

流形学习中非线性维数约简方法概述

黄启宏, 刘 钊

(电子科技大学 电子工程学院, 成都 610054)

摘 要: 较为详细地回顾了流形学习中非线性维数约简方法, 分析了它们各自的优势和不足。与传统的线性维数约简方法相比较, 可以发现非线性高维数据的本质维数, 有利于进行维数约简和数据分析。最后展望了流形学习中非线性维数方法的未来研究方向, 期望进一步拓展流形学习的应用领域。

关键词: 维数约简; 流形学习; 多维尺度; 等距映射; 拉普拉斯特征映射; 局部线性嵌入; 局部切空间排列

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2007) 11-0019-07

Overview of nonlinear dimensionality reduction methods in manifold learning

HUANG Qi-hong, LIU Zhao

(School of Electronic Engineering, University of Electronic Science & Technology of China, Chengdu 610054, China)

Abstract: A detailed retrospection was made on nonlinear dimensionality reduction methods in manifold learning, whose advantages and defects were pointed out respectively. Compared with traditional linear method, nonlinear dimensionality reduction methods in manifold learning could discover the intrinsic dimensions of nonlinear high-dimensional data effectively, help researcher to reduce dimensionality and analyzer data better. Finally, the prospect of nonlinear dimensionality reduction methods in manifold learning was discussed, so as to extend the application area of manifold learning.

Key words: dimensional reduction; manifold learning; multidimensional scaling(MDS); isomap; Laplacian eigenmap; locally linear embedding(LLE); local tangent space alignment(LTSA)

维数约简方法的基本原理是将样本从输入空间通过线性或非线形映射到一个低维空间, 从而获得一个关于原数据集紧致的低维表示。传统的线性维数约简方法具有简单、易解释和可延展等优点, 使其在高维数据处理中成为一个主要研究方向。已有的线性维数约简方法主要包括主成分分析(principal component analysis, PCA)^[1]、独立成分分析(independent component analysis, ICA)^[2]、Fisher 判别分析(Fisher discriminant analysis, FDA)^[3]、主曲线(principal curves)^[4]、投影寻踪(projection pursuit, PP)^[5]、局部线性投影(local linear projection, LLP)^[6]以及基于神经网络的自组织映射(self-organizing map, SOM)^[7]等。这些方法实际是在不同优化准则之下, 寻求最佳线性模型的方法, 这也是线性维数约简方法的共性。然而, 随着信息时代的到来, 不可避免地出现大量的高维非线性数据。传统的线性维数约简方法难以直接用于分析来源于真实世界的高维和非线性数据。其主要原因有膨胀的维数导致计算量迅速上升; 高维导致样本数相对较少, 使得某些统计上的渐近性质受到破坏; 传统方法在处理高维数据时不满足稳健性要求等。因此, 研究高维非线性数据面临诸多困难^[8]。这主要是高维带来了数据的稀疏和维数灾难, 而非线性使得现有的快速成熟的线性模型不再适用。目前主要存在两类非线性维数约简方法, 即基于核的方法和基于流形的方法。前者利用 Mercer 核及其对应的再生核希尔伯特空间(reproduction kernel Hilbert space, RKHS), 不用创建复杂的假设空间, 通过定义 Mercer 核隐式地定义特征空间。因此大部分线性维数约简方法都有其

对应的基于核的非线性维数约简方法, 如 KPCA^[9]、KICA^[10]、KFDA^[11]、KSOM^[12]、核特征映射^[13]等。然而, 基于核的方法其难点在于如何选择一个合适的核函数。一个好的核函数可以使数据在特征空间上线性可分或者近似线性可分, 但并不是每个核函数对于每一种数据集都适用。核函数的选择反映了人们对问题的先验, 在实际的应用中往往是凭经验选择某种核函数, 如径向基函数(radial basis function, RBF)。同时, 在使用核函数时不必知道具体的特征空间, 使得核函数方法缺乏物理直观, 这也是核函数方法的一个缺点。后者就是近年发展起来的基于流形学习的非线性维数约简方法, 主要包括多维尺度方法^[14]、等距映射方法^[15]、局部线性嵌入方法^[16]、拉普拉斯特征映射法^[17]、局部切空间排列方法^[18]等。

流形学习中非线性维数约简方法与线性维数约简方法相比的一个显著特点是分析中的局部性。对数据集的内蕴结构而言有如下特性: 由泰勒定理, 任何可微函数在一点的充分小的邻域内满足线性条件, 形象地说, 认为曲面流形由大小不一的局部线性块拼接而成; 数据流形经常是由许多可分割的子流形所组成; 数据流形的本征维数沿着流形不断地发生变化, 只有局部性才能抓住其根本特性。

1 流形定义和流形学习

流形是现代数学的概念, 可以将流形简单地理解为由 $f(x) = 0, x \in R^d$ 确定的 R^N ($d \ll N$) 空间中的 d 维曲面。下面是流形的严格数学定义:

收稿日期: 2006-09-27; 修返日期: 2006-12-18

作者简介: 黄启宏(1974-), 男(回族), 湖南隆回人, 博士研究生, 主要研究方向为图像边缘提取和图像噪声压抑(y97hqb@ sina. com. cn); 刘钊(1942-), 男, 教授, 博导, 主要研究方向为图像采集/识别/压缩/存储、图像传输系统及纠错编码、数字视频和多媒体技术。

定义 1 设 M 是一个 Hausdorff 拓扑空间, 若对每一点 $p \in M$, 都有 p 的一个开邻域 U 和 R^d 中的一个开子集同胚, 则称 M 是 d 维拓扑流形, 简称 d 维流形。假定 $\varphi(U) \subset R^d$ 与 U 同胚, 则 (U, φ) 称做流形 M 的坐标卡, $\varphi(p)$ 在 R^d 中的坐标称为点 p 在 M 上的坐标。

有了对流形的定义, 就可以概括流形学习中维数约简过程: 假设数据是均匀采样于一个高维欧式空间中的低维流形, 流形学习就是从高维采样数据中恢复低维流形结构, 即找到高维空间中的低维流形, 并求出相应的嵌入映射, 以实现维数约简或者数据可视化。流形学习是从观测到的现象中寻找事物的本质, 找到产生数据的内在规律。

维数约简数学定义如下:

定义 2 给定 n 个数据点: $x = \{x_i | i = 1, \dots, n, x_i \in R^D\}$, 给出映射 $f: R^D \rightarrow R^d, d \ll D$; 或者给出弱一些的结果 $Y = \{y_i | i = 1, \dots, n, y_i \in R^D\}, d \ll D$ 。

自从 Tenenbaum、Roweis、Saul 和 Balasubramanian 等人^[15, 16, 19]相继发表关于流形学习的初步研究成果后, 利用流形特性进行维数约简的研究就成为热门的研究话题。早在 1996 年, Nayar 等人就提出了高维图像具有低的内在维数的现象^[20]。Seung 等人从神经生理学的角度提出了感知以流形的形式存在^[21], 从而从认知学的角度支持了图像空间是高维空间中的流形这一观点。Donoho 等人还拓展了 LLE 方法, 提出 Hessian LLE 方法^[22], 能够发现流形上局部的潜在等距映射参数。张长水等人在 LLE 的基础上提出一种从低维嵌入空间向高维空间映射的方法^[23], 并在多姿态人脸图像的重构实验中得到有效的验证, 进一步完善了非线性维数约简方法。Weinberger 等人在 LLE 方法的基础上, 结合流形的微分几何特征, 发展出基于半定规划的流形方法^[24~26], 适用于不同类型的流形。Belkin 等人提出了 Laplacian eigenmaps^[17], 并成功地将其应用于半监督的 Riemannian 流形学习^[27]、流形规则化^[28]和部分标记分类^[29]等方面。Zhu 等人将半监督思想与 Gaussian random field 结合^[30], 显示了半监督流形学习的巨大发展前景。Roweis 等人意识到以 isomap、LEE、Laplacian eigenmaps 为代表的维数约简方法只是对训练集中的样本给出嵌套空间中的位置, 缺少从高维空间到低维空间的映射, 并且依赖于点集之间的关系, 对噪声非常敏感, 特征值分解又加剧了这种不稳定性。因此, 提出了在局部空间学习一个线性映射, 并使用 EM 算法解决优化问题^[31]。Brand^[32]继续发展了这一思想, 对流形上的邻域用 Gaussian 分布建模, 并给出了闭合解的形式, 避免了效率低下的爬山方法。这一领域的巨大发展不仅引起了计算机科学研究人员的注意, 同时也引起了数学家和物理学家的兴趣。普林斯顿大学的 Chigirev 等人将数据的维数约简看做为数据压缩的过程, 应用率失真定理将信息理论引入流形学习框架, 提出了对流形的最优描述, 不仅完成了维数约简过程, 同时揭示了流形的内部结构^[33]。由于已有的流形学习方法对噪声和参数都比较敏感, 詹德川等人^[34]针对 isomap 方法, 通过引入集成学习技术扩大了可以产生有效可视化结果的输入参数范围, 并且降低了对噪声的敏感性。另外, 赵连伟等人^[35]完善了 isomap 的理论基础, 给出了连续流形与其低维参数空间等距映射的存在性证明, 并区分了嵌入空间维数、高维数据的固有维数与流形维数这些容易混淆的概念; 证明如果高

维数据空间存在环状流形, 流形维数要小于嵌入空间维数; 同时, 还给出了一种有效的环状流形发现方法, 以得到正确的低维参数空间。何力等人^[36]提出了一种从方法因子和延伸方向两方面显示出观测空间的高维数据与维数约简后的低维数据之间的联系方法, 并且比较了 isomap 和 LLE 方法的性能。Zhang 等人提出了一种局部切空间排列方法^[18], 通过逼近每一样本点的切空间来构建低维流形的局部几何, 然后利用局部切空间排列求出整体低维嵌入坐标。由于用于特征值分解的矩阵阶数等于样本数, 样本集较大时将无法处理; 此外, 该方法不能有效处理新来的样本点。一种基于划分的局部切空间排列方法^[37]被提出以改善这些缺点。它建立在向量量化主成分分析算法和 LTSA 方法的基础上, 解决了向量量化主成分分析算法不能求出整体低维坐标和 LTSA 中大规模矩阵的特征值分解问题, 且实验证明能够有效处理新来的样本点。

2 流形学习中非线性维数约简方法

在流形学习中, 许多非线性维数约简方法属于谱方法 (spectral method)。谱方法主要利用流形的二阶特征来描述流形的内在结构, 主要包括 MDS、isomap、LLE、Hessian LLE、Laplacian eigenmaps、LSTA 等。谱方法将流形结构信息存放在一个邻接矩阵中, 并通过对矩阵的谱分解得到高维数据的低维嵌入。谱方法的缺点在于通过已知数据集推广到未知数据 (out-of-sample) 非常困难。另外, 这一类方法大多需要手动选择邻域大小和内在维数, 这两个参数往往关系到流形结构是否能被真实恢复。

2.1 MDS

MDS^[14]是一种传统的寻求保持数据点之间差异性 (或相似性) 的维数约简方法, 它使得在原数据集中点经过变换后, 仍旧保留数据点之间的距离。通过适当定义准则函数来体现在低维空间中对高维距离的重建误差, 对准则函数用梯度下降法求解, 对于某些特殊的距离可以推导出解析解法。

经典 MDS 方法的主要步骤如下:

a) 计算 $S \in M_n$, 若 $D = (d_{ij})$, $s = (s_{ij})$, 则 $s_{ij} = d_{ij}^2$ 。

b) 取矩阵 $H = (h_{ij}) \in M_{nn}$ 满足 $h_{ij} = \delta_{ij} - 1/n$ 。其中: δ_{ij} 是 Kronecker 符号。

c) 计算 $\tau(D) = -HSH/2 \in M_{nn}$ 。这步计算称为双中心化 (double centering)。

d) 对 $\tau(D)$ 作奇异值分解 (singular value decomposition, SVD)。因为矩阵 $\tau(D)$ 对称, 即有 $\tau(D) = U^T \Lambda U$ 。其中: U 是一个正交矩阵; Λ 是一个奇异值为从大到小排列的对角阵, 且对角线元素非负。

e) 计算 $Z \in M_{dn}$, 取 U 的前 d 行和前 n 列组成的 U_d , Λ 的前 n 行前 n 列组成的 $\Lambda_d \in M_{dd}$, 并对 Λ_d 开算术平方 (对角线元素依次取算术平方根) 得 $\Lambda_d^{1/2}$, 而 $Z = U_d \Lambda_d^{1/2}$ 。

2.2 Isomap

Isomap^[15]方法是一个利用非线性的局部变量度量信息学习数据集的全局几何结构的方法, 其突出的特点是用测地距离来度量高维空间中样本点之间的距离。Isomap 的主要思想是首先计算流形上的测地线距离; 然后利用 MDS 方法, 发现嵌入在高维空间的低维空间的低维坐标。这样 isomap 就通过数据

间的测地线距离,保留了数据固有的几何分布结构。

Isomap 方法的假设前提如下: 高维数据所在的低维流形与欧氏空间的一个子集是整体等距的; 与数据所在的流形等距的欧氏空间的子集是一个凸集。

- Isomap 的主要步骤如下:
- a) 构造局部邻域, 对数据集 $X = \{x_1, x_2, \dots, x_n \in R^n\}$ 中的每一个数据点 x_i 的近邻点 x_j (k 邻域) ($i = 1, \dots, N; j = 1, \dots, k$), 得到邻接矩阵 G 。
 - b) 计算近邻图中任意两个样本之间的测地距离。计算该图上两点之间的最短路径 $d_G(i, j)$, 作为样本点之间的测地距离的近似(这里 i 和 j 表示两个样本, 下标 G 是 Graph 缩写)。图上任意两点最短路径可利用经典的 Floyd 迭代算法。

c) 用 MDS 方法获得低维投影。图距离矩阵定义为 $D_G = \{d_G(i, j)\}$, 用 MDS 方法构造一个保持本质几何结构的 d 维欧氏空间。令 $S = (S_{ij}) = (D_{ij}^2)$, $H = (H_{ij}) = (\delta_{ij} - 1/N)$, $\tau(D) = -HSH/2$, 低维嵌入是 $\tau(D)$ 的第 $2 \sim d+1$ 小的特征值所对应的特征向量。

Isomap 给出了确定维数约简的准则。衡量维数约简误差的残差: $e_d = 1 - R^2(D_G, D_Y)$ 。其中: D_Y 是 d 维空间的欧氏距离矩阵; R^2 表示线性相关系数。一般地, 维数约简 d 越高, 残差越小。在下面两种情况下 isomap 可以确定维数约简 d : 当残差曲线出现拐点; 残差已经小到一定的阈值。

Balasubramanian 等人^[19]指出, 对于有噪声数据, isomap 方法用于可视化时不稳定、取较大的邻域会产生短路现象。如果为了追求整体结构不发生变化, 而选取较小的邻域就会大量出现空洞或使最短路径算法重构的图不连通。因此, 采用以下两个约束来选取合适的邻域: 第一个约束是最大连通分支中的样本数占样本总数的比例尽可能大, 这一约束会使绝大部分样本出现在同一个连通区域中, 从而可以抑制不连通区域的出现; 第二个约束使残差尽可能小, 直观上来说, 当出现短路现象时, 残差会比较大, 因此这一约束会抑制短路现象的出现。然而, 残差描述的是进行 MDS 过程中信息的不一致, 虽然噪声会干扰 MDS 过程, 但是噪声更大程度上影响了近邻的选取, 进而会影响到距离重构过程。残差描述的其实只是 MDS 而不是整个 isomap 学习过程中信息的不一致, 所以使用此约束未必能奏效。

2.3 LLE

LLE^[16] 是一种依赖于局部线性的非线性维数约简方法。它的提出极大地拓展了关于维数约简的认识, 人们开始关注数据集所包含的内蕴特征, 通过内蕴特征的探讨来进一步研究维数约简问题。根据泰勒定理, 可微函数具有良好的局部线性, 即每点的微小邻域总可以用线性模型来较好地近似。同样, 对于任意光滑流形, 其微小局部一定程度上也具有线性的特征, 只要可以清晰地描述这一局部特征, 就抓住了数据流形的根本所在。因此, LLE 方法就是基于流形的局部线性特性的非线性维数约简方法。LLE 方法的主要步骤如下:

- a) 构造局部邻域, 对数据集 $X = \{x_1, x_2, \dots, x_n \in R^n\}$ 中的每一个数据点 x_i 的近邻点 x_j (k 邻域) ($i = 1, \dots, N; j = 1, \dots, k$), 构成邻域 $X_i = \{x_{i1}, x_{i2}, \dots, x_{ik}\}$ 。
- b) 对每个点 x_i 及其构成的邻域, 计算重建系数 W_{ij} , 使得

- $\|x_i - \sum_j W_{ij} x_j\|$ 最小, 并且 $\sum_j W_{ij} = 1$ 。
- c) 计算 $Y = \{y_1, y_2, \dots, y_n \in R^d\}$ 使得代价函数 $\phi(Y) = \sum_i \|x_i - \sum_j W_{ij} y_j\|^2 = \text{trace}(Y^T(I - W)^T(I - W)Y)$ 最小。这一优化问题可以通过对 $M = (I - W)^T(I - W)$ 的特征值分解得到。令 M 的最小 $d+1$ 个特征值及对应的特征向量为 $\langle \lambda_i, v_i \rangle$, $i = 1, \dots, d$, 第 i 个样本嵌入的第 j 个分量即为 $\sqrt{\lambda_i} v_{ji}$ 。 v_{ji} 为 v_j 的第 i 个分量。

LLE 方法可以学习任意维的局部线性的低维流形。与 isomap 方法一样只有两个待定系数 k 和 d 。同时由重构代价函数最小化得到的最优权值遵循对称特性, 每个点的近邻权值在平移、旋转、伸缩变化下保持不变。LLE 方法有解析的全局最优解, 不需要迭代, 低维嵌入的计算归结为稀疏矩阵特征值的计算, 这样计算复杂度相对较小。然而, LLE 方法要求所学习的流形只能是不闭合的且在局部是线性的, 还要求样本在流形上是稠密采样的。另外, 该方法的参数选择不确定, 对样本中的噪声很敏感。

2.4 Laplacian eigenvaps

拉普拉斯特征映射法^[17]的基本思想是, 在高维空间中离得很近的点投影到低维空间中的像也应该离得很近, 最终求解归结到图拉普拉斯算子的广义特征值问题。

- 拉普拉斯算子(Laplacian operator)定义如下:
- 定义 3 设 M 是光滑的黎曼流形, f 是 M 上的光滑函数, ∇f 是 f 的梯度, 则称线性映射 $\Delta: C^\infty(M) \rightarrow C^\infty(M)$, $\Delta f = -\text{div}(\nabla f)$ 为 M 上的拉普拉斯算子。其中 div 是散度算子。

图上的拉普拉斯算子定义如下:

定义 4 设 G 是一个图, v 是它的顶点, d_v 是 v 的自由度, $w(u, v)$ 是连接顶点 u, v 的边的权值, 则有

$$I(u, v) = \begin{cases} d_v - w(u, v) & u = v \\ -w(u, v) & u, v \text{ is connected} \\ 0 & \text{others} \end{cases}$$

令 $L = -(1/2)IT^{-1/2}$ 。其中: T 是对角矩阵, 对角线的元素为 $\sum_{u \sim v} w(u, v)$, 则称 L 为图 G 上的拉普拉斯算子。

拉普拉斯特征映射法的主要步骤如下:

- a) 构造局部邻域, 对数据集 $X = \{x_1, x_2, \dots, x_n \in R^n\}$ 中的每一个数据点 x_i 的近邻点 x_j (k 邻域) ($i = 1, \dots, N; j = 1, \dots, k$), 得到邻接矩阵 G 。
- b) 建立加权邻接矩阵 W , 使用热核有 $w_{ij} = \exp(-d_{ij}^2/\sigma^2)$ 。如果 $G_{ij} = 1$, 否则 $w_{ij} = 0$; 使用 $0/1$ 核有 $w_{ij} = 1$, 如果 $G_{ij} = 1$, 否则 $w_{ij} = 0$ 。
- c) 计算图拉普拉斯算子 $L = D - W$ 的广义特征向量。其中: D 为对角矩阵, 且 $D_{ij} = \sum_j w_{ji}$, 得到 $d+1$ 个最小特征值对应的特征向量 $v_0, v_1, \dots, v_d$, 使用特征向量 $v_1, v_2, \dots, v_d$ 作为嵌入空间的坐标。

拉普拉斯特征映射法是局部的非线性方法, 其突出特点与谱图理论有着很紧密的联系。从方法描述中可以看出, 拉普拉斯特征映射与 LLE 方法类似: 待定参数相同; 求解的是稀疏矩阵的广义特征值问题。它能使输入空间中离得很近的点在低维空间也离得很近, 故可用于分类和聚类。

Belkin 等人^[28]还基于 Laplacian-Beltrami 算子的意义, 对

Laplacian eigenmaps 作了半监督情况下的推广。其主要思想是,用样本点(有标记的和无标记的)来估计流形的结构或者先验概率 $p(x)$,用于标记的样本点进一步估计类条件概率密度 $p(x|y=j)$ 。 L 算子的可数特征函数 $v_1, v_2, \cdots, v_\infty$ 构成 M 上的二次可微函数的基函数,并且它们是完备的,则分类函数 $f: x \mapsto \{-1, +1\}$ 可以展开为以特征函数 $v_1, v_2, \cdots, v_\infty$ 为基的 Fourier 级数(为方便起见,这里只考虑两类情况)。在离散情况下,特征函数变为特征向量。利用 $L - W$ 的特征向量 $v_1, v_2, \cdots, v_d$ 对已有标签的样本 $x_i (i=1, 2, \cdots, s)$ 拟合分类器 $f(x_i) = \sum_{j=1}^d a_j v_{ij}$ 。这样对未标记的样本 $x_i (i=s+1, s+2, \cdots, n)$, 有如下

分类器:
$$f(x_i) = \begin{cases} 1 & \sum_{j=1}^d a_j v_{ij} \geq 0 \\ -1 & \sum_{j=1}^d a_j v_{ij} < 0 \end{cases}。$$

2.5 局部切空间排列法

LTSA^[18] 是一种基于切空间的流形学习方法。它通过逼近每一样本点的切空间来构建低维流形的局部几何;然后利用局部切空间排列求出整体低维嵌入坐标。LTSA 方法的主要思想是找出每个数据点的邻近点,用邻域中低维(如 d 维)切空间的坐标近似表示局部的非线性几何特征;再通过变换矩阵将各数据点邻域切空间的局部坐标映射到一个统一的全局坐标上,从而实现数据维数约简,即从 n 维数据中得到一个全局的 d 维坐标。LTSA 方法的主要步骤如下:

- a) 构造局部邻域,对数据集 $X = \{x_1, x_2, \cdots, x_n \in R^n\}$ 中的每一个数据点 x_i 的近邻点 x_j (k 邻域) ($i=1, \cdots, N; j=1, \cdots, k$), 构成邻域 $X_i = \{x_{i1}, x_{i2}, \cdots, x_{ik}\}$ 。
- b) 局部线性拟合,在每个数据点 x_i 的邻域内选择一组正交基 Q_i 构成 x_i 的 d 维切空间,计算邻域中的每个点 $x_{ij} (j=1, 2, \cdots, k)$ 到切空间上的正交投影 $\theta_j^{(i)} = Q_i^T (x_{ij} - \bar{x}_i)$ 。其中: $\bar{x}_i$ 为邻域内的均值; Q_i 通常取 X_i 的前 d 个最大的左奇异向量。
- c) 将局部坐标转换为局部坐标,对每个邻域的局部切空间坐标 $\theta = [\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(N)}] (i=1, \cdots, N)$, 构造转换矩阵 $L_i = T_i \theta^+ \in R^{d \times d}$ 。通过 $\min \sum_i \|T_i(I - ee^T/k) - L_i \theta\|_2^2$ 的求解,得到行向量为正交的低维全局坐标映射 $T = [\tau_1, \tau_2, \cdots, \tau_n] \in R^{d \times n}$ 。其中: θ^+ 是 θ 的广义 Moor-Penrose 逆; $T_i = [\tau_1, \tau_2, \cdots, \tau_n]$, 对应的下标集合 $\{i_1, i_2, \cdots, i_n\}$ 由数据点 x_i 的邻域确定; e 为全 1 向量。

Min 等人证明了局部切空间可以用局部样本协方差矩阵的特征向量来表示^[38]。因此可以把求取表示局部信息的样本点在局部切空间中的投影坐标问题转换为局部主成分分析问题。

3 常见应用

随着非线性维数约简方法理论研究的深入,其广泛应用于医学、图像处理、语音信号处理、文本数据处理、手写体数据处理、天文数据处理以及金融数据处理等领域。本文对非线性维数约简方法的具体应用进行综述。

- 1) 医学方面 主要应用于基因序列表示^[39]、蛋白质检测^[40]、疾病分类^[41]、fMRI 数据处理^[42]、微阵列数据处理^[43]、弥散张量成像(diffusion tensor image, DTI) 处理^[44]、神经生理

学^[45] 等方面。

- 2) 图像方面 主要应用于人脸表示^[46]、人脸识别^[47]、人脸分类^[48]、纹理合成^[49]、图像恢复^[50]、运动图像检测^[51,52]、超谱图像处理^[53]、三维动画处理^[54]、树木表面检测^[55] 以及图像去噪^[56] 等方面。

- 3) 语音信号处理^[57]
- 4) 文本数据处理^[47,58]
- 5) 手写体数据处理^[59,60]
- 6) 天文数据处理^[61]
- 7) 金融数据处理^[62]

4 最新进展和发展方向

4.1 非线性维数约简新方法

- 1) 局部判别嵌入方法

Isomap 方法与 LLE 方法仍有其限制之处,其主要通过维数约简对数据集的可视化进行处理,着重于训练数据集的表示而不是分类,且有样本外问题,即当有新的样本数据时,需要重新学习。Chen 等人^[63] 提出一种监督式的局部判别嵌入(local discriminant embedding, LDE) 方法,主要针对有标记的数据集,以分类为目标的最优处理。LDE 的特点是构建两个数据样本邻近关系图,分别用于记录样本间的类别关系以及邻近几何关系,并可将最优问题变成特征值求解问题。因此最佳维数约简空间可用简单的特征值和特征向量的计算获得。此外, LDE 没有样本外问题,可直接处理新的测试样本。

- 2) 局部规则嵌入

目前的非线性维数约简方法都是基于数据集中点与点之间的局部关联。由于在维数约简过程中没有考虑非同性数据集集合所具有的自然类结构,使得对非同性数据集效果不佳,即在恢复数据流形的同时,不能很好地进行数据分类。谭璐等人^[64] 引入拓扑邻域、拓扑结构以及规则的概念,构造了相应的结构规则性度量,提出了一种局部规则嵌入方法(locally regular embedding, LRE)。该方法由于较好地保留了数据集拓扑结构,使得维数约简后能较好地保留数据集的类特性,更好地揭示了数据集的本征结构。

- 3) 局部不变投影

结合近年来提出的 LLE 方法和 Laplacian eigenmap,谭璐等人^[65] 提出了局部不变投影方法。该方法在高维数据的维数约简处理中具有保持数据集几何结构和拓扑结构不变的性质,而且继承了线性方法计算方便、快捷的优点。通过分析可以看出,非线性和线性维数约简方法,两者并不矛盾,而是相互关联的,它们之间存在结合的可能性,并且通过两者的综合可将非线性方法的特性有效地移植到线性方法中。He 等人^[66] 提出了一种 Laplacian eigenmap 的线性化版本的局部不变投影(local preserving projection, LPP) 方法。LPP 可以看做一种子空间方法与流形学习的结合。

- 4) 其他方法

Kouropiteva 等人^[60] 和 Ridder 等人^[67] 提出的监督局部线性嵌入方法(supervised locally linear embedding, SLLE) 以及 Pang 等人^[68] 提出的邻域保留投影方法(neighborhood preserving projections, NPP) 等非线性维数约简方法。

4.2 稳健性研究

LLE 方法在求解模型过程中可能出现矩阵病态, 导致求解过程对于数据点噪声十分敏感, 使结果受噪声影响较大。针对 LLE 方法的不足, 从分析噪声对数据集局部特性的影响入手, 提出了稳健局部线性嵌入方法^[69], 不仅较好地解决了噪声对数据集的影响, 而且对邻域的选择也有较好的适应性, 可更好地挖掘数据的本征特性, 具有更强的数据可视化能力。

现有的非线性维数约简方法很大部分是基于谱分解的。由于谱分解固有的稳定性问题, 使得已有的方法对噪声和参数都比较敏感, 噪声的存在使得输入参数更加难以选择, 参数较小的变化会导致差异显著的学习结果。因此, 提高非线性维数约简方法的抗噪性成为亟待解决的问题。通过将非线性维数约简方法与已有的成熟机器学习方法(如核技术^[70])相结合, 研究各种噪声模型对非线性维数约简方法的影响方式和影响度。杨建宏等人^[56]也认为, 带噪的时间序列在高维的相空间中其本质特征隐含在一个低维的主流形中, 利用 LTSA 方法提取其主流形, 再根据主流形对时间序列进行重构, 就可以达到去除噪声的目的。与现有的非线性时间序列噪声去除方法不同, 基于流形的方法更强调时间序列的整体结构。

4.3 核技术应用

主流的流形学习方法都可以看做是在数据相关的核函数下的 KPCA。其核函数从数据中学习获得, 而不是事先给定。因此, 很多研究者希望将流形学习的方法统一到对核函数的研究框架下来。Choi 等人^[70]针对 isomap 提出了核 isomap 方法。Yan 等人^[71]提出了 graph embedding 框架, 子空间与流形学习都被统一到此框架下。Weinberger 等人^[72]给出了 KPCA 和流形方法之间的联系, 即采用半定规划, 可以获得一个具有最大的迹(如特征空间变量等)的半正定核矩阵 Ψ 。其中: Ψ 满足约束条件 $|\Psi(x_i) - \Psi(x_j)|^2 = |x_i - x_j|^2$; x_i, x_j 为邻域。Ham 等人^[73]对核技术用于流形学习中的维数约简方法进行了综述。

4.4 可视化度量研究

非线性维数约简方法作为一种可视化手段, 自然需要某种可视化效果的度量, 如坐标相关性^[34]。坐标相关性不仅关注投影后样本间的距离信息, 同时考虑了样本投影后的位置信息。但它的作用对象一般集中在本真结构已知的数据集, 研究开发更广泛意义上的度量可视化效果的方法也可以成为一个努力的方向。

4.5 约简效果度量研究

N 维高维空间的样本数据, 一般不可能弥漫于整个 R^N 空间, 否则就不会有什么信息。数据实际上处于一个高维空间的低维流形上, 即一个维数约简后的曲面上。该流形的维数即为数据的本征维数(intrinsic dimensionality), 而 N 只是数据的表象维数。因此, 可以考虑形成一整套的非线性维数约简机制。对给定的非线性高维数据集, 首先有效判断其是否存在流形; 然后估计固有维数和潜在空间维数, 确定约简后的维数提供有效的参数, 避免为了得到残差与维数的关系进行多次实验造成不必要的开销以及综合判断时的主观误差。在流形学习的主要处理完成后, 制订指标衡量观测空间的高维数据与维数约简后低维数据间的定量关系。这样有利于对数据规律的深入探

索, 可直观比较不同流形方法的维数约简效果。何力等人^[36]提出的放大因子和延伸方向已被证明是两个有效的指标。

4.6 交叉流形问题

Souvenir 等人^[51]提出了流形学习应用范围扩展的新问题, 即多个交叉流形分类和参数化。当流形出现重叠现象时, 先前的非线性维数约简方法会失效。针对这一问题, 提出了两种解决方案, 即节点加权的 MDS 和对低阶矩阵近似的快速方法, 并且通过一个混合拓扑结构和维数的交叉流形以及人体运动捕捉数据的实例展示了该方法的应用。

4.7 互逆问题

能否在得到同胚映射的前提下进行流形学习的互逆过程^[75], 在本真维数空间生成有效数据后, 向高维空间投影合成新的训练数据, 或者利用同胚映射关系判断高维空间合成数据的有效性和完备性, 解决训练样本稀疏的问题。另外, 对已经获得本真维数的数据集, 如何确定各个独立自由度的具体意义, 便于对应用的实践指导, 同样需要思考。

4.8 其他问题

现有的非线性维数约简方法多数是无监督的, 如何将其推广到半监督^[27, 76]以及有监督^[48]的情况下, 着眼提高方法的泛化能力, 也是一个有价值的研究课题。

现有非线性维数约简方法大多基于小的邻域学习, 期望通过在小邻域上的学习得到一个全局的坐标, 这往往是不现实的。如何将全局与局部数据学习结合起来^[77]是一个非常有意义的话题。此外, 非线性维数约简方法在空时扩展^[78]、快速算法^[79]、局部线性光滑性问题^[80]等方面也值得关注。

5 结束语

尽管非线性维数约简方法及其应用在过去几年中已经取得了丰硕的成果, 但由于其数学理论基础较为复杂, 以及多学科间的交叉、融合, 对高维数据中有意义的低维结构的研究依然有很多值得进一步探讨的问题。本文比较系统地回顾了流形学习中非线性维数约简方法的基本原理、算法流程、应用及发展方向等问题, 尤其对国内研究者在这一领域的研究、发展和应用等工作用了很大篇幅进行阐述分析。目前非线性维数约简方法还有很多问题处于研究中, 如现有的方法能在多大程度上逼近流形的内在结构、流形学习与核函数关系、流形学习的泛化能力等。这些问题也是机器学习领域中非常基础且非常具有挑战性的问题, 一旦在理论上取得突破性的进展, 有可能会引发整个机器学习领域的一次革命。

参考文献:

[1] OLLIFFE I T. Principal component analysis[M]. New York: Springer-Verlag, 1986.

[2] HYVARINEN A, OJA E, KARHUNEN J. Independent component analysis[M]. New York: Wiley, 2001.

[3] GORDON A D. Classification: methods for the exploratory analysis of multivariate data[M]. New York: Wiley, 1977.

[4] K GL B, KRZYZAK A, LINDER T, *et al.* Learning and design of principal curves[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(3) : 281-297.

[5] HUBER P J. Projection pursuit[J]. Annals of Statistics, 1985, 13

(2) : 435-475.

[6] HUO Xiao-ming, CHEN Ji-hong. Local linear projection (LLP) [C] //Proc of the Workshop on Genomic Signal Processing and Statistics. 2002: 1183-1186.

[7] KOHNONE T. Self-organizing maps[M] . 3rd ed. Berlin: Springer-Verlag, 2001.

[8] HASTIE T, TIBSHIRANI R, FRIEDMAN J. The element of statistical learning: data mining, inference, and predication[M] . New York: Springer, 2001.

[9] SCH LKOPF B, SMOLA A J, M LLER K R. Nonlinear component analysis as a kernel eigenvalue problem[J] . Neural Computation, 1998, 10(5) : 1299-1319.

[10] BACH F R, JORDAN M I. Kernel independent component analysis [J] . Journal of Machine Learning Research, 2002, 3: 1-48.

[11] MIKA S, RATSCH G, WESTON J, *et al.* Fisher discriminant analysis with kernels[C] //Proc of Neural Networks for Signal Processing Workshop. Madison: [s. n.] , 1999: 41-48.

[12] MacDONALD D, FYFE C. The kernel self-organising map [EB/OL] . [2006-12-16] . <http://cis.paisley.ac.uk/fyfe-ci0/kernelk-means.ps>.

[13] BRAND M. Continuous nonlinear dimensionality reduction by kernel eigenmaps[C] //CHAPELLE O, SCH LKOPF B, ZIEN A. Semi-supervised learning. Acapulco: MIT Press, 2006: 547-554.

[14] BORG I, GROENEN P. Modern multidimensional scaling: theory and application[M] . New York: Springer-Verlag, 1997.

[15] TENENBAUM J B, SILVAM V D, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction[J] . Science, 2000, 290(12) : 2319-2323.

[16] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J] . Science, 2000, 290 (5500) : 2323-2326.

[17] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation[J] . Neural Computations, 2003, 15(6) : 1373-1396.

[18] ZHANG Zhen-ye, ZHA Hong-yuan. Principal manifolds and nonlinear dimension reduction via local tangent space alignment[J] . SIAM Journal of Scientific Computing, 2004, 26(1) : 313-338.

[19] BALASUBRAMANIAN M, SCHWARTZ E L, TENENBAUM J B. The isomap algorithm and topological stability[J] . Science, 2002, 295(5552) : 7.

[20] NAYAR S K, NENE S A, MURASE H. Subspace methods for robot vision[J] . IEEE Transactions on Robotics and Automation, 1996, 12(5) : 750-758.

[21] SEBASTIAN H S, LEE D D. The manifold ways of perception[J] . Science, 2000, 290(12) : 2268-2269.

[22] DONOHO D L, GRIMES C. Hessian eigenmaps: new locally linear embedding techniques for high-dimensional data[J] . Proceedings of the National Academy of Sciences of the United States of American, 2003, 100(10) : 5591-5596.

[23] ZHANG Chang-shui, WANG Jun, ZHAO Nan-yuan, *et al.* Reconstruction and analysis of multi-pose face images based on nonlinear dimensionality reduction[J] . Pattern Recognition, 2004, 37 (1) : 325-336.

[24] WEINBERGER K Q, PACKER B, SAUL L K. Nonlinear dimensionality reduction by semidefinite programming and kernel matrix factorization[C] //COWELL R G, GHAMRANI Z. Proc of the 10th International Workshop on Artificial Intelligence and Statistics. Barbados: Society for Artificial Intelligence and Statistics, 2005: 381-388.

[25] WEINBERGER K Q, SAUL L K. Unsupervised learning of image manifolds by semidefinite programming[C] //Proc of the IEEE Conference on Computer Vision and Pattern Recognition. 2004: 988-995.

[26] WEINBERGER K Q, SAUL L K. An introduction to nonlinear dimensionality reduction by maximum variance unfolding[EB/OL] . [2006-12-16] . <http://www.seas.upenn.edu/~kilianw/publications/PDFs/AAAI0620WeinbergerK.pdf>.

[27] BELKIN M, NIYOGI P. Semi-supervised learning on riemannian manifolds[J] . Machine Learning, 2004, 56(1) : 209-239.

[28] BELKIN M, NIYOGI P, SINDHWANI V. On manifold regularization [C] //Proc of the 10th International Workshop on Artificial Intelligence and Statistics. 2005: 17-24.

[29] BELKIN M, NIYOGI P. Using manifold structure for partially labelled classification[C] //Proc of Advances in Neural Information Processing Systems. 2003: 929-936.

[30] ZHU Xiao-jin, GHAMRANI Z, LAFFERTY J. Semi-supervised learning using gaussian fields and harmonic functions[C] //Proc of the 20th International Conference on Machine Learning. Washington, DC: [s. n.] , 2003: 58-65.

[31] ROWEIS S, SAUL L, HINTON G. Global coordination of local linear models[C] //Proc of Advances in Neural Information Processing Systems. 2002: 889-896.

[32] BRAND M. Minimax embeddings[C] //Proc of Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2004: 505-512.

[33] CHIGIREV D, BIALEK W. Optimal manifold representation of data: an information theoretic approach[EB/OL] . [2006-12-16] . http://books.nips.cc/papers/files/nips16/NIPS2003_AA21.ps.gz.

[34] 詹德川, 周志华. 基于集成的流形学习可视化[J] . 计算机研究与发展, 2005, 42(9) : 1533-1537.

[35] 赵连伟, 罗四维, 赵艳敞. 高维数据的低维嵌入及嵌入维数研究 [J] . 软件学报, 2005, 16(8) : 1423-1430.

[36] 何力, 张军平, 周志平. 基于放大因子和延伸方向研究流形学习算法[J] . 计算机学报, 2005, 28(12) : 2000-2009.

[37] 杨剑, 李伏欣, 王珏. 一种改进的局部切空排列算法[J] . 软件学报, 2005, 16(9) : 1584-1589.

[38] MIN Wan-li, LU Ke, HE Xiao-fei. Locality pursuit embedding[J] . Pattern Recognition, 2004, 37(4) : 781-788.

[39] NILSSON J. Nonlinear dimensionality reduction of gene expression data[D] . Lund, Sweden: Lund University, 2006.

[40] 徐志节, 杨杰, 王猛. 利用非线性降维方法预测膜蛋白类型[J] . 上海交通大学学报, 2005, 39(2) : 279-293.

[41] 翁时锋, 张长水, 张学工. 非线性降维在高维医学数据处理中的应用[J] . 清华大学学报: 自然科学版, 2004, 44(4) : 485-488.

[42] SHEN Xi-lin, MEYER F G. Nonlinear dimension reduction and activation detection for fMRI dataset[C] //Proc of Computer Vision and Pattern Recognition Workshop. 2006: 144-151.

[43] DAWSON K, RODRIGUEZ, MALYJ W. Sample phenotype clusters in high-density oligonucleotide microarray data sets are revealed using isomap, a nonlinear algorithm[J] . BMC Bioinformatics, 2005, 6 (1) : 195.

[44] KHURD P, VERMA R, DAVATZIKOS C. On characterizing and analyzing diffusion tensor images by learning their underlying manifold structure[C] //Proc of Computer Vision and Pattern Recognition

Workshop. 2006: 61-68.

[45] LASKARIS N A, IOANNIDES A A. Semantic geodesic maps: a unifying geometrical approach for studying the structure and dynamics of single trial evoked responses[J]. Clinical Neurophysiology, 2002, 113(8) : 1209-1226.

[46] HI C, CHANG Y, FERIS R, *et al.* Manifold based analysis of facial expression[J]. Image and Vision Computing, 2006, 24(6) : 605-614.

[47] SAUL L, ROWEIS S. Think globally, fit locally: unsupervised learning of nonlinear manifolds[J]. Journal of Machine Learning Research, 2003, 4: 119-155.

[48] GENG Xin, ZHAN De-chuan, ZHOU Zhi-hua. Supervised nonlinear dimensionality reduction for visualization and classification[J]. IEEE Transactions on Systems, Man, and Cybernetics: Part B, 2005, 35(6) : 1098-1107.

[49] LIU C B, LIN R S, NARENDRA A, *et al.* Dynamic textures synthesis as nonlinear manifold learning and traversing[C] //Proc of the 17th British Machine Vision Conference. 2006: 859-868.

[50] CHANG H, YEUNG D Y. Locally linear metric adaptation with application to semi-supervised clustering and image retrieval[J]. Pattern Recognition, 2006, 39(7) : 1253-1264.

[51] SOUVENIR R, PLESS R. Manifold clustering[C] //Proc of IEEE International Conference on Computer Vision. Beijing: [s. n.], 2005: 648-653.

[52] MORARIU V I, OCTAVIA I O I. Modeling correspondences for multi-camera tracking using nonlinear manifold learning and target dynamics[C] //Proc of the IEEE Conference on Computer Vision and Pattern Recognition. 2006: 545- 552.

[53] BACHMANN C M, AINSWORTH T L, FUSINA R A. Exploiting manifold geometry in hyperspectral imagery[J]. IEEE Transactions Geoscience Remote Sensing, 2005, 43(3) : 441-454.

[54] SEWARD A E, BODENHEIMER B. Using nonlinear dimensionality reduction in 3D figure animation[C] //Proc of the 43th ACM SE Regional Conference. 2005: 388-392.

[55] NISKANEN M, SILVN O. Comparison of dimensionality reduction methods for wood surface inspection[C] //Proc of the 6th International Conference on Quality Control by Artificial Vision. Gatlinburg: [s. n.], 2003: 178-188.

[56] 杨建宏, 徐金梧, 杨德斌, 等. 基于相重构和主流形识别的非线性时间序列降噪方法[J]. 北京科技大学学报, 2005, 27(5) : 631-634.

[57] JAIN V, SAUL L K. Exploratory analysis and visualization of speech and music by locally linear embedding[C] //Proc of the International Conference of Speech, Acoustics, and Signal Processing. 2004: 984-987.

[58] HE Xiao-fei, CAI Deng, LIU Hai-feng, *et al.* Locality preserving indexing for document representation[C] //Proc of the 27th Annual ACM SIGIR Conference. Sheffield: [s. n.], 2004: 96-103.

[59] BELKIN M, NIYOGI P. Laplacian eigenmaps and spectral techniques for embedding and clustering[C] //Proc of Neural Information Processing Systems. Cambridge: [s. n.], 2003: 1373-1396.

[60] KOUROPTOVA O, OKUN O, PIETIK INEN M. Classification of handwritten digits using supervised locally linear embedding algorithm and support vector machine[C] //Proc of the 11th European Symposium on Artificial Neural Networks. Belgium: [s. n.], 2003: 229-234.

[61] 许馨, 吴福朝, 胡占义, 等. 一种基于非线性降维求正常星系红移的新方法[J]. 光谱学与光谱分析, 2006, 26(1) : 182-186.

[62] LIOU C Y, KUO Y T. Economic states on neuronc maps[C] //Proc of the 9th International Conference on Neural Information Processing. 2002: 787-791.

[63] CHEN H T, CHANG Huang-wei, LIU T L. Local discriminant embedding and its variants[C] //Proc of Conference on Computer Vision and Pattern Recognition. San Diego: IEEE Computer Society, 2005: 846-853.

[64] 谭璐, 吴翊. 局部规则嵌入[J]. 计算机应用, 2005, 25(4) : 817-819.

[65] 谭璐, 易东云, 冯国柱, 等. 局部不变投影[J]. 自然科学进展, 2004, 14(3) : 282-287.

[66] HE Xiao-fei, NIYOGI P. Locality preserving projections[C] //Proc of Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2004: 153-160.

[67] De RIDDER D, KOUROPTOVA O, OKUN O, *et al.* Supervised locally linear embedding[C] //Lecture Notes in Computer Science. [S. l.] : Springer, 2003: 333-341.

[68] PANG Yan-wei, ZHANG Lei, LIU Zheng-kai. Neighborhood preserving projections (NPP) : a novel linear dimension reduction method [C] //Lecture Notes in Computer Science. 2005: 117-125.

[69] CHOI H, CHOI S. Kernel isomap on noisy manifold[C] //Proc of IEEE International Conference on Development and Learning. 2005: 208-213.

[70] CHOI H, CHOI S. Kemel isomap[J]. Electronics Letters, 2004, 40(25) : 1612-1613.

[71] YAN Shui-cheng, XU Dong, ZHANG Hong-jiang, *et al.* Graph embedding: a general framework for dimensionality reduction[C] //Proc of Conference on Computer Vision and Pattern Recognition. 2005: 830-837.

[72] WEINBERGER K Q, SHA F, SAUL L K. Learning a kernel matrix for nonlinear dimensionality reduction[C] //Proc of the 21st International Conference on Machine Learning. Banff: [s. n.], 2004: 839-846.

[73] HAM J, LEE D D, MIKA S, *et al.* A kernel view of the dimensionality reduction of manifolds[C] //Proc of the 21st International Conference on Machine Learning. Banff: [s. n.], 2004: 369-376.

[74] 谭璐, 吴翊, 易东云. 稳健局部线性嵌入方法[J]. 国防科技大学学报, 2004, 26(4) : 91-95.

[75] 谭璐. 高维数据的降维理论及应用[D]. 长沙: 国防科学技术大学, 2005.

[76] YANG Xin, FU Hao-ying, ZHA Hong-yuan, *et al.* Semi-supervised nonlinear dimensionality reduction[C] //Proc of the 23th International Conference on Machine Learning. 2006: 1065-1072.

[77] BENGIO Y, MONPERRUS M. Non-local manifold tangent learning [J]. Advances in Neural Information Processing Systems, 2005, 17: 129-136.

[78] JENKINS O, MATARIC M. A spatio-temporal extension to isomap nonlinear dimension reduction[C] //Proc of the 21st International Conference on Machine Learning. 2004: 56-64.

[79] ZHANG Zhen-yue, ZHA Hong-yuan. A domain decomposition method for fast manifold learning[C] //Advances in Neural Information Processing Systems. 2006: 1625-1632.

[80] ZHANG Zhen-yue, ZHA Hong-yuan. Local linear smoothing for nonlinear manifold learning, CSE- 03- 003 [R]. Pennsylvania: Penn State University, 2003.