

一种基于离散 Morse 理论的优化模型及其应用

张建萍, 刘希玉

(山东师范大学 管理科学与工程学院, 济南 250014)

摘要: 根据 Forman 的离散 Morse 理论的特点, 提出一种基于离散 Morse 理论的优化模型. 该模型在 3 维及以上空间点构建离散 Morse 函数进行最优化, 得到了问题的最优解或近似最优解. 同时, 证明了所构建的函数确实是复形上的离散 Morse 函数. 利用 4 个典型的测试函数进行仿真实验, 结果表明了该模型的有效性, 且该模型尤其适用于解决大数据量的优化问题. 从聚类的过程即目标函数的优化过程这一角度考虑, 尝试将优化模型应用于聚类分析. 仿真实验结果表明, 所提出的算法能较好地划分数据点重叠区域的聚类形状, 验证了所提出算法的可行性和有效性.

关键词: 单纯复形; 离散 Morse 函数; 离散梯度向量场; 最优化; 密度聚类

中图分类号: TP273

文献标志码: A

Optimization model based on discrete Morse theory and its application

ZHANG Jian-ping, LIU Xi-yu

(School of Management Science and Engineering, Shandong Normal University, Ji'nan 250014, China.

Correspondent: ZHANG Jian-ping, E-mail: love_fly_20@sina.com)

Abstract: An optimization model based on the discrete Morse theory is presented for the characteristics of Forman's theory, which can obtain an optimal value or approximate optimal one by the algorithm of constructing a discrete Morse function in 3D space and above. The experimental results show that this model is effective for applications and especially for the optimization problem of a large amount of data. Based on an optimization model based on discrete Morse theory, the clustering framework is proposed. The experimental results on synthetic and UCI data sets show that the proposed algorithm can successfully divide data points overlapping into the feature space to many correct clusters, and also show the feasibility and effectiveness of the algorithm.

Key words: simplicial complex; discrete Morse function; discrete gradient vector field; optimization; density-based clustering

0 引言

聚类^[1]就是将数据对象分成多个类或簇, 在一个簇中对象相似度高, 而在不同簇中对象的相似度低. 将空间样本数据进行聚类转换成基于聚类目标函数的优化问题, 由此产生了很多计算智能技术——用于聚类分析的融合算法^[2-5], 如模糊控制、神经网络、进化计算、群体智能、人工生命和 DNA 计算. 基于计算智能的聚类分析模型对处理目标的特性具有良好的适应能力, 弥补了传统聚类算法的缺点及不足, 从而取得了良好的效果. 但是, 随着数据挖掘技术应用领域的不断扩展, 数据挖掘系统通常面对的是更为复杂、任意分布、任意大小和密度的数据集, 因此, 提出一种新的技术应用于聚类分析具有重要意义.

受文献[6-7]中离散 Morse 理论的启发, 本文尝试在 3 维及以上的离散空间 K 对任意给定的 $f: K \rightarrow R$ 进行优化分析, 构造最优离散 Morse 函数, 产生尽可能少的临界单元, 从而得到函数的最优值或接近最优值. 根据上述思想构建一种基于离散 Morse 理论的优化模型, 以解决离散 Morse 理论的优化问题, 并将该优化模型应用于聚类分析, 提出一种基于离散 Morse 优化模型的聚类算法. 该算法采用基于核密度估计的层次聚类的思想, 根据离散 Morse 优化模型得到密度函数的极值, 同时根据构造的离散梯度向量场得到以极值点为聚类中心的数据集的初始划分; 然后, 通过临界单元的抵消算法对初始聚类进行合并, 产生不同层次的划分模式. 实验分别在人工数据集

收稿日期: 2012-04-26; 修回日期: 2013-09-24.

基金项目: 国家自然科学基金项目(60873058, 60743010); 山东省“泰山学者”建设工程专项项目.

作者简介: 张建萍(1979-), 女, 讲师, 博士, 从事数据挖掘、遗传算法的研究; 刘希玉(1964-), 男, 教授, 博士生导师, 从事数据挖掘、计算智能等研究.

和 UCI 数据库中的 Haberman's Survival 数据集上进行. 理论分析和仿真实验结果均表明, 该算法能够发现任意形状、大小和密度的聚类, 能较好地划分数据点重叠区域的聚类形状, 并验证了所提出算法的可行性和有效性.

1 离散 Morse 理论优化模型

受离散 Morse 理论中离散梯度向量场这一特点的启发, 提出一种离散结构的优化模型. 单纯复形是一种很重要的研究离散结构拓扑属性的工具.

首先, 给出本文所用到的相关符号表示及术语. 有限数据集为 $X = \{x_1, x_2, \dots, x_n\} \subset R^n$; 数据集的映射函数 $h: X \rightarrow R^1$; 离散曲面为 $S = \{(x_i, h(x_i)) | i = 1, 2, \dots, n\} \subset R^{n+1}$; 每个几何无关点组 $\sigma = (a_0, a_1, \dots, a_q)$ 决定的最小凸集称为 E^n 中的一个 q 维单形, 记为 $\sigma^{(q)}$, $\sigma^{(q)}$ 的连接点 $a_i \notin \sigma^{(q)}$ 形成的凸集即为 $q+1$ 维单形 $\sigma^{(q+1)}$; 定义 $\max h(\sigma) = \max_{0 \leq j \leq i} \{h(a_j)\}$.

下面给出顶点 x 的邻域的定义.

定义 1^[8] x 的较小星域 $S(x)$ 由 x 的所有依附面组成, 包括 x 本身且满足 $g(x) = \max g(\alpha)$, 即

$$S(x) = \{\alpha \in K | x \in \alpha, g(x) = \max_{y \in \alpha} g(y)\}.$$

定义 2^[8] x 较小链域 $L(x)$ 由 $S(x)$ 的所有单形中不包含 x 的面组成, 且满足 $g(x) > \max g(v)$, 即

$$L(x) = \{v \in K | v \subseteq \alpha \in S(x), v \cap x = \emptyset\}.$$

1.1 离散结构——单纯复形

曲面的离散化是通过基于凸包的三角剖分实现的, 对曲面的三角剖分形成单纯复形. $\text{hull}(X)$ 表示数据集 $X = \{x_1, x_2, \dots, x_n\} \subset R^n$ 形成的凸包. 对于凸包内的点 $x_i \in X$, 对应于曲面上点 $(x_i, h(x_i))$ 存在其邻域同胚于 R^n ; 对于凸包上的点 $x_i \in X$, 对应于曲面上的点 $(x_i, h(x_i))$ 存在其邻域同胚于 R^n 的半球, 故曲面的三角剖分是一个有边界的 n 维流形. 对于曲面的三角剖分, 分两步得到其单纯复形. 首先, 在数据集 $X = \{x_1, x_2, \dots, x_n\} \subset R^n$ 上进行 Delaunay 三角剖分. 一个 $q+1$ 维单形是通过 q 维单形与点 $x_i \in X$ 连接形成, 且 $q+1$ 维单形的外接球内不存在其他点. 然后, 对每个数据点 x_i 提升 $h(x_i)$, 即可得到曲面的单纯复形. 本文以 F_1 为例随机产生 30 000 个坐标点生成曲面单纯复形(部分), 如图 1 所示. 表 1 为运算实例(30 000 个坐标点)建立单纯复形的时间.

表 1 运算实例(30 000 个坐标点)
建立单纯复形的时间

实例	F_1	F_2	F_3	F_4
CPU/s	0.028	0.032	0.038	0.05

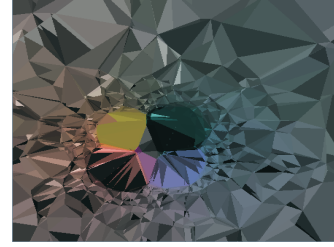


图 1 剖分后生成的曲面单纯复形(部分)

1.2 离散 Morse 理论

定义 3^[6](离散 Morse 函数) 将复形 K 的每个单形映射为一个实数函数 $f: K \rightarrow R$, 称为离散 Morse 函数. 此函数满足: 对于每个单形 $\alpha^{(p)} \in K$, 有

$$\#\{\tau^{(p+1)} > \alpha^{(p)} : f(\tau) \leq f(\alpha)\} \leq 1, \quad (1)$$

$$\#\{v^{(p-1)} < \alpha^{(p)} : f(v) \geq f(\alpha)\} \leq 1. \quad (2)$$

定义 4^[6](临界单形) 设 $f: K \rightarrow R$ 是一个离散 Morse 函数. 若单形 $\alpha^{(p)}$ 是一个临界单形, 则需满足

$$\#\{\tau^{(p+1)} > \alpha^{(p)} : f(\tau) \leq f(\alpha)\} = 0, \quad (3)$$

$$\#\{v^{(p-1)} < \alpha^{(p)} : f(v) \geq f(\alpha)\} = 0. \quad (4)$$

不满足上述条件的单形称为规则单形.

定义 5(离散梯度向量) 离散梯度向量场 V 是复形 K 中单形的有序对 $\langle \alpha^{(p)}, \beta^{(p+1)} \rangle$ 集合, α, β 满足 $\alpha < \beta$ 且 $f(\beta) \leq f(\alpha)$.

性质 1 若离散 Morse 函数 f 是在离散梯度向量场 V 上构建的, 则函数 f 沿着 V -path 方向是下降的.

由性质 1 可以看出, Morse 函数 f 沿着下降流方向是递减的, 这便为 Morse 理论提供了优化基础, 但在一个给定复形上直接构建离散 Morse 函数是一个 NP-hard 问题^[9]. 本文从构建离散梯度向量场的角度出发构建 Morse 函数.

1.3 构造离散梯度向量场算法

本文基于修正的 Hasse 图^[9]中的子图 R_i ($i = 1, 2, \dots, \dim K$) 考虑构造离散梯度向量场. 离散梯度向量场(DVF)算法包含两部分: ConstructDVF(K, g) 和 CancelCriticalCell(K, g, j). 其中: ConstructDVF(K, g) 产生未经调整的 A, B, C 和 r ; CancelCriticalCell(K, g, j) 修正已产生的 A, B, C 和 r .

1) ConstructDVF(K, g).

Step 1: 输入有限单形 K , 映射函数 $h: K_0 \rightarrow R$.

Step 2: 初始化相关参数. 设定 A, B, C 为空; 设定映射 $r: B \rightarrow A$, 其中 $r(\sigma)$ 是 σ 的一个面.

Step 3: 取顶点 $x \in K_0$, 求 x 的较小链域 K' . 若 K' 为空, 则将 x 放入 C 中 (x 为局部最小值); 否则将 x

放入 A 中, 在 K' 上重新定义 $h' : K'_0 \rightarrow R$ 作为 h 的约束条件, 再转 Step 2 重新执行, 直至 K' 上的单形分别存入 A', B', C' .

Step 4: 找到 $w_0 \in C'_0$, 使得 $h'(w_0)$ 最小; 将 $[x, w_0]$ 放入 B 中, 定义 $r([x, w_0]) = x$; 对于每个 $\sigma \in C' - w_0$, 将 $x * \sigma$ 放入 C 中; 对于每个 $\sigma \in B'$, 将 $x * \sigma$ 放入 B , 将 $x * r'(\sigma)$ 放入 A 中, 定义 $r(x * \sigma) = x * r'(\sigma)$.

Step 5: 转 Step 3, 直到 K_0 中的顶点取完为止.

h' 采用下式定义:

$$h'(w) = (h(w) - h(x)) / l([x, w]),$$

其中 $l([x, w])$ 表示边 $[x, w]$ 的欧氏距离.

为了得到尽可能少的临界单元, 产生最优的离散 Morse 函数, 在已产生的离散梯度向量场中, 如发现存在如下路径: 以 $\tau \in C_j$ 为起点, $\sigma \in C_{j-1}$ 为终点, 且梯度路径上的每个单形 β 的 $h(\beta) < \delta$ (若数据集无噪声点, 则通常取 0), 则选择 $\sigma = \max\{\max h(\tau) - \max h(\sigma)\}$, 执行算法 2).

2) CancelCriticalCell(K, g, j).

Step 1: 取 1 个 $\tau \in C_j$.

Step 2: 找到所有梯度路径 $\tau = \tau_{i1} \rightarrow \tau_{i2} \rightarrow \dots \rightarrow \tau_{il} \in C_{j-1}$. 若从 τ 的面到 $\sigma \in C_{j-1}$ 的梯度路径只有一条, 则令 $m_i = \max\{h(\tau_{ki})\}$.

Step 3: m_i 至少被赋值 1 次, 选择 $m_j = \max\{m_i\}$. 若梯度路径上每个单形 β 的 $h(\beta) < \delta$, 则执行 Step 4, 否则执行 Step 1.

Step 4: 找到惟一的梯度路径 $\tau = \tau_1 \rightarrow \sigma_1 \rightarrow \tau_2 \rightarrow \sigma_2 \rightarrow \dots \rightarrow \sigma_j = \sigma \in C_{j-1}$, 其中 $V(\sigma_i) = \tau_{i+1}$, σ_i 是 τ_i 的一个面且 $\sigma_i \neq \sigma_{i+1}$.

Step 5: 从 C 中删除 σ 和 τ , 翻转 τ 到 σ 的方向, 即 $V(\sigma_i) = \tau_i$.

Step 6: 重复 Step 1, 直到取完为止.

1.4 构造离散 Morse 函数

文献 [10] 提出了一种将顶点函数 h 扩展成离散 Morse 函数 f 的算法, 并给出了证明.

定理 1^[10] 给定 $\varepsilon > 0$, 构造的 Morse 函数 f 满足 $|f(\tau) - \max h(\tau)| \leq \varepsilon$, 其中 $\max h(\tau) = \max\{h(x) : x \text{ 是 } \tau \text{ 的顶点}\}$.

由于 $K = S(x_1) \cup S(x_2) \cup \dots \cup S(x_n)$, 考虑在单个 $S(x_i)$ 上构造离散 Morse 函数后扩展到整个 K 区域. 若 $r(\delta) = x_i$, 则认为 δ 在 x_i 之前, 即 $f(\delta) \leq f(x_i)$, 由此根据进入 A, B 或 C 的顺序得到 $S(x_i)$ 所

有单元的一个序列 $\delta, x_i, \alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}$. 受定理 1 启发, 结合本文构造离散梯度向量场的方法, 下面给出具体 Morse 函数的定义: 给定 $\varepsilon > 0$, 有

$$\begin{cases} f(\delta) = h(x_i) - \varepsilon, \\ f(x_i) = h(x_i), \\ f(\alpha_{ij}) = h(x_i) + j\varepsilon, \end{cases} \quad (5)$$

其中 $j > 0$, 且 j 尽可能的小.

1) 若 $S(x_i)$ 含单元数 $k = 1$, 则只有顶点 x_i, x_i 存入 C 中, 此时 $f(x_i) = h(x_i)$ 成立.

2) 若 $S(x_i)$ 含单元数 $k = 2$, 则会出现配对 $r(\delta) = x_i, \delta$ 存入 B 中, x_i 存入 A 中, 此时 $(f(\delta) = h(x_i) - \varepsilon) < (f(x_i) = h(x_i))$ 成立.

3) 若 $S(x_i)$ 含单元数 $k = n (n > 2)$, 且规则单元 $(\alpha^{(p)}, \beta^{(p+1)}) \in V$, 则 $f(\beta) < f(\alpha)$, 显然满足 f , 由算法 1 可知, β 的其他面之前已存入 r 或 C 中, α 是 β 的唯一自由面. 若 $\gamma \in C$, 则表明 γ 的所有面之前已存入 r 或 C 中, 而其依附面会随后加进来, 因此, Morse 函数产生的临界单元 γ 的条件 f 也得到满足.

定理 2 若 x 是 K 的顶点且 $h(x)$ 最小, 则 DVF 算法会将 x 作为一个临界单元输出.

临界顶点是 h 取得局部最小值的顶点, 但不是每个局部最小值顶点都是临界顶点. 因为如果取得局部最小值的顶点 w 与 h 的鞍点^[11]相连接, 则 w 可能不会产生临界顶点. 同样, 临界 2 维单形 e 与 h 的局部最大值 m 相邻接, 但若 m 与 h 的鞍点相连接, 则 e 可能不会是临界单元.

2 基于离散 Morse 的密度聚类算法

在密度聚类算法^[12]的总体框架下, 本文采用离散 Morse 理论作为工具进行聚类. 这是一种新颖的方法. 基于 Morse 理论的聚类算法是一种基于图论的聚类方法^[13]. 每棵树代表一个簇, 树的根节点代表簇的中心点, 叶子节点表示位于低密度区的数据点. 聚类产生由多棵树构成的森林. 通过离散梯度路径进行聚类的过程就是有向树的形成过程. 有向树的形成是通过查找节点的父节点来完成. 由离散梯度向量场的构造方法可知, 查找节点 $x \in X$ 的父节点是以 x 为起点沿着下降最快方向路径得到的点 y . 利用该方法, 数据点很快就划分到相应的簇中, 每个簇就是一棵有向树, 每个簇是由离散梯度向量场构成的.

基于离散 Morse 理论的聚类算法总体分为两步进行: 1) 构造数据集曲面的三角剖分; 2) 在曲面构成的单纯复形上进行聚类. 下面给出基于离散 Morse 理

论的密度聚类算法的总体框架及步骤.

Step 1: 输入数据集 $X = \{x_1, x_2, \dots, x_n\}$;

Step 2: 对每个点 x_i , 计算其概率密度 $f(x_i), i = 1, 2, \dots, n$;

Step 3: 根据 1.1 节的方法构造曲面的单纯复形 K ;

Step 4: 根据算法 1 得到初始聚类划分;

Step 5: 利用层次聚类^[14]的思想, 根据算法 2 对满足条件的聚类进行合并, 得到最终聚类结果.

本文中的所有计算都是在密度函数生成的离散曲面上完成, 因此最终聚类的结果要还原到数据点的聚类, 通过水平集确定数据点所属类别. 算法框架如图 2 所示.

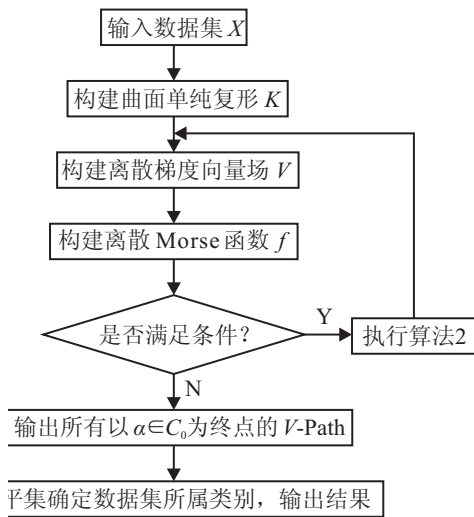


图 2 聚类算法流程

3 实验结果及分析

3.1 离散 Morse 优化模型实验及分析

下面对本文所提出的算法进行实验验证和分析. 算法由 Visual c++ 2005 实现. 实验采用表 1 中的 4 个标准测试函数^[15]进行一系列实验. 这 4 个测试函数具有不同的特点, 可充分测试算法对不同类型问题的优化性能.

表 2 标准测试函数

测试函数	x_i
$F_1 = \sum_{i=1}^n x_i^2$	$[-100, 100]$
$F_2 = \sum_{i=1}^{n-1} (100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	$[-2.048, 2.048]$
$F_3 = \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i) + 10)$	$[-5.12, 5.12]$
$F_4 = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	$[-600, 600]$

F_1 和 F_2 是连续的单峰函数, 其中 F_2 是一个经典

的复杂优化问题, 取值区间内走势平坦, 收敛到全局最优点的机会微乎其微; F_3 和 F_4 是复杂的非线性多峰函数, 存在大量局部极值, 可检验算法的全局搜索和逃离局部极值的能力.

实验中测试函数维数 $\dim = 3$, 数据规模 $\text{size} = 30\,000$, 所有数据均为测试函数在其取值范围内随机产生. 表 3 详细记录了算法循环 100 次的结果.

表 3 优化模型的实验结果

测试函数	最优 Morse 函数	平均寻优结果	与最优值差值
F_1	(1, 0, 0, 0)	1e-261	0
F_2	(2, 5, 2, 0)	4.22e+01	-4.22e+01
F_3	(1, 1, 0, 0)	1.01e+01	-1.01e+01
F_4	(1, 0, 0, 0)	7e-02	-4e-02

通过将表 3 中的最优值与文献 [15] 中的最优值进行比较可以看出, 本文构建的离散 Morse 函数能够取得或接近取得函数的最优值. 由此可知, 通过构建最优离散 Morse 函数进行函数优化的思路可行, 且能取得较好的效果.

对于单峰函数, 利用 DVF 算法所得最优解的收敛精度较高. 对于多峰函数, 此算法有可能会陷入局部最优, 但可以尝试通过调整参数的方法尽可能避免收敛于某一局部极值点. 通过大量的模拟实验发现, 利用该方法进行优化问题求解结果的准确性很大程度上取决于问题空间的取值类型. 取值类型是整型的求解结果效果要好于浮点型的求解效果, 这是因为在顶点的低值链域上 h' 定义为两个顶点坐标的欧氏距离所致.

3.2 基于离散 Morse 优化模型聚类算法实验及分析

为便于聚类结果的比较以及结果的可视化, 图形分别在 Matlab 7.0 及 Geomview 中显示. 算法采用人工数据集 Dataset1 及 UCI 数据库中 Haberman's Survival 数据集进行仿真实验.

3.2.1 参数设置

在基于离散 Morse 理论的密度聚类算法中需要设置的参数为窗宽 σ . 由于引进层次聚类中的合并思想, σ 的选取降低了对全局密度函数结果的影响. 在所有的实验中设定 $\sigma = 0.5 \sim 2$.

3.2.2 人工数据集实验结果及分析

Dataset1 包含 300 个数据点, 数据分布呈现 2 个不规则形状的聚类, 如图 3(a) 所示. 在图 3(b) 上采用基于离散 Morse 理论的聚类算法, σ 在 [0.1, 1] 范围

内取值, 对样本集进行 10 次实验, 每次都能得到图 3(c) 的聚类结果. 所用的整个处理时间是 1.3 s. 图 3 中, “+”属于一个簇, “·”属于另一个簇.

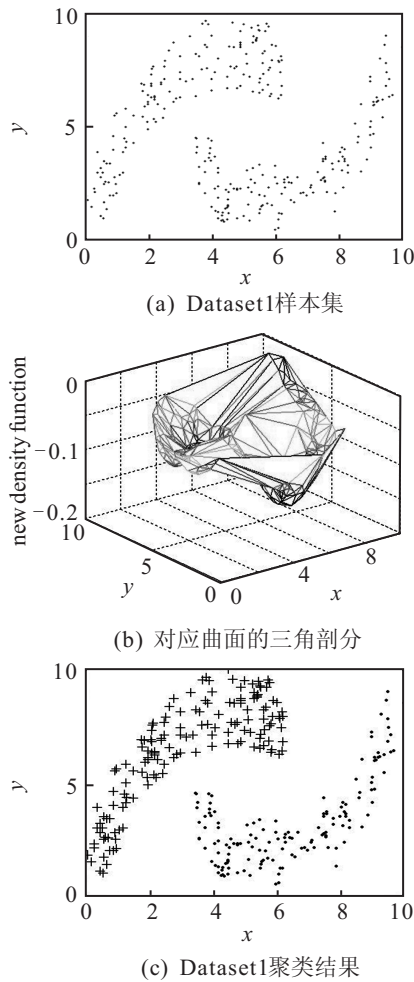


图 3 本文聚类算法对 Dataset1 进行聚类

3.2.3 Haberman's Survival 数据集实验结果及分析

Haberman's Survival 是一个 3 维数据集, 包含 306 个数据点, 数据点分为 2 类, 每个样本点由 3 个属性组成. 数据分布呈现不规则形状的聚类, 数据点在其属性空间中交叉重叠. 图像采用 Gemview 显示, 在窗口中包含数据点和大小不同的球, 顶点和球之间通过线连接. 通过线连接的单形是临界单形, 球包着单形的重心, 线连接着它的相关面的重心, 最大的球代表临界顶点.

Haberman's Survival 数据集根据 1.1 节得到对应曲面的 Delaunay 三角剖分, 如图 4(a) 所示. 执行算法 1 后得到的临界单元为 (4, 11, 8, 0): 4 个 0-临界单形, 11 个 1-维单形, 8 个 2-维临界单形, 0 个 3-维临界单形, 如图 4(b) 所示. 图 4(b) 中, 最大球表示 0-维临界单形, 较小球代表 1-维临界单形, 最小球表示 2-维临界单形. 图 4(c) 显示了 0-1 层梯度路径 (灰色线表示), 图 4(d) 显示了 1-2 层梯度路径 (灰色线表示). 执行算法

2 CancelCriticalCell(K, g, j), 消除 1-2 层后临界单元数分别为 (4, 3, 0, 0), 梯度路径 (灰色线表示) 如图 4(e) 所示. 对于 0-1 层临界单元则按照由大到小依次消除. 0-维临界单元的个数即为聚类的簇数, 沿着离散梯度路径达到 $\alpha \in C_0$ 的所有数据点属于以 α 为密度吸引点的簇.

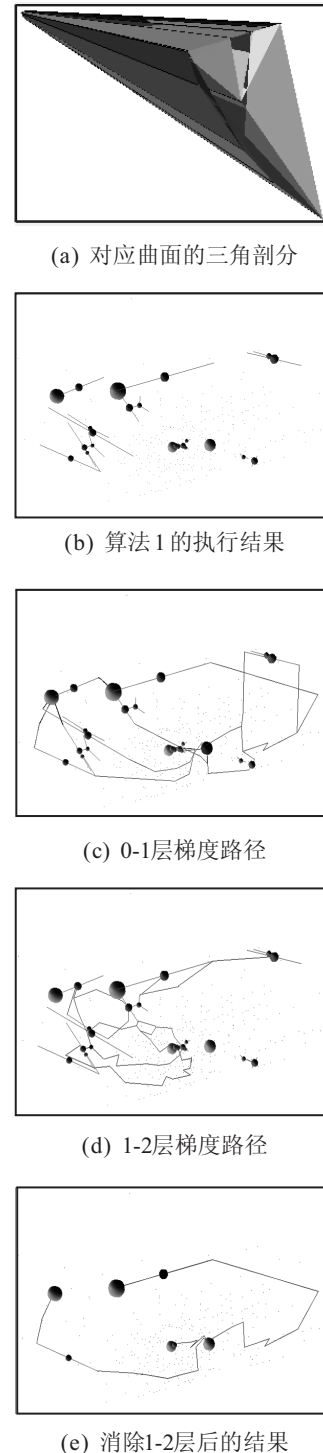


图 4 本文聚类算法对 Haberman's Survival 进行聚类

在 Haberman's Survival 数据集中, 在 [0.5 ~ 2] 范围内改变 σ 的取值, 进行 10 次实验. 由于簇间数据点的交叉重叠, 聚类平均正确率达到 98%, 所用的整个处理时间是 4.3 s.

由以上实验结果可知, 基于离散 Morse 理论的聚类算法能够产生令人满意的聚类结果, 即使在数据点交叉重叠的情况下, 该算法也能够得到正确的聚类。

4 结 论

本文提出了一种构建离散 Morse 函数求最优解的算法, 证明了所构建的函数确实是 K 上的离散 Morse 函数, 并得到了问题的最优解或近似最优解, 这是一个全新的尝试. 同时构建了一种基于离散 Morse 理论的优化模型, 实验结果验证了该模型的有效性. 从聚类的过程就是目标函数的优化过程这一角度考虑, 本文尝试将优化模型应用于聚类分析. 仿真实验结果表明, 该算法能较好划分数据点重叠区域的聚类形状, 并验证了所提出算法的可行性及有效性.

需要说明的是, 在对多峰函数构建最优离散 Morse 函数时, 还需要对参数进行人工调整才能尽可能避免过早陷入局部最优. 另外, 为了提高取得最优值的精度, 对问题的解空间数据类型需进行一定的限制, 这是下一步需要解决的问题.

参考文献(References)

- [1] Han J W, Kamber M. Data mining concepts and techniques[M]. 2nd ed. Francisco: Morgan Kaufmann Publishers, 2006: 383-385.
- [2] Graves D, Pedrycz W. Performance of kernel-based fuzzy clustering[J]. Electronics Letters, 2007, 43(25): 1445-1446.
- [3] Pal N R, Bezdek J C, Tsao E C K. Generalized clustering networks and Kohonen's self-organization[J]. IEEE Trans on Neural Networks, 1993, 4(4): 549-557.
- [4] Babu G P, Murty M N. Clustering with evolution strategies[J]. Pattern Recognition, 1994, 2(27): 321-329.
- [5] Bader J S, Chaudhuri A, Rothberg J M, et al. Gaining confidence in high-throughput protein interaction networks[J]. Nature Biotechnology, 2004, 22(1): 78-85.
- [6] Forman R. Morse theory for cell complexes[J]. Advances in Mathematics, 1998, 134(1): 90-145.
- [7] Forman R. A user's guide to discrete Morse theory[J]. Seminaire Lotharingien de Combinatoire, 2002, 48(2): 1-35.
- [8] Edelsbrunner H, Harer J, Natarajan V, et al. Morse-smale complexes for piecewise linear 3-manifolds[C]. Proc of 19th Annual Symposium on Computational Geometry. San Diego, 2003: 361-370.
- [9] Lewiner T, Lopes H, Tavares G. Toward optimality in discrete morse theory[J]. Experimental Mathematics. 2003, 12(3): 271-286.
- [10] Henry King, Kevin Knudson. Generating discrete Morse function from point data[J]. Experimental Mathematics, 2005, 14(4): 435-444.
- [11] Milnor J W. Morse theory[M]. New Jersey: Princeton University Press, 1963:5-15.
- [12] Tran T N, Wehrens R, Buydens L M C. KNN-kernel density-based clustering for high-dimensional multivariate data[J]. Computational Statistics & Data Analysis, 2006, 51(2): 513-525.
- [13] Fukunaga K. Introduction to statistical pattern recognition[J]. 2nd ed. Boston: Academic Press, 1990: 181-397.
- [14] Wang Xiaochun, Wang Xiali, Mitchell Wilkes D. A divide-and-conquer approach for minimum spanning tree-based clustering[J]. IEEE Trans on Knowledge and Data Engineering, 2009, 21(7): 945-958.
- [15] Huang Han, Qin Hu, Hao Zhifeng, et al. Example-based learning particle swarm optimization for continuous optimization [J]. Information Sciences, 2012, 182(1): 125-138.

(责任编辑: 滕 蓉)