

一个新的填充函数在优化核聚类中的应用

夏俊杨

浙江师范大学, 数学与计算机科学学院, 浙江 金华

收稿日期: 2022 年 12 月 28 日; 录用日期: 2023 年 1 月 21 日; 发布日期: 2023 年 1 月 31 日

摘要

针对约束优化问题, 本文提出一个新的单参数填充函数, 并设计相应算法。然后将此填充函数应用于对聚类问题中的权重优化, 结合聚类模型的过程给出相应算法步骤, 运用 Python 对数据集进行聚类分析并得出权重参数。最后, 实验结果表明该填充函数算法的可行性。

关键词

填充函数, 权重优化, 聚类

A New Filled Function and the Application in Optimizing Kernel Clustering Problem

Junyang Xia

College of Mathematics and Computer Science, Zhejiang Normal University, Jinhua Zhejiang

Received: Dec. 28th, 2022; accepted: Jan. 21st, 2023; published: Jan. 31st, 2023

Abstract

In this paper, we propose a new one parameter filled function for constrained opti-

mization and design the corresponding algorithm. Then this filled function is applied to the weight optimization of clustering problem, the corresponding algorithm steps are given in combination with the process of clustering model, and Python is used to cluster the data set and obtain the weight parameters. Finally, the experimental results show that the filled function algorithm is feasible.

Keywords

Filled Function, Weight Optimization, Clustering

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



1. 引用

填充函数法是求解非线性全局优化问题的一种确定性算法, 由西安交通大学葛仁溥 [1] 教授首先提出, 随后张连生教授等 [2-4] 对填充函数的定义进行了改进, 不少学者又相继提出了双参数填充函数法 [5]、单参数填充函数法 [6, 7] 和无参数填充函数法 [8]. 在文献 [9] 中针对无 Lipschitz 连续而在目标函数满足强制条件时, 提出了一类新的单参数填充函数来解决一般无约束最优化问题. 但在现实中, 大多数的全局优化问题都有约束条件存在并且不一定满足强制性条件. 在本文中避开了强制性这个条件从而使得极小点可以在边界上达到. 随后提出了一个新的单参数填充函数法, 数值实验结果表明该算法是有效可行的, 并通过结合聚类模型的实例验证了该填充函数的实用性.

2. 假设与定义

考虑如下不等式约束的全局优化问题 (P):

$$\min f(x), s.t. \quad x \in S = \{x \in R^n | g_i(x) \leq 0, i = 1, \dots, m\}.. \quad (1)$$

其中 $f(x), g_i(x), i = 1, \dots, m : R^n \rightarrow R$, 是连续可微函数.

设 x_k^* 为问题 (P) 的当前局部极小点, 针对问题 (P) 给出填充函数所满足的定义如下.

定义 1 [6] 函数 $p(x, x_k^*, A)$ 如果满足以下条件, 则称它为 $f(x)$ 在局部极小点 x_k^* 处的填充函数.

(1) 在 R^n 空间中, x_k^* 是 $p(x, x_k^*, A)$ 的一个严格局部极大点;

(2) 对任意的 $x \in S_1 \cap S$ 且 $x \neq x_k^*$ 或 $x \in R^n \setminus S$ 有 $\nabla p(x, x_k^*, A) \neq 0$, 这里 $S_1 = \{x \in R^n | f(x) \geq f(x_k^*), g_i(x) \leq 0, i = 1, \dots, m\}$;

(3) 若 x_k^* 不是全局极小点, 那么 $p(x, x_k^*, A)$ 一定在 $S_2 = \{x \in S | f(x) < f(x_k^*), g_i(x) \leq 0, i = 1, \dots, m\}$ 上有局部极小点.

以上定义的填充函数的意义为: 对于带约束的最优化问题, 首先要考虑的就是所求出来的点是否可行. 由 (2) 知, 要找的点一定在可行域内; 其次, 用填充函数解决全局最优化问题, 目的是找到比当前局部极小点更好的点, 由 (2) 知, 要找的点一定不会再比当前局部极小点高的水平集上得到. 若 x_k^* 是局部极小点但不是全局极小点, 由 (1) 和 (3), 则可以从 x_k^* 的邻域中的任意一点出发, 用求带约束最优化问题的极小化方法极小化该填充函数, 总能找到填充函数的局部极小点 x_0^* , 由 (3) 知, $f(x_0^*) < f(x_k^*)$, 再由 (2), 对 $x \in S_1 \cap S$ 且 $x \neq x_k^*$ 或 $x \in R^n \setminus S$ 有 $\nabla p(x, x_k^*, A) \neq 0$, 得知 $x_0^* \in S_2$.

3. 介绍

基于定义 1, 本文构造如下填充函数:

$$p(x, x_k^*, A) = \frac{e^{A \cdot \varphi(f(x) - f(x_k^*))}}{1 + \|x - x_k^*\|^2} + \frac{1}{A} \{\min [0, \max (f(x) - f(x_k^*), g_i(x), i = 1, \dots, m)]\}. \quad (2)$$

其中

$$\varphi(t) = \{1, t \geq 0, 0, t < 0\}. \quad (3)$$

这里 x_k^* 是问题 (1) 的当前局部极小点, $A > 0$ 是参数. 由以下定理易知函数 $p(x, x_k^*, A)$ 是满足定义 1 的一个填充函数.

定理 1: 设 x_k^* 是 $f(x)$ 的一个局部极小点, 当 $A > 0$ 时, x_k^* 是填充函数 $p(x, x_k^*, A)$ 的一个严格局部极大点.

证明 因为 x_k^* 是 $f(x)$ 的一个局部极小点, 存在它的一个邻域 $N(x_k^*, \delta)$, ($\delta > 0$), 对任意的 $x \in N(x_k^*, \delta) \cap S$ 且 $x \neq x_k^*$ 或 $x \in R^n \setminus S$, 有 $f(x) \geq f(x_k^*)$. 于是 $\varphi(x) = 1$. 则有

$$\min [0, \max (f(x) - f(x_k^*), g_i(x), i = 1, \dots, m)] = 0,$$

于是

$$p(x, x_k^*, A) = \frac{e^{A \cdot \varphi(f(x) - f(x_k^*))}}{1 + \|x - x_k^*\|^2} = \frac{e^A}{1 + \|x - x_k^*\|^2}.$$

由于 $A > 0$, 得到

$$\frac{e^A}{1 + \|x - x_k^*\|^2} < 1 < e^A = P(x_k^*, x_k^*, A).$$

所以对任意的 $x \in (x_k^*, \delta) \cap S, x \neq x_k^*$, 或 $x \in R^n \setminus S$, 有 $P(x, x_k^*, A) < P(x_k^*, x_k^*, A)$, 故填充函数 $p(x, x_k^*, A)$ 的一个严格局部极大点在 x_k^* 处取得.

定理 2: 设 x_k^* 是 $f(x)$ 的一个局部极小点, 对任意的 $A > 0$, $p(x, x_k^*, A)$ 在 S_1 且 $x \neq x_k^*$ 或 $R^n \setminus S$ 上有 $\nabla P(x, x_k^*, A) \neq 0$ 成立.

证明 对任意的 $x \in S_1$, 且 $x \neq x_k^*$ 或 $R^n \setminus S$ 有

$$\min [0, \max (f(x) - f(x_k^*), g_i(x), i = 1, \dots, m)] = 0.$$

此时

$$p(x, x_k^*, A) = \frac{e^{A \cdot \varphi(f(x) - f(x_k^*))}}{1 + \|x - x_k^*\|^2} = \frac{e^A}{1 + \|x - x_k^*\|^2},$$

从而

$$\nabla p(x, x_k^*, A) = \frac{-2e^A \cdot (x - x_k^*)}{(1 + \|x - x_k^*\|^2)^2} \neq 0.$$

故对任意的 $x \in S_1$, 对任意 $A > 0$ 且 $x \neq x_k^*$ 或 $R^n \setminus S$, 有 $\nabla P(x, x_k^*, A) \neq 0$.

定理 3: 如果 x_k^* 不是 $f(x)$ 的全局极小点, 且 $\text{clint } S = \text{cl}S$, 则当 $A > 0$ 充分大时, 填充函数 $P(x, x_k^*, A)$ 在 S_2 上一定存在局部极小点 x_0^* , 其中 $S_2 = \{x \in S | f(x) < f(x_k^*), g_i(x) \leq 0, i = 1, \dots, m\}$ 非空.

证明 因为 x_k^* 是 $f(x)$ 的局部极小点, 不是 $f(x)$ 的全局极小点, 故 $S_2 \neq \emptyset$, 存在 $f(x)$ 的另一个局部极小点 x_1^* , 使得 $f(x_1^*) < f(x_k^*), g_i(x_1^*) \leq 0, i = 1, \dots, m$.

由于 $f(x), g_i(x), i = 1, \dots, m$ 是连续函数, 且 $\text{clint } S = \text{cl}S$, 则一定存在 $x_2^* \in R^n$, 使得 $f(x_2^*) < f(x_k^*), g_i(x_2^*) < 0, i = 1, \dots, m$,

从而 $f(x_2^*) - f(x_k^*) < 0$, 故 $\varphi[f(x_2^*) - f(x_k^*)] = 0, \max(f(x_2^*) - f(x_k^*), g_i(x_2^*), i = 1, \dots, m) < 0$. 则

$$\begin{aligned} p(x_2^*, x_k^*, A) &= \frac{e^{A \cdot \varphi(f(x_2^*) - f(x_k^*))}}{1 + \|x_2^* - x_k^*\|^2} + \frac{1}{A} \{ \min [0, \max (f(x_2^*) - f(x_k^*), g_i(x_2^*), i = 1, \dots, m)] \} \\ &= \frac{1}{1 + \|x_2^* - x_k^*\|^2} + \frac{1}{A} \max (f(x_2^*) - f(x_k^*), g_i(x_2^*), i = 1, \dots, m). \end{aligned}$$

另一方面, 对任意的 $x \in \partial S$, 至少存在一个指标 $i \in \{1, \dots, m\}$, 使得 $g_{i_1}(x) = 0$.

(i) $x \in \partial S$, 当 $f(x) \geq f(x_k^*)$ 时, 则 $f(x) - f(x_k^*) \geq 0$, 因此, $\min [0, \max (f(x) - f(x_k^*), g_i(x),)] = 0$

(ii) $x \in \partial S$, 当 $f(x) < f(x_k^*)$ 时, 则 $f(x) - f(x_k^*) < 0$, 因此, $\min [0, \max (f(x) - f(x_k^*), g_i(x),)] = 0$

综上 $\forall x \in \partial S$ 都有 $\min [0, \max (f(x) - f(x_k^*), g_i(x), i = 1, \dots, m)] = 0$.

所以 $\forall x \in \partial S$ 有

$$p(x, x_k^*, A) = \frac{e^{A \cdot \varphi(f(x) - f(x_k^*))}}{1 + \|x - x_k^*\|^2} + \frac{1}{A} \cdot 0 = \frac{e^{A \cdot \varphi(f(x) - f(x_k^*))}}{1 + \|x - x_k^*\|^2},$$

故当 $A > 0$ 时, 有 $\varphi[f(x) - f(x_k^*)] \geq 0$, 则有 $e^{A \cdot \varphi[f(x) - f(x_k^*)]} \geq 1$.

所以, 当 A 充分小时, 对 $\forall x \in \partial S$, 我们有 $P(x_2^*, x_k^*, A) < \frac{1}{1 + \|x - x_k^*\|^2} \leq P(x, x_k^*, A)$. 因此存在一点 $x_0^* \in S \setminus \partial S$ 使得

$$\min_{x \in S} P(x, x_k^*, A) = \min_{x \in S \setminus \partial S} P(x, x_k^*, A) = P(x_0^*, x_k^*, A) \leq P(x_2^*, x_k^*, A),$$

由于 $S \setminus \partial S$ 是开集, 于是 A 充分小时存在一点 $x_0^* \in S \setminus \partial S$ 是 $p(x, x_k^*, A)$ 的局部极小点, 又由定理 2 知 $x_0^* \in S_2$, 定理得证.

综上所述, 函数 $p(x, x_k^*, A)$ 符合填充函数的定义.

4. 算法和数值实验

根据上述填充函数 $p(x, x_k^*, A)$ 的理论分析, 提出以下关于问题 (1) 的填充函数算法:

步骤 1 初始步. 选取初始点 $x_1 \in S$, 选取 $A_L > 0$ 作为可接受的终止参数; 令 $A = 1, k = 1, 0 < r < 1$, 一般取 $r = 0.1$;

步骤 2 以 x_k 为初始点, 运用已有的局部下降算法得到问题 (1) 的一个局部极小点, 记作 x_k^* ;

步骤 3 选取初始点列 $\{x_k^i | i = 1, \dots, m\}$, 使得对某个 $\delta_k > 0$, 有 $x_{k+1}^i \in S \setminus N(x_k^*, \delta_k)$;

步骤 4 令填充函数为 $p(x, x_k^*, A), i = 1$;

步骤 5 若 $i \leq m$, 令 $x := x_{k+1}^i$, 转步骤 6; 否则, 转步骤 8;

步骤 6 若 $f(x) < f(x_k^*)$, 且 $x \in S$, 则以 x 为初始点, 运用已有的局部下降算法得到问题 (1) 的局部极小点 x_{k+1}^* , 使得 $f(x_{k+1}^*) < f(x_k^*)$, 令 $x_k^* = x_{k+1}^*, k = k + 1$, 转步骤 3; 否则, 转步骤 7;

步骤 7 沿方向 $D = -p(x, x_k^*, A)$, 通过局部下降算法找到填充函数的一个新的点 x , 使得 $p(x, x_k^*, A)$ 能下降到一定程度使得 $p(x, x_{k+1}^*, A) < p(x, x_k^*, A)$. 若在极小化过程中, x 超出 S 的边界, 则令 $i = i + 1$, 转步骤 5, 否则, 转步骤 6;

步骤 8 令 $A = r \cdot A$. 若 $A > A_L$, 转步骤 4; 否则, 算法终止, 视 x_k^* 为问题 (1) 的全局极

小点.

步骤 3 中对当前的局部极小点的初始点集选取方式为: 令 $m = 2n$, 并且初始点对称. 例如, 当 $n = 2$ 时, 初始点为:

$$x_k^* + \delta(1, 0), x_k^* + \delta(0, 1), x_k^* - \delta(1, 0), x_k^* - \delta(0, 1),$$

这里 $\delta > 0$ 是预先给定的步长.

由定理 3 可知: 步骤 6 中若出现 $f(x) < f(x_k^*)$, 则 x 一定落在 S_2 中, 由此以 x 为初始点极小化目标函数 $f(x)$, 得到比当前的局部极小值更小的局部极小值.

由定理 1 得知若要使 $p(x, x_k^*, A)$ 的局部极小点在 x_k^* 得到, 则参数 A 应该充分小. 在步骤 8 中, 减小 A 的值. 若对所有的初始点已进行计算且 $A < A_L$, 我们可以认为, 根据现有的初始点, 找不到更小的局部极小点, 所以把当前的局部极小点看作全局极小点, 算法终止.

选取文献 [8] 中的无参数填充函数算例, 用 Python 对以上算法进行编程, 以便于计算结果的比较. 取 $m = 4$, 记初始点为 x_1 , 迭代次数为 k , 全局极小点为 x^* , 全局极小值为 f^* .

算例 1 Simplified Rosenbrock's 函数

$$\begin{aligned} \min f(x) &= 0.5(x_1^2 - x_2)^2 + (x_1 - 1)^2 \\ \text{s.t.} \quad & -3 \leq x_1 \leq 3, \\ & -3 \leq x_2 \leq 3. \end{aligned}$$

已知该函数的全局极小点 $x^* = (1, 1)^T$, 全局极小值 $f^* = 0$.

Table 1. The running result of example 1

表 1. 算例 1 的运行结果

| | x_1 | k | x^* | f^* |
|--------|-------------|-----|--------------------------------|--------------------------|
| 本文 | $(-2, 2)^T$ | 2 | $(1, 1)^T$ | 0 |
| | $(-3, 3)^T$ | 2 | $(0.999999999, 0.999999998)^T$ | 1.6264×10^{-16} |
| 文献 [8] | $(-2, 2)^T$ | 5 | $(1, 1)^T$ | 1.23×10^{-10} |
| 文献 [8] | $(-3, 3)^T$ | 6 | $(1, 1)^T$ | 0 |

算例 2

$$\begin{aligned} \min f(x) &= -25(x_1 - 2)^2 - (x_2 - 2)^2 - (x_3 - 1)^2 - (x_4 - 4)^2 - (x_5 - 1)^2 - (x_6 - 4)^2 \\ \text{s.t.} \quad & (x_3 - 3)^2 + x_4 \geq 4, \quad (x_5 - 3)^2 + x_6 \geq 4, \quad x_1 - 3x_2 \leq 2, \\ & -x_1 + x_2 \leq 2, \quad x_1 + x_2 \leq 6, \quad x_1 + x_2 \geq 2, \quad 0 \\ & \leq x_1, x_4 \leq 6, \quad 0 \leq x_2 \leq 8, \quad 1 \leq x_3, x_5 \leq 5, \quad 0 \leq x_6 \leq 10. \end{aligned}$$

已知该函数的全局极小点 $x^* = (5, 1, 5, 0, 5, 10)^T$, 全局最优值 $f(x^*) = -310$.

Table 2. The running result of example 2**表 2.** 算例 2 的运行结果

| | x_1 | k | x^* | f^* |
|---------|---|-----|---|-----------|
| 本文 | (4.9402, 1.0598, 4.8895, 2.0000, 5.0000, 10.0000) ^T | 2 | (4.9999, 1.0000, 5.0000, 0.0000, 4.9998, 10.0000) ^T | -309.9891 |
| | (5.0000, 1.0000, 5.0000, 1.7864, 5.0000, 9.9999) ^T | 2 | (5.0000, 1.0000, 5.0000 0.0000, 4.9998, 9.9998) ^T | -297.9349 |
| 文献 [10] | (4.9402, 1.0598, 4.8895 2.0000, 5.0000, 10.0000) ^T | 2 | (4.9680, 1.0320, 5.0000 1.9045, 4.9136, 10.0000) ^T | -292.8629 |
| 文献 [10] | (5.0000, 1.0000, 5.0000 | 4 | (5.0000, 1.0000, 5.0000 | -297.9349 |

表 1 和表 2 数值算例结果表明本文所构造的单参数填充函数在求解不等式约束的优化问题时是有效可行的, 对比文献 [10] 中的结果, 本文的填充函数法计算精度相对较高. 在算例 2 的结果表明: 利用填充函数进行优化后, 目标函数能找到全局极小点, 与文献 [10] 相比, 函数值更优.

5. 填充函数在聚类中的应用

5.1. 聚类模型概述

聚类是指把相似的数据划分到一起, 目标就是把相似的数据聚合到一起, 聚类是机器学习一种无监督学习方法. 支持向量机 (SVM) 是一种监督式学习方法, 广泛应用于统计分类以及回归分析中, 通常是通过构造核映射将样本数据进行分类处理. 本文的聚类模型表述如下.

聚类过程中通常会遇到原始样本空间过于复杂从而使得聚类模型不精确的情况, 为了克服这一难题, 本文首先建立一个合适的核映射, 将原始样本中的线性不可分问题通过核映射的方式转化为高维特征空间中的线性可分问题, 使得之后在特征空间中的聚类更精确.

给定一组数据 (X_i, y_i) , $i = 1, \dots, n$. $A_1 \times A_2 \times \dots \times A_n$ 是数据在具有多种属性类型的 m 维空间, 某个区域内存在用来构造 K 类分类器的训练数据点对集

$\hat{S} = \langle X_1, y_1 \rangle, \langle X_2, y_2 \rangle, \dots, \langle X_n, y_n \rangle$, 其中 $X_i = (x_{i1}, x_{i2}, \dots, x_{im}) \in A_1 \times A_2 \times \dots \times A_m (i = 1, \dots, n)$ 表示第 i 个训练数据点对在样本空间中的位置; $y_i \in \{1, \dots, K\}$ 表示第 i 个训练数据点对的所属类别标号. 在聚类、分类研究中, 为了体现出每一个属性对分类预测的不同作用, 为了描述数据点 X_i 和 X_j 之间的差异, 会设定一个带权重的“距离”度量, 即: $D(X_i, X_j) = (\sum_{k=1}^m (w_k \times d_k^2(x_{ik}, x_{jk})))^{\frac{1}{2}}$, 其中 w_k 表示样本空间中的属性权重参数, 为构造一个能拟合训练数据点对集 $\hat{S} = \{\langle X_1, y_1 \rangle, \langle X_2, y_2 \rangle, \dots, \langle X_n, y_n \rangle\}$ 的最优分类器, 可以按照每个数据点对所属的类别来划分, 从而得到多个样本数据点子集的集合 $\bar{S} = S_1, S_2, \dots, S_K$, 其中第 i 个数据点子集 S_i 属于第 i 类; 然后构造一个从样本空间到特征空间的映射 $\Phi: A_1 \times A_2 \times \dots \times A_m \rightarrow R^H$,

其中 R^H 是高维特征空间, 样本空间中的线性不可分问题通过核映射转换为高维特征空间中的线性可分问题来使得之后的聚类更精确. 因此构造核映射应该尽量满足两个条件, 即: 特征空间中每个子聚类内数据点之间的距离尽可能小, 特征空间中不同聚类之间数据点之间的距离足够大. 构造这样一个映射后, 使样本空间中属于第 i 类的数据点子集 S_i 的所有元素在特征空间的核映射下所形成的像形成一个紧密的子聚类.

为了满足上述使得特征空间中每个子聚类内数据点之间的距离尽可能小的目的, 可以建立如下的优化问题将特征空间的聚类优化问题和原始样本空间的分类问题联系起来. 为寻找一组适合样本空间中的数据点对集的属性权重组, 我们提出了一种基于核映射的填充函数法来优化属性权重.

优化属性权重向量 $w = (w_1, w_2, \dots, w_m)^T$ 可描述为如下的优化问题:

$$\begin{aligned} \min F(w) &= \sum_{h=1}^K \sum_{\Phi(X_i), \Phi(X_j) \in F_h} \left(D_{\tilde{F}}(\Phi(X_i), \Phi(X_j)) \right)^2 \\ &= \sum_{h=1}^K \sum_{X_i, X_j \in S_h} \left(2 - 2e^{-\frac{\sum_{k=1}^m (w_k \cdot d_k^2(x_{ik}, x_{jk}))}{2\sigma^2}} \right), \\ \text{s.t.} \quad &\sum_{k=1}^m w_k \leq m, 0 \leq w_k \leq 1, \quad (k = 1, 2, \dots, m). \end{aligned} \quad (4)$$

其中, $F(w)$ 表示同一个聚类块内, 像向量最大限度地接近以保证数据点尽可能地接近聚类中心. \tilde{F} 为特征空间的映射集, F_h 为映射集第 h 个子聚类. $\Phi(X_i), \Phi(X_j)$ 是第 h 个子聚类的两个像向量.

为了满足上述使得特征空间中每个子聚类数据点之间的距离尽可能大的目的, 可以建立如下的优化问题将特征空间的聚类优化问题和原始样本空间的分类问题联系起来.

$$\begin{aligned} \max G(w) &= \sum_{F_h \neq F_r} \sum_{\Phi(X_i) \in F_h, \Phi(X_j) \in F_r} \left(D_{\tilde{F}}(\Phi(X_i), \Phi(X_j)) \right)^2 \\ &= \sum_{S_h \neq S_r} \sum_{X_i \in S_h, X_j \in S_r} \left(2 - 2e^{-\frac{\sum_{k=1}^m (w_k \cdot d_k^2(x_{ik}, x_{jk}))}{2\sigma^2}} \right), \\ \text{s.t.} \quad &\sum_{k=1}^m w_k \leq m, 0 \leq w_k \leq 1, \quad (k = 1, 2, \dots, m). \end{aligned}$$

上式中, $\Phi(X_i)$ 和 $\Phi(X_j)$ 分别是特征空间中第 h 个子聚类 $F_h (h = 1, \dots, K)$ 和第 r 个子聚类 $F_r (r = 1, \dots, K)$ 的两个像向量. σ 表示高斯核函数的宽度, 本文用固定核宽度表示. $d_k(x_{ik}, x_{jk})$ 根据第 k 个属性的类型决定其具体定义式. 如果第 k 个属性是无序类别属性, 则定义为

$$d_k(x_{ik}, x_{jk}) = \{0, x_{ij} = x_{jk}, 1, x_{ik} \neq x_{jk}.$$

如果第 k 个属性是有序类别属性, 则定义为 $d_k(x_{ik}, x_{jk}) = |x_{ik} - x_{jk}|$. 一般设定 $\sum_{k=1}^m w_k \leq m$, $0 \leq w_k \leq 1 (k = 1, \dots, m)$ 是为了将权重归一化.

为了找到更优配置的属性权重向量, 使建立在加权后的“距离”度量下的分类器的分类效果更佳, 可构造 $F(W)$ 和 $G(W)$ 的一个最小化混合目标函数:

$$\begin{aligned} \min E(W, \lambda) &= \lambda \cdot F(W) - (1 - \lambda) \cdot G(W) \\ \text{s.t. } \sum_{k=1}^m w_k &\leq m, 0 \leq w_k \leq 1, \quad (k = 1, 2, \dots, m). \end{aligned} \quad (5)$$

5.2. 用填充函数法求解聚类模型

根据上节填充函数 $P(x, x^*, A)$ 的理论依据, 提出了求解上述优化问题 (4) 的算法.

步骤 1 初始步. 选取初始点 $w_1 \in S$, 其中 S 是问题 (4) 中 w_k 的取值范围. 选取 $A_L > 0$ 作为终止参数; 令 $A = 1, k = 1, 0 < r < 1$, 一般取 $r = 0.1$;

步骤 2 以 w_k 为初始点, 应用梯度下降算法求得问题 (4) 的一个局部极小点, 记作 w_k^* ;

步骤 3 选取初始点列 $\{w_k^i | i = 1, 2, \dots, m\}$, 使得对某个 $\delta_k > 0$, 有 $w_{k+1}^i \in S \setminus N(w_k^*, \delta_k)$;

步骤 4 令填充函数为 $p(w, w_k^*, A), i = 1$;

步骤 5 若 $i \leq m$, 令 $w := w_{k+1}^i$, 转步骤 6; 否则, 转步骤 8;

步骤 6 若 $F(w) < F(w_k^*)$, 且 $w \in S$, 则以 w 为初始点, 应用梯度下降算法求问题 (4) 的局部极小点 w_{k+1}^* , 使得 $F(w_{k+1}^*) < F(w_k^*)$, 令 $w_k^* = w_{k+1}^*, k = k + 1$, 转步骤 3; 否则, 转步骤 7;

步骤 7 沿方向 $D = -p(w, w_k^*, A)$, 使用梯度下降算法找到填充函数的一个新的点 w , 使得 $p(w, w_k^*, A)$ 能下降到一定程度使得 $p(w, w_{k+1}^*, A) < p(w, w_k^*, A)$. 若在极小化过程中, w 超出 S 的边界, 则令 $i = i + 1$, 转步骤 5, 否则, 转步骤 6;

步骤 8 令 $A = r \cdot A$. 若 $A > A_L$, 转步骤 4; 否则, 算法终止, 视 w_k^* 为问题 (4) 的全局极小点.

步骤 3 中的初始点集选取为: $m = 2n$, 对当前的局部极小点, 假设初始点是对称的. 例如, 当 $n = 4$ 时, 初始点为:

$$w_k^* + \alpha(1, 0, 0, 0)^T, x_k^* + \alpha(0, 1, 0, 0)^T, x_k^* - \alpha(1, 0, 0, 0)^T, x_k^* - \alpha(0, 1, 0, 0)^T, \dots,$$

这里 $\alpha > 0$ 是预先给定的步长.

对于步 6, 我们首先比较 $F(w)$ 与 $F(w_k^*)$ 的大小, 若 $F(w) < F(w_k^*)$ 直接进入极小化目标函数阶段, 不需要在该点处极小化填充函数, 用 Python 编码上述算法, 选取不同的初始点和步长代入算法中, 代码运行得到最终结果后发现当选取适当的初始点和步长时, 得到的最小值为全局最优解.

5.3. 模型结果的比较和分析

通过对文献 [11] 中鸢尾花数据进行聚类分析, 已知有四个属性共同决定鸢尾花类别时, 为了使得在同一聚类中的像向量距离最小, 在不同聚类中的像向量距离最大, 即使得聚类块更紧密,

通过填充函数优化混合目标函数的极小值得到的权重向量. 图 1 表示未用填充函数进行优化时得到的函数值. 图 1 表示在运用填充函数对上述问题进行优化后得出的最优值.

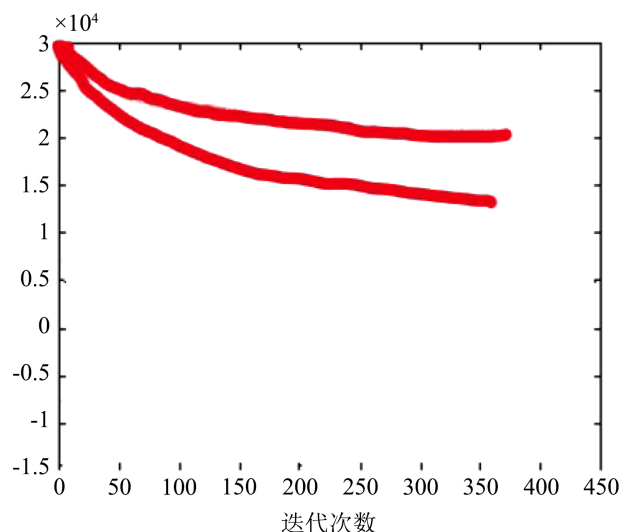


Figure 1. Objective function values

图 1. 目标函数值

通过以上过程可以看出当 $W = (0.34294512, 0.13697411, 0.06632191, 0.45375887)^T$ 时得到图 1 的最优权重参数, 此时目标函数值为 15000. 图 1 在未用填充函数法对目标函数优化情况下得到的函数值为 20000. 由此可以看出利用填充函数后得到的优化效果更好.

6. 总结

本文提出了一类单参数填充函数, 新设计的填充函数算法解决了不等式约束 [12] 的优化问题, 由于不需要考虑强制性条件的情况, 并且运用填充函数进行优化时可以使得原目标函数跳出当前局部极小点, 找到全局极小点, 因此该方法更具有精确性.

通过引入基于核映射的属性权重自适应优化模型, 结合填充函数法进行求解, 通过分析虹膜数据集来证实了算法在聚类问题中的可用性.

参考文献

- [1] Ge, R.P. (1990) The Filled Function Transformations for Constrained Global Optimization. *Applied Mathematics and Computation*, **39**, 1-20. [https://doi.org/10.1016/0096-3003\(90\)90118-M](https://doi.org/10.1016/0096-3003(90)90118-M)
- [2] 张连生. 求解全局优化的填充函数法的进展 [C]//中国运筹学会. 2001 年全国数学规划及运筹研讨会论文集. 贵阳: 贵州大学学报编辑部, 2001: 154-156.
- [3] 王伟祥, 尚有林, 张连生. 约束全局优化问题的一个单参数填充函数方法 [J]. 工程数学学报, 2008, 25(5): 795-803.

-
- [4] 朱文兴. 一类不依赖于局部极小解个数的填充函数法 [J], 统计科学与数学, 2002, 22(4): 406-413.
- [5] 孔敏. 一类改进的非光滑规划的填充函数法 [J], 统计科学与数学, 2004, 20(4): 149-154.
- [6] 钟以维, 徐应涛, 张莹. 用填充函数法改进的人脸识别算法 [J]. 计算机技术与发展, 2009, 19(8): 78-81.
- [7] Wang, W.X., Shang, Y.L. and Zhang, L.S. (2007) A Filled Function Method for Global Optimization Problem. *Operations Research Transactions*, **11**, 43-50.
- [8] 吴波, 高岳林. 全局优化问题的一个无参数填充函数算法 [J]. 数学的实践与认识, 2017, 47(4): 170-174.
- [9] Zhang, Y., Zhang, L.S. and Xu, Y.T. (2010) A One-Parameter Filled Function for Nonsmooth Global Optimization and Its Application. *Journal of Systems Science and Complexity*, **23**, 1195-1209. <https://doi.org/10.1007/s11424-010-7199-5>
- [10] 赵德芬, 王薇. 一类求解约束全局优化问题的填充函数 [J]. 运筹学学报, 2010, 14(2): 119-126.
- [11] 陈新泉. 一种基于核映射的自适应优化配置属性权重组的方法 [J]. 数值计算与计算机应用, 2008, 29(2): 105-118.
- [12] 陈佳利, 张莹, 王胜刚, 谢笑盈. 一个新的填充函数及其在数据拟合问题中的应用 [J]. 运筹学学报, 2021, 25(1): 81-88.