

# 流形学习概述

徐 蓉,姜 峰,姚鸿勋

(哈尔滨工业大学 计算机科学与技术学院,黑龙江 哈尔滨 150001)

**摘 要:**流形学习是一种新的非监督学习方法,近年来引起越来越多机器学习和认知科学工作者的重视.为了加深对流形学习的认识和理解,该文由流形学习的拓扑学概念入手,追溯它的发展过程.在明确流形学习的不同表示方法后,针对几种主要的流形算法,分析它们各自的优势和不足,然后分别引用 Iso map 和 LLE 的应用示例.结果表明,流形学习较之于传统的线性降维方法,能够有效地发现非线性高维数据的本质维数,利于进行维数约简和数据分析.最后对流形学习未来的研究方向做出展望,以期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;等距离映射算法;局部线性嵌入算法;交叉流形

**中图分类号:** TP181 **文献标识码:** A **文章编号:** 1673-4785(2006)01-0044-08

## Overview of manifold learning

XU Rong, JIANG Feng, YAO Hong-xun

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

**Abstract:** As a new unsupervised learning method, manifold learning is capturing increasing interests of researchers in the field of machine learning and cognitive sciences. To understand manifold learning better, the topology concept of manifold learning was presented firstly, and then its development history was traced. Based on different representations of manifold, several major algorithms were introduced, whose advantages and defects were pointed out respectively. After that, two kinds of typical applications of Iso map and LLE were indicated. The results show that compared with traditional linear method, manifold learning can discover the intrinsic dimensions of nonlinear high-dimensional data effectively, helping researchers to reduce dimensionality and analyze data better. Finally the prospect of manifold learning was discussed, so as to extend the application area of manifold learning.

**Key words:** dimensionality reduction; manifold learning; Iso map; LLE; intersecting manifold

随着信息时代的到来,科研工作者在研究过程中不可避免地会遇到大量的高维数据,如全球气候模型、人类基因分布、文本聚类中的词频等,所以经常会面临维数约简的问题,维数约简的目的是要找出隐藏在低维数据中的低维结构.人脑也面临着同样的问题:人脑在每天的感知活动中要从3万个听觉神经的输入和1万个视觉神经的输入中抽取有意义的信息.

从几何的观点来看,维数约简可以看成是挖掘嵌入在高维数据中的低维线性或非线性流形.这种嵌入保留了原始数据的几何特性,即在高维空间中

靠近的点在嵌入空间中也相互靠近.

现有的维数约简方法中,独立分量分析<sup>[1]</sup>、主成分分析<sup>[2]</sup>对具有线性结构的数据集处理效果很好,而小波、傅里叶变换、Had2mard变换处理图像<sup>[3]</sup>的结果也是令人满意的.然而,独立分量分析假定数据集由内在多个信源产生的信号叠加而成,根据信息论最小化互信息来获得数据的线性结构,分析中忽略了数据在观测空间的全局与局部性质.主成分分析寻找数据二阶统计性质来发现数据集的线性结构,但对于高度非线性分布的数据集并不能找到真正的分布结构.傅里叶变换本质上是将数据集变换到频域进行约简,小波变换增加了时域信息,但它们缺乏几何上的直观解释.因此,基于数据分布的内在

收稿日期:2006-03-01.

基金项目:国家自然科学基金资助项目(60332010,60533030).

维数来分析数据是机器学习和多元数据分析的重要研究方向,流形学习方法提供了一种新的研究途径.

流形学习旨在发现高维数据集分布的内在规律性,其基本思想是:高维观测空间中的点由少数独立变量的共同作用在观测空间张成一个流形,如果能有效地展开观测空间卷曲的流形或发现内在的主要变量,就可以对该数据集进行降维.

## 1 流形和流形学习

拓扑学最基本的研究对象是拓扑空间.所谓拓扑空间就是一个集合,上面赋予了拓扑结构,拓扑空间之间可以定义连续映射.设  $X, Y$  是 2 个拓扑空间,  $f: X \rightarrow Y$  是 1 个连续映射,如果  $f$  有逆映射,而且逆映射也是连续的,那么就说  $f$  是一个同胚映射,并且说  $X$  与  $Y$  同胚.比如,不同大小的 2 个球面就是同胚的,它们跟凸多面体的表面也是同胚的.从根本上说,拓扑学就是研究几何体在同胚映射下的不变性质的一个数学分支.

拓扑空间  $M$  在满足以下条件时,称  $M$  为  $m$  维流形,即:

1)  $M$  为豪斯多夫空间.豪斯多夫空间是数学拓扑学中的一个分离空间,满足分离定理:对于拓扑空间  $M$  中任意 2 个不同的点  $x$  和  $y$ ,存在  $x$  的邻域  $U$  和  $y$  的邻域  $V$ ,满足  $U \cap V = \emptyset$ .

2) 对于任意一点  $p \in M$ ,存在包含  $p$  的  $m$  维坐标邻域  $(U, \phi)$ ,坐标邻域是拓扑空间中的开集与其在欧几里德空间中的映射  $\phi$  的有序对.

流形是拓扑学中的概念,其表示一个局部处为欧几里德的拓扑空间.局部欧几里德特性意味着对于空间上任一点都有一个邻域,在这个邻域中的拓扑与  $R^m$  空间中的开放单位圆相同,  $R^m$  表示  $m$  维欧式空间.也就是说,流形是一个局部可坐标化的拓扑空间,从拓扑空间的一个开集(邻域)到欧氏空间的开子集的同胚映射,使得每个局部可坐标化.它的本质是分段线性处理,计算时要考虑开集和同胚映射的选择问题.

有了对流形的定义,就可以形式化地概括流形学习这一维数约简过程:

假设数据是均匀采样于一个高维欧氏空间中的低维流形,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.它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.

用数学语言可以这样描述,令  $Y \subset R^d$  且  $f:$

$Y \subset R^d$  是一个光滑的嵌套,  $D > d$ .流形学习的目标是基于  $R^D$  上的一个给定被观测数据集  $\{x_i\}$  去恢复  $Y$  与  $f$ .在  $Y$  中隐藏的数据  $\{y_i\}$  被随机地产生,然后被  $f$  映射到观测空间,使得  $\{x_i = f(y_i)\}$ .

## 2 流形学习的产生与发展

20 世纪 80 年代末期,在 PAMI 上就已经有流形模式识别的说法.2000 年,美国《Science》上发表 3 篇论文<sup>[4-6]</sup>,从认知上讨论了流形学习,并使用了 manifold learning 的术语,强调认知过程的整体性.

王守觉提出自然界中同类事物整体的连续性规律<sup>[7]</sup>,即同类中的 2 个不同个体之间总是存在一个渐变的过程,且这个渐变过程中间的各事物都是属于同一类的,它也是量变引起质变这一哲学思想的体现;认知科学家认为同一事物在随着时间、空间等因素连续发生变化时形成了一个低维的流形,并且人的强大的认知能力正是基于对这一稳定状态流形的视觉记忆的<sup>[4]</sup>.在 2 者基础上,假定事物本身所在空间里离得近的物体具有相同的类别属性就更为合理,因为不同的特征提取方法会造成在表示空间或者特征空间中离得很近的 2 个样本并不一定是在事物空间本身中离得很近.

机器学习的一个大目标是要从数据中学习其相关的规律性.因此,近年来流形学习在机器学习领域成为一个新的热点问题,它试图把人的认知流形规律引入到机器学习研究中,使机器能更有效地进行学习.流形学习的本质是当采样数据所在的空间为一低维光滑流形时,要从采样数据学习出低维流形的内在几何结构或者内在规律.这就意味着流形学习比传统的维数约简方法更能体现事物的本质,更利于对数据的理解和进一步处理,进而更好地解决一些以前在机器学习领域完成得不好或者无法解决的问题.

近年来,流形学习领域产生了大量的研究成果<sup>[8]</sup>.LLE<sup>[6]</sup>和 Isomap<sup>[5]</sup>是 2 种有代表性的非线性降维方法.Roweis 和 Saul 提出的 LLE 算法能够实现高维输入数据点映射到一个全局低维坐标系,同时保留了邻接点之间的关系,这样固有的几何结构就能够得到保留.此算法不仅能够有效地发现数据的非线性结构,同时还具有平移、旋转等不变特性.Tenenbaum 等人提出的 Isomap 算法首先使用最近邻图中的最短路径得到近似的测地线距离,代替不能表示内在流形结构的 Euclidean 距离,然后输入到多维尺度分析(MDS)中处理,进而发现嵌入在高维空间的低维坐标.在人脸和手势的实验中, Iso-

map 发现了存在于高维空间中的潜在低维参数空间. Donoho 等人利用人工合成的数据用 Isomap 算法进行测试实验, 实验结果表明, Isomap 能够准确地发现图像流形潜在的参数空间, 并在自然图像(人脸图像)中不同姿态和亮度等潜在的未知参数下也可得到较好的结果. Donoho 等人还拓展了 LLE 算法, 提出 HLL 算法, 能够发现流形上局部的潜在等距映射参数. 张长水等人<sup>[9]</sup>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, 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降维方法. 由于已有的流形学习算法对噪音和算法参数都比较敏感, 詹德川、周志华<sup>[10]</sup>针对 Isomap 算法, 提出了一种新方法, 通过引入集成学习技术, 扩大了可以有效可视化结果的输入参数范围, 并且降低了对噪音的敏感性. 另外, 赵连伟<sup>[11]</sup>等人完善了 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, 流形维数则要小于嵌入空间维数. 他们还给出一种有效的环状流形发现算法, 以得到正确的低维参数空间. 特别地, 周志华等<sup>[12]</sup>提出了一种方法, 可以从放大因子和延伸方向两方面显示出观测空间的高维数据与降维后的低维数据之间的联系, 并实验比较了 2 种著名流形学习算法 (Isomap 和 LLE) 的性能.

目前, 大部分流形学习算法都是用于非线性维数约简或数据可视化的, 如等距映射算法<sup>[6]</sup> (Isomap), 局部线性嵌入算法<sup>[7]</sup> (LLE), 拉普拉斯特征映射算法<sup>[13]</sup> (laplacian eigenmap), 局部切空间排列算法<sup>[14]</sup> (LTSA) 等等, 并且已经在图像处理如人脸图像、手写数字图像、语言处理方面取得了较好的效果. 而切空间距离 (tangent distance) 分类器算法<sup>[15]</sup> 则将流形学习思想引入到监督学习中用于模式识别, 并在 MNIST 手写数字数据集上效果显著. 但它在求切距离时假设已知对样本原型点进行的各种变换操作, 这就需要很强的领域知识和对数据集本身结构的清楚认识. 而对于机器学习研究者来说, 这 2 个条件都是较难具备的, 所以在这方面的应用中取得的成果有限.

### 3 流形学习的分类和主要算法

#### 3.1 流形学习的分类

按照对低维流形或嵌入映射的不同约束或限制方式, 流形学习可以大致划分为 4 类<sup>[16]</sup>:

1) 投影法: 寻找通过数据内部的主表面, 比如主曲线法<sup>[17-18]</sup> (principal curves). 但由于几何直觉, 这种方法很难将弧长参数这样的全局变量推广到高维空间.

2) 生成法: 采用生成拓扑模型<sup>[19-21]</sup>. 假定观测数据是从均匀分布在低维的隐结点中生成的, 然后对观测空间和隐空间建立映射关系. 但是由于所用期望最大化算法本身的缺陷, 生成模型容易陷入局部最小并且收敛速度较慢.

3) 嵌入法: 分为全局嵌入法和局部嵌入法. Isomap<sup>[6]</sup> 就是全局嵌入法, 假设在观察空间与仿射变换下的本质嵌入空间中容量特征能够得以保留. LLE<sup>[7]</sup> 和 Laplacian 特征映射<sup>[13]</sup> 则侧重于保证局部邻域结构. 全局嵌入法的优点在于能够更忠实地表达数据的全局结构, 易于从理论角度理解度量的保留. 而局部嵌入法拥有 2 方面的优势: 计算量上的优势, 只包含多项式数量级的稀疏矩阵运算; 良好的表达能力, 即全局结构为非欧式空间的情况下, 局部几何结构接近于欧式空间.

4) 公有信息法: 将公有信息作为观察空间与嵌入空间概率分布差异度的衡量. 典型的算法应用是随机最近邻法 SNE<sup>[22]</sup> 和流形制图<sup>[23]</sup>.

#### 3.2 流形学习的主要算法

Isomap 方法和 LLE 方法作为流形学习的突出代表, 以其各自的优势引起了众多理论和应用上的关注, 2 种算法的主要假设是: 数据在观测空间中的距离关系仅在局部可使用欧氏度量. 拉普拉斯特征映射与 LLE 方法都是局部嵌入法, 是以把高维空间中离得很近的点映射到低维空间中也离得很近为目的, 利用图拉普拉斯算子的性质进行求解的一种流形学习算法. LTSA 是一种新的流形算法, 它假定样本集采样于某个参数流形且含有噪声的无序数据点集, 然后用近似的局部切空间来表示样本点所在流形的局部几何.

##### 3.2.1 Isomap 方法

Isomap 算法建立在多维尺度变换<sup>[24]</sup> (MDS) 的基础上, 力求保持数据点的内在几何性质, 即保持 2 点间的测地距离. 如图 1 所示, (a) 中样本分布于 swiss-roll 上, 两点间的欧式距离 (虚线) 不能表征 2 点的实际距离, 分布于流形面上的曲线是 2 点的测地线距离. 流形未知时可以通过最短路径算法对邻域内的距离进行近似地重构两点间的测地线距离, 见图 1(b). (c) 是 Isomap 降维后 2 点和 2 条路径 (测地线和短程拼接) 的投影结果.

Isomap 算法的关键是利用样本向量之间的欧

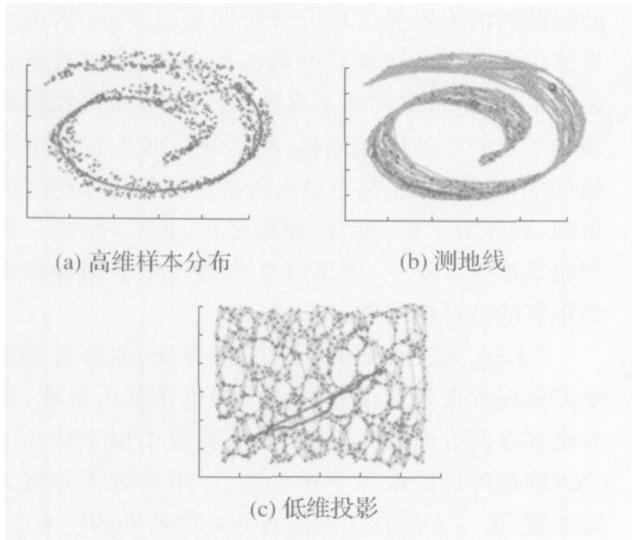


图 1 Isomap 的基本思想  
Fig.1 The basic idea of Isomap

式距离  $d_x(i, j)$  计算出样本之间的测地距离  $d_c(i, j)$ , 真实地再现高维数据内在的非线性几何结构. 然后使用经典 MDS 算法构造一个新的  $d$  维空间  $Y$  ( $d$  是降维后空间的维数), 最大限度地保持样本之间的欧式距离  $d_Y(i, j)$  与  $d_c(i, j)$  误差最小, 以达到降维的目的, 算法描述见图 2.

输入:输入样本  $X = \{x_1, x_2, \dots, x_n \in R^N\}$ , 样本本真维数  $d$ , 邻域参数  $k$  (邻域);  
输出:低维嵌入  $Y = \{y_1, y_2, \dots, y_n \in R^d\}$   
算法:计算每个点  $x_i$  的近邻点 ( $k$  邻域),  $1 \leq i \leq n$ ;  
/\* 构造近邻图 \*/ 若两点  $x_i, x_j$  互为近邻点, 则相应的边值设为欧式距离  $d_x(i, j)$ , 否则为 0;  
/\* 求得最短路径矩阵  $D_G = \{d_G(i, j)\} \in R^{n \times n}$  \*/  
计算任意 2 点  $x_i, x_j$  间的最短路径  $d_G(i, j)$ ;  
/\* 用 MDS 求低维嵌入流形 \*/  
 $S = (S_{ij}) = (D_{ij}^2)$ ,  $H = (H_{ij}) = (d_{ij} - 1/N)$ ,  $(D) = -HS^{-1}H/2$ ,  
低维嵌入是  $(D)$  最小的第 2 个到第  $d+1$  个特征值对应的特征向量.

图 2 Isomap 算法  
Fig.2 The algorithm of Isomap

Isomap 是一种非线性的学习方法, 它适用于学习内部平坦的低维流形, 但不适于学习有较大内在曲率的流形. 它的算法中需要确定 2 个参数: 1 个是邻域的大小, 1 个是降维的维数. 在噪声干扰下, Isomap 用于可视化会有不稳定现象, 取较大的邻域会产生短路现象, 即低维流形中不同邻域部分的点投影后出现明显的混杂现象. 选取较小的邻域, 虽然能够保证整体结构的稳定, 但低维投影结果会产生大量“空洞”, 或使最短路径算法重构的图不连通. 降维

维数的确定通常是在本质维数未知的情况下进行的, 经多次实验绘制残差曲线观察得到. Isomap 算法计算图上 2 点间的最短距离可采用 Dijkstra 算法, 但执行起来仍然比较慢.

3.2.2 LLE 方法

LLE 算法认为在局部意义下, 数据的结构是线性的, 或者说局部意义下的点在一个超平面上, 因此任取一点, 可以使用它的邻近点的线性组合来表示. 图 3 中 (b) 的 3 维数据由 (a) 中的 2 维流形中采样而来, 彩色带表明了 LLE 映射处理后保留的邻域带情况. (b) 与 (c) 中用黑线圈出了单个点的邻域部分. 该文在图 4 中给出了 LLE 的算法描述.

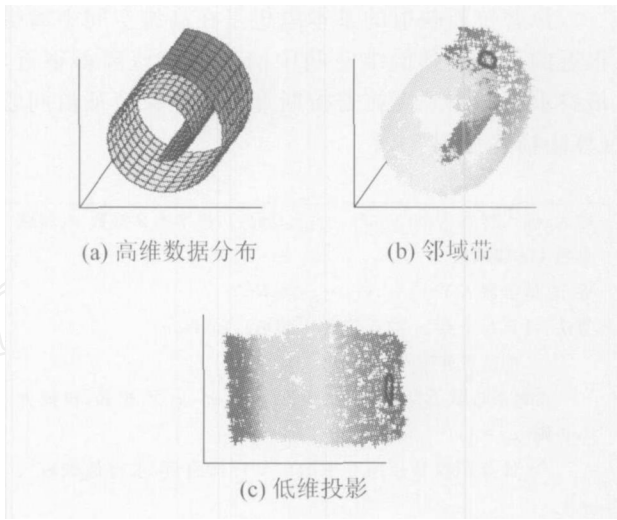


图 3 LLE 的基本思想  
Fig.3 The basic idea of LLE

输入:输入样本  $X = \{x_1, x_2, \dots, x_n \in R^N\}$ , 样本本真维数  $d$ , 邻域参数  $k$  (邻域);  
输出:低维嵌入  $Y = \{y_1, y_2, \dots, y_n \in R^d\}$   
算法:计算每个点  $x_i$  的近邻点 ( $k$  邻域),  $1 \leq i \leq n$ ;  
/\* 求解权值矩阵 \*/  
若  $x_i, x_j$  不是近邻点, 则  $w_{ij} = 0$  且  $\sum_j w_{ij} = 1$ ,  
用最小二乘法最小化重构成本函数  $g(w) = \sum_i \|x_i - \sum_j w_{ij} x_j\|^2$ ,  
解得  $w_{ij} = (G_{jk}^{-1})_{ik} / \sum_l (G_{lk}^{-1})_{il}$ ,  
其中  $G_{jk} = (x_i - x_j)(x_i - x_k)$ ,  $j$  和  $k$  是  $x_i$  的近邻点;  
/\* 权值矩阵  $W = \{w_{ij}\}$  不变 \*/  
最小化嵌入成本函数  $\phi(Y) = \sum_i \|y_i - \sum_j w_{ij} y_j\|^2$ , 使低维重构误差最小,  
 $M = (I - W)^T(I - W)$ , 低维嵌入是  $M$  最小的第 2 个到第  $d+1$  个特征向量.

图 4 LLE 算法的描述  
Fig.4 The algorithm of LLE

LLE 算法可以学习任意维的局部线性的低维流形,和 Isomap 方法一样只有 2 个待定系数  $k$  和  $d$ . 同时由重构成本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、伸缩变换下是保持不变的. LLE 算法有解析的全局最优解,不需迭代,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 算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,该算法的参数选择不确定,对样本中的噪音很敏感.

### 3.2.3 拉普拉斯映射法

拉普拉斯映射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(算法描述见图 5).

输入:输入样本  $X = \{x_1, x_2, \dots, x_n \in R^N\}$ , 样本本真维数  $d$ , 邻域参数  $k$  (邻域);

输出:低维嵌入  $Y = \{y_1, y_2, \dots, y_n \in R^d\}$

算法:计算每个点  $x_i$  的近邻点 ( $k$  邻域),  $1 \leq i \leq n$ ;

/\*构造近邻图\*/

给每条边赋予权值  $w_{ij}$ , 如果两个点  $x_i, x_j$  不相邻, 权值为 0, 否则  $w_{ij} = 1$ ;

/\*计算图拉普拉斯算子的广义特征向量, 求得低维嵌入\*/

$D$  为对角矩阵, 且  $D_{ij} = \sum_j w_{ji}$ ,  $L = D - W$ ,  $L$  是近邻图上的拉普拉斯算子,

求解广义特征值问题,  $Lf = \lambda f$ .

图 5 拉普拉斯特征映射的算法描述

Fig. 5 The algorithm of Laplacian eigenmaps

拉普拉斯映射法是局部的非线性方法,其突出特点是与谱图理论有着很紧密的联系. 从算法描述中可以看到,拉普拉斯映射和 LLE 算法类似,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,故可用于聚类.

### 3.2.4 局部切空间排列法

局部切空间排列算法(local tangent space alignment, L TSA)是一种基于切空间的流形学习算法,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.

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<sup>[25]</sup>,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

的问题转换为局部主成分分析问题. L TSA 用  $K$ -近邻求出每一个样本点的局部散布矩阵,对其进行特征值分解,求出局部主成分,通过向主成分投影得到样本点的低维局部坐标. 算法进一步假定样本点的低维整体嵌入坐标可以由局部坐标经过一些变换得到,具体为平移、旋转、伸缩变换. 最后,经过一系列数学推导, L TSA 把求解整体嵌入坐标问题转换为矩阵的特征值问题.

L TSA 是一种新的流形学习算法,能够有效地学习体现数据集低维流形结构的整体嵌入坐标,但它也存在两方面的不足:一方面算法中用于特征值分解的矩阵的阶数等于样本数,样本集较大时将无法处理;另一方面算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算法<sup>[26]</sup>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 L TSA 算法的基础上,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 L TSA 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.

## 4 应用举例

### 4.1 Isomap 算法在医学数据处理中的一个应用

医学研究中的数据往往是非线性的高维数据,传统降维和聚类方法由于自身局限性,使得一些本质维数较低的高维数据无法投影到较低的观察空间去. Isomap 方法提供了这样一个平台来获得低维流形,便于人们进行可视化操作.

乳腺癌是临床上比较常见的一种癌症,特别是女性患者比较多. 区分乳腺癌的良性肿瘤和恶性肿瘤在临床上具有很重要的意义. 以下是对乳腺癌病理数据的处理结果,并且和传统的 PCA 方法进行了比较,可见 Isomap 方法有更好的降维能力<sup>[27]</sup>.

定义残差  $e_d$  来衡量降维的误差,以确定降维的维数  $d$ ,定义残差如下式:

$$e_d = 1 - R^2(D_G, D_Y).$$

式中:  $D_Y$  是  $d$  维空间中的欧式距离矩阵,  $D_G$  是最短路径矩阵,  $R^2$  代表相关系数.

一般来说,降维的维数  $d$  越大,残差越小. 确定  $d$  有 2 种情况,一是残差曲线出现拐点,二是残差小于一定的阈值.

图 6 是邻域  $k = 7$  的情况下构造近邻图后获得的残差曲线,可见在维数是 3 时有拐点出现,并且残差的绝对值小于 0.05.

图 7(a)是 PCA 前 3 维的投影结果,2 类样本在空间中有很大的重叠;而图 7(b)中良性肿瘤和恶性

肿瘤得以良好的分开,很直观地看到数据分布的特性.同时通过统计样本的类内距离发现,Isomap 的效果也很明显: Isomap 投影后,良性和恶性肿瘤的类内距离分别是:0.105 4 和 0.330 1,而 PCA 投影得到的类内距离则是:0.321 5 和 0.381 2.

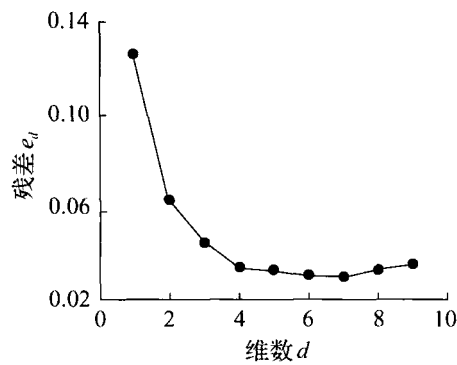
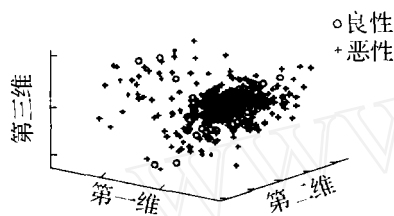
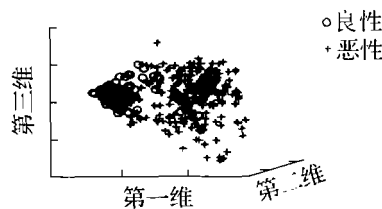


图 6 乳腺癌数据的残差曲线

Fig. 6 The variance fraction curve of breast cancer data



(a) 乳腺癌病理数据的 PCA 投影图



(b) 乳腺癌病理数据的 Isomap 投影结果

图 7 处理结果

Fig. 7 The results

4.2 LLE 算法在人脸图像处理中的一个应用

图 8 中的例子是嵌入在一个具有较高维数空间中的 2 维流形. 样本产生的过程是:在具有较大随机噪声的背景中平移单个人脸的图像. 样本之间的噪声是不相关的,人脸质心参数化的 2 维流形描述了结果图像唯一的一致性结构. LLE 的输入是 961 幅灰度图像,每幅图像包含一张大小为 28 ×20 的人脸,其添加在 59 ×51 的噪音背景中. 因此,进行可视化操作之前,平移人脸的图像是处于用像素坐标值表示的非线性高维(3 009)空间中. 图 8 的下面部分是 LLE 发现的前两维分量,每个数据点采用 4 个近

邻点. 作为对比,上面的部分显示了 PCA 发现的前两维分量. 显而易见, LLE 更好地表达了此例中的流形结构. LLE 把人脸居于边角位置的图像映射到二维嵌入空间的边角位置,而 PCA 却不能保留这些邻域结构.

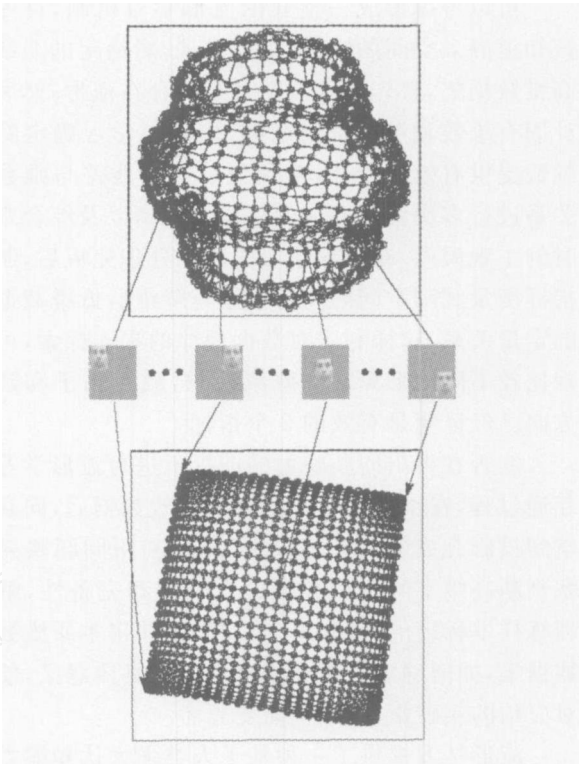


图 8 投影结果:PCA 结果(上)和 LLE 结果(下)

Fig. 8 The results of PCA (top) and LLE (bottom)

5 未来研究方向

LLE, Isomap, Laplacian Eigenmap 之所以有效,是因为它们既是非参数的方法,又是非线性的方法,这意味着无需对流形作很多参数假设,就能真实再现数据的内在维数和聚类特性. 它们的求解简单,都转化为求解特征值问题,而不需要用迭代算法. LTSA 方法是对非线性空间的局部线性信息的全局整合,局部应用 PCA 方法,其优化问题也等价地转化为一个导致特征值解法的简单问题.

目前已有的流形算法对噪声和算法参数都比较敏感,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,提高流形学习的抗噪性成为亟待解决的问题. 希望通过将流形算法与已有的成熟机器学习方法相结合,研究各种噪声模型对流形学习的影响方式和影响度,来对这个问题的解决提供帮助.

流形学习作为一种可视化手段,自然需要某种可视化效果的度量,比如坐标相关性<sup>[10]</sup>. 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.

可以考虑形成一整套的流形学习机制,首先应该知道嵌入空间的维数,也就是说,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,为流形学习确定降维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的不必要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,制定指标衡量观测空间的高维数据与降维后低维数据间的定量关系,这样利于对数据规律的深入探索,可直观比较不同流形算法的降维效果,放大因子和延伸方向已经证明是有效的2个指标<sup>[12]</sup>.

能否在得出同胚映射的前提下进行流形学习的互逆过程,在本质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,对已经获得本质维数的数据集,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.

流形学习提供了一种基于人类强大认知能力模型的工具,然而现实世界的情况依旧是复杂的、多变的, Richard Souvenir 和 Robert Pless 就提出了流形学习应用范围扩展的新问题,即多个交叉流形的分类和参数化.图9(a)既可以嵌入到一个最小误差意义下的二维流形中,又可以嵌入为2个一维流形;(b)则由2个不同维数的分流形组成.当流形出现重叠现象的时候,先前的 Iso map, LLE 方法将会失效.

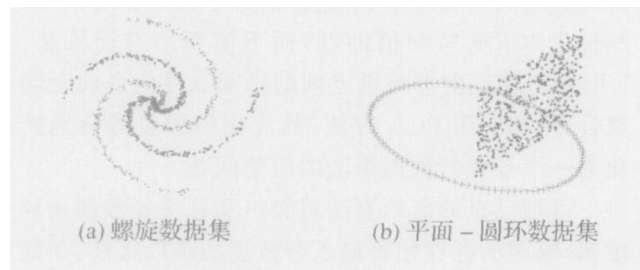


图9 交叉流形

Fig. 9 Intersecting manifold

针对这一问题,提出2种技术解决方案:节点加权的MDS,对秩-1矩阵的低秩矩阵近似的快速算法,并且通过一个具有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.

## 6 结束语

流形学习的过程是有章可循的,首先对嵌入映射或者低维流形做出某种特定的假设,或者以保持高维数据的某种性质不变为目标,然后将问题转化为求解优化问题并且提供有效算法的支持.

尽管流形学习的算法和应用在过去的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.这同时也意味着流形学习具有广阔的发展空间,既是机遇又是挑战,会促进新思路的产生.

## 参考文献:

- [1] HYVRINEN A. Survey on independent component analysis[J]. Neural Computing Surveys, 1999, 2(4): 94 - 128.
- [2] TURK M, PENTLAND A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71 - 86.
- [3] GONZALEZ R C, WOODS R E. Digital image processing: 2nd ed[M]. Beijing: Publishing House of Electronics Industry, 2003.
- [4] SEUNG H S, LEE D D. The manifold ways of perception[J]. Science, 2000, 290(5500): 2268 - 2269.
- [5] TENENBAUM J, SILVA D D, LANGFORD J. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000, 290(5500): 2319 - 2323.
- [6] ROWEIS S, SAUL L. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323 - 2326.
- [7] 王守觉. 仿生模式识别(拓扑模式识别)——一种模式识别新模型的理论与应用[J]. 电子学报, 2002, 30(10): 1417 - 1420.  
WANG Shoujue. Bionic (Topological) pattern recognition—a new model of pattern recognition theory and its applications[J]. Acta Electronica Sinica, 2002, 30(10): 1417 - 1420.
- [8] 张军平, 曹存根. 神经网络及其应用[M]. 北京: 清华大学出版社, 2004.  
ZHANG Junping, CAO Cungen. Neural network and applications[M]. Beijing: Tsinghua University Press, 2004.
- [9] ZHANG C S, WANG J, ZHAO N Y, ZHANG D. Reconstruction and analysis of multi-pose face images based on nonlinear dimensionality reduction[J]. Pattern Recognition, 2004, 37(1): 325 - 336.
- [10] 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.
- ZHAN Dechuan, ZHOU Zhihua. Ensemble-based manifold learning for visualization[J]. Journal of Computer Research and Development, 2005, 42 (9): 1533 - 1537.
- [11] 赵连伟, 罗四维, 赵艳敞, 等. 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.
- ZHAO Lianwei, LUO Siwei, ZHAO Yanchang, et al. Study on the low-dimensional embedding and the embedding dimensionality of manifold of high-dimensional data[J]. Journal of Software, 2005, 16 (8): 1423 - 1430.
- [12] 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.
- HE Li, ZHANG Junping, ZHOU Zhiping. Investigating manifold learning algorithms based on magnification factors and principal spread directions[J]. Chinese Journal of Computers, 2005, 28 (12): 2000 - 2009.
- [13] BEL KIN M, NIYOGE P. Laplacian eigenmaps for dimensionality reduction and data representation [J]. Neural Computation, 2003, 15 (6): 1373 - 1396.
- [14] ZHANG Z Y, ZHA H Y. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment [J]. SIAM Journal of Scientific Computing, 2005, 26 (1): 313 - 338.
- [15] SIMARD P Y, LECUN Y A, DENKER J S. Efficient pattern recognition using a new transformation distance [J]. Advances in Neural Information Processing Systems, 1993 (5): 50 - 58.
- [16] ZHANG J P, LI S Z, WANG J. Intelligent Multimedia Processing with Soft Computing [C]. Heidelberg: Springer-Verlag, 2004.
- [17] HASTIE T, STUETZLE W. Principia Curves [J]. Journal of the American Statistical Association, 1988, 84 (406): 502 - 516.
- [18] K E G L B, KRZYSAK A, LINDER T, ZEGER K. Learning and design of principal curves [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22 (3): 281 - 297.
- [19] BISHOP C M, SEVENSEN M, WILLIAMS C K I. GTM: The generative topographic mapping[J]. Neural Computation, 1998, 10 (1): 215 - 234.
- [20] CHANG K, GHOSH J. A unified model for probabilistic principal surfaces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23 (1): 22 - 41.
- [21] SMOLA A J, MIKA S. Regularized principal manifolds[A]. In Computational Learning Theory: 4th European Conference [C]. New York: Springer-Verlag, 1999.
- [22] HINTON G, ROWEIS S. Stochastic neighbor embedding [A]. Neural Information Proceeding Systems: Natural and Synthetic[C]. Vancouver, Canada, 2002.
- [23] BRAND M, MERL. Charting a manifold[A]. Neural Information Proceeding Systems: Natural and Synthetic [C]. Vancouver, Canada, 2002.
- [24] BORG I, GROENEN P. Modern multidimensional scaling: theory and application [M]. New York: Springer-Verlag, 1997.
- [25] MIN W L, LU L, HE X F. Locality pursuit embedding[J]. Pattern Recognition, 2004, 37 (4): 781 - 788.
- [26] 杨 剑, 李伏欣, 王 珏. 一种改进的局部切空间排列算法[J]. 软件学报, 2005, 16 (9): 1584 - 1589.
- YANG Jian, LI Fuxin, WANG Jue. A better scaled local tangent space alignment algorithm[J]. Journal of Software, 2005, 16 (9): 1584 - 1589.
- [27] 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.
- WENG Shifeng, ZHANG Changshui, ZHANG Xuegong. Nonlinear dimensionality reduction in the analysis of high dimensional medical data[J]. Journal of Tsinghua University (Sci & Tech), 2004, 44 (4): 485 - 488.

#### 作者简介:



徐 蓉,女,1982年生,哈尔滨工业大学在读硕士研究生,主要研究方向为模式识别、机器学习、手语识别。

E-mail: sinuozhimeng @163.com



姜 峰,在职博士研究生,讲师。主要研究方向为模式识别、机器学习、自然人机交互技术、多媒体技术、数字版权保护等。



姚鸿勋,教授,博导。主要研究方向为自然人机交互技术、多媒体技术、图像处理及模式识别、信息隐藏与检测、数字版权保护、生物特征识别技术等。已发表学术论文 60 余篇,其中被 SCI、EI、ISTP 检索收录 30 余篇,另获发明专利 1 项。