, 20, 14, 25]:

$$\begin{aligned} \mathrm{SPD}_n &= \{Y : \mathrm{M}_n(\mathbb{R}) : \forall X \in \mathbb{R}^n : X^T Y X > 0\}. \\ T_Y\ \mathrm{SPD}_n &= \{Y : \mathrm{M}_n(\mathbb{R}) : Y = Y^T\} = \mathrm{Sym}_n\,. \end{aligned} \tag{3.2}$$

赋予 SPD 流形黎曼度量后，它便成为黎曼流形。SPD 流形上可以考虑多种黎曼度量，其中一些最重要的度量汇总于表（2）。

表 2. SPD 流形上的黎曼度量。

度量	**描述** ($Y \in \mathrm{SPD}n$ & $V \in T_Y\mathrm{SPD}_n \simeq$ 对称矩阵空间n)	**Bib**
欧几里得空间	$g_Y^E(V,\ V) = \|\|V\|\|_2^2$	–
幂欧几里得	$g_Y^{PE}(V,\ V) = \|\|d_Y \mathrm{pow}_\tau(V)\|\|_2^2$	[14]
对数欧几里得	$g_Y^{LE}(V,\ V) = \|\|d_Y \log(V)\|\|_2^2$	[22]
乔列斯基	$g_Y^C(V,V) = \|\|d_Y \mathrm{L}(V)\|\|_2^2$	[14]
对数-乔列斯基	$g_Y^{LC}(V,V) = \|\|\lfloor \mathrm{L}(Y)(\mathrm{L}(Y)^{-1} V \mathrm{L}(Y)^{-T})_{1/2} \rfloor\|\|_2^2 + \|\|\mathbb{D}((\mathrm{L}(Y)^{-1} V \mathrm{L}(Y)^{-T})_{1/2})\|\|_2^2$	[25]
仿射不变	$g_Y^{AI}(V,V) = \mathrm{tr}((Y^{-1}V)^2)$	[11]

SPD 流形具有多种可根据待解决问题选用的黎曼度量。例如，在分析大型数据集时，对数-乔列斯基度量和对数欧几里得度量都很适用。然而，有时需要考虑合同不变性，此时仿射不变度量更为合适。

3.3. 等距黎曼流形 ***(IRM)*** 与欧几里得拉回度量 ***(EPM)***

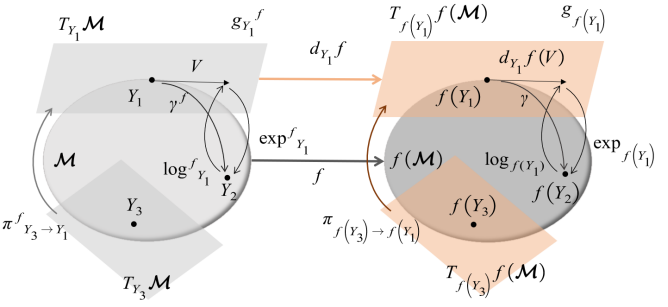


图 ***1.*** 配备拉回度量的流形上黎曼运算之间关系的可视化表示。

在各种实际情形中，可以通过微分同胚 f将流形 $\mathrm{M}n$ 与一个较简单的流形 $f(\mathrm{M}n)$ 联系起来。例如，对 $\mathrm{SPD}n$ 取对数可得到对称矩阵流形 $\mathrm{Sym}n$，可表示为 $\log(\mathrm{SPD}n) = \mathrm{Sym}_n$ [22]。这种情形有利于将一种度量从一个流形推广到另一个流形，或简化问题。一般而言，给定一个微分同胚 $f\colon \mathrm{M}_n \to f(\mathrm{M}n)$，可以利用 $(f(\mathrm{M}n), g)$ 在流形 $\mathrm{M}n$ 上定义一个新度量 g^f，使其成为黎曼流形。这种度量称为拉回度量 [9]:

定义 3.1 (拉回度量与方向导数)。设 $\mathcal{M}n$ 为一个流形，且 $(f(\mathcal{M}n), g)$ 为一个黎曼流形。考虑微分同胚 $f\colon \mathcal{M}_n \to f(\mathcal{M}n)$。由 f 诱导的 g 在 $\mathcal{M}n$ 上的拉回度量定义如下。对于 $Y \in \mathcal{M}n$ 和 $V \in T_Y\mathcal{M}n$，有

$$g_Y^f(V,V) = gf(Y)(d_Yf(V), d_Yf(V)), \tag{3.3}$$

其中，$d_Yf(V)\colon T_Y\mathcal{M}n \to T_{f(Y)}f(\mathcal{M}n)$ 是 f 在 Y 处沿方向 V 的方向导数。也就是说，

$$d_Yf(V) = \lim_{\epsilon\to 0} \epsilon^{-1}\left(f(Y+\epsilon V) - f(Y)\right). \tag{3.4}$$

该定义使得我们能够将所有黎曼运算从一个流形推广到另一个流形，见表（3）。为说明这一概念，我们可以考虑配备对数欧几里得度量的对称正定流形。根据定义（3.1），我们可以很快验证，对数欧几里得度量是弗罗贝尼乌斯度量 [22]。

表 *3.* 配备拉回度量的流形中黎曼运算之间的关系[26]。

黎曼运算	$\mathcal{M}$	$f(\mathrm{M})$	**关系** $(Y,\ Y_1,\ Y_2 \in \mathcal{M},\ V \in T_Y\mathcal{M},\ t \in \mathbb{R})$
度量	g_Y^f	$g_{f(Y)}$	$g_Y^f(V,V) = g_{f(S)}(d_Yf(V),\ d_Yf(V))$
测地线	$\gamma^f_{Y,V}$	$\gamma_{f(Y),d_Yf(V)}$	$\gamma_{Y,\ V}(t) = f^{-1}\left(\gamma_{f(Y),\ d_Yf(V)}(t)\right)$
指数映射	$\exp_Y^f$	$\exp f(Y)$	$\exp_Y^f(V) = f^{-1}\left(\exp_{f(Y)}(d_Yf(V))\right)$
对数映射	$\log_Y^f$	对数 $f(Y)$	$\log_{Y_1}^f(Y_2) = d_{f(Y_1)}\ f^{-1}\left(\log_{f(Y_1)}\ f(Y_2)\right)$
平行移动	$\pi^f_{Y_1\to Y_2}$	$\pi_{f(Y_1)\to f(Y_2)}$	$\pi^f_{Y_1\to Y_2}V = d_{f(Y_2)}\ f^{-1}\left(\pi_{f(Y_1)\to f(Y_2)}d_{Y_1}f(V)\right)$
距离	dist_{g^f}	dist_g	$\mathrm{dist}_{g^f}(Y_1,Y_2) = \mathrm{dist}_g(f(Y_1),\ f(Y_2))$

一般而言，当两个黎曼流形在上述意义下具有等价的几何结构时，借助微分同胚 f，我们可以将这两个黎曼流形中的每一个视为另一个经过变形后的结果。这就是所谓的等距黎曼流形，或者更准确地说:

定义 3.2 (等距黎曼流形)。 设 $(\mathcal{M}_n, g^f)$ 和 $(f(\mathcal{M}_n), g)$ 是两个黎曼流形。微分同胚 $f: \mathcal{M}_n \to f(\mathcal{M}_n)$ 如果保持黎曼度量不变，则称为等距映射，即 g^f 是 g 在 f 下的拉回度量。如果两个黎曼流形 $\mathcal{M}_n$ 和 $f(\mathcal{M}_n)$ 之间存在等距映射，则称它们是等距的。

等距黎曼流形的一个例子是这对 (对称正定$_n$, g^{LE}) 和 (对称矩阵空间$_n$, g^E)。在这种情况下，函数 log 是它们之间的等距映射。这意味着 log(对称正定$_n$) 等于 对称矩阵空间$_n$，而黎曼度量 g^{LE} 则通过使用 g^E 并借助 log 函数将其拉回而得到 [22]。换言之，这两个流形具有相同的底层几何结构，可以在不改变点之间的距离或曲线之间的角度的情况下相互转换。

本文讨论的一个重要概念是欧几里得拉回度量（EPM）。这种黎曼度量源自通过微分同胚对欧几里得度量进行拉回。换言之，给定两个黎曼流形之间的微分同胚，EPM 通过将欧几里得度量从一个流形“拉回”到另一个流形上而得到。这使我们能够使用标准度量比较不同流形中点之间的距离，这在特定应用中可能非常有用。

定义 3.3 (欧几里得拉回度量（EPM）)。设 $(\mathcal{M}_n, g^f)$ 是一个黎曼流形。若存在一个微分同胚 g，使得 f，则 g^f 是由映射 *3.1* 所定义的弗罗贝尼乌斯度量的拉回度量 g^E。 f。

如果空间 M_n 上的黎曼度量 g 可以表示为欧几里得度量经微分同胚 f 的拉回，那么流形 (M_n, g) 和 $(f(\mathrm{M}_n), g^E)$ 都是等距黎曼流形。这是因为 $f: \mathrm{M}_n \to f(\mathrm{M}_n)$ 是一个微分同胚，且 g 是欧几里得度量 g^E 经 f 的拉回。实践中，用于研究对称正定流形性质的许多度量 g 都可以表示为欧几里得拉回度量（EPM）。引理（3.4）总结了对称正定流形上最重要的 EPM。证明见附录（B）。

引理 3.4。 表（g *4*）中的所有黎曼度量均为EPM，这是由微分同胚 f：对称正定$_n \to f$(对称正定$_n$)

表 *4*。 在对称正定流形上诱导欧几里得拉回度量的等距变换 f g .

度量 g	$f(Y)$, $(Y \in$ 对称正定$_n)$	f(对称正定$_n$)
幂 - 欧几里得空间	$\mathrm{pow}_\tau(Y)$	SPD_n
对数-欧几里得空间	$\log(Y)$	Sym_n
乔列斯基	$\mathcal{L}(Y)$	LT_n
对数-乔列斯基	$\lfloor \mathrm{L}(Y) \rfloor + \log(\mathbb{D}(\mathrm{L}(Y)))$	LT_n

3.4. 多元回归的泰勒展开

将 LPR 表述为加权最小二乘优化问题需要对回归函数 [1, 2, 3] 进行泰勒展开。我们感兴趣的是研究这样一类回归问题：协变量属于欧几里得空间，而响应变量属于黎曼矩阵流形；这些流形是 $\mathrm{M}_{n,\,l}(\mathbb{R})$ 的非线性子空间。这意味着回归函数 $m(\cdot)$ 是一个向量值矩阵函数，因此需要一个良好的泰勒展开 [33, 34]。由于存在多种等价的方式来定义向量值矩阵函数关于向量的偏导数 [35, 36, 32]，文献中可以找到多种对 $m(X)$ 进行展开的方案。在这些方案中，我们选择使用 [34] 中给出的泰勒展开，因为它更加自然且直观。该展开使用了附录（A.1）中给出的向量导数的定义。根据该定义，每个向量值函数 $m(X) \in \mathrm{M}_{n,\,l}(\mathbb{R})$ 若在 $\mathbb{R}^p$ 中的 x 处 $k+1$ 次可微，则在 x 的邻域内具有泰勒展开

$$m(X) - m(x) \approx \sum_{j=1}^{k} \alpha_j \left(I_l \otimes (X - x)^{\otimes j} \right), \tag{3.5}$$

$$j!\alpha_j = \partial_x^j m(X), \quad j = 1, 2, \cdots, k, \tag{3.6}$$

其中， $\alpha_j \in \mathrm{M}n,\ mp^j(\mathbb{R})$ 是在 x [34]处计算的 j阶偏导数。处理偏导数时需要使用的一个重要工具是引理（3.5）。该结果已在 [32, 34]中得到证明，并且对于证明本研究的主要结论非常有帮助。

引理 3.5。若 $x \in \mathbb{R}^p,\ A(x) \in \mathrm{C}(\infty)(\mathbb{R}^p,\ M_n(\mathbb{R}))$ 且 $B(x) \in \mathrm{C}(\infty)(\mathbb{R}^p,\ M_n(\mathbb{R}))$ 则

$$\partial_x \left(A(x)B(x)\right) = \partial_x A(x)\left(B(x) \otimes I_p\right) + A(x)\partial_x B(x). \tag{3.7}$$

特别地，若 $B(x) = B$ 为常数，则

$$\partial_x^j \left(A(x)B\right) = \partial_x^j A(x)\left(B \otimes I_p^{\otimes j}\right). \tag{3.8}$$

4. 等距黎曼流形上的多元内在局部多项式回归（ILPR-IRM）

4.1. 欧几里得空间上的多元局部多项式回归（*LPR*）

黎曼矩阵流形记作 $(\mathrm{M}n(\mathbb{R}),\ g^E)$，可视为配备欧几里得度量的欧几里得空间。该流形与配备欧几里得距离的经典欧几里得空间 $\mathbb{R}^{n^2}$ 具有若干相同的几何性质。vec 算子在二者之间建立了同构关系。由于这些相似性， $\mathbb{R}^{n^2}$ 中的回归方法可以轻松推广到 $\mathrm{M}n(\mathbb{R})$。因此， $\mathbb{R}^{n^2}$ [1, 2, 3] 中多元局部多项式回归（LPR）的经典定义可以直接应用于 $\mathrm{M}n(\mathbb{R})$。

以下是在欧几里得空间 $(\mathrm{M}n(\mathbb{R}),\ g^E)$ 中点 x 处，给定来自总体 $(X,\ Y)$ 的一组独立同分布样本点 $\mathbf{\Sigma} = \{(X_i,\ Y_i) \in \mathbb{R}^p \times \mathrm{M}n(\mathbb{R}),\ i = 1,\ 2,\ \cdots,\ N\}$ 时，次数为 k 的经典多元 LPR 的正式定义：

定义 4.1（**多元局部多项式回归（LPR）**）。设 $\mathbf{\Sigma}$ 如上所定义。对于次数为 k 的经典多元局部多项式回归（*LPR*），在 $(\mathrm{M}n(\mathbb{R}),\ g^E)$ 中的 x 处，估计回归函数 $m(x) = \mathbb{E}[Y|X = x]$ 及其导数 $\partial_x^j(m(X) - m(x))$，其中 $j = 1$、2、$\cdots$、k，并假设 m 在 $X = x$ 处是一个 $k+1$ 次可微函数。

定义 4 1 x t X =。的假设意味着, 在 的某个邻域内, m X m x $k+1$ 向量值矩阵函数 ()()() 是一个 次可微函数。因此，我们可以使用第 3.5 小节中介绍的泰勒展开，得到：

$$\begin{aligned} t &: \mathbb{R}^p \to \mathrm{M}n(\mathbb{R}), \\ t(X) &= m(X) - m(x), \\ t(X) &\approx \sum_{j=1}^{k} \frac{1}{j!} \partial_x^j t(X)(I_n \otimes (X-x)^{\otimes j}), \end{aligned} \tag{4.1}$$

其中， $\partial_x^j t(X)/\,j! \in \mathrm{M}n,\ np^j(\mathbb{R})$ 是 $t(X)$ 的泰勒系数。该多项式通过加权最小二乘回归问题在局部拟合：

$$\begin{aligned} \widehat{\alpha}(x,k,h) &= \arg\min_{\alpha} N^{-1} \sum_{i=1}^{N} K_h(X_i - x) \left\| m_{\alpha,x}^{[k]}(X_i) - Y_i \right\|_2^2, \\ m_{\alpha,x}^{[k]}(X) &= \alpha_0 + t_{\alpha,x}^{[k]}(X), \\ t_{\alpha,x}^{[k]}(X) &= \sum_{j=1}^{k} \alpha_j \left(I_n \otimes (X-x)^{\otimes j}\right), \end{aligned} \tag{4.2}$$

其中， $\alpha_0 = m(x)$、$\alpha_j = \partial_x^j t(X)/\,j! \in \mathrm{M}n,\ np^j(\mathbb{R})$，其中 $j = 1,\ 2,\ \cdots,\ k$，而 $Kh(\cdot)$ 是带宽为h的核函数。需要特别指出的是， $\widehat{\alpha}(x;\ k,\ h)$ 是一个分块矩阵，更准确地说

$$\begin{aligned} \widehat{\alpha}(x;k,h) &= (\widehat{\alpha}_0(x;k,h), \widehat{\alpha}_1(x;k,h), \cdots, \widehat{\alpha}_k(x;,k,h)), \\ \widehat{\alpha}_j(x;k,h) &\in \mathrm{M}_{n,np^j}(\mathbb{R}). \end{aligned} \tag{4.3}$$

观察泰勒展开式（4.1）可清楚看出， $j!\ \widehat{\alpha}_j(x;\ k,\ h)$ 是 $\partial_x^j m(X)$ 的估计量。对于标量协变量向量和响应 Y_i，优化问题式（4.2）可归结为经典单变量局部多项式回归问题 [2, 3]。

4.2. 任意黎曼矩阵流形上的多元内在局部多项式回归（ILPR）

如引言中所述，在处理响应位于黎曼流形上的数据时，LPR 方法等统计方法可能并不适用。此类流形的几何结构要求采用一种能够考虑其非线性的统一方法。内在局部多项式回归（ILPR）旨在将 LPR 推广到符合黎曼值数据几何结构的形式[12]。ILPR 最初是针对 SPD 流形提出的，但可以轻松适用于任意黎曼流形。因此，我们提出了 ILPR 的修订定义，使其能够适应涉及等距黎曼流形中更一般协变量向量的情形。

定义 4.2（内在局部多项式回归（ILPR））。设 $\mathbf{\Sigma} = \{(X_i, Y_i) \in \mathbb{R}^p \times M_n, i = 1, 2, \cdots, N\}$ 是来自总体 (X, Y) 的一组独立同分布样本，且 $\omega \in M_n$。在点 $(x, \omega) \in \mathbb{R}^p \times M_n$ 处，黎曼流形 (M_n, g) 上的 k 次 ILPR 是一种用于估计回归函数 $m(x)$ 的统计方法，使得 $\mathbb{E}[\log_{m(x)} Y | X = x] = O_n$，且导数 $\partial_x^j \left(\pi_{m(x)\to\omega}\left(\log_{m(x)} m(X)\right)\right)$, $j = 1\ 2 \quad k \quad m \quad k+1 \quad X = x \quad \pi \quad , , \cdots$, , 假设 在 处是 次可微函数。这里， 和 $\log$ 分别表示由度量 g 导出的平行移动和对数映射。

与原始内在局部多项式回归（ILPR）定义相比，定义（4.2）更加灵活。原始定义见 [12]，将其应用于 SPD 流形时，需要从单位矩阵 I_n进行平行移动；而该定义不需要这一操作，因此适用于更广泛的黎曼矩阵流形。在对数据集 $\mathbf{\Sigma}$进行回归分析时，这种灵活性的提升使我们能够考察等距变换的影响，从而简化回归问题的数学复杂性。值得注意的是，当 n 时，在 SPD 流形上，ILPR 的原始定义和新定义会得到相同的结果 $\omega = I_n$。

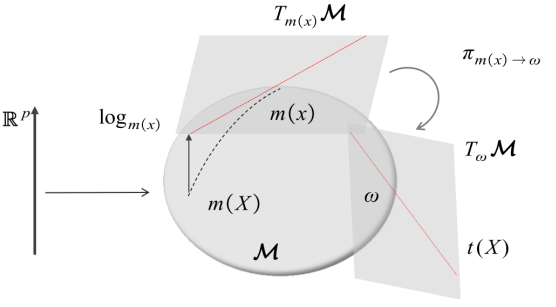


图 2. 函数 $t(X)$的可视化表示。

对于黎曼矩阵流形，函数 $t(X) = \pi_{m(x)\to\omega}\left(\log_{m(x)} m(X)\right)$ 在切空间 $T_\omega M_n$中有像；该切空间是欧几里得空间，可识别为 $M_n(\mathbb{R})$ 的向量子空间。因此，这是一个向量值矩阵函数，并具有第（3.4）小节所述的泰勒展开。因此，定义（4.2）中的假设意味着，在 x 的邻域内，我们可以用一个多变量泰勒多项式来近似函数 $t(X) = \pi_{m(x)\to\omega}\left(\log_{m(x)} m(X)\right)$

$$
\begin{aligned}
t &: \mathbb{R}^p \to T_\omega \mathcal{M}_n \subseteq \mathrm{M}_n(\mathbb{R}), \\
t(X) &= \pi_{m(x)\to\omega}\left(\log_{m(x)} m(X)\right), \\
t(X) &\approx \sum_{j=1}^{k} \frac{1}{j!} \partial_x^j t(X)(I_n \otimes (X - x)^{\otimes j}),
\end{aligned}
\tag{4.4}
$$

其中 $\partial_x^j t(X)/j! \in \mathrm{M}_{n, np^j}(\mathbb{R})$ 是 $t(X)$的泰勒系数。与欧几里得情形类似，该级数展开使我们能够将定义（4.2）中的内在局部多项式回归改写为一个加权回归问题，从而确定内在局部估计量

定义 4.3 (内在局部估计量 (ILE))。 设 $(\mathcal{M}_n, g)$ 为黎曼流形，数据集为 $\mathbf{\Sigma} = \{(X_i, Y_i) \in \mathbb{R}^p \times \mathcal{M}_n, i = 1, 2, \cdots, N\}$。则在 $(\mathcal{M}_n, g)$ 中，点 $(x, \omega) \in \mathbb{R}^p \times \mathcal{M}_n$ 处的 k 次 ILPR 估计量

对于 d 数据 Σ, $\widehat{\alpha}(x;\ k, h) = \mathrm{ILE}(x, \omega, k, h, \Sigma, g)$ 是以下优化问题的解 :

$$
\begin{aligned}
\widehat{\alpha}(x;k,h) &= \arg\min_{\alpha} N^{-1}\sum_{i=1}^{N} K_h(X_i - x)\,\mathrm{dist}_g^2(m_{\alpha,x}^{[k]}(X_i), Y_i) \\
m_{\alpha,x}^{[k]}(X) &= \exp_{\alpha_0}\left(\pi_{\omega\to\alpha_0}\left(t_{\alpha,x}^{[k]}(X)\right)\right), \\
t_{\alpha,x}^{[k]}(X) &= \sum_{j=1}^{k}\alpha_j\left(I_n \otimes (X_i - x)^{\otimes j}\right).
\end{aligned}
\tag{4.5}
$$

其中， $Kh(\cdot)$ 是带宽为 h 的核函数，而 $dist_g$、exp 和 π 分别表示由度量 g诱导的测地距离、指数映射和平行移动。我们将 $\widehat{\alpha}(x;\ k, h) = (\widehat{\alpha}_0\ x\ k\ \ \widehat{\alpha}_1\ x\ k\ \ \widehat{\alpha}_n\ x\ k\ \ \widehat{\alpha}_j\ x\ k\ \ X \in {}_{n\ np^j}\ \mathbb{R}\ \ j = 1\ 2\ \ k\ \ x\ k\ \ m\ x$ $(;, h), (;, h), \cdots, (;, h))$ 记为……，则 $!\,(;, h)$ 是针对 $,, \cdots$ 的 $\partial_x^j t(\,)\,M, (\,)$ 的估计量，而 $\widehat{\alpha}_0(;, h)$ 是 $(\,)$ 的估计量。

在定义（4.3）中，为简化记号，我们将 $\widehat{\alpha}(x;\ k, h) = \mathrm{ILE}(x, \omega, k, h, \Sigma, g)$记作。然而，在本文中每次使用 $\widehat{\alpha}$时，该值都是 x、ω、k、h、Σ 以及 g的函数。

本文使用留一交叉验证（LOOCV）方法进行带宽选择 [1, 2, 3, 4]。最优带宽是留一交叉验证得分 $CV(h)$ 的最小值

$$
\begin{aligned}
\widehat{h} &= \arg\min_{h} CV(h), \\
CV(h) &= N^{-1}\sum_{i=1}^{N}\mathrm{dist}_g^2\left(\widehat{\alpha}_0^{(-i)}(X_i;k,h), Y_i\right), \\
\widehat{\alpha}^{(-i)}(X_i;k,h) &= \mathrm{ILE}(X_i, \omega, k, h, \Sigma^{(-i)}, g), \\
\Sigma^{(-i)} &= \Sigma - (X_i, Y_i).
\end{aligned}
\tag{4.6}
$$

忽略数据的几何信息时 Σ，即将响应的黎曼流形（$\mathcal{M}_n$）视为 M 的欧几里得子空间$n(\mathbb{R})$。定义（4.5）中的内在局部多项式回归将简化为定义（4.1）中的 LPR。忽略几何信息的影响在数学上体现为对式（4.5）进行以下变更

$$
\begin{aligned}
g &\to g^F, \\
\mathrm{dist}_g^2(m_{\alpha,x}^{[k]}(X_i), Y_i) &\to \left\|m_{\alpha,x}^{[k]}(X_i) - Y_i\right\|_2^2, \\
m_{\alpha,x}^{[k]}(X) = \exp_{\alpha_0}\left(\pi_{\omega\to\alpha_0}\left(t_{\alpha,x}^{[k]}(X)\right)\right) &\to m_{\alpha,x}^{[k]}(X) = \alpha_0 + t_{\alpha,x}^{[k]}(X).
\end{aligned}
\tag{4.7}
$$

4.3. 黎曼流形上内蕴局部多项式回归问题之间的关系

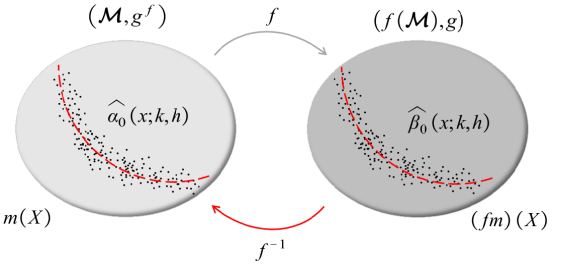


图*3*。 $(\mathcal{M},\ g^f)$和$(f(\mathcal{M}),\ g)$通过等距映射 f 成为等距黎曼流形。流形 $\mathcal{M}$和 $f(\mathcal{M})$上的点分别表示样本 Σ和 $f(\Sigma)$。不连续线分别表示其黎曼流形上的内在局部估计量 $\widehat{\alpha}_0(x;\ k,\ h)$和 $\widehat{\beta}_0(x;\ k,\ h)$。这些内在局部估计量通过等距映射 f 相关联。

在文献中，对数据集 $\Sigma = (X_i,\ Y_i) \in \mathbb{R}^p \times \mathrm{M}_n$，$i = 1,\ 2,\ \cdots,\ N$进行内蕴线性预测回归（ILPR）的标准方法（其响应位于黎曼矩阵流形 $(\mathrm{M}_n,\ g)$ 上）包括

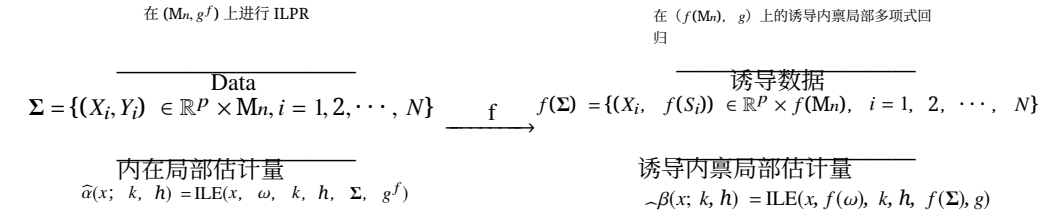
直接计算内在局部估计量 $\widehat{\alpha}(x;\ k,\ h) = \mathrm{ILE}(x,\ \omega,\ h,\ k,\ \mathbf{\Sigma},\ g)$，而不借助其他流形。这需要求解一个加权最小二乘优化问题（定义（4.3））， 该问题取决于黎曼流形 $(\mathcal{M}n,\ g)$的几何结构。然而，如果黎曼度量 g^f 所诱导的几何结构较为复杂，在没有辅助信息的情况下，直接搜索解可能不可行，因此需要采用数值方法。与其外在等价方法相比，这种方法可能会增加计算成本，并降低 ILPR 的准确性。

在实际应用中，通常可以获得有关黎曼流形几何结构的额外信息。这有助于我们找到一个等距映射 f，将一个复杂的黎曼流形 $(\mathrm{M}n, g^f)$ 映射到一个较简单的黎曼流形 $(f(\mathrm{M}n), g)$。为简化内在局部多项式回归方法，我们提出一种名为 ILPR-IRMs 的新方法，该方法利用等距黎曼流形中包含的额外信息，尤其是等距映射 f。

考虑一个数据集 $\mathbf{\Sigma}$，其中包含定义在黎曼流形 $(\mathrm{M}n, g^f)$ 上的响应，且需要使用内在局部多项式回归进行分析。假设 $(\mathrm{M}n, g^f)$ 和 $(f(\mathrm{M}n), g)$ 通过等距映射 f构成等距黎曼流形，这表明 $\mathrm{M}n$ 与 $f(\mathrm{M}n)$ 相似，并且 g^f 是通过 f从 g 导出的度量，这具有重要意义。

首先，等距映射 f将数据的响应空间从 $\mathrm{M}n$ 转移到 $f(\mathrm{M}n)$，从而生成新的数据 $f(\mathbf{\Sigma}) = \{(X_i, f(Y_i)) \in \mathbb{R}^p \times f(\mathrm{M}n), i = 1, 2, \cdots, N\}$（见图（3））。其次，利用具有等距黎曼流形等距映射 f的黎曼度量来确定几何结构，限制更少，也更易于分析。

因此，利用 IRM 的等距映射 f可以简化数据集分析：我们可以针对诱导数据 $f(\mathbf{\Sigma})$（其响应位于黎曼流形 $(f(\mathrm{M}n), g)$ 上）执行第二个 ILPR 问题。这个新问题与原问题等价，但其几何结构更为简单。



在此语境下，尚不清楚 $\widehat{\alpha}(x; k, h)$ 与 $\widehat{\beta}(x; k, h)$之间是否存在关系。本节将探讨这一问题，主要结果见定理（4.5），该结果在特定假设下给出了肯定答案。该定理的证明见附录（D），并依赖于引理（3.5）及后续引理。

引理 4.4。 设 $L: \mathcal{A} \subseteq \mathrm{M}n(\mathbb{R}) \to \mathrm{M}n(\mathbb{R})$ 是一个可逆线性映射。那么存在两个可逆矩阵 $A, B \in \mathrm{GL}n(\mathbb{R})$，使得对所有 $Z \in \mathcal{A}$， $L(Z) = AZB$均成立。

该引理揭示了可逆线性映射在 $n \times n$ 实矩阵子集上的结构。它指出，任意线性映射 L 都可以利用两个可逆矩阵 A 和 B 表示为 $L(Z) = AZB$，且对 Z 子集中的所有元素都成立；该子集记为 $\mathcal{A}$。引理（4.4）的证明见附录（C）。

定理 4.5 （等距黎曼流形上内禀局部多项式回归之间的关系）。 在以下条件下：

(a) (M_n, g^f) 和 $(f(M_n), g)$ 是通过等距映射 f 相关联的等距黎曼流形（*IRM*）。*(b)* 对所有 $(x, \omega) \in \mathbb{R}^p \times M_n$，内在局部估计量（*ILE*）映射 $\widehat{\alpha}(x; k, h) = ILE(x, \omega, k, h, \mathbf{\Sigma}, g^f)$ 以及 $\widehat{\beta}(x; k, h) = ILE(x, f(\omega), k, h, f(\mathbf{\Sigma}), g)$ 均存在。*(c)* 存在唯一矩阵 $A, B \in GL_n(\mathbb{R})$，使得对于所有 $V \in T_{f(\omega)} f(M_n)$，$d_{f(\omega)} f^{-1}(V)$。

*ILE*映射之间满足以下关系：

$$\widehat{\alpha}_j(x; k, h) = \delta_{j0} f^{-1}(\widehat{\beta}_j(x; k, h)) + (1 - \delta_{j0}) A \widehat{\beta}_j(x; k, h)(B \otimes I_p^{\otimes j}), \quad for\ j = 0, 1, 2, \ldots, k, \tag{4.8}$$

$_{wh}$其中 δ_{j0} 是克罗内克δ， $I_p^{\otimes j}$ 是 j 重张量积，由 $p \times p$ 个单位矩阵组成 $_{x.}$

定理（4.5）是求解等距黎曼流形上的 ILPR 问题的重要工具，因为它能够将 ILPR 问题从一个黎曼流形转移到另一个黎曼流形，最终在最方便的流形上求解，如图（3）所示。本文将利用该定理，推导配备 EPM 的黎曼流形上 ILE 的闭式解析表达式。

本节得到了流形 $\mathcal{M}n$ 上多元内在局部多项式回归的精确解析形式，该流形配备了属于欧几里得拉回度量（EPM）族的度量 g^f，其定义见式（3.3）。本节的主要结果是定理（4.6），它是定理（4.5）的推论。

定理 4.6（具有欧几里得拉回度量的内禀局部多项式回归的精确解析解）。 在以下条件下

（a）（$\mathcal{M}n$、 g^f）是一个配备了由等距映射 g^f诱导的欧几里得拉回度量（EPM）的黎曼流形， e 等距映射 f。（b）对于所有（x、 ω） $\in \mathbb{R}^p \times Mn$，内在局部估计量（$ILE$） $\widehat{\alpha}$（x； k， h） $=ILE$（x、 ω、 k， h， $\mathbf{\Sigma}$、 g^f）以及 $\widehat{\beta}$（x； k， h） $=ILE$（x、 f（ω）、 k， h， f（$\mathbf{\Sigma}$）、 g^E）均存在。

则， $\widehat{\alpha}_0(x;\ k,\ h)$ 和 $\beta_0(x;\ k,\ h)$ 的闭式解析表达式为：

$$\widehat{\alpha}_0(x;k,h) = f^{-1}\left(\widehat{\beta}_0(x;k,h)\right), \tag{4.9}$$

$$\widehat{\beta}_0(x;k,h) = f(\mathbf{Y})\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^- e_0)\otimes I_n\right), \tag{4.10}$$

其中

$$\begin{aligned}
e_0 &= (1,0,\cdots,0)^T \in \mathrm{M}_{\sum_{j=0}^k p^j,1}(\mathbb{R}),\\
f(\mathbf{Y}) &= (f(Y_1), f(Y_2),\cdots, f(Y_N)) \in \mathrm{M}_{n,nN}(\mathbb{R}),\\
\mathbf{X} &= \begin{bmatrix}
1 & 1 & \cdots & \cdots & 1\\
(X_1-x)^{\otimes 1} & (X_2-x)^{\otimes 1} & \cdots & \cdots & (X_N-x)^{\otimes 1}\\
(X_1-x)^{\otimes 2} & (X_2-x)^{\otimes 2} & \cdots & \cdots & (X_N-x)^{\otimes 2}\\
\cdots & \cdots & \cdots & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots\\
(X_1-x)^{\otimes j} & (X_2-x)^{\otimes j} & \cdots & \cdots & (X_N-x)^{\otimes j}\\
\cdots & \cdots & \cdots & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots\\
(X_1-x)^{\otimes k} & (X_2-x)^{\otimes k} & \cdots & \cdots & (X_N-x)^{\otimes k}
\end{bmatrix} \in \mathrm{M}_{\sum_{j=0}^k p^j,N}(\mathbb{R}),\\
\mathbf{W} &= \mathrm{diag}(K_h(X_1-x), K_h(X_2-x),\cdots, K_h(X_N-x)) \in \mathrm{M}_{N,N}(\mathbb{R}).
\end{aligned} \tag{4.11}$$

定理（4.6）的证明E详见附录。定理（4.9、4.10）中的一个重要观察是，定理（4.6）指出

$$\widehat{\alpha}_0(x;k,h) = f^{-1}\left(f(\mathbf{Y})\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^- e_0)\otimes I_n\right)\right). \tag{4.12}$$

这一结果给出了在配备欧几里得拉回度量（EPM）族中的度量的任意黎曼流形上，多元内在局部多项式回归（ILPR）的精确解析表达式。

举例来说，假设我们希望求得局部常数估计量 $\widehat{\alpha}(x;0,\mathrm{h})$ $=\mathrm{ILE}(x,\omega,0,\mathrm{h},\mathbf{\Sigma},g^f)$，其中 g^f是由等距映射 f诱导的欧几里得拉回度量。定理（4.6）指出，为了计算 $\widehat{\alpha}_0(x;0,\mathrm{h})$，我们可以计算 $\widehat{\beta}_0(x;0,\mathrm{h})$，它是由等距映射 f诱导的局部估计量。接着，定理（4.6）指出，这类估计量具有由公式（4.10）确定的闭式解析表达式，其中 $f(\mathbf{Y})$ 和 W 按公式（4.11）定义，而 e_0 和 X 则在局部常数回归中特定表达式

$$e_0 = 1,\ \mathbf{X} = 1_{N\times 1}^T, \tag{4.13}$$

可由该表达式计算$\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^- e_0$ 即为

$$\begin{aligned}
\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^- e_0 &= \mathbf{W}1_{N\times 1}(1_{N\times 1}^T\mathbf{W}1_{N\times 1})^- 1\\
&= \frac{(K_h(X_1-x),\cdots,K_h(X_N-x))^T}{\sum_{i=1}^N K_h(X_i-x)},
\end{aligned}$$

现在我们可以按以下方式计算局部常数估计量 $\beta_0(x;0,h)$

$$\begin{aligned}\widehat{\beta}_0(x;0,h) &= f(\mathbf{Y})\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^{-}e_0)\otimes I_n\right)\\ &= (f(Y_1),f(Y_2),\cdots,f(Y_N))\left(\frac{(K_h(X_1-x),\cdots,K_h(X_N-x))^T}{\sum_{i=1}^N K_h(X_i-x)}\otimes I_n\right)\\ &= (f(Y_1),f(Y_2),\cdots,f(Y_N))\frac{(K_h(X_1-x)I_n,\cdots,K_h(X_N-x)I_n)^T}{\sum_{i=1}^N K_h(X_i-x)}\\ &= \frac{\sum_{i=1}^N K_h(X_i-x)f(Y_i)}{\sum_{i=1}^N K_h(X_i-x)}.\end{aligned} \tag{4.14}$$

最后，根据式（4.9），我们可以如下计算 $\widehat{\alpha}(x;0,h)$

$$\begin{aligned}\widehat{\alpha}(x;0,h) &= f^{-1}(\widehat{\beta}_0(x;0,h))\\ &= f^{-1}\left(\frac{\sum_{i=1}^N K_h(X_i-x)f(Y_i)}{\sum_{i=1}^N K_h(X_i-x)}\right).\end{aligned} \tag{4.15}$$

值得注意的是，欧几里得空间 $\beta_0(x;0,h)$ 上的诱导局部估计量 $f(\mathcal{M}n),g^E)$ 的表达式与 Nadaraya-Watson 估计量相似，这正如预期的那样。一般而言，欧几里得空间 $\widehat{\beta}_0(x;k,h)$ 上的估计量 $f(\mathcal{M}n),g^E)$ 与 LPR 相似。具体而言，对于单变量协变量的情况（$p=1$），将式（4.10）向量化后，即可得到欧几里得空间中 LPR 的精确表达式。

针对响应取值于任意配备欧几里得拉回度量的黎曼矩阵流形上的样本，计算多元内在局部多项式回归的算法如伪代码（4.4）所示。值得注意的是，该伪代码包含 Tikhonov 正则化过程，可增强回归过程的数值稳定性。完整算法的 MATLAB 实现可在 GitHub 上获取，网址如下：
https://github.com/ronald1129/Multivariate-Intrinsic-Local-Polynomial-Regression-on-Isometric-Riemannian-Manifolds.git。

算法 1 具有欧几里得拉回度量的内禀局部多项式回归（ILPR-EPM） 输入： $\mathbf{\Sigma}=\{(X_i,Y_i)\in\mathbb{R}^p\times\mathrm{M}n,i=1,2,\cdots,N\},x\in\mathbb{R}^p,k,\mathrm{h}$, $K\mathrm{h}(\cdot)$, f输出： $\widehat{\alpha}_0(x;k,\mathrm{h})$ 1: $\lambda\leftarrow 10^{-3}$，Tikhonov 正则化参数 2: pdim $\leftarrow\sum_{j=0}^k p^j$3: $f(\mathrm{Y})\leftarrow[\ ]$4: X $\leftarrow[\ ]$ 5: W $\leftarrow O_N$6: 对于 $i=1:\ N$ 执行 7: $f(\mathrm{Y})\leftarrow[f(\mathrm{Y}),\ f(Y_i)]$8: $\mathrm{W}(i,i)\leftarrow K\mathrm{h}(X_i-x)$ 9: $\mathrm{Xcol}_i\leftarrow[1]$10: 对于 $j=2:\ k+1$ 执行 11: $\mathrm{Xcol}_i\leftarrow\left[\mathrm{Xcol}_i;(X_i-x)^{\otimes(j-1)}\right]$12: 结束 13: $\mathrm{X}\leftarrow[\mathrm{X},\mathrm{Xcol}_i]$14: 结束 15: $\widehat{\beta}_0(x;k,\mathrm{h})\leftarrow f(\mathrm{Y})\left((\mathrm{W}\mathrm{X}^T(\mathrm{X}\mathrm{W}\mathrm{X}^T+\lambda^2 I\mathrm{pdim})^{-}e_0)\otimes I_n\right)$16: $\widehat{\alpha}_0(x;k,\mathrm{h})\leftarrow f^{-1}(\widehat{\beta}_0(x;k,\mathrm{h}))$ 17: 返回 $\widehat{\alpha}_0(x;k,\mathrm{h})$

配备 EPM 的对称正定流形是定理（4.6）涵盖的一个特例。因此，我们可以有效利用该定理的闭式解析公式，计算对称正定流形上 EPM 对应的内在局部多项式回归。这尤其有益，因为该定理为 EPM 族中的无穷多种度量提供了精确的解析表达式

，其中包括对数-乔列斯基、对数欧几里得、幂欧几里得和乔列斯基等重要度量。这些度量在小节（3.3）中进行了讨论，并已由引理（3.4）证明属于 EPM。

关于黎曼流形 $(\mathrm{M}n, g^f)$ 上估计量 $\widehat{\alpha}_0(x;\ k, \mathrm{h})$ 的渐近性质，利用欧几里得空间 $(f(\mathrm{M}n),\ g^E)$ 上估计量 $\widehat{\beta}_0(x;\ k, \mathrm{h})$ 与 LPR 估计量之间的相似性大有裨益。基于此，我们建立了定理 (4.7)，证明了针对 EPM 的局部常数和局部线性内在局部多项式回归具有一致性，具体而言，即 $\widehat{\alpha}_0(x;\ k, \mathrm{h})$ 对于 $k = 0, 1$。定理 (4.7) 的详细内容见附录 (G)，其中给出了估计量 $\widehat{\beta}_0(x; 1, \mathrm{h})$ 渐近偏差的精确公式。该公式表明估计量 $\widehat{\alpha}_0(x; 1, \mathrm{h})$ 具有一致性，这是统计推断所期望的性质。

定理 4.7。 设 x 是 $\mathrm{supp}(F_X)$ 内部的一个固定元素。假设正则性条件 $(R1) - (R3)$ 成立。那么

$$(a)\ \mathbb{E}\left[\widehat{\beta}_0(x; 1, h) - (fm)(x) | X_1, \cdots, X_N\right] = \frac{1}{2}h^2\mu_2(K)\partial_x^2(fm)(X)\,\mathrm{vec}(I_p)\otimes I_n + o_p(h^2 I_n).$$
$$(b)\ \mathbb{E}\left[\widehat{\alpha}_0(x; 1, h) - m(x) | X_1, \cdots, X_N\right] \to_{N\to\infty} O_n.$$

5. 对称正定流形上的模拟

在对称正定流形上，采用仿射不变度量的外蕴局部多项式回归（LPR）被广泛用于开展非参数回归 [20, 6, 19, 21]。该方法需要在流形上选择一个参考点，例如内蕴均值（也称为 Karcher 均值）或单位矩阵 I_n，然后以该参考点为基准，将数据局部嵌入高维欧几里得空间。由于仿射不变度量具有合同不变性，在脑机接口、计算机视觉和机器学习等领域有诸多实际应用，因此常被采用 [20, 7]。然而，使用该度量进行非参数推断的计算成本可能很高，需要进行大量对数运算 [25]。此外，对于高度扭曲的数据，这种方法的外蕴性质并不合适。所谓高度扭曲的数据，是指各元素与参考点相距甚远的数据；将这类数据准确地嵌入能够反映流形底层几何结构的高维欧几里得空间颇具挑战性。

前文的理论结果（第4.4小节）表明，为流形配备 EPM 后，即可使用闭式解析表达式计算 ILPR 估计量。本研究将这一通用理论应用于 SPD 流形上的 ILPR，并明确采用对数-乔列斯基度量和对数欧几里得度量；二者均为 EPM。这些度量所需计算的矩阵对数较少，因此计算效率高于仿射不变度量，但不具备一般的合同不变性 [25, 14]。我们通过模拟，将采用对数-乔列斯基度量和对数欧几里得度量的 ILPR 估计量与采用仿射不变度量的外禀 LPR 进行比较，以评估其性能。后一种度量是文献中的标准方法。我们旨在评估各方法在流形维度和协变量维度变化时，回归误差与计算时间之间的权衡。

为评估各回归方法的性能，我们进行了蒙特卡洛模拟，其中流形维度和协变量维度限制为 $(p,\ n) \in ([1, 6] \times [3, 18]) \cap \mathbb{N}^2$。对于每组协变量维度和流形维度 $(p,\ n)$，我们生成了 100 次模拟回归问题的实现，包括真实数据、含噪数据、估计数据；其中响应位于 SPD (n) 流形上，协变量位于 $\mathbb{R}^p$。对于每次实现和每种回归方法——ILPR 对数-乔列斯基、ILPR 对数-欧几里得和 LPR 仿射不变——我们都计算了均方根误差（RMSE）和计算时间（CPU 时间）。

为生成真实数据响应，使用以下函数 $m(X, p, n)$

$$(\log m(X, p, n))_{ij} = \delta_{ij}\frac{1_p^T X}{2} + (1 - \delta_{ij})\frac{\lambda_{ij}^T X}{||\lambda_{ij}||},\ \ i, j = 1, 2, \cdots, n \tag{5.1}$$
$$\lambda_{ij} = \lambda_{ji},\ \ \lambda_{ij} = \mathrm{U}_p(0, 1).$$

其中，δ_{ij}为克罗内克δ，向量 $1_p \in \mathbb{R}^p$的所有分量均等于1。协变量向量根据模型 $X_i \sim \mathrm{N}^p(0_{p\times 1}, I_p) \in \mathbb{R}^p$生成，其中 $i = 1, \cdots, N$，这里 N表示每次模拟的样本维度。因此，参数为(p, n)的模拟中，真实数据由$\mathbf{\Sigma}_p^{TRUE}{}_{,n} = \{(X_i, m(X_i, p, n) \in \mathbb{R}^p \times$对称正定$n\colon i = 1, \cdots, N\}$给出。这里，$m(X_i, p, n)$表示与维度为 p的协变量 X_i相对应的真实协方差矩阵，其流形维度为 n。

为获得数据 +噪声数据集，将真实数据的响应 $\mathbf{\Sigma}_{p\ n}^{TRUE}$，使用黎曼对数正态噪声模型 [12, 37]加入噪声。文献 [12, 38]中提出了其他将噪声加入 SPD 流形上的样本的方法，例如对数正态噪声模型和莱斯噪声模型。对于

简单起见，本文仅使用黎曼对数正态模型。然而，此处介绍的过程可以推广到任意噪声模型，这是一个值得进一步研究的课题。黎曼对数正态模型定义如下：

$$\begin{aligned} m^{NOISE}(X_i, p, n) &= L \exp(\mathcal{E}_i) L^T, \\ L &= \mathcal{L}(m(X_i, p, n)), \\ \mathcal{E}_i &= \text{vech}^{-1}\left(\mathcal{N}_{n(n+1)/2}(0_{(n(n+1)/2)\times 1}, (0.5)^2 I_{n(n+1)/2})\right), \end{aligned} \tag{5.2}$$

其中 vech^{-1} 是半向量化算子的逆算子 [29]。因此，数据 $+$噪声为 $\mathbf{\Sigma}_p^{NOISE}$，$n$ $=\{(X_i,\ m^{NOISE}(X_i,$ $p,\ n) \in \mathbb{R}^p \times$对称正定$n$：$i = 1,\ \cdots,\ N\}$。

如前所述，估计数据是通过对含噪数据进行估计得到的 $\mathbf{\Sigma}_{p\ n}^{NOISE}$，仅采用三种不同方法的局部线性情形：ILPR 对数-乔列斯基、ILPR 对数-欧几里得和外在 LPR 仿射不变。为此，使用算法（4.4）计算对数-乔列斯基度量和对数欧几里得度量对应的 ILPR 估计量，并在 X_i 进行评估，以获得每种度量对应的等距映射 f，如表（4）所示。

$$\begin{aligned} f^{\text{LC}}(S) &= \lfloor \mathcal{L}(S) \rfloor + \log \mathbb{D}(\mathcal{L}(S)), \\ f^{\text{LE}}(S) &= \log(S). \end{aligned} \tag{5.3}$$

每种回归方法的最优带宽 $\widehat{\ }h$ 均通过留一交叉验证计算，所用公式为（4.6），并采用各度量生成的内在黎曼距离 dist^{g^f} 和高斯核函数 [2, 3]。模拟仅关注各回归方法的局部线性情形（$k = 1$）。因此，采用每种度量（对数-乔列斯基、对数欧几里得、仿射不变）得到的估计数据 g^f $\mathbf{\Sigma}_{p\ n\ G}^{EST\ IM} = X_i\ \widehat{\alpha}_0\ X_i\ k \in$„{(，(；，$\mathbb{R}^p \times$ $i = 1$ N h)) 对称正定n：，$\cdots$，}。

为评估每次模拟中各回归方法的性能，我们针对每个度量 g^f，计算真实数据与估计数据之间的均方根误差（RMSE）。该计算的精确公式如下：

$$\text{RMSE}_{AI}^2 = \frac{1}{N} \sum_{i=1}^{N} \text{dist}_{AI}^2 \left(\widehat{\alpha}_0(X_i; k, h), m(X_i, p, n)\right). \tag{5.4}$$

为计算计算时间（CPU 时间），我们使用 Matlab 的一个内部函数，测量各回归方法计算估计数据所用的时间。

蒙特卡洛模拟的结果分为两个独立的小节进行展示。第一小节（5.1）重点介绍流形 SPD(3) 上的结果；该流形对应于扩散张量成像（DTI）中的扩散张量场 [22, 11]。在此情形下，回归问题将采用数据的椭球表示 [39, 40]。与此同时，小节（5.2）展示了改变流形维度与协变量维度时，不同回归方法在定量和定性方面的表现。

DTI 中常用的椭球方法已不足以表示高维（SPD(n), $n >> 3$）数据。为解决这一问题，近期一篇论文提出采用两种互补的降维技术 Rie-SNE 和线性化 PGA，来可视化高维回归问题 [27, 41]。

Rie-SNE 是一种先进版本的流行 t-SNE 可视化技术，可应用于具有黎曼度量的任意流形 [27, 42, 43]。该方法能够将数据非线性嵌入到低维空间，同时保留“邻域”条件并遵循流形的底层几何结构。为了判断两个元素是否属于同一邻域，Rie-SNE 需要计算它们之间的黎曼距离以及用于嵌入的度量张量行列式。然而，这些计算的开销可能很大，并且会随着数据元素数量的增加而迅速攀升 [27]。对于大数据集而言，快速高效的 Rie-SNE 版本不可或缺，因此在执行 Rie-SNE 之前，选择合适的黎曼度量至关重要。本文提出在对称正定流形上使用对数-乔列斯基度量，以实现 Rie-SNE 在该流形上的预期效率。该度量无需进行大量矩阵对数运算。在其他问题中，它的表现与对数欧几里得度量和仿射不变度量相似，因此适用于大型数据集 [25]。本文将此技术称为 Log-Cholesky Rie t-SNE。

线性化 PGA 是另一种已知技术，它将经典 PCA 从欧几里得空间扩展到黎曼流形，以可视化黎曼值数据。它提取流形中最具信息量的方向，用以描述数据的方差，因此是可视化高维数据结构与变异性的有用工具。论文使用线性化 PGA 来比较通过 Log-Cholesky Rie t-SNE 得到的回归结果。

论文预期，合适的回归方法能够使估计数据在 Log-Cholesky Rie t-SNE 和线性化 PGA 可视化技术中都接近真实数据。因此，数据+噪声被隔离开来，回归方法也就可被视为成功捕捉了高维数据的底层结构。

5.1. 模拟 *I*：*DTI* 类型数据

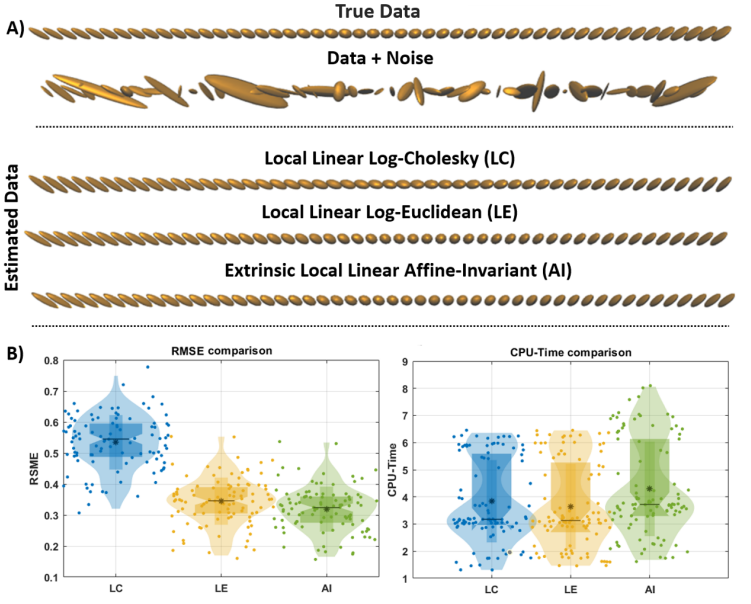


图 *4.* 图 *4* 比较了通过蒙特卡洛模拟在 *SPD(3)* 流形上应用三种非参数回归方法的性能。**A)** 椭球表示展示了真实数据和含+噪声数据的一次实现。所有方法处理噪声时的表现相近。**B)** 图 *4* 使用箱线图比较了各方法的均方根误差和 *CPU* 时间。对于 *DTI* 模拟数据，对数欧几里得方法在误差与非线性回归耗时之间实现了最佳权衡。

本文本节研究了所提出的非参数回归方法在三维 SPD 流形上的性能，并将其与文献中已有的方法进行比较。正如上一节所述，该流形源于弥散张量成像分析。在 DTI 中，对称正定矩阵表示每个体素对应的水分子扩散情况。为了评估这些非参数回归方法的性能，我们针对标量协变量，在 SPD(3) 流形上进行了蒙特卡洛模拟。具体而言，我们生成了$N = 100$ 个真实数据实现，并据此按照上一节所述，计算了各非参数回归方法的均方根误差和 CPU 时间。

在图 (4. A) 中，我们使用椭球表示展示了一组真实数据和数据+噪声的单次实现。真实数据根据模型方程 (5.1) 生成，其中 $n = 3$ 且 $p = 1$；而数据+噪声则基于真实数据，使用噪声模型方程 (5.2) 生成。下面三行对应于使用本节开头讨论的三种非参数回归模型得到的估计数据：内蕴局部线性对数-乔列斯基、内蕴局部线性对数欧几里得，以及外蕴局部线性仿射不变。

如图（4. A）所示，每种回归方法得到的估计数据在视觉上都比数据+噪声更接近真实数据。所有回归方法都捕捉到了复杂的非线性模式，并有效去除了数据+噪声，表明它们非常适用于此类数据。

图 (4. B)pro　　进一步揭示回归方法的性能。这里，我们比较
e 均方根误差和 CPU
使用箱线图展示各方法所需的时间。我们观察到，局部线性对数-乔列斯基（Local Linear Log-Cholesk　　y

估计量的均方根误差最高，而使用仿射不变度量的外在方法的均方根误差最低。有趣的是，对数欧几里得估计量的表现与仿射不变估计量相似。

此外，图（4. B）还展示了每种方法计算估计数据所需的计算时间，即 CPU 时间。我们观察到，内蕴局部线性对数欧几里得方法所需的计算时间最少，而外在局部线性仿射不变方法所需的计算时间最多。在所有用于执行非线性回归的方法中，对数欧几里得方法在误差与时间之间实现了最佳权衡。该方法的性能相对较好，同时所需的计算资源少于其他方法。

5.2. 模拟 *II*。高维数据

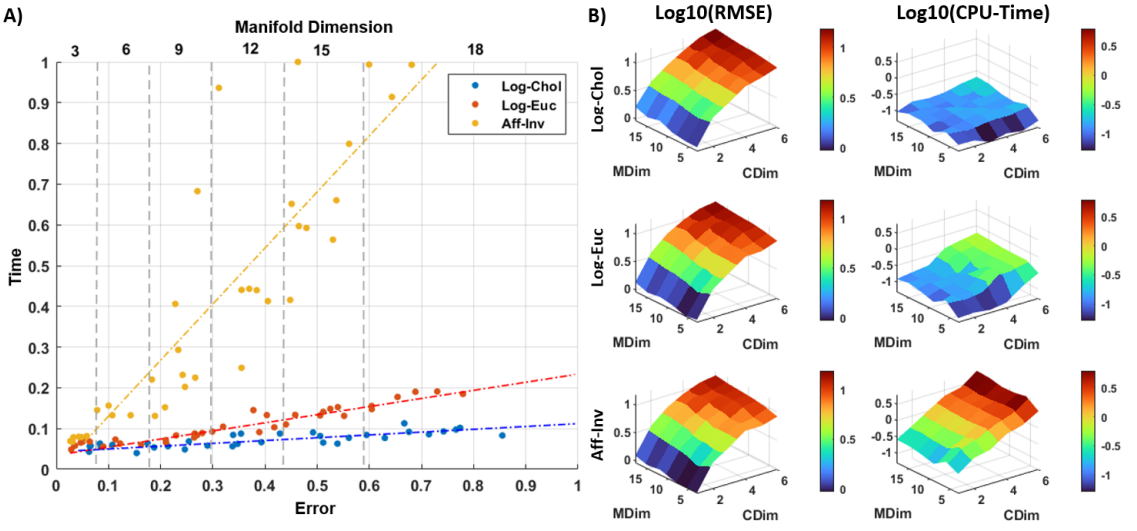


图 *5*。 散点图（**A**）和热图（**B**）展示了随着流形与协变量维度增加，非参数回归方法的性能权衡。**A**）在 *SPD* 流形上使用对数-乔列斯基和对数欧几里得度量的内在局部线性估计量，以及使用仿射不变度量的外在局部线性估计量，其归一化 *CPU* 时间和均方根误差值分别以蓝色、红色和黄色点表示。**B**）热图的第一列以对数尺度显示均方根误差值，第二列同样以对数尺度显示 *CPU* 时间值。

在本节中，我们考察随着流形维度和协变量维度同时增加，非参数回归方法的性能如何变化。为此，我们针对范围 $(p, n) \in ([1, 6] \times [3, 18]) \cap \mathbb{N}^2$ 内的每种流形维度与协变量维度组合进行了蒙特卡洛模拟。

图（5. A）展示了蒙特卡洛模拟中所考察的所有非参数回归方法的计算性能散点图。蓝色和红色圆点分别对应于在 SPD 流形上使用 Log-Cholesky 度量和对数欧几里得度量的内在局部线性估计量。黄色圆点对应于使用仿射不变度量的外在局部线性估计量。每个圆点表示在特定协变量维度下，所有方法的 CPU 时间和均方根误差的归一化值。协变量的取值范围标注在图表顶部。我们通过计算蒙特卡洛模拟后的数值与各特征在所有回归方法中的最大值之比，对 CPU 时间和均方根误差这两项特征进行归一化。从图（5. A）可以看出，流形维度增加会使所有回归方法的计算时间（即 CPU 时间）增加。尤其是，外在仿射不变方法的计算开销最大。相比之下，内在 Log-Cholesky 方法的计算时间随流形维度增长得最慢。

从误差（即均方根误差）来看，我们可以得出结论：所有回归方法的取值都处于同一范围内。外在仿射不变方法的表现略优于其他方法。尽管如此，各方法的误差值总体上仍处于相近的范围。

图（5. B）为图（5. A）中的数据提供了另一种表示方式。这次，图中给出了用于评估性能的各项特征值：均方根误差（第一列）和 CPU 时间（第二列），均采用对数尺度。从第一列可以看出，在流形维度（MDim）和协变量维度（Cdim）变化时，各回归方法的均方根误差在定性和定量上表现相似。所有方法的均方根误差总体趋势是：随着流形维度增加而略微上升，随着协变量数量增加而大幅上升。这种随协变量维度增加而快速增长的现象，可能是回归文献中所称的已知现象“维数灾难”所导致的 [46]。

第二列显示了给定流形和协变量维度时，进行一次蒙特卡洛模拟所需的计算时间。从第二列可以看出，与其他方法相比，内在 Log-Cholesky 方法的计算时间随规模增长的幅度非常小。另一方面，外在仿射不变回归方法的计算时间会随着流形维度和协变量维度的增加而迅速增长，这意味着该回归方法的时间性能最差。同时，内在对数欧几里得方法的时间性能介于另外两种回归方法之间。对于所有回归方法而言，总体趋势都是计算时间随着流形维度和协变量维度的增加而增加。这些回归方法在时间性能方面呈现出的定性趋势符合预期，其原因与每种度量所需的矩阵对数计算量有关 [25, 22]。

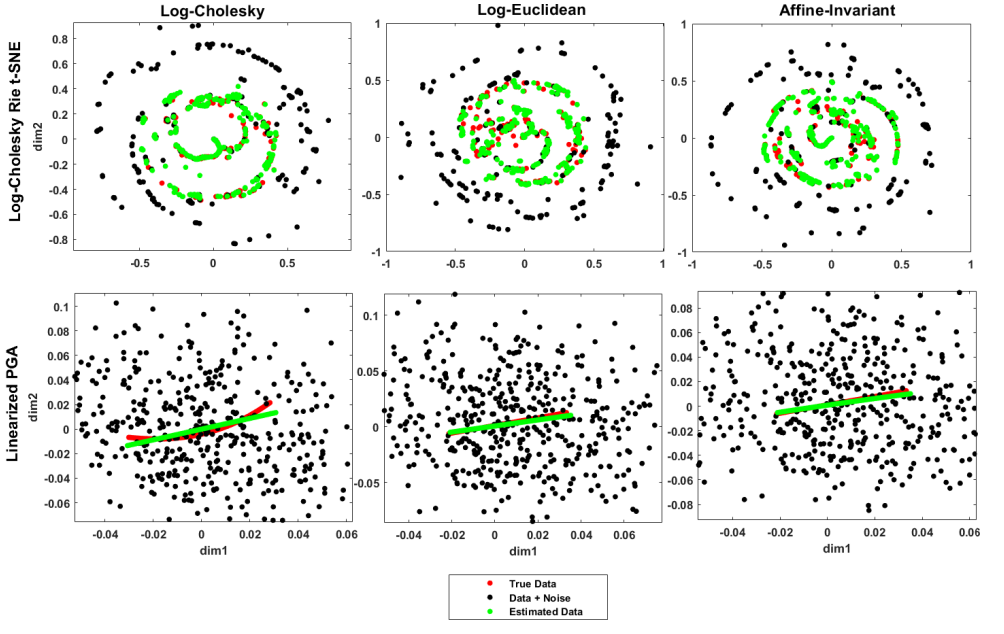


图 **6.** 使用 *Log-Cholesky Rie t-SNE* 和线性化 *PGA* 对具有标量协变量的 *SPD(18)* 流形上的一次蒙特卡洛模拟实现进行可视化。*Log-Cholesky Rie t-SNE* 能够有效捕捉非线性，并在含噪情况下揭示出引人入胜的螺旋模式，而线性化 *PGA* 则呈现出类似的定性行为。对于所有回归方法，真实数据（红色数据）均接近估计数据（绿色数据），表明这些方法能够有效处理含噪数据。

视觉探索通常用于深入了解估计量的性能。然而，椭球表示对于高维对称正定数据已不再有效。因此，采用了降维技术，例如 Log-Cholesky Rie t-SNE 和线性化 PGA。我们的团队开发了一种名为 Log-Cholesky Rie t-SNE 的新型可视化技术，该技术在配备对数乔列斯基度量的 SPD 流形上执行 Rie SNE。该方法能够快速计算 Rie-SNE 流程中的多个特征。

图（6）展示了在带有标量协变量的 SPD(18) 流形上进行蒙特卡洛模拟时，某次实现中采用 Log-Cholesky Rie t-SNE（第一行）和线性化 PGA（第二行）所得到的结果。第一行展示了模拟中考察的各种回归技术通过 Log-Cholesky Rie t-SNE 得到的结果。我们观察到，无论采用哪种回归方法，真实数据（红色数据）都靠近

估计数据（绿色数据），且距离数据+噪声（黑色数据）更远。这表明所有回归方法都能有效处理含噪数据。采用线性化 PGA 时，定性表现也相似。不过，即使在数据+噪声情形下，Log-Cholesky Rie t-SNE 仍能处理数据的非线性，并捕捉到引人入胜的螺旋模式。

6. 讨论与结论

本文呈现了理论和实践两方面的结果。理论结果包括通用公式，可用于计算或表达最适宜的等距黎曼流形（IRM）上的多元内在局部多项式回归（ILPR）估计量。与以往研究相比，我们对 ILPR 的定义更加形式化，从而能够纳入多元协变量，并在估计问题中处理等距变换。我们确立了一个关键定理，将等距黎曼流形上的多元 ILPR 估计量联系起来，使得在几何结构更简单的黎曼流形上表达 ILPR 问题变得更加容易。

我们的研究得出了一个重要定理，为配备欧几里得拉回度量（EPM）的黎曼流形上的多元 ILPR 估计量建立了闭式解析表达式。这些度量是通过将欧几里得度量从一个流形拉回到另一个流形而得到的，涵盖了一个无限的黎曼度量族。EPM 下多元 ILPR 估计量的闭式表达式与欧几里得空间上的局部多项式回归（LPR）估计量有若干相似之处 [1, 2, 3]。例如，内在局部常数估计量的表达式与欧几里得空间中的局部常数估计量相似，后者也称为 Nadaraya-Watson 估计量 [2, 3]。

基于这些与非参数估计经典结果的相似之处，我们建立了诱导估计量情形下渐近偏差的公式（第 4.3 小节），并证明了在正则性条件下估计量的一致性（附录 F）。这些结果至关重要，因为许多重要的黎曼度量实际上都是 EPM。例如，我们证明，对数-乔列斯基度量 [25] 与对数欧几里得、幂欧几里得和 SPD 流形上的乔列斯基度量一样，都是 EPM；这意味着，我们在正则性条件下得到的结果可直接推出针对这些度量的多元内在局部线性估计量的一致性。

我们还编写了一个伪代码算法，用于在配备 EPM 的黎曼流形上实现多元 ILPR 估计量。该算法概述了所提出的方法，可作为后续实现的初稿。该算法以及其他用于进行模拟的算法可在 GitHub 页面上以 MATLAB 代码的形式获取。为展示该方法的应用，我们在配备不同度量的 SPD 流形上进行了蒙特卡罗模拟。在 [20, 6, 19, 21] 线性情形下，我们将采用对数-乔列斯基和对数欧几里得黎曼度量的多元 ILPR 的性能，与采用仿射不变度量的成熟外禀 LPR 进行了比较。

在三维对称正定矩阵流形上对 DTI 模拟数据进行模拟表明，所考察的所有回归方法都能够处理含噪数据。例如，我们从图（4）的椭球表示中直观地观察到了相似性。另一方面，通过对三维对称正定矩阵流形上标量协变量进行蒙特卡洛模拟来评估性能，结果表明，对于 DTI 类型数据，在误差（均方根误差）与时间（CPU 时间）之间权衡最佳的回归方法，是采用对数欧几里得度量的内在局部线性回归。这一结果符合文献 [22]。均方根误差最低的方法是仿射不变回归方法，但代价是计算时间较长。

随后，我们考察了这些发现能否推广到更高的流形维度和协变量维度。为此，我们建立了一个数学模型，用于生成模拟高维回归问题，并将流形协变量维度纳入考量。对于指定范围内的每种流形维度与协变量维度组合，我们都进行了蒙特卡洛模拟，并在每次迭代结束时计算均方根误差和 CPU 时间（第 5 小节）。图（5）中的结果表明，先前针对 DTI 模拟数据得出的发现无法推广到更高的流形维度和协变量维度。相比之下，在这些更一般的情形下，对数乔列斯基度量实现了更优的时间与误差权衡。最后，通过使用 Log-Cholesky Rie t-SNE 和线性化 PGA，我们展示了：与 DTI 模拟数据的椭球表示一样，所有回归方法都能处理 SPD 流形上的含噪数据，即使在更高维度下，尤其是在 SPD(18) 中也是如此。

此前的研究，例如 Yuan [12], 的研究，考察了 ILPR 估计量在 SPD 流形上的特性，但仅限于对数欧几里得度量和仿射不变度量，并侧重于单变量协变量。然而，该方法缺乏对多变量情形下 ILPR 的严格形式化定义，而且依赖于将数据平行移动到单位矩阵处。在单变量情形下，[12] 中的研究仅为对数欧几里得度量下的 ILPR 估计量推导出了闭式解析表达式。尽管他们进行了模拟，并证明对数欧几里得度量下的 ILPR 估计量优于仿射不变度量下的 ILPR，但他们也推导出了一个

渐近偏差公式。他们证明了该估计量的一致性和正态性。然而，考虑到等距变换的存在，仍需进一步研究这一课题，以探讨一般黎曼流形上多变量 ILPR 估计量更为一般的性质。

通过我们的研究，我们在理解多元 ILPR 估计量的理论和实际性质方面取得了显著进展。我们给出的正式定义使本研究区别于该主题的既有工作，并使我们能够解决原始表述中固有的一些问题。基于这些发现，我们推导出了 EPM度量族（一个无限的度量集合）下多元内在局部多项式回归的闭式解析表达式。此外，我们证明了局部线性情形下该估计量的一致性，并得到了其在 EPM度量族下的渐近偏差公式。这些结果代表着一项重大进展，拓展了此前的研究发现，并将其推广到更复杂的情形。

在 SPD 流形上的蒙特卡罗模拟采用了与 [12]类似的方法。对于 $n = 3$的 DTI模拟数据，对数欧几里得度量在时间与误差之间展现出最佳权衡，这与此前的研究发现一致。然而，随着流形与协变量维度增加，这种性能并未保持。相比之下，我们提出的新方法在局部线性情形下使用对数乔列斯基度量进行多元 ILPR 估计，表现更为出色。这项进展为解决实际数据问题增添了一种有价值的估计量。理解回归问题如何随流形与协变量维度扩展，对功能连接、EEG交叉谱或计算机视觉等实际应用至关重要；在这些领域中，针对 SPD(n) 的回归问题通常涉及 $n >> 3$。

必须承认，本研究的理论和实践结果存在一些局限性。本文建立的定理依赖于较强的假设，因此其有效性在某些情况下可能受到限制。此外，多元 ILPR 估计量的闭式解析表达式仅适用于 EPM 度量。估计量一致性的证明仅在局部常数和局部线性情形下的严格正则性条件下成立。研究在更宽松的假设下验证估计量的有效性，并将一致性证明扩展到更高阶回归（$k > 1$），将会很有意义。此外，正态性是任何估计量的一项关键统计性质，因此推导并证明多元 ILPR 估计量的这一渐近性质将很有价值。对于 EPM，由于该估计量的闭式解析表达式与欧几里得空间中的 LPR 相似，因此可以利用使用FFT分箱的思想推广其他更先进的回归技术 [47]。

关于 SPD 流形，我们证明了一组有限的黎曼度量是 EPM。然而，除了引理（3.4）中所考虑的度量外，许多其他度量也可能是 EPM [14, 48]。并非所有黎曼度量都能成为 EPM，但我们可以将这些结果与黎曼几何中的经典结果（例如纳什嵌入定理）结合起来 [49], ，以研究任意黎曼度量的一般性质。模拟研究仅针对有限的一组度量，测试了使用仿射不变度量进行外蕴局部多项式回归时多元 ILPR 估计量的性能。所使用的模型可能无法准确反映现实世界的情形，因此有必要使用真实数据或其他模型生成真实数据，开展更为一般的研究。有必要更详细地研究随着流形维数增加，回归方法的表现。例如，功能连接矩阵上的典型回归问题要求在 SPD(n）上进行回归，其中包含$n >> 18$。

模拟采用固定的噪声水平（$\sigma^2 = 0.5$），以便与 [12], 中报告的结果进行比较，但需要注意的是，噪声水平的选择可能会影响结果。本模拟研究仅限于方程（5.2）中的一种噪声模型，而该模型可能只能反映某些应用中的噪声特征。为确保结果稳健，有必要考察估计量在不同噪声水平下对不同黎曼噪声模型的敏感性，其中包括扩散张量成像中的瑞利噪声模型等专用模型 [37, 38, 12]。此外，用于带宽选择的最小一留交叉验证方法可能需要充分优化，并且可能对离群值敏感 [2, 3]。因此，对于大型数据集，可能需要采用稳健的交叉验证方法。探索计算效率更高的带宽选择方法也会有所帮助，例如 k 折交叉验证 [2], 、广义交叉验证 [50, 51], 或赤池信息准则 [52]。最后，我们介绍了 Log-Cholesky Rie t-SNE，但仅聚焦于其在高维对称正定数据中的应用。我们的介绍较为简单且有限，因此有必要进一步探索该方法在更复杂任务中的性能 [27]。

在这项研究中，研究人员在更一般的情形下考察了多元 ILPR 估计量——一种用于对对称正定（SPD）矩阵进行回归分析的统计方法。在特定情况下，对数-乔列斯基度量的性能优于其他广泛使用的度量。此外，研究强调，理解回归问题如何随流形与协变量维度的变化而扩展十分重要，因为这对计算机视觉和神经科学等领域具有实际意义。然而，该研究也存在局限，例如依赖较强的假设，且在模拟中测试的度量种类有限

。总而言之，该研究为运用多元 ILPR 估计量研究对称正定矩阵提供了宝贵见解，但仍需开展更多研究，才能全面了解其潜力与局限。

致谢。 感谢 Maria Luisa Bringas Vega 提供技术协助。

经费声明。 本研究由成都市科学技术局项目资助，资助编号为 2022-GH02- 00042-HZ。

利益冲突。 无利益冲突

数据可用性声明。 https://github.com/ronald1129/Multivariate-Intrinsic-Local-Polynomial-Regression-on-Isometric-Riemannian git

A. 关于向量的偏导数的定义

定义 A.1 (矩阵对向量的偏导数)。标量 $\sigma(X)$ 对向量 $X \in \mathbb{R}^p$、向量 $s(X) \in \mathrm{M}_{n,1}(\mathbb{R})$ 以及矩阵 $S(X) \in \mathrm{M}_{n,m}(\mathbb{R})$ 的偏导数分别定义为

$$\begin{aligned}
\partial_X \sigma(X) &= \begin{bmatrix} \frac{\partial \sigma(X)}{\partial x_1} & \frac{\partial \sigma(X)}{\partial x_2} & \cdots & \frac{\partial \sigma(X)}{\partial x_p} \end{bmatrix} \in \mathrm{M}_{1,p}(\mathbb{R}), \\
\partial_X s(X) &= \begin{bmatrix} \partial_X \operatorname{row}_1(s(X)) \\ \partial_X \operatorname{row}_2(s(X)) \\ \cdots \\ \partial_X \operatorname{row}_n(s(X))) \end{bmatrix} \in \mathrm{M}_{n,p}(\mathbb{R}), \\
\partial_X S(X) &= \begin{bmatrix} \partial_X \operatorname{col}_1(S(X)) & \partial_X \operatorname{col}_2(S(X)) & \cdots & \partial_X \operatorname{col}_m(S(X)) \end{bmatrix} \in \mathrm{M}_{n,mp}(\mathbb{R}).
\end{aligned} \tag{A.1}$$

B. 引理（3.4）的证明

证明。 由表（2）可知，对于黎曼度量、对数欧几里得、幂欧几里得、弗罗贝尼乌斯度量和乔列斯基度量，引理（3.4）显然成立。因此，只需证明对数-乔列斯基度量也是 EPM，即证明以下恒等式成立： $f(Y) = \lfloor \mathcal{L}(Y) \rfloor + \log \mathbb{D}(\mathcal{L}(Y))$。

$$\begin{aligned}
g_Y^{LC}(V,V) &= ||d_Y f(V)||_2^2, \\
f(Y) &= \lfloor \mathcal{L}(Y) \rfloor + \log \mathbb{D}(\mathcal{L}(Y)).
\end{aligned} \tag{B.1}$$

其中 $Y \in$ 对称正定n, $V \in T_Y$ 对称正定n，且 g^{LC} 为表（2）中定义的对数乔列斯基度量。证明该恒等式需要用到文献中的先前结果 [25]

$$d_Y \mathcal{L}(V) = \mathcal{L}(Y)(\mathcal{L}(Y)^{-1} V \mathcal{L}(Y)^{-T})_{1/2}, \tag{B.2}$$

$$\mathcal{L}(Y + \epsilon V) = \mathcal{L}_Y(V) + \epsilon d_Y \mathcal{L}(V) + \mathrm{O}(\epsilon^2), \tag{B.3}$$

$$(Y)_{1/2} = \lfloor Y \rfloor + \mathbb{D}(Y)/2, \tag{B.4}$$

$$\mathbb{D}(L_1 + L_2) = \mathbb{D}(L_1) + \mathbb{D}(L_2), \tag{B.5}$$

$$\mathbb{D}(L_1 L_2) = \mathbb{D}(L_1)\mathbb{D}(L_2), \tag{B.6}$$

$$\mathbb{D}(D_1)\mathbb{D}(D_2) = \mathbb{D}(D_2)\mathbb{D}(D_1) \tag{B.7}$$

$$\log(\mathbb{D}(D_1)\mathbb{D}(D_2)) = \log \mathbb{D}(D_1) + \log \mathbb{D}(D_2), \tag{B.8}$$

$$\lfloor L_1 + L_2 \rfloor = \lfloor L_1 \rfloor + \lfloor L_2 \rfloor, \tag{B.9}$$

其中与之前一样， $Y \in$ 对称正定n、 $V \in T_Y$对称正定n，且 $D_1, D_2 \in$ 对角矩阵n、 $L_1, L_2 \in$ 下三角矩阵n。由此恒等式可以计算微分 $d_Y f(V)$

$$\begin{aligned}
d_Y f(V) &= \lim_{\epsilon \to 0} \epsilon^{-1} \left(\lfloor \mathcal{L}(Y + \epsilon V) \rfloor + \log \mathbb{D}(\mathcal{L}(Y + \epsilon V)) - \lfloor \mathcal{L}(Y) \rfloor - \log \mathbb{D}(\mathcal{L}(Y)) \right) \\
&= \lim_{\epsilon \to 0} \epsilon^{-1} \left(\lfloor \mathcal{L}(Y + \epsilon V) \rfloor - \lfloor \mathcal{L}(Y) \rfloor + \log \mathbb{D}(\mathcal{L}(Y + \epsilon V)) - \log \mathbb{D}(\mathcal{L}(Y)) \right).
\end{aligned} \tag{B.10}$$

为了继续进行一阶泰勒展开，需要考虑增量 L ( ) L ( ) 和 log (L ( )) log (L ( ))。

$$\begin{aligned}
\lfloor \mathcal{L}(Y + \epsilon V) \rfloor - \lfloor \mathcal{L}(Y) \rfloor &=^{\text{by B.9}} \lfloor \mathcal{L}(Y + \epsilon V) - \mathcal{L}(Y) \rfloor \\
&\approx^{\text{by B.3}} \lfloor \mathcal{L}(Y) + \epsilon d_Y \mathcal{L}(V) - \mathcal{L}(Y) \rfloor \\
&=^{\text{by B.2}} \lfloor \mathcal{L}(Y) + \epsilon \mathcal{L}(Y)(\mathcal{L}(Y)^{-1} V \mathcal{L}(Y)^{-T})_{1/2} - \mathcal{L}(Y) \rfloor \\
&= \epsilon \lfloor \mathcal{L}(Y)(\mathcal{L}(Y)^{-1} V \mathcal{L}(Y)^{-T})_{1/2} \rfloor.
\end{aligned} \tag{B.11}$$

另一方面，$\log \mathbb{D}(\mathrm{L}(Y+\epsilon V))-\log \mathbb{D}(\mathrm{L}(Y))$ 的一阶泰勒展开为

$$\begin{aligned}
&\log\mathbb{D}(\mathcal{L}(Y+\epsilon V))-\log\mathbb{D}(\mathcal{L}(Y)) =^{\text{by B.8}} \log\left(\mathbb{D}(\mathcal{L}(Y+\epsilon V))\mathbb{D}(\mathcal{L}(Y))^{-1}\right)\\
&=^{\text{by B.3}} \log\left(\mathbb{D}\left(\mathcal{L}(Y)+\epsilon d_Y\mathcal{L}(V)+\mathrm{O}(\epsilon^2)\right)\mathbb{D}(\mathcal{L}(Y))^{-1}\right)\\
&=^{\text{by B.2}} \log\left(\mathbb{D}\left(\mathcal{L}(Y)+\epsilon\mathcal{L}(Y)(\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2}+\mathrm{O}(\epsilon^2)\right)\mathbb{D}(\mathcal{L}(Y))^{-1}\right)\\
&=^{\text{by B.5}} \log\left(\left(\mathbb{D}(\mathcal{L}(Y))+\epsilon\mathbb{D}(\mathcal{L}(Y)(\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2})+\mathrm{O}(\epsilon^2)\right)\mathbb{D}(\mathcal{L}(Y))^{-1}\right)\\
&=^{\text{by B.6}} \log\left(\left(\mathbb{D}(\mathcal{L}(Y))+\epsilon\mathbb{D}(\mathcal{L}(Y))\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2})+\mathrm{O}(\epsilon^2)\right)\mathbb{D}(\mathcal{L}(Y))^{-1}\right)\\
&=^{\text{by B.7}} \log\left(I_n+\epsilon\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2})+\mathrm{O}(\epsilon^2)\right)\\
&\approx \epsilon\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2}).
\end{aligned} \tag{B.12}$$

将式（B.11、B.12）代入式（B.10）得

$$\begin{aligned}
d_Y f(V) &= \lim_{\epsilon\to 0}\epsilon^{-1}\left(\lfloor\mathcal{L}(Y+\epsilon V)-\mathcal{L}(Y)\rfloor+\log\mathbb{D}(\mathcal{L}(Y+\epsilon V))-\log\mathbb{D}(\mathcal{L}(Y))\right)\\
&= \lim_{\epsilon\to 0}\epsilon^{-1}\left(\epsilon\lfloor\mathcal{L}(Y)(\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2}\rfloor+\epsilon\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2})\right)\\
&= \lfloor\mathcal{L}(Y)(\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2}\rfloor+\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2}).
\end{aligned} \tag{B.13}$$

由此，我们终于可以证明恒等式式（B.1）

$$\begin{aligned}
||d_Y f(V)||_2^2 &= |||\lfloor\mathcal{L}(Y)(\mathcal{L}(Y)^{-1}V\mathcal{L}(Y))^{-T})_{1/2}\rfloor+\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2})||_2^2\\
&= |||\lfloor\mathcal{L}(Y)(\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2}\rfloor||_2^2+||\mathbb{D}((\mathcal{L}(Y)^{-1}V\mathcal{L}(Y)^{-T})_{1/2})||_2^2\\
&= g_Y^{LC}(V,V).
\end{aligned} \tag{B.14}$$

□

C. 引理（4.4）的证明

证明。 若 $L:\mathrm{A}\subset\mathrm{M}_n(\mathbb{R})\to\mathrm{M}_n(\mathbb{R})$ 是一个线性可逆映射，则根据线性变换的矩阵表示，存在一个矩阵 $M\in\mathrm{GL}_{n^2}(\mathbb{R})$，使得

$$\mathrm{vec}(L(Z))=M\,\mathrm{vec}(Z),\quad \forall Z\in\mathcal{A}. \tag{C.1}$$

另一方面，根据克罗内克积分解 [53, 54],，存在两个可逆矩阵 A 和 $B\in\mathrm{GL}_n(\mathbb{R})$，使得

$$M=B^T\otimes A, \tag{C.2}$$

因此

$$\begin{aligned}
L(Z) &= \mathrm{vec}^{-1}(\mathrm{vec}(L(Z)))\\
&= \mathrm{vec}^{-1}(M\,\mathrm{vec}(Z))\\
&= \mathrm{vec}^{-1}\left((B^T\otimes A)\,\mathrm{vec}(Z)\right)\\
&= \mathrm{vec}^{-1}(\mathrm{vec}(AZB))\\
&= AZB.
\end{aligned} \tag{C.3}$$

□

D. 定理证明 (4.5)

证明。记由黎曼矩阵流形 (M_n,g^f) 上的度量 G 所诱导的指数映射、对数映射、平行移动和距离分别为 $\exp^f_Y(\cdot)$、$\log^f_Y(\cdot)$、$\pi^f_{Y_1\to Y_2}$和 dist_{g^f}。同样，记黎曼矩阵流形 $(f(\mathrm{M})_n,g)$ 上由度量 g 所诱导的等价运算为 $\exp_{f(Y)}(\cdot)$、$\log_{f(Y)}(\cdot)$、$\pi_{f(Y_1)\to f(Y_2)}$ 和 dist_g。由于根据假设，f是等距映射，g^f是由 f诱导的 g的拉回度量。因此有以下关系，见表（3）

$$\exp^f_Y(V)=f^{-1}\left(\exp_{f(Y)}(d_Yf(V))\right), \tag{D.1}$$

$$\log^f_{Y_1}(Y_2)=d_{f(Y_1)}f^{-1}\left(\log_{f(Y_1)}f(Y_2)\right), \tag{D.2}$$

$$\mathrm{dist}_{g^f}(Y_1,Y_2)=\mathrm{dist}_g(f(Y_1),f(Y_2)), \tag{D.3}$$

$$\pi^f_{Y_1\to Y_2}V=d_{f(Y_2)}f^{-1}\left(\pi_{f(Y_1)\to f(Y_2)}(d_{Y_1}f(V))\right), \tag{D.4}$$

其中 Y、Y_1、$Y_2 \in \mathrm{M}n$, $f(Y_1)$、$f(Y_1)$、$f(Y_2) \in f(\mathrm{M}n)$ 以及 $V \in T_Y \mathrm{M}n$。为了理解证明的推导过程，我们必须牢记，这些应用

$$\begin{aligned} d_Y f &: T_S \mathcal{M}_n \to T_{f(S)} f(\mathcal{M}_n), \\ d_{f(Y)} f^{-1} &: T_{f(Y)} f(\mathcal{M}_n) \to T_Y \mathcal{M}_n, \\ \pi^f_{Y_1 \to Y_2} &: T_{Y_1} \mathcal{M}_n \to T_{Y_2} \mathcal{M}_n, \\ \pi_{f(Y_1) \to f(Y_2)} &: T_{f(Y_1)} f(\mathcal{M}_n) \to T_{f(Y_2)} f(\mathcal{M}_n), \end{aligned} \tag{D.5}$$

都是满足以下性质的可逆线性映射

$$\begin{aligned} (d_Y f)^{-1} &= d_{f(Y)} f^{-1}, \\ (\pi^f_{Y_1 \to Y_2})^{-1} &= \pi^f_{Y_1 \to Y_2}, \\ (\pi_{f(Y_1) \to f(Y_2)})^{-1} &= \pi_{f(Y_2) \to f(Y_1)}. \end{aligned} \tag{D.6}$$

根据假设 b)，ILE $\widehat{\alpha}(x;\ k,\ h)$ 存在，且定义为以下优化问题定义（4.3)的解

$$\begin{aligned} \widehat{\alpha}(x; k, h) &= \arg\max_{\alpha} G_1(\alpha, x, \omega, k, h, \Sigma, \mathcal{M}_n, g^f), \\ G_1(\alpha, x, \omega, k, h, \Sigma, \mathcal{M}_n, G) &= \sum_{i=1}^{N} K_{i,h}(x) \operatorname{dist}^2_{g^f}\left(m^{[k]}_{\alpha,x}(X_i), Y_i\right), \end{aligned} \tag{D.7}$$

其中 $K_{i,\ h}(x) = K_h(X_i - x)$ 对某个带宽为 h 的核函数 $K_h(\cdot)$ 成立，dist^{g^f}是与g^f, $\alpha = (\alpha_0,\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_k) \in \mathrm{M}n \times \mathrm{M}n,\ np(\mathbb{R}) \times \cdots \times \mathrm{M}n,\ np^k(\mathbb{R})$ 相关的黎曼距离，而 $m^{[k]}_{\alpha,\ x}(X)$ 是文中所述的内在多项式

$$\begin{aligned} m^{[k]}_{\alpha,x}(X) &=^{\text{by 4.3}} \exp^f_{\alpha_0}\left(\pi^f_{\omega \to \alpha_0} t^{[k]}_{\alpha,x}(X)\right), \\ t^{[k]}_{\alpha,x}(X) &= \sum_{j=1}^{k} \alpha_j (I_n \otimes (X - x)^{\otimes j}). \end{aligned} \tag{D.8}$$

根据假设 b)，ILE $\widehat{\beta}$ 存在，且定义为以下优化问题定义（4.3)

$$\begin{aligned} \widehat{\beta} &= \arg\max_{\beta} G_2(\beta, x, f(\omega), k, h, f(\Sigma), f(\mathcal{M}_n), g), \\ G_2(\beta, x, f(\omega), k, h, f(\Sigma), f(\mathcal{M}_n), g) &= \sum_{i=1}^{N} K_{i,h}(x) \operatorname{dist}^2_{g}\left((f m)^{[k]}_{\beta,x}(X_i), f(Y_i)\right), \end{aligned} \tag{D.9}$$

而之前 $K_{i,h}(x) = K_h(X_i - x)$ 是针对某个带宽为 h 的核函数 $K_h(\cdot)$，d 是与 g 相关的黎曼距离，$\beta = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k) \in f(\mathrm{M}n) \times \mathrm{M}n,np(\mathbb{R}) \times \cdots \times \mathrm{M}n,np^k(\mathbb{R})$，而 $(f m)^{[k]}_{\beta\ x},(X)$ 是下文所述的内在多项式

$$\begin{aligned} (f m)^{[k]}_{\beta,x}(X) &=^{\text{by 4.3}} \exp_{\beta_0}\left(\pi_{f(\omega) \to \beta_0} t^{[k]}_{\beta,x}(X)\right), \\ t^{[k]}_{\beta,x}(X) &= \sum_{j=1}^{k} \beta_j (I_n \otimes (X - x)^{\otimes j}). \end{aligned} \tag{D.10}$$

为得到定理式 (4.8) 的结论，我们必须找到一种方法来关联式 (D.7, D.9) 中的问题，这意味着我们必须找到一个变量变换 $\alpha = \phi(\beta)$，使得 ϕ: $f(\mathrm{M}n) \times \mathrm{M}n,np(\mathbb{R}) \times \cdots \times \mathrm{M}n,np^k(\mathbb{R}) \to \mathrm{M} \times \mathrm{M}n,np(\mathbb{R}) \times \cdots \times \mathrm{M}n,np^k(\mathbb{R})$ 是一个微分同胚，由此我们保证 $\widehat{\alpha} = \phi(\widehat{\beta})$。要构造这样的微分同胚，首先我们注意到，式 (D.3, D.1, D.8) 意味着我们可以按如下方式改写式 (D.7) 中的优化函数 $G_1(\alpha, x, \omega, k, \mathrm{h},\ \Sigma, \mathrm{M}n, g^f)$

$$\begin{aligned} G_1(\alpha, x, \omega, k, h, \Sigma, \mathcal{M}_n, g^f) &= \sum_{i=1}^{N} K_{i,h}(x) \operatorname{dist}^2_{g^f}\left(m^{[k]}_{\alpha,x}(X_i), Y_i\right) \\ &=^{\text{by D.3}} \sum_{i=1}^{N} K_{i,h}(x) \operatorname{dist}^2_{g}\left(f\left(m^{[k]}_{\alpha,x}(X_i)\right), f(Y_i)\right) \\ &=^{\text{by D.1,D.8}} \sum_{i=1}^{N} K_{i,h}(x) \operatorname{dist}^2_{g}\left(\exp_{f(\alpha_0)}\left(d_{\alpha_0} f\, \pi^f_{\omega \to \alpha_0} t^{[k]}_{\alpha,x}(X)\right), f(Y_i)\right), \end{aligned} \tag{D.11}$$

现在，如果我们仔细观察最后一个表达式，可以注意到该表达式重新构成了优化函数$G_2(\beta,\ x,\ f(\omega),\ k,\ h,\ f(\Sigma),\ f(\mathcal{M}n),\ g)$，见公式（D.9），然而，由于距离项中的第一个项涉及指数映射等因素，因此两者并不完全匹配。另一方面，如果可能的话，我们可以通过选择一种变换使这些表达式相匹配

变量 $\alpha = \phi(\beta)$ 使得方程 (D.11) 中的问题项恰好为内在多项式 $(fm)^{[k]}_{\beta,x}(X)$ 在方程 (D.10) 中，即

$$\alpha_0 = f^{-1}(\beta_0), \tag{D.12}$$

$$\exp_{f(\alpha_0)}\left(d_{\alpha_0} f\, \pi^{f}_{\omega\to\alpha_0} t^{[k]}_{\alpha,x}(X)\right) = \exp_{\beta_0}\left(\pi_{f(\omega)\to\beta_0} t^{[k]}_{\beta,x}(X)\right), \quad \forall X \in \mathbb{R}^p. \tag{D.13}$$

由此，我们已经知道 α 和 β 的第一组分量如何通过 ϕ 进行变换。现在还需确定其余分量 α_j 和 β_j（其中 $j = 1、2、\cdots、k$）之间的关系。利用公式（D.12）和公式（D.13），我们可以消去指数算子，剩下公式（D.13）的参数，即

$$d_{\alpha_0} f\, \pi^{f}_{\omega\to\alpha_0} t^{[k]}_{\alpha,x}(X) = \pi_{f(\omega)\to\beta_0} t^{[k]}_{\beta,x}(X), \quad \forall X \in \mathbb{R}^p, \tag{D.14}$$

这意味着，我们所寻找的变量变换 ϕ 满足公式（D14、D.12）。现在考虑公式（D.4、D.12、D.6）的性质，可得

$$\begin{aligned}
\left(d_{\alpha_0} f\, \pi^{f}_{\omega\to\alpha_0}\right)^{-1} \pi_{f(\omega)\to\beta_0} &=^{\text{by D.4}} \left(\pi_{f(\omega)\to f(\alpha_0)} d_{\omega} f\right)^{-1} \pi_{f(\omega)\to\beta_0} \\
&=^{\text{by D.12,D.6}} \left(d_{f(\omega)} f^{-1} \pi_{\beta_0\to f(\omega)}\right) \pi_{f(\omega)\to\beta_0} \\
&= d_{f(\omega)} f^{-1} \left(\pi_{\beta_0\to f(\omega)} \pi_{f(\omega)\to\beta_0}\right) \\
&= d_{f(\omega)} f^{-1}.
\end{aligned} \tag{D.15}$$

利用公式（D.15），我们可以将公式.（D.14）改写为

$$\begin{aligned}
t^{[k]}_{\alpha,x}(X) &= d_{f(\omega)} f^{-1} t^{[k]}_{\beta,x}(X) \\
&= L\left(t^{[k]}_{\beta,x}(X)\right), \quad \forall X \in \mathbb{R}^p,
\end{aligned} \tag{D.16}$$

其中 $L:\ T_{f(\omega)}\, f(\mathrm{M}n) \to T_{\omega}\mathrm{M}n$ 是可逆线性映射，因为它是可逆线性映射的复合映射

$$L = d_{f(\omega)} f^{-1}. \tag{D.17}$$

由于（$\mathrm{M}n,\ g^{f}$）和（$f(\mathrm{M}n),\ g$）是黎曼矩阵流形，$f(\mathrm{M}n) \subseteq \mathrm{M}n(\mathbb{R})$以及切空间 $T_{\omega}\mathrm{M}n$ 和 $T_{f(\omega)}\, f(\mathrm{M}n)$是 $\mathrm{M}n(\mathbb{R})$ 的向量子空间，因此 L是 $\mathrm{M}n(\mathbb{R})$ 的子空间之间的可逆线性映射。这意味着，引理（4.4）的假设得到满足，因此存在两个可逆矩阵 $A、\ B \in \mathrm{GL}n(\mathbb{R})$，使得

$$L(Z) = AZB, \tag{D.18}$$

结合结果式（D.18、D.16），我们得到

$$t^{[k]}_{\alpha,x}(X) = A t^{[k]}_{\beta,x}(X) B, \quad \forall X \in \mathbb{R}^p. \tag{D.19}$$

为了找出系数 α_l β_l 19 3 5, 和 之间的关系，只需对式（D.）求导,[3 5] 并考虑方程（..）

$$\begin{aligned}
\alpha_l &=^{\text{by 3.5}} \frac{1}{l!}\partial^{l}_{x} t^{[k]}_{\alpha,x}(X) \\
&=^{\text{by D.19}} \frac{1}{l!}\partial^{l}_{x}\left(A t^{[k]}_{\beta,x}(X) B\right) \\
&=^{\text{by 3.5}} A\left(\frac{1}{l!}\partial^{l}_{x}\left(\sum_{j=1}^{k}\beta_j (I_n \otimes (X-x)^{\otimes j})\right)\right)\left(B \otimes I_p^{\otimes l}\right) \\
&=^{\text{by 3.5}} A\beta_l \left(B \otimes I_p^{\otimes l}\right), \quad \forall l = 1, 2, \cdots, k.
\end{aligned} \tag{D.20}$$

方程（D.20、D.12）意味着映射 ϕ

$$\begin{aligned}
\phi : f(\mathcal{M}_n) \times \mathrm{M}_{n,np}(\mathbb{R}) \times \cdots \times \mathrm{M}_{n,np^k}(\mathbb{R}) &\to \mathcal{M}_n \times \mathrm{M}_{n,np}(\mathbb{R}) \times \cdots \times \mathrm{M}_{n,np^k}(\mathbb{R}) \\
\beta = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k) &\to \alpha = (\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_k), \\
\alpha_j = \delta_{j0} f^{-1}(\beta_j) + (1-\delta_{j0}) A\beta_j \left(B \otimes I_p^{\otimes j}\right), &\quad \forall j = 0, 1, 2, \cdots, k
\end{aligned} \tag{D.21}$$

其中 A 和 B 是式（D.18）中定义的 δ_{ij} 可逆矩阵，且 δ_{ij} 是克罗内克δ；该映射是微分同胚，满足

$$\begin{aligned}
\widehat{\alpha} &= \phi(\widehat{\beta}), \\
G_1(\phi(\widehat{\beta}), x, \omega, k, h, \Sigma, \mathcal{M}_n, g^{f}) &= G_2(\widehat{\beta}, x, f(\omega), k, h, f(\Sigma), f(\mathcal{M}_n), g), \\
d_{\widehat{\alpha}_0} f\, \pi^{f}_{\omega\to\widehat{\alpha}_0} t^{[k]}_{\widehat{\alpha},x}(X) &= \pi_{f(\omega)\to\widehat{\beta}_0} t^{[k]}_{\widehat{\beta},x}(X), \quad \forall X \in \mathbb{R}^p.
\end{aligned} \tag{D.22}$$

因此，我们得出该定理的主要论点

$$\widehat{\alpha}_j(x; k, h) = \delta_{j0} f^{-1}(\widehat{\beta}_j(x; k, h)) + (1-\delta_{j0}) A\widehat{\beta}_j(x; k, h)\left(B \otimes I_p^{\otimes j}\right), \quad \forall j = 0, 1, 2, \cdots, k \tag{D.23}$$

□

E. 定理（4.6）的证明

证明。 为了使证明更容易理解，我们将其分为两部分，分别证明定理中的各项陈述。

E.1. ILP 方程（4.9）的表达式

根据假设，黎曼矩阵流形 $(\mathcal{M}n, g^f)$ 和 $(f(\mathcal{M}n), g^E)$ 通过等距映射构成等距黎曼流形 f，这意味着根据定义（3.3），g^f 是欧几里得度量的拉回度量。因此，定理（4.5）的假设成立。特别地，将该定理应用于方程（$j = 0$ 4.8）可得

$$\widehat{\alpha}_0(x;k,h) = f^{-1}(\widehat{\beta}_0(x;k,h)). \tag{E.1}$$

E.2. 方程（4.11）中 ILE 的精确解析表达式

根据假设 $\widehat{\beta}(x;\ k,\ h) = \text{ILE}(x,\ f(\omega),\ k,\ h,\ f(\mathbf{\Sigma}),\ f(\mathcal{M}n),\ g^E)$ 存在，这意味着 $\widehat{\beta}$ 根据定义（4.3）是以下优化问题的解

$$\begin{aligned}\widehat{\beta}(x;k,h) &= \arg\max_{\beta} G_2(\beta, x, f(\omega), k, h, f(\mathbf{\Sigma}), f(\mathcal{M}_n), g^E),\\ G_2(\beta, x, f(\omega), k, h, f(\mathbf{\Sigma}), f(\mathcal{M}_n), g^E) &= \sum_{i=1}^{N} K_{i,h}(x)\,\text{dist}^2_{g^E}\left((fm)^{[k]}_{\beta,x}(X_i), f(Y_i)\right),\end{aligned} \tag{E.2}$$

现在，由假设可知 dist^{g^E} 是由 Frobenius 范数诱导的距离 $g^E(\text{dist}^2_{g^E}(x,y) = ||x-y||_2^2)$。现在，我们可以将优化问题式（E.2）化简为以下问题

$$\begin{aligned}\widehat{\beta} &= \arg\max_{\beta} \sum_{i=1}^{N} K_{i,h}(x)\,\text{dist}^2_{g^E}\left((fm)^{[k]}_{\beta,x}(X_i), f(Y_i)\right)\\ &= \arg\max_{\beta} \sum_{i=1}^{N} K_{i,h}(x)||-f(Y_i) + \sum_{j=0}^{k}\beta_j X^j_{x,n}(X_i)||_2^2,\end{aligned} \tag{E.3}$$

其中，为简洁起见，我们记 $\widehat{\beta} = \widehat{\beta}(x;\ k,\ h)$ 和 $X^j_{x,n}(X_i) = I_n \otimes (X_i - x)^{\otimes j}$。接下来，我们只需计算式（E.3）中优化函数的梯度，因此记 $Z_i = -f(Y_i) + \sum_{j=0}^{k}\beta_j X^j_{x,n}(X_i)$，则

$$d_{\beta_l} Z_i = d(\beta_l X^l_{x,n}(X_i)), \tag{E.4}$$

利用这一事实，我们计算优化函数的微分

$$\begin{aligned}d_{\beta_l}\left(\sum_{i=1}^{N} K_{i,h}(x)||-f(Y_i) + \sum_{j=0}^{k}\beta_j X^j_{x,n}(X_i)||_2^2\right) &= d_{\beta_l}\left(\sum_{i=1}^{N} K_{i,h}(x) Z_i : Z_i\right)\\ &= 2\sum_{i=1}^{N} K_{i,h}(x) d_{\beta_l}(Z_i : Z_i)\\ &= 2\sum_{i=1}^{N} K_{i,h}(x) Z_i : d_{\beta_l} Z_i\\ &=^{\text{by E.4}} 2\sum_{i=1}^{N} K_{i,h}(x) Z_i : d(\beta_l X^l_{x,n}(X_i))\\ &= 2\sum_{i=1}^{N} K_{i,h}(x)(Z_i X^l_{x,n}(X_i)^T) : d\beta_l\\ &= 2\left(\sum_{i=1}^{N} K_{i,h}(x)\left(-f(Y_i) + \sum_{j=0}^{k}\beta_j X^j_{x,n}(X_i)\right) X^l_{x,n}(X_i)^T\right) : d\beta_l,\end{aligned} \tag{E.5}$$

因此，优化函数的梯度为

$$\frac{\partial}{\partial \beta_l}\left(\sum_{i=1}^{N} K_{i,h}(x)||-f(Y_i) + \sum_{j=0}^{k}\beta_j X^j_{x,n}(X_i)||_2^2\right) = 2\sum_{i=1}^{N} K_{i,h}(x)\left(-f(Y_i) + \sum_{j=0}^{k}\beta_j X^j_{x,n}(X_i)\right) X^l_{x,n}(X_i)^T, \tag{E.6}$$

其中 $l=0,\ 1,\ 2,\ \cdots,\ k$。最后，为了求出 $\widehat{\beta}$，我们只需令梯度等于零，由此得到以下 $k+1$ 矩阵方程组

$$\sum_{i=1}^{N} K_{i,h}(x) f(Y_i) X_{x,n}^{l}(X_i)^T = \sum_{j=0}^{k} \widehat{\beta}_j \sum_{i=1}^{N} K_{i,h}(x) X_{x,n}^{j}(X_i) X_{x,n}^{l}(X_i)^T,\ \ l=0,1,2,\cdots,k. \tag{E.7}$$

或更明确的形式

$$\sum_{i=1}^{N} K_{i,h}(x) f(Y_i) \left(I_n \otimes (X_i-x)^{\otimes l}\right)^T = \sum_{i=1}^{N}\sum_{j=0}^{k} \widehat{\beta}_j K_{i,h}(x) \left(I_n \otimes (X_i-x)^{\otimes j}\right)\left(I_n \otimes (X_i-x)^{\otimes l}\right)^T, \tag{E.8}$$

其中 $l=0,1,2,\cdots,\ k$。

为了证明该定理的最后一个结论，我们需要求解 $k+1$ 个关于 $\widehat{\beta}$ 式（E.8）的矩阵方程组。为实现这一目标，我们将使用以下线性代数恒等式

$$\sum_{i=1}^{N} \lambda_i \mathrm{colb}_i(A_{r\times m})\mathrm{rowb}_i(B_{m\times s}) = A_{r\times m}\left(I_{m/N} \boxtimes \mathrm{Diag}(\lambda_1,\lambda_2,\cdots,\lambda_N)\right) B_{m\times s}, \tag{E.9}$$

当 m/N 是整数，且 A 的分块列中的元素个数等于 B对应分块行中的元素个数时，该恒等式对所有分块矩阵 A 和 B 均成立。另一方面，根据 $\widehat{\beta}$的定义、定理末尾的 $f(\mathbf{Y})$和$\mathbf{X}$，我们可以得出：

$$\mathrm{colb}_j(\widehat{\beta}) = \widehat{\beta}_j \in \mathrm{M}_{n,np^j}(\mathbb{R}),\ \ j=1,2,\cdots,k, \tag{E.10}$$

$$\mathrm{colb}_i(f(\mathbf{Y})) = f(Y_i) \in \mathrm{M}_{n,n}(\mathbb{R}),\ \ i=1,2,\cdots,N, \tag{E.11}$$

$$\mathrm{colb}_i(\mathbf{X}) = (\mathrm{rowb}_i(\mathbf{X}^T))^T \in \mathbb{R}^{\sum_{j=0}^{k} p^j},\ \ i=1,2,\cdots,N. \tag{E.12}$$

采用矩阵记法，可将 $k+1$ 个矩阵方程组成的方程组简化为一个方程

$$\sum_{i=1}^{N} K_{h,i}(x) f(Y_i) \begin{bmatrix} I_n\otimes(X_i-x)^{\otimes 0} \\ I_n\otimes(X_i-x)^{\otimes 1} \\ \cdots \\ \cdots \\ I_n\otimes(X_i-x)^{\otimes k} \end{bmatrix}^T = \sum_{i,j}^{N,k} \widehat{\beta}_j K_{h,i}(x) (I_n\otimes(X_i-x))^{\otimes j} \begin{bmatrix} I_n\otimes(X_i-x)^{\otimes 0} \\ I_n\otimes(X_i-x)^{\otimes 1} \\ \cdots \\ \cdots \\ I_n\otimes(X_i-x)^{\otimes k} \end{bmatrix}^T. \tag{E.13}$$

为了进一步简化这个表达式，我们可以注意到，根据 X 的定义

$$\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T) = \begin{bmatrix} I_n\otimes(X_i-x)^{\otimes 0} \\ I_n\otimes(X_i-x)^{\otimes 1} \\ \cdots \\ \cdots \\ I_n\otimes(X_i-x)^{\otimes k} \end{bmatrix}^T, \tag{E.14}$$

$$\mathrm{rowb}_j(\mathrm{colb}_i(I_n \boxtimes \mathbf{X})) = I_n\otimes(X_i-x)^{\otimes j}, \tag{E.15}$$

因此

$$\sum_{i=1}^{N} K_{i,h}(x) f(Y_i)\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T) =^{\text{by E.14}} \sum_{i,j}^{N,k} \widehat{\beta}_j K_{i,h}(x)\left(I_n\otimes(X_i-x)^{\otimes j}\right)\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T)$$

$$\sum_{i=1}^{N} K_{h,i}(x)\mathrm{colb}_i(f(\mathbf{Y}))\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T) =^{\text{by E.11}} \sum_{i=1}^{N} K_{i,h}(x)\left(\sum_{j=0}^{k}\widehat{\beta}_j\left(I_n\otimes(X_i-x)^{\otimes j}\right)\right)\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T)$$

$$f(\mathbf{Y})(I_n \boxtimes \mathbf{W})(I_n \boxtimes \mathbf{X}^T) =^{\text{by E.9}} \sum_{i=1}^{N} K_{i,h}(x)\left(\sum_{j=0}^{k}\widehat{\beta}_j\left(I_n\otimes(X_i-x)^{\otimes j}\right)\right)\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T)$$

$$f(\mathbf{Y})(I_n \boxtimes \mathbf{W})(I_n \boxtimes \mathbf{X}^T) =^{\text{by E.15}} \sum_{i=1}^{N} K_{i,h}(x)\left(\sum_{j=0}^{k}\mathrm{colb}_j(\widehat{\beta})\mathrm{rowb}_j(\mathrm{colb}_i(I_n \boxtimes \mathbf{X}))\right)\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T)$$

$$f(\mathbf{Y})(I_n \boxtimes \mathbf{W})(I_n \boxtimes \mathbf{X}^T) =^{\text{by E.9}} \sum_{i=1}^{N} K_{i,h}(x)\widehat{\beta}\mathrm{colb}_i(I_n \boxtimes \mathbf{X})\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T)$$

$$f(\mathbf{Y})(I_n \boxtimes \mathbf{W})(I_n \boxtimes \mathbf{X}^T) = \widehat{\beta}\sum_{i=1}^{N} K_{i,h}(x)\mathrm{colb}_i(I_n \boxtimes \mathbf{X})\mathrm{rowb}_i(I_n \boxtimes \mathbf{X}^T)$$

$$f(\mathbf{Y})(I_n \boxtimes \mathbf{W})(I_n \boxtimes \mathbf{X}^T) =^{\text{by E.9}} \widehat{\beta}(I_n \boxtimes \mathbf{X})(I_n \boxtimes \mathbf{W})(I_n \boxtimes \mathbf{X}^T)$$

$$f(\mathbf{Y})(I_n \boxtimes (\mathbf{W}\mathbf{X}^T)) = \widehat{\beta}(I_n \boxtimes (\mathbf{X}\mathbf{W}\mathbf{X}^T))$$

$$\widehat{\beta} = f(\mathbf{Y})(I_n \boxtimes (\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-)).$$

现在，由于 $\widehat{\beta}_0$ 是 $\widehat{\beta}=(\widehat{\beta}_0,\cdots,\widehat{\beta}_k)$ 的第一个分量，因此我们必须将 $\widehat{\beta}$ 乘以 $I_n\boxtimes e_0$，

$$\begin{aligned}\widehat{\beta}_0&=\widehat{\beta}(I_n\boxtimes e_0)\\&=f(\mathbf{Y})(I_n\boxtimes(\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-))(I_n\boxtimes e_0)\\&=f(\mathbf{Y})(I_n\boxtimes(\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0))\\&=f(\mathbf{Y})((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\otimes I_n).\end{aligned}\tag{E.16}$$

□

F. 正则性条件

R.1 核 K 是一个紧支撑、有界核，使得 $\int K_h(u)u^{\otimes 2}du=\mu_2(K)\,\mathrm{vec}(I_p)$，其中 $\mu_2(K)\neq 0$ 是一个标量。此外，K 的所有奇阶矩均为零。

R.2 点 x 属于 $\mathrm{supp}(F_X)$。在 x 处，F_X 连续可微，且 $(fm)(X)=(f\circ m)(X)$ 的所有二阶导数均连续。其中，f 是由度量 g^f 诱导的等距映射，F_X 是 X 的密度分布。此外，$(fm)(x)>0$。

R.3 带宽序列 h 满足 $n^{-1}h\to 0$，其中 $n\to\infty$。

G. 定理证明（4.7）

证明。为了证明 $\widehat{\alpha}_0(x;1,h)$ 的一致性，我们采用一种策略：首先推导 $\widehat{\beta}_0(x;1,h)$ 的渐近偏差公式，然后利用连续映射定理 [55] 来建立前一个估计量的一致性。这是可行的，因为这两个估计量通过微分同胚 f 相互关联。

G.1. β_0的渐近偏差$(x;1,h)$

根据定理（4.6），估计量 $\widehat{\beta}_0(x;1,h)$ 具有以下闭式解析表达式

$$\widehat{\beta}_0(x;1,h)=f(\mathbf{Y})\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\otimes I_n\right),\tag{G.1}$$

因此可得

$$\mathbb{E}\left[\widehat{\beta}_0(x;1,h)|X_1,\cdots,X_N\right]=(fm)(X_1,\cdots,X_N)\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\otimes I_n\right),\tag{G.2}$$

其中，$(fm)(X_1,\cdots,X_N)=(f(m(X_1)),\cdots,f(m(X_N)))\in\mathbb{M}_{n,nN}(\mathbb{R})$。每个分量 $(fm)(X_i)$ 均可利用式 (3.5) 中的二阶泰勒展开进行展开，因此

$$(fm)(X_i)=(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_i-x))+\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_i-x)^{\otimes 2})+r_i,\tag{G.3}$$

其中 r_i 是误差项。这个泰勒展开意味着，我们可以用以下近似来表示 $(fm)(X_1,\cdots,X_N)$

$$\begin{aligned}(fm)(X_1,\cdots,X_N)&=\begin{bmatrix}(fm)(X_1)\\(fm)(X_2)\\\cdots\\(fm)(X_N)\end{bmatrix}^T\\&=\begin{bmatrix}(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_1-x))+\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_1-x)^{\otimes 2})\\(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_2-x))+\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_2-x)^{\otimes 2})\\\cdots\\(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_N-x))+\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_N-x)^{\otimes 2})\end{bmatrix}^T+R\\&=\begin{bmatrix}(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_1-x))\\(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_2-x))\\\cdots\\(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_N-x))\end{bmatrix}^T+\begin{bmatrix}\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_1-x)^{\otimes 2})\\\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_2-x)^{\otimes 2})\\\cdots\\\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_N-x)^{\otimes 2})\end{bmatrix}^T+R,\end{aligned}\tag{G.4}$$

为了继续，我们必须考虑以下恒等式。

$$\begin{aligned}\begin{bmatrix}(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_1-x))\\(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_2-x))\\\cdots\\(fm)(x)+\partial_x(fm)(X)(I_n\otimes(X_N-x))\end{bmatrix}^T&=\begin{bmatrix}(fm)(x)&\partial_x(fm)(X)\end{bmatrix}\begin{bmatrix}I_n&\cdots&I_n\\I_n\otimes(X_1-x)&\cdots&I_n\otimes(X_N-x)\end{bmatrix}\\&=\begin{bmatrix}(fm)(x)&\partial_x(fm)(X)\end{bmatrix}(I_n\boxtimes\mathbf{X}),\end{aligned}\tag{G.5}$$

and

$$\begin{aligned}\begin{bmatrix}\frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_1-x)^{\otimes 2})\\ \frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_2-x)^{\otimes 2})\\ \cdots\\ \frac{1}{2}\partial_x^2(fm)(X)(I_n\otimes(X_N-x)^{\otimes 2})\end{bmatrix}^T &= \frac{1}{2}\partial_x^2(fm)(X)\left[I_n\otimes(X_1-x)^{\otimes 2},\cdots,I_n\otimes(X_N-x)^{\otimes 2}\right]\\ &= \frac{1}{2}\partial_x^2(fm)(X)\left(I_n\boxtimes[(X_1-x)^{\otimes 2},\cdots,(X_N-x)^{\otimes 2}]\right)\\ &= \frac{1}{2}\partial_x^2(fm)(X)\left(I_n\boxtimes\Delta_x(X_1,\cdots,X_N)\right),\end{aligned} \tag{G.6}$$

其中，我们将 $\Delta_x(X_1,\cdots,X_N)=[(X_1-x)^{\otimes 2},\cdots,(X_N-x)^{\otimes 2}]\in \mathrm{M}_{p^2,N}(\mathbb{R})$记为。因此，由这些恒等式可知，

$$(fm)(X_1,\cdots,X_N)\sim[(fm)(x)\ \partial_x(fm)(X)](I_n\boxtimes\mathbf{X})+\frac{1}{2}\partial_x^2(fm)(X)\left(I_n\boxtimes\Delta_x(X_1,\cdots,X_N)\right). \tag{G.7}$$

从此处起，我们忽略误差项 R，因为右乘 $((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\otimes I_n)$后，所得项相对于$(fm)(X_1、\cdots、X_N)$在式（G.7）中泰勒展开的第二项阶数较低。接下来，我们推导泰勒展开式（G.7）中的第一项受到的影响。我们将其乘以 $((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\otimes I_n)$

$$\begin{aligned}&[(fm)(x)\ \partial_x(fm)(X)](I_n\boxtimes\mathbf{X})\left((\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\otimes I_n\right)\\ &=[(fm)(x)\ \partial_x(fm)(X)](I_n\boxtimes\mathbf{X})\left(I_n\boxtimes(\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\right)\\ &=[(fm)(x)\ \partial_x(fm)(X)]\left(I_n\boxtimes(\mathbf{X}\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\right)\\ &=[(fm)(x)\ \partial_x(fm)(X)](I_n\boxtimes e_0)\\ &=(fm)(x).\end{aligned} \tag{G.8}$$

因此，我们得出结论：

$$\mathbb{E}\left[\widehat{\beta}_0(x;1,h)-(fm)(x)|X_1,\cdots,X_N\right]\sim\frac{1}{2}\partial_x^2(fm)(X)\left(I_n\boxtimes(\Delta_x(X_1,\cdots,X_N)\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0)\right). \tag{G.9}$$

现在我们需要计算表达式 $N^{-1}\Delta_x(X_1、\cdots、X_N)\mathbf{W}\mathbf{X}^T$ 和 $N^{-1}(\mathbf{X}\mathbf{W}\mathbf{X}^T)^-e_0$。先来计算 $N^{-1}\Delta_x(X_1、\cdots、X_N)\mathbf{W}\mathbf{X}^T$

$$\begin{aligned}N^{-1}\Delta_x(X_1,\cdots,X_N)\mathbf{W}\mathbf{X}^T &= N^{-1}[(X_1-x)^{\otimes 2},\cdots,(X_N-x)^{\otimes 2}]\,\mathrm{diag}(K_h(X_1-x),\cdots,K_h(X_N-x))\begin{bmatrix}1 & (X_1-x)^T\\ \cdots & \cdots\\ 1 & (X_N-x)^T\end{bmatrix}\\ &= N^{-1}[(X_1-x)^{\otimes 2},\cdots,(X_N-x)^{\otimes 2}]\begin{bmatrix}K_h(X_1-x) & K_h(X_1-x)(X_1-x)^T\\ \cdots & \cdots\\ K_h(X_N-x) & K_h(X_N-x)(X_N-x)^T\end{bmatrix}\\ &= N^{-1}\left[\textstyle\sum_{i=1}^N K_h(X_i-x)(X_i-x)^{\otimes 2},\sum_{i=1}^N K_h(X_i-x)(X_i-x)^{\otimes 2}(X_i-x)^T\right],\end{aligned} \tag{G.10}$$

and

$$\begin{aligned}N^{-1}\sum_{i=1}^N K_h(X_i-x)(X_i-x)^{\otimes 2}(X_i-x)^T &= \int K_h(u)(hu)^{\otimes 2}(hu)^T F_X(x+hu)du+o_p(h^3 1_{p^2\times p})\\ &= o_p(h^3 1_{p^2\times p}).\\ N^{-1}\sum_{i=1}^N K_h(X_i-x)(X_i-x)^{\otimes 2} &= \int K_h(u)(hu)^{\otimes 2}F_X(x+hu)du+o_p(h^2 1_{p^2\times 1})\\ &= F_X(x)\int K_h(u)(hu)^{\otimes 2}du+o_p(h^2 1_{p^2\times 1})\\ &= F_X(x)h^2\mu_2(K)\,\mathrm{vec}(I_p)+o_p(h^2 1_{p^2\times 1}).\end{aligned} \tag{G.11}$$

其中，$1_{n,\ m}$ 是 $\mathrm{M}_{n,\ m}(\mathbb{R})$ 上的一个矩阵，其所有元素均为 1。因此

$$N^{-1}\Delta_x(X_1,\cdots,X_N)\mathbf{W}\mathbf{X}^T=[F_X(x)h^2\mu_2(K)\,\mathrm{vec}(I_p)+o_p(h^2 1_{p^2\times 1}),o_p(h^3 1_{p^2\times p})]. \tag{G.12}$$

另一方面，文献中的一个著名结果是 [1]

$$(N^{-1}\mathbf{X}\mathbf{W}\mathbf{X}^T)^-=\begin{bmatrix}F_X(x)^{-1}+o_p(1) & J_F^T F_X(x)^{-2}+o_p(1_{1\times p})\\ -J_F F_X(x)^{-2}+o_p(1_{p\times 1}) & (h^2\mu_2(K)F_X(x)I_p)^{-1}+o_p(h^{-2}I_p)\end{bmatrix}, \tag{G.13}$$

其中 J_F 是密度分布 $F(X)$ 的梯度。因此

$$\begin{aligned}(\mathbf{X}\mathbf{W}\mathbf{X}^T)^- e_0 &= \frac{1}{N}\begin{bmatrix} F_X(x)^{-1}+o_p(1) & J_F^T F_X(x)^{-2}+o_p(1_{1\times p}) \\ -J_F F_X(x)^{-2}+o_p(1_{p\times 1}) & (h\mu_2(K)F_X(x)I_p)^{-1}+o_p(h^{-1}I_p)\end{bmatrix}\begin{bmatrix}1\\ O_{p\times 1}\end{bmatrix} \\ &= \frac{1}{N}\begin{bmatrix} F_X(x)^{-1}+o_p(1) \\ -J_F F_X(x)^{-2}+o_p(1_{p\times 1})\end{bmatrix}.\end{aligned} \tag{G.14}$$

因此

$$\begin{aligned}&\mathbb{E}\left[\widehat{\beta}_0(x;1,h)-(fm)(x)|X_1,\cdots,X_N\right] \sim \frac{1}{2}\partial_x^2(fm)(X)\left(I_n \boxtimes (\Delta_x(X_1,\cdots,X_N)\mathbf{W}\mathbf{X}^T(\mathbf{X}\mathbf{W}\mathbf{X}^T)^- e_0)\right) \\ &= \frac{1}{2}\partial_x^2(fm)(X)\left(I_n \boxtimes \left(\left[F_X(x)h^2\mu_2(K)\,\mathrm{vec}(I_p)+o_p(h^2 1_{p^2\times 1}),\ o_p(h^3 1_{p^2\times p})\right]\begin{bmatrix} F_X(x)^{-1}+o_p(1) \\ -J_F F_X(x)^{-2}+o_p(1_{p\times 1})\end{bmatrix}\right)\right) \\ &= \frac{1}{2}\partial_x^2(fm)(X)\left(I_n \boxtimes \left(h^2\mu_2(K)\,\mathrm{vec}(I_p)+o_p(h^2 1_{p^2\times 1})\right)\right) \\ &= \frac{1}{2}\partial_x^2(fm)(X)\left(h^2\mu_2(K)\,\mathrm{vec}(I_p)\otimes I_n + o_p(h^2 1_{p^2\times 1}\otimes I_n)\right) \\ &= \frac{1}{2}h^2\mu_2(K)\partial_x^2(fm)(X)\,\mathrm{vec}(I_p)\otimes I_n + o_p(h^2 I_n).\end{aligned} \tag{G.15}$$

由此，我们可以得出定理的结论。

$$\mathbb{E}\left[\widehat{\beta}_0(x;1,h)-(fm)(x)|X_1,\cdots,X_N\right] = \frac{1}{2}h^2\mu_2(K)\partial_x^2(fm)(X)\,\mathrm{vec}(I_p)\otimes I_n + o_p(h^2 I_n). \tag{G.16}$$

G.2. 估计量 $\widehat{\alpha}_0(x;1,h)$ 的一致性

对公式（G.16）取极限后，可推知估计量 $\widehat{\beta}_0(x;1,h)$ 是一致的。此外，根据定理（4.6），该估计量与 $\widehat{\alpha}_0(x;1,h)$ 通过以下表达式相关联：

$$\widehat{\alpha}_0(x;1,h) = f^{-1}(\widehat{\beta}_0(x;1,h)), \tag{G.17}$$

其中，f^{-1} 是一个微分同胚，这是因为 f 既是等距映射，也是微分同胚。根据连续映射定理 [55] 和定理（4.6），结合 f^{-1} 的连续性，我们可以证明估计量 $\widehat{\alpha}_0(x;1,h)$ 的一致性。

□

参考文献

[1] David Ruppert 和 Matthew P Wand。多变量局部加权最小二乘回归。《统计年鉴》，第 1346–1370 页，1994年。[2] Jianqing Fan 和 Irene Gijbels。局部多项式建模及其应用。劳特利奇出版社，2018年。[3] Matt P Wand 和 M Chris Jones。核平滑。CRC 出版社，1994年。[4] Matthew Avery。局部多项式回归文献综述。未发表手稿，2013年。[5] Natasha Lepore、Caroline Brun、Yi-Yu Chou、Ming-Chang Chiang、Rebecca A Dutton、Kiralee M Hayashi、Eileen Luders、Oscar L Lopez、Howard J Aizenstein、Arthur W Toga 等。利用变形张量上的多变量统计对 HIV/AIDS 进行广义张量形态测量。《*IEEE* 医学影像汇刊》，27(1)：129–141，2007年。[6] Min Li、Ying Wang、Carlos Lopez-Naranjo、Shiang Hu、Ronaldo César García Reyes、Deirel Paz-Linares、Ariosky Areces-Gonzalez、Aini Ismafairus Abd Hamid、Alan C Evans、Alexander N Savostyanov 等。协调统一的多国 qEEG 常模（HARMNqEEG）。《神经影像》，256：119190，2022年。[7] Alexandre Barachant、Stéphane Bonnet、Marco Congedo 和 Christian Jutten。面向脑机接口应用、基于黎曼核的协方差矩阵分类。《神经计算》，112：172–178，2013年。[8] Oncel Tuzel、Fatih Porikli 和 Peter Meer。基于黎曼流形分类的行人检测。《*IEEE* 模式分析与机器智能汇刊》，30(10)：1713–1727，2008年。[9] John M Lee。黎曼流形导论，第 176 卷。施普林格，2018年。[10] Kisung You 和 Hae-Jeong Park。重新审视对称正定矩阵的黎曼几何及其在功能连接分析中的应用。《神经影像》，225：117464，2021年。[11] Xavier Pennec、Pierre Fillard 和 Nicholas Ayache。张量计算的黎曼框架。《国际计算机视觉杂志》，66(1)：41–66，2006年。[12] Ying Yuan、Hongtu Zhu、Weili Lin 和 James Stephen Marron。对称正定矩阵的局部多项式回归。《英国皇家统计学会杂志：*B*辑（统计方法学）》，74(4)：697–719，2012年。[13] Alexander Petersen 和 Hans-Georg Müller。具有欧几里得预测变量的随机对象 Fréchet 回归。《统计年鉴》，47(2)：691–719，2019年。

[14] Ian L Dryden、Alexey Koloydenko 和 Diwei Zhou。协方差矩阵的非欧几里得统计及其在扩散张量成像中的应用。《应用统计年鉴》，3(3)：1102–1123，2009年。[15] Lizhen Lin、Brian St Thomas、Hongtu Zhu 和 David B Dunson。流形值数据的外在局部回归。《美国统计学会杂志》，112(519)：1261–1273，2017年。[16] Xiongtao Dai 和 Hans-Georg Müller。黎曼流形与球面上的函数数据主成分分析。《统计年鉴》，46(6B)：3334–3361，2018年。[17] Zhenhua Lin 和 Fang Yao。内在黎曼函数数据分析。《统计年鉴》，47(6)：3533–3577，2019年。[18] Hwiyoung Lee。流形值数据的稳健外在回归分析。arXiv 预印本 arXiv:2101.11872，2021年。[19] Lizhen Lin、Mu Niu、Pokman Cheung 和 David Dunson。用于流形上回归与分类的外在高斯过程。《贝叶斯分析》，14(3)：887–906，2019年。[20] Alexandre Barachant、Stéphane Bonnet、Marco Congedo 和 Christian Jutten。应用于脑机接口分类的黎曼几何。载于：国际潜变量分析与信号分离会议，第 629–636 页。施普林格，2010年。[21] David Sabbagh、Pierre Ablin、Gaël Varoquaux、Alexandre Gramfort 和 Denis A Engemann。MEG/EEG 的预测回归建模：从源功率到信号与认知状态。《神经影像》，222：116893，2020年。[22] Vincent Arsigny、Pierre Fillard、Xavier Pennec 和 Nicholas Ayache。用于扩散张量快速简便计算的对数欧几里得度量。《医学共振》，56(2)：411–421，2006年。[23] Hyunwoo J Kim、Nagesh Adluru、Maxwell D Collins、Moo K Chung、Barbara B Bendlin、Sterling C Johnson、Richard J Davidson 和 Vikas Singh。黎曼流形上的多变量一般线性模型（MGLM）及其在扩散加权图像统计分析中的应用。载于：IEEE 计算机视觉与模式识别会议论文集，第 2705–2712 页，2014年。[24] Zhenhua Lin、Hans-Georg Müller 和 Byeong U Park。对称正定矩阵与李群的加性模型。《生物计量学》，asac055，2022年。[25] Zhenhua Lin。通过乔列斯基分解研究对称正定矩阵的黎曼几何。《SIAM 矩阵分析与应用杂志》，40(4)：1353–1370，2019年。[26] Yann Thanwerdas 和 Xavier Pennec。对称正定矩阵上混合欧几里得度量的几何。arXiv 预印本 arXiv:2111.02990，2021年。[27] Andri Bergsson 和 Søren Hauberg。利用 Rie-SNE 可视化黎曼数据。arXiv 预印本 arXiv:2203.09253，2022年。[28] Antoni A Kosinski。微分流形。Courier Corporation，2013年。[29] Ruud H Koning、Heinz Neudecker 和 Tom Wansbeek。分块克罗内克积与 vecb 算子。《线性代数及其应用》，149：165–184，1991年。[30] Charles F Van Loan。无处不在的克罗内克积。《计算与应用数学杂志》，123(1-2)：85–100，2000年。[31] Roger A Horn 和 Charles R Johnson。矩阵分析。剑桥大学出版社，2012年。[32] Nguyen Van Khang。机械系统动力学中矩阵函数偏导数的一致定义。《机构与机器理论》，45(7)：981–988，2010年。[33] José E Chacón 和 Tarn Duong。利用向量化导数进行高阶微分分析。arXiv 预印本 arXiv:2011.01833，2020年。[34] Nguyen Van Khang、Dinh Cong Dat 和 Nguyen Thai Minh Tuan。利用克罗内克积对向量变量的矩阵函数进行泰勒展开。[35] Jan R Magnus 和 Heinz Neudecker。矩阵微分演算及其在普通积、Hadamard 积和克罗内克积中的应用。《数学心理学杂志》，29(4)：474–492，1985年。[36] Jan R Magnus。矩阵导数的概念。《多变量分析杂志》，101(9)：2200–2206，2010年。[37] Armin Schwartzman。高角分辨率扩散成像中的黎曼正态性与随机旋转不变统计。《统计年鉴》，34(4)：1996–2015，2006年。[38] H. Zhu、H. Zhang、J. G. Ibrahim 和 B. G. Peterson。扩散加权磁共振图像数据中扩散张量的统计分析（附讨论）。《美国统计学会杂志》，102(480)：1085–1102，2007年。[39] C-F Westin、Stephan E Maier、Hatsuho Mamata、Arya Nabavi、Ferenc A Jolesz 和 Ron Kikinis。扩散张量 MRI 的处理与可视化。《医学图像分析》，6(2)：93–108，2002年。[40] Yoshitaka Masutani、Shigeki Aoki、Osamu Abe、Naoto Hayashi 和 Kuni Otomo。MR 扩散张量成像：扩散张量可视化的近期进展与新技术。《欧洲放射学杂志》，46(1)：53–66，2003年。[41] P Thomas Fletcher、Conglin Lu、Stephen M Pizer 和 Sarang Joshi。用于研究形状非线性统计的主测地线分析。《IEEE 医学影像汇刊》，23(8)：995–1005，2004年。[42] Laurens Van der Maaten 和 Geoffrey Hinton。使用 t-SNE 可视化数据。《机器学习研究杂志》，9(11)，2008年。[43] Martin Wattenberg、Fernanda Viégas 和 Ian Johnson。如何有效使用 t-SNE。《Distill》，1(10)：e2，2016年。[44] Markus Ringnér。什么是主成分分析？《自然生物技术》，26(3)：303–304，2008年。[45] Sasan Karamizadeh、Shahidan M Abdullah、Azizah A Manaf、Mazdak Zamani 和 Alireza Hooman。主成分分析概述。《信号与信息处理杂志》，4(3B)：173，2013年。[46] Gery Geenens。维数灾难及非参数函数回归中的相关问题，2011年。

[47] Ying Wang、Min Li、Deirel Paz-Linares、Maria L Bringas Vega 和 Pedro A Valdés-Sosa。Fkreg：用于快速多元核回归的 MATLAB 工具箱。*arXiv* 预印本 *arXiv:2204.07716*，2022年。[48] Mehrtash T Harandi、Mathieu Salzmann 和 Richard Hartley。从流形到流形：针对对称正定矩阵的几何感知降维。载于计算机视觉—*ECCV 2014*：第*13*届欧洲会议，瑞士苏黎世，*2014*年*9*月*6*日至*12*日，论文集，第*II*卷第*13*部分，第17–32页。施普林格，2014年。[49] BY Chen。我们能如何运用纳什嵌入定理。《苏州数学杂志》，30(3):303–338，2004年。[50] Donald WK Andrews。异方差误差回归中广义 CL、交叉验证和广义交叉验证的渐近最优性。《计量经济学杂志》，47(2-3):359–377，1991年。[51] Ker-Chau Li。岭回归中 CL 和广义交叉验证的渐近最优性及其在样条平滑中的应用。《统计年鉴》，第1101–1112页，1986年。[52] Clifford M Hurvich、Jeffrey S Simonoff 和 Chih-Ling Tsai。使用改进的赤池信息准则选择非参数回归中的平滑参数。《英国皇家统计学会会刊*B*辑：统计方法学》，60(2):271–293，1998年。[53] Fuxiang Liu。新的克罗内克积分解及其应用。*Research Inventy*：《国际工程与科学杂志》，1:25–30，2012年。[54] Youssouf Mezzar 和 Kacem Belghaba。克罗内克和分解及其应用。[55] Aad W Van der Vaart。《渐近统计学》，第3卷。剑桥大学出版社，2000年。