

# 基于梯度下降法选取高斯径向基函数形状参数

谢志超, 王玲\*, 龚佃选, 孙建, 田炜印

华北理工大学理学院, 河北 唐山

收稿日期: 2023年3月24日; 录用日期: 2023年4月18日; 发布日期: 2023年4月26日

## 摘要

高斯径向基函数是径向基函数中常用的函数, 对于高斯径向基函数插值形状参数的取值, 常规的方法一般是通过人工修改形状参数的取值, 计算成本较高, 插值精度较低。本文基于梯度下降算法, 通过设定针对性的目标函数, 通过迭代的方式得到形状参数, 有更好的插值效果, 为高斯径向基函数的应用提供了更加便捷的途径。

## 关键词

梯度下降法, 高斯径向基函数, 插值逼近, 形状参数

# Gaussian Radial Basis Function Shape Parameter Is Selected Based on Gradient Descent Method

Zhichao Xie, Ling Wang\*, Tongxuan Gong, Jian Sun, Weiyin Tian

College of Science, North China University of Science and Technology, Tangshan Hebei

Received: Mar. 24<sup>th</sup>, 2023; accepted: Apr. 18<sup>th</sup>, 2023; published: Apr. 26<sup>th</sup>, 2023

## Abstract

Gaussian radial basis function is a commonly used function in radial basis function. For the value of shape parameters interpolated by Gaussian radial basis function, the conventional method is generally to manually modify the value of shape parameters, which has high calculation cost and low interpolation accuracy. In this paper, based on the gradient descent algorithm, shape parameters are obtained through iteration by setting targeted objective functions, which has better interpolation effect and provides a more convenient way for the application of Gaussian radial ba-

\*通讯作者。

sis function.

## Keywords

Gradient Descent Method, Gaussian Radial Basis Function, Interpolation Approximation, Shape Parameter

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/>



Open Access

## 1. 引言

在实际生活中,在不同的领域会产生各种散乱数据的拟合问题,选择一个良好的方法是处理问题的关键。径向基函数有着良好的逼近效果,是处理散乱数据问题的一种较优方法[1]。径向基函数利用简单的一元函数作为基函数,使得自身可以有效的减少计算最优值点所需要的运算次数,在一定程度上为找到全局最优值节约了大量计算成本。

高斯径向基函数方法是插值多维散乱数据的有力工具,是在处理大规模散乱数据时常用的方法。目前,对于高斯径向基函数的最优形状参数取值有众多国内外的学者取得了不错的成果,使得高斯径向基函数在实际问题的数值模拟中插值逼近的精度或速度有一定的提升,在一定程度上推动了高斯径向基函数实际应用的范围。

秦伶俐[2]在径向基函数在无单元法应用中:探讨了求解微分方程基于径向基函数的无单元法,总结出了高斯的径向基函数和 Multi-Quadric 的径向基函数中的形状参数与微分方程求解精度的关系曲线,以及节点均匀分布时最佳取值的计算公式,在保证求解精度的前提下,验证出自由分布的节点适用于均匀分布节点下的自由参数形状公式,证明出在求解微分方程时基于径向基函数的无单元法对节点的分布位置不敏感。

徐咏梅[3]在高斯径向基核函数参数的 GA 优化方法中提出:利用 GA 的非线性优化的性质确定高斯径向基核函数参数的取值,通过实验证明使用该方法所确定的高斯径向基核函数参数应用于 KFDA 中具有良好的效果,模式分类的可分性测试度优于其他方法的分类效果。

高斯径向基函数是一种常见的径向基函数,形状参数  $c$  决定了其插值精度,不同的插值问题对应着不同的形状参数。目前来讲,对于各种高斯径向基函数插值形状参数选取问题还没有通用的方法,故对高斯径向基函数的形状参数选取方法进行研究具有重要的科研价值。

## 2. 径向基函数的定义

径向基函数(RBF) [4]就是一个取值仅仅依赖于离原点距离的实值函数  $\phi$ , 即  $\phi(x) = \phi(\|x\|)$ 。如果满足:  $\|x_1\| = \|x_2\|$ , 那么满足  $\phi(x_1) = \phi(x_2)$  的实值函数  $\phi$  就是径向函数,是仅仅由  $r = \|x\|$  控制的函数(其中的范数是 Euclid 范数)。

高斯径向基函数[5]的常用表法式为:

$$\phi(r) = e^{-c^2 r^2} \quad (1)$$

(1)式中:  $c$  是基函数为高斯函数的形状参数。

在二维空间中具体可表示为:

$$Gauss = \exp \left[ -\frac{(x-\mu_x)^2}{2\delta_x^2} - \frac{(y-\mu_y)^2}{2\delta_y^2} \right] \quad (2)$$

在三维空间中具体表示为:

$$Gauss = \exp \left[ -\frac{(x-\mu_x)^2}{2\delta_x^2} - \frac{(y-\mu_y)^2}{2\delta_y^2} - \frac{(z-\mu_z)^2}{2\delta_z^2} \right] \quad (3)$$

上述(2) (3)两式中:  $(x, y, z)$  是网格中心坐标;  $(\mu_x, \mu_y, \mu_z)$  是径向基函数中心坐标;  $\delta_x, \delta_y, \delta_z$  是径向基函数分布半径。

这里介绍一些常见的径向基函数:

1) Kriging 方法的 Gauss 分布函数:

$$\phi(r) = e^{-c^2 r^2} \quad (4)$$

2) Kriging 方法的 Markoff 分布函数:

$$\phi(r) = e^{-c|r|} \quad (5)$$

3) Hardy 的 Multi-Quadric 函数:

$$\phi(r) = (c^2 + r^2)^\beta \quad (6)$$

(其中  $\beta$  为正实数)

4) Hardy 的逆 Multi-Quadric 函数:

$$\phi(r) = (C^2 + r^2)^{-\beta} \quad (7)$$

(其中  $\beta$  为正实数)

5) Duchon 的薄板样条:

$d$  为偶数时:

$$\phi(r) = r^{2k-d} \log r \quad (8)$$

$d$  为奇数时:

$$\phi(r) = r^{2k-d} \quad (9)$$

6) 紧支柱正定径向基函数:

紧支柱正定径向基函数是: 全局形的函数通过一个多项式在给定的连续性条件和给定维数的自变量空间下截断产生的, 同时满足正定和连续, 一般情况下表示为:

$$\phi_{d,k}(r) = (1-r)_+^k p(r), (k \geq 0) \quad (10)$$

其中  $p(r)$  为某一给定的多项式,

当  $(0 \leq r \leq 1)$  时,  $(1-r)_+^k = (1-r)^k$ ;

当  $(1 \geq r)$  时,  $(1-r)_+^k = 0$ 。

### 3. 梯度下降法

梯度下降法(Gradient Descent) [6]是机器学习中常见的迭代法, 是一种一阶最优化算法, 也称为最速

下降法。梯度下降法可以用来处理任何非约束优化问题，普遍应用于神经网络参数的优化。

一般的梯度下降算法步骤为：

Step1: 给定待优化连续可微的目标函数  $f(x) = f(x_1, \dots, x_n)$ ，学习率  $\eta$  或步长  $d$ ，以及函数的一组初始值  $(x_1, \dots, x_n)$

Step2: 计算待优化函数梯度方向:  $\frac{\partial f}{\partial x_1}(x^{(i)})$

Step3: 更新迭代:  $x_1^{(i+1)} = x_1^{(i)} - \eta \cdot \frac{\partial f}{\partial x_1}(x^{(i)})$

Step4: 更新参数(梯度方向、学习率)

Step5: 如果到达收敛停止迭代输出最值，否则回到 Step2 继续迭代

由于梯度下降法是一阶优化算法[3]，因此可以从导数的定义去理解梯度下降法的梯度。在微积分中，函数  $f(x)$  在点  $x_0$  上的导数定义为：

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad (11)$$

导数的几何意义就是函数  $f(x)$  在  $x_0$  上的切线方向。从 Taylor 级数的角度来说， $f(x)$  在  $x_0$  附近的 Taylor 级数是：

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + O(|x - x_0|^3) \quad (12)$$

对于临界点  $x_0$ ，满足  $f'(x_0) = 0$ 。当  $f''(x_0) > 0$  时，可以得到  $x_0$  是  $f(x)$  的局部最小值；当  $f''(x_0) < 0$  时，此时  $x_0$  是  $f(x)$  的局部最大值。因此，当函数可导时，梯度方向就是函数的一阶导。

#### 4. 基于改进的梯度下降法确定形状参数

对于高斯径向基函数最优形状参数确定，在选取初始形状参数后，并不能保证所选取的参数是当前的最优形状参数。因此，可以利用梯度下降法寻找到高斯径向基函数最优形状参数。故在处理最优形状参数选取问题时，需要针对性的对梯度下降法进行改进。改进主要针对目标函数、梯度方向、学习率三个方面进行：

##### 1) 目标函数

基于最佳一致逼近原则当绝对最大误差(MAXERROR)最小时，此时的形状参数称为当前高斯径向基函数的最优形状参数。因此，基于最优形状参数与误差函数之间的关系，可以定义这样的目标函数：

$$F(x) = \max |f(x) - \phi(x)| \quad (13)$$

(40)式中:  $f(x)$  为原函数，具体表示为：

$$f(x) = \alpha (a_i \sin(k\pi x + \theta)) \quad (14)$$

(40)式中:  $\phi(x)$  为高斯径向基函数， $\alpha$  为常数，下式中  $c$  为形状参数，具体表示为：

$$\phi(x) = e^{-c^2 x^2} \quad (15)$$

##### 2) 梯度

用差商公式近似代替导数 Euler 法：

初值问题的形式为：

$$\frac{dy}{dx} = f(\lambda, y), x \in [a, b], y(a) = y_0 \tag{16}$$

首先将[a,b]区间等分即取等间距的插值间距，用向前差商近似代替代数：

$$\frac{dy}{dx} \approx \frac{y(x_n - 1) - y(x_n)}{h} \tag{17}$$

可得： $y_n \approx y(x_n)$ 。

由于构建的目标函数无法直接进行求导或求偏导，但是目标函数在固定区间上是连续函数，因此可以利用差商来代替导数，根据 Euler 法的差商公式改进后的梯度方向表示为：

$$\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} \tag{18}$$

### 3) 学习率

一般来说，需要有合适的学习率才能获得最优的算法模型，得出目标函数的最小值。在设定学习率时需要满足如下的条件：

$$\sum_{i=1}^{\infty} \varepsilon_i = \infty, \sum_{i=1}^{\infty} \varepsilon_i^2 < \infty \tag{19}$$

其中  $i$  表示第  $i$  次更新。

对于学习率的具体数值没有参考标准，根据二分线性搜索(Bisection Line Search)的方法，针对本文讨论的问题进行学习率的数值实验。

#### 数值实验 1:

实验中目标函数为：

$$f(x) = \sin(\pi x) + \sin(0.1\pi x)$$

初始形状参数、插值间距、插值范围均取相同值，学习率、误差及运行次数实验结果如下：

**Table 1.** Learning rate verification results

**表 1.** 学习率验证结果

$\eta$	Max error	Count times
0.2	3140	29
0.1	null	null
0.05	0.00035	23
0.01	0.00035	15
0.005	0.00118	97
0.001	0.00253	205

由于高斯径向基函数对形状参数及其敏感，在梯度下降时，若梯度下降过快，会造成权重矩阵奇异，导致最终无法插值拟合出函数。根据上述数值实验可以看出，当学习率稍大时，插值拟合效果较差，误差较大，针对高斯径向基函数形状参数敏感问题，对迭代下降的步长设定一个阈值，使其避免因迭代过快，造成无法拟合的情况。设定步长阈值的迭代公式表示为：

$$x_1^{(i+1)} = x_1^{(i)} - d, d = \eta \cdot \frac{\partial f}{\partial x_1}(x^{(i)}) \tag{20}$$

当  $d > 0.1$  时, 令  $d = 0.1$ , 否则  $d = \eta \cdot \frac{\partial f}{\partial x_1}(x^{(i)})$  继续更新迭代, 在此基础上, 对学习率问题再次进行数值实验。

### 数值实验 2:

实验中目标函数为:

$$f(x) = \sin(\pi x) + \sin(0.1\pi x)$$

初始形状参数、插值间距、插值范围均取相同值, 学习率、误差及运行次数实验结果如下:

**Table 2.** Verification of the improved learning rate

**表 2.** 改进后的学习率验证

$\eta$	Max error	Count times
0.2	0.00087	126
0.1	0.00046	83
0.05	0.00035	46
0.01	0.00035	32
0.005	0.00118	127
0.001	0.00253	258

根据表 1、表 2 数值实验的结果, 当学习率  $\eta = 0.01$  时, 误差函数相对最小, 拟合效果相对最优, 计算速度相对最快。

在对一般梯度下降法的目标函数、梯度、学习率针对性改进后, 高斯径向基函数最优形状参数输出算法为:

Step1: 既定的目标函数为  $F(x) = \max|f(x) - \phi(x)|$ , 学习率  $\eta = 0.01$ , 初始值设置为  $x_i, x_{i+1}$

Step2: 计算待优化函数梯度方向:  $\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$

Step3: 更新迭代:  $x_1^{(i+1)} = x_1^{(i)} - d, d = \eta \cdot \frac{\partial f}{\partial x_1}(x^{(i)}), d > 0.1 \rightarrow d = 0.1$

Step4: 更新梯度方向:  $\frac{f(x_{i+2}) - f(x_{i+1})}{x_{i+2} - x_{i+1}}$

Step5: 若  $|F(x_i)_{\min} - F(x_{i+1})_{\min}| \leq 1 \times 10^{-8}$ , 输出当前形状参数, 否则转 Step2 继续迭代

## 5. 数值实验验证

### 数值实验 3:

实验函数:  $f(x) = \sin(\pi x) + \sin(0.1\pi x)$ , 随机取形状参数  $c = 100$ , 基于改进后的梯度下降法得到的形状参数为  $c = 1.5348$ 。

函数插值拟合图像为图 1, 误差函数图像为图 2。

### 数值实验 4:

实验函数:  $f(x) = 100 \sin(0.01\pi x)$ , 随机取形状参数  $c = 0.1$ , 基于改进后的梯度下降法得到的形状参数为  $c = 80.435$ 。

函数插值拟合图像为图 3, 误差函数图像为图 4。

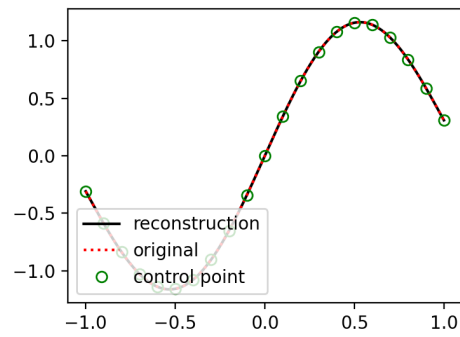


Figure 1. Effect of interpolation  
图 1. 插值效果图

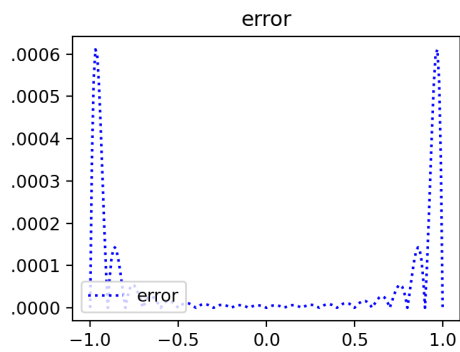


Figure 2. Error diagram  
图 2. 误差图

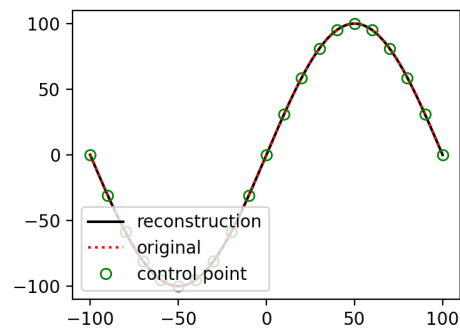


Figure 3. Effect of interpolation  
图 3. 插值效果图

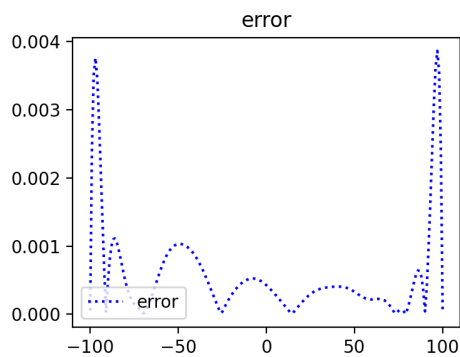


Figure 4. Error diagram  
图 4. 误差图

## 6. 结论

本文的主要工作是对高斯径向基插值形状参数的选取提供一个便捷的方法。本文首先对高斯径向基函数和梯度下降法的基本理论进行介绍,并针对目标函数、梯度方向、学习率三个方面对梯度下降法进行改进。经过数值实验,基于梯度下降算法得到的形状参数在对原函数进行插值后,通过图 1、图 3 可以看出函数的拟合效果较好。通过图 2、图 4 可以看出,插值函数与原函数之间的整体误差较小。说明改进后的梯度下降法选取高斯径向基插值形状参数的有效性。此方法的建立,为高斯径向基函数形状参数的选取提供了便捷的方法,为高斯径向基函数的实际应用推广奠定良好基础。

## 基金项目

本工作受到国家自然科学基金(No. 11601151)和河北省青年拔尖人才支持计划项目支持。

## 参考文献

- [1] 张淮清, 陈玉, 聂鑫. 基于径向基函数插值的积分方程求解[J]. 应用数学与计算数学学报, 2017, 31(3): 275-289.
- [2] 秦伶俐. 无网格方法在结构计算上的改进[D]: [博士学位论文]. 北京: 中国农业大学, 2005.
- [3] 徐咏梅, 柳桂国, 柳贺. 高斯径向基核函数参数的 GA 优化方法[J]. 电力自动化设备, 2008(6): 52-55.
- [4] 王书博. 基于改进梯度下降法求解多元线性回归方程[J]. 数学的实践与认识, 2022, 52(10): 167-172.
- [5] 孙娅楠, 林文斌. 梯度下降法在机器学习中的应用[J]. 苏州科技大学学报(自然科学版), 2018, 35(2): 26-31.
- [6] 齐静. 径向基函数插值若干问题研究[D]: [硕士学位论文]. 重庆: 重庆师范大学, 2016.