DOI: 10.3901/JME.2013.11.079

一种基于多流形局部线性嵌入算法的故障诊断方法*

谢小欣 胡建中 许飞云 贾民平
(东南大学机械工程学院 南京 211189)

摘要: 故障类别的多样性往往导致原始样本数据在特征空间中呈间断性分布, 针对传统 k 近邻的邻域构建方法难以保证数据集几何结构完整性的问题, 提出一种新的非线性最小二乘约束-局部线性嵌入算法。通过非负线性最小二乘约束搜索边界点, 并结合第一主成分直线寻找其邻域样本点, 重新构造关于边界点的邻域图, 用经典的局部线性嵌入算法机理发现数据内在分布和几何结构, 根据得到的低维嵌入采用 KNN 分类器进行故障模式识别; 仿真数据分析与试验验证结果表明该算法高度保持了原始数据的几何拓扑结构, 增强了低维嵌入的有效性, 提高了故障识别精度。

关键词: 局部线性嵌入 多流形 非负线性最小二乘 故障诊断

中图分类号: TH17

A Fault Diagnosis Method Using Multi-manifold Learning Based on Locally Linear Embedding

XIE Xiaoxin HU Jianzhong XU Feiyun JIA Minping
(School of Mechanical Engineering, Southeast University, Nanjing 211189)

Abstract: The original data sets are disjunctively distributed in the feature space frequently because of the diversity of faults; the completeness of the geometric structure cannot be keep up by the classical neighborhood structure with k -nearest neighbor. A new approach nonnegative linear least squares locally linear embedding (NLLS-LLE) is presented to solve this problem. The boundary points are searched with the nonnegative linear least squares constraints, the neighborhoods of boundary points are determined by the first principal straight line. The neighborhood graphs of boundaries are reconstructed, and the intrinsic distribution and geometry structure of data set is discovered by the classical locally linear embedding, and the fault modes are recognized accurately by KNN classifiers where the low dimensional spaces projection used as inputs. The simulation data analysis and experiment show that the proposed algorithm highly keeps geometry and topology of original data sets, enhances the validity of low dimensional spaces, and improves the classification correctness of fault.

Key words: Locally linear embedding Multi-manifold Nonnegative linear least squares Fault diagnosis

0 前言

故障特征提取过程中, 原始数据具有高维数、非线性等特性, 剔除无用冗余数据, 提取有效信息, 以便于后续的模式识别, 这一过程称为“特征提取”^[1]。将流形学习方法引入故障诊断领域是近几年来研究的热点且取得显著的成果。如蒋全胜等人将流形学习方法应用于旋转机械的故障识别^[2], 杨庆等^[3]提出增量式局部切空间排列算法成功应用于滚动轴承的故障诊断。WANG 等^[4]提出基于核的

LDA 算法并用于故障模式识别。局部线性嵌入 (Locally linear embedding, LLE) 算法^[5]具有处理高维度、非线性、大批量数据的能力^[6]。在故障诊断过程中, 由于故障类别的多样性, 原始数据集在空间中呈间断性分布, 用 k 近邻构建邻域图时, 若数据集类间距较大, 各流形未能相互连接, 则导致得到的低维投影不能保持原始数据几何结构的完整性^[7]。针对此类多流形的数据集, WILLIAM 等^[8]通过求被间断流形间的最短欧氏距离对应的数据点并将这些数据点设置为互为邻域点, 然后用经典 ISOMAP 算法找到嵌入结果; 马瑞等^[9]首先对数据集进行聚类, 然后分析每一类数据的低维流形的本质维数以及流形空间的构成。这两种方法在构建邻

* 国家自然科学基金资助项目(51075069、51075070)。20120621 收到初稿, 20130116 收到修改稿

域图时都引入了无监督聚类,调整数据间的欧式距离,从而破坏了数据的拓扑结构,而且参数较多,若选择不当,则难以得到理想效果。PAN等^[10]将cam加权分布用于LLE算法中进行邻域选取,有效地解决了多流形连接问题,但是对于相距较长的两个流形部分,则无法连接,得不到连通的邻域图。现实生活中高维数据中包含了多个子流形,如果直接用LLE对多流形数据降维不能得到连通的邻域图,导致降维的失效。

本文提出一种新的非线性最小二乘约束的LIE(Nonnegative linear least squares-LIE, NLLS-LIE)算法,构建邻域图时,将部分的处于不同流形中的样本点互为邻域点,使邻域图连通,然后再用LLE算法对数据降维得到低维特征矢量,以提高LLE流行学习方法对此类多流形高维数据降维的有效性和正确性。试验结果表明,本文提出的算法能够有效地挖掘高维数据的几何拓扑结构,提高了故障识别率。

1 局部线性嵌入法

局部线性嵌入算法是由ROWEIS等^[4]提出的一种流形学习方法,通过对数据进行非线性维数约简,能够使数据在低维空间中保持原有的几何拓扑结构。

LLE算法在选取邻域时,由于数据分布的不确定性,基本选用 k 近邻构建数据集的邻域图。一般情况下,对于单一流形的数据集,该方法能够得到较好的邻域图^[5]。而对于有两个或两个以上流形、在空间中呈间断性分布的数据集,则得不到连通的邻域图,破坏了低维投影数据的几何结构。如图1所示。

对于此类多流形的数据集,在用 k 近邻构建邻域图时,整个数据集分为两个部分,构成两个子邻域图,如图1b所示。这是因为在LLE算法中,邻域值 k 的合适与否,直接关系到降维结果的正确与否;若 k 取值较大,重构权的计算复杂度为 $O(k^3N)$,则耗时巨大,且有时亦引入“短路边”,造成局部扭曲^[11];若 k 取值较小,则得不到单一的连通邻域图,导致低维投影失去数据集几何结构的完整性,如图1c所示。因此本文针对如何在 k 取值适当的情况下仍能够得到单一连通的邻域图这一问题提出新的方法。

2 非负线性最小二乘约束LLE算法

针对LLE算法的局限性,本文提出了一种新的非负线性最小二乘约束的LLE算法(NLLS-LLE),该算法引入非负线性最小二乘问题找到边界点,并

(a) 三维空间中的多流形“S”曲面

(b) 用 k 邻近构建的邻域图

(c) 采用LLE算法($k=8, d=2$)得到低维投影

图1 多流形数据LLE算法分析投影图

结合第一主成分直线重新构建其邻域,最后用经典的LLE算法得到数据的内在结构。具体描述:假设给定的输入数据为 $(x_1, x_2, \dots, x_N)$,其中 $x_i \in \mathbf{R}^D$,相应的低维嵌入输出数据集为 $(y_1, y_2, \dots, y_N)$, $y_i \in \mathbf{R}^d$ (一般 $d \ll D$),算法步骤如下。

(1) 构造邻域图:选取邻域参数 k ,本文利用降维后测地线距离与估计的测地线距离之间的残差来选择最优参数 k ^[12]。求出每个样本点 x_i 对应的邻域点集并将这些邻域样本点按到 x_i 的距欧氏距离的由近及远依次排列,得到邻域集 $\{G(x_i) | (x_{i1}, x_{i2}, \dots, x_{ij})\}$,其中 $j=1, 2, \dots, k$ 。

(2) 搜寻边界点:对于每个样本点 x_i ,用其邻域样本 $x_{i1}, x_{i2}, \dots, x_{ij}$ 去线性逼近、构造,计算每个邻域样本的权值 w_{ij} ,最终得到所有样本的重构权值矩阵 $\mathbf{w}$ 。在样本的重构过程中遵循非负线性约束最小

二乘算法^[12]，其公式如下所示

$$E(x_i) = \left| x_i - \sum_{j=1, j \neq i}^k w_{ij} x_{ij} \right|^2$$

$$\text{s.t.} \sum_{j=1, j \neq i}^k w_{ij} = 1 \quad w_{ij} \geq 0 \quad (1)$$

(3) 得到每个样本点的逼近误差 $E(x_i)$ 。首先假设边界点初始集合为 $S = \{\emptyset\}$ ，从第一个样本开始判断其逼近误差是否为 0，如果不是 0，则表示其为边界点，将其加入边界点集合 S ，依次判断后面的样本点是否为边界点，如果是，则加入边界点集合，最终得到包含所有边界点的集合 $S = \{x_i, x_j, \dots, x_l, 1 \leq i < j \leq N\}$ 。

(4) 对集合 S 中每个边界点 x_i ，重新构造其邻域图 $G'(x_i)$ ：① 构造出每个边界点 x_i 邻域样本 $x_{i1}, x_{i2}, \dots, x_{ij}$ 的第一主成分直线 H_i ^[13]，并得到边界点 x_i 在直线 H_i 上的投影 x'_i 以及方向矢量 $v'_i = x_i - x'_i$ ；② 计算所有样本点中除边界点 x_i 以及其所对应的初始 k 邻域样本点之外的样本点 $(x_1, x_2, \dots, x_{N-k-1})$ 与边界点 x_i 的矢量差得到矩阵 $v = (v_1, v_2, \dots, v_{N-k-1})$ ，其中 $v_i \in \mathbf{R}^D$ ，再计算矢量 v'_i 与矩阵 v 中每个矢量 v_i 的点积，得到点积矢量 $\text{dot}(i) = [\text{dot}(i_1), \text{dot}(i_2), \dots, \text{dot}(i_j), j = N - k - 1]$ ；③ 提取点积矢量中 $\text{dot}(ij) > 0$ 所对应的样本点 x_j 放入边界点 x_i 的备选邻域集 $G(x_i)$ 中，在邻域集 $G(x_i)$ 中找出离边界点最近的 $k/2$ 个新样本代替边界点初始邻域样本 $\{G(x_i) | (x_{i1}, x_{i2}, \dots, x_{ij})\}$ 中最远的 $k/2$ 个样本点；如果备选邻域集 Gx_i 中样本的个数 $n < k/2$ ，则用 Gx_i 中所有的样本代替初始邻域样本 $\{G(x_i) | (x_{i1}, x_{i2}, \dots, x_{ij})\}$ 中最远的 n 个样本，最终得到新的邻域图 $G'(x_i)$ ；④ 重复步骤③，直至将所有边界点的邻域图更新完成，得到新的邻域图 $G'(X)$ 。

(5) 由每个样本点的近邻点计算出该样本点的局部重建权值矩阵 w'_{ij} ，使得

$$\Phi(w') = \sum_{i=1}^n \left| x_i - \sum_{j=1, j \neq i}^k w'_{ij} x_{ij} \right|^2 \quad (2)$$

(6) 保持权值不变，在低维空间中重构原数据点 x_i 在低维空间的像 y_i ，使得低维空间重构误差最小

$$\Psi(w') = \sum_{i=1}^n \left| y_i - \sum_{j=1, j \neq i}^k w'_{ij} y_{ij} \right|^2 \quad (3)$$

且要满足两个条件

$$\sum_{i=1}^n y_i = 0 \quad \frac{1}{n} \sum_{i=1}^n y_i y_i^T = I \quad (4)$$

式中， I 是单位矩阵。

这一优化问题可以通过对矩阵 $M = (I - w)^T (I - w)$ 的特征值分解得到。要使损失函数值达到最小，则取 M 的最小 d 个非零特征值所

对应的特征矢量得到 x 在低维嵌入空间的坐标 y 。在处理过程中，将 M 的特征值从小到大排列，通常取第 $2, 3, \dots, d+1$ 间的特征值所对应的特征矢量作为输出结果(第一个特征值几乎等于零，舍去)，得到原始数据在低维空间中的投影 $(y_1, y_2, \dots, y_N)$ 。

NLLS-LLE 算法与经典的 LLE 算法有着相同的可调参数，即邻域参数 k 和嵌入维数 d ，算法参数较少，保证了该算法的健壮性。

数据仿真：本文构造了两种多流形数据：双峰高斯数据和双 S-curve 数据集以及用 NLLE-LLE 算法得到低维投影，其在空间中的分布如图 2 第一列所示。传统的 LLE 算法在构建邻域时未能得到单一连通的邻域图(图 1b)，将数据集划分为两个间断的部分，构成两个子邻域图，破坏了原始数据集几何结构的完整性，导致低维空间投影的失败；而本文算法在构建邻域图时采用非负线性最小二乘约束将处于不同流形的部分边界点互设为邻域点(图 2b、图 2d)，保证在 k 取值适当的情况下仍然能够得到完整、单一的邻域连通图，而避免了 k 取值过大引入“短路边”造成局部扭曲的后果。

NLLS-LLE 算法与经典的 LLE 算法之间的区别与联系：① 两种算法的可调参数都一样，即邻域参数 k 和低维嵌入空间 d ；② 如果选择的邻域参数在整个数据集上形成一个邻域图，那么此时 NLLE-LLE 算法与 LLE 算法基本等效，当数据集在空间中呈间断性分布时，即体现出了本文算法的优越性；因此可以说经典的 LLE 算法是 NLLE-LLE 算法的一个特例，而 NLLE-LLE 算法是经典 LLE 算法的扩充。

(a) 三维空间中的双峰高斯数据

(b) NLLE-LLE 算法对双峰高斯数据构造的邻域图

(c) NLLE-LLE 算法对双峰高斯数据降维得到的低维投影 ($k=8, d=2$)

(d) NLLE-LLE 算法对原始双 S-curve 数据构造的邻域图

(e) NLLE-LLE 算法对 S-curve 数据降维得到的低维投影 ($k=8, d=2$)

图 2 多流形数据 NLLE-LLE 算法分析投影图

仿真试验表明该算法在构建邻域图时保持数据集的拓扑结构,得到的低维投影能够反映原始高维数据内在几何结构(图 2c、图 2e)。

3 工程实例分析

以某化纤化工厂高速离心压缩机中转子振动监测采样样本为例。数据集包含了三种故障状态类别:油膜涡动、轴承松动、不平衡。依据旋转机械故障诊断的特点,选择 9 个特征频段($0.01 \sim 0.39f$, $0.40 \sim 0.49f$, $0.50f$, $0.51 \sim 0.99f$, $1.00f$, $2.00f$, $3.00 \sim 2.00f$, 奇数 f , $>5.00f$)的振幅作为原始特征空间,故

障样本的特征维数为 9,油膜涡动故障样本数为 86 个,轴承松动故障样本数为 76 个,不平衡故障样本数为 148 个,数据集呈现较强的非线性特征。

分别采用 LLE 算法和 NLLS-LLE 算法得到低维投影(本文分别选取 $d=2,3,4$ 进行比较分析,邻域大小设为 8),进行信号的特征提取,然后从低维特征信号中抽取每一类故障类别的 $2/3$ 数据作为训练数据,余下的 $1/3$ 作为测试数据。最后采用最近邻分类器(K -nearest neighbors classifier, KNNC)^[15]对测试样本进行类别划分,本文选取邻近 $K=5$ 。

(a) LLE 算法得到的低维投影

(b) NLLS-LLE 算法得到的低维投影

• 不平衡Train样本 • 油膜涡动Train样本
• 轴承松动Train样本 • 所有Tset样本

图 3 两种方法所得到的低维投影对比

嵌入维数 d 比较小时,用坐标图可以直观看到将数据的分布情况,达到可视化效果,对比图 3 a 与图 3b 可以看出,图 3b 的低维嵌入更具数据原始分布,表明 NLLS-LLE 算法得到的低维嵌入的分布基本可以保持原始数据在高维空间中的内在几何结构,相对于 LLE 算法而言,得到的低维嵌入数据具有更高的结构准确性。

利用 NLLS-LLE 算法将高维特征矢量约简为低维特征矢量,并输入到最近邻分类器中进行故障识别。下表是经特征约简后的故障诊断结果。

表 两种算法特征约简后各状态故障识别结果

方法	不平衡识别精度 (%)	油膜涡动识别精度 (%)	轴承松动识别精度 (%)	平均识别精度 (%)
$d=2$	LLE 算法	94.00	82.14	88.46
	NLLS-LLE 算法	98.00	92.86	92.31
$d=3$	LLE 算法	94.00	92.86	84.62
	NLLS-LLE 算法	88.00	100.00	96.15
$d=4$	LLE 算法	96.00	82.14	92.31
	NLLS-LLE 算法	96.00	100.00	100.00

对比表中关于三种故障的识别精度可知:结合 NLLS-LLE 算法与 KNNC 分类器的故障识别模型识别精度高于传统的 LLE 算法,说明本文提出的算法

得到的低维嵌入有效性优于传统的 LLE 算法, 提高了故障识别精度; 特别的当 $d=4$ 时, NLLS-LLE 算法的识别精度达到 98.08%, 充分表明了该算法的有效性。

4 结论

(1) 提出了 NLLS-LLE 算法, 引入非负线性最小二乘约束找到各部分数据的边界点, 结合第一主成分直线得到每个边界点的邻域样本点, 重新构造的邻域图, 再用经典的 LLE 算法得到低维嵌入。关于双峰高斯数据和双 S-curve 数据的仿真结果表明由此方法得到的低维嵌入投影呈光滑分布, 基本上保持了数据的拓扑结构, 没有导致数据集的高度扭曲, 在低维空间中达到了可视化效果。

(2) 分别采用 NLLS-LLE 算法和 LLE 算法对转子故障数据进行分析对比, 结合 KNN 分类器随机选取一部分低维投影数据作为训练数据, 余下的用做测试, 识别结果表明 NLLS-LLE 算法相对于经典的 LLE 算法具有更优的分类效果。

参 考 文 献

- [1] 钟秉林, 黄仁, 贾民平, 等. 机械故障诊断学[M]. 3 版. 北京: 机械工业出版社, 2007.
ZHONG Binglin, HUANG Ren, JIA Minping, et al. Introduction to machine fault diagnosis[M]. 3rd ed. Beijing: China Machine Press, 2007.
- [2] JIANG Quansheng. Modified Laplacian eigenmap method for fault diagnosis[J]. Chinese Journal Mechanical Engineering, 2008, 21(3): 90-93.
- [3] 杨庆, 陈桂明, 童兴民, 等. 增量式局部切空间排列算法在滚动轴承故障诊断中的应用[J]. 机械工程学报, 2012, 48(5): 81-86.
YANG Qing, CHEN Guiming, TONG Xingmin, et al. Application of incremental local tangent space alignment algorithm to rolling bearings fault diagnosis[J]. Journal of Mechanical Engineering, 2012, 48(5): 81-86.
- [4] WANG Guangbin, HUANG Liangpei. Kernel orthogonal local fisher discrimination for rotor fault diagnosis[C]// International Conference on Image Processing and Pattern Recognition in Industrial Engineering. USA: SPIE, 2010.
- [5] ROWEIS S, SAUL L. Nonlinear dimensionality analysis by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326
- [6] XIANG Shiming, NIE Feiping. Regression reformulations of LLE and LTSA with locally linear transformation[J]. IEEE Trans Actions on Systems, 2011, 41(5): 1250-1262.
- [7] LIU Xiangyang, LU Hongtao, LI Wenbin. Multi-manifold modeling for head pose estimation[C]// The Fundamental Research Funds for the Central Universities, Proceedings of 2010 IEEE 17th International Conference on Image Processing, 2010: 3277-3280.
- [8] WILLIAM K A, CHEN Guangliang, MAURO M. Multi-scale geometric methods for data sets II: Geometric multi-resolution analysis[J]. Applied and Computational Harmonic Analysis, 2012, 32: 435-462.
- [9] 马瑞, 王家, 宋亦旭. 基于局部线性嵌入(LLE)非线性降维的多流形学[J]. 清华大学学报, 2008, 48(4): 583-586.
MA Rui, WANG Jia, SONG Yixu. Multi-manifold learning using locally linear embedding (LLE) nonlinear dimensionality reduction[J]. Journal of Tsinghua University, 2008, 48(4): 583-586.
- [10] PAN Yaozhang, GE Shuzhi, ABDULLAH A M. Weighted locally linear embedding for dimension reduction[J]. Pattern Recognition, 2009, 42(5): 798-811.
- [11] JOSHUA B T, VIN De S, JOHN C L. The Isomap algorithm and topological stability[J]. Science, 2002, 295(4): 5552.
- [12] SAMKO O, MARSHALL A D, ROSIN P L. Selection of the optimal parameter value for the Isomap algorithm[J]. Pattern Recognition Letters, 2006, 27(9): 968-979
- [13] LAW S C, HAN S R. Solving least squares problems[M]. New York: Prentice Hall, 1974.
- [14] TURK M, PENTLAD A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86
- [15] HANE H, GERGE K, VIPLN K, et al. Text categorization using weight adjusted k -nearest neighbor classification [R]. Minnesota: University of Minnesota, 2000.

作者简介: 谢小欣, 男, 1988 年出生。主要研究方向为故障诊断与模式识别。

E-mail: suxiexiaoxin@sina.com

胡建中(通信作者), 男, 1971 年出生, 博士, 副教授, 硕士研究生导师。主要研究方向为故障诊断、人工智能与模式识别。

E-mail: hjz@seu.edu.cn

许飞云, 男, 1969 年出生, 博士, 教授, 博士研究生导师。主要研究方向为智能测试与信号处理、动态系统建模与辨识和状态监测与智能故障诊断。

E-mail: fyxu@seu.edu.cn

贾民平, 男, 1960 年出生, 博士, 教授, 博士研究生导师。主要研究方向为智能诊断理论、动态信号处理和故障机理及其特征提取。

E-mail: mpjia@seu.edu.cn