

求解区间优化问题的一种谱梯度法

郭二威

青岛大学数学与统计学院, 山东 青岛

收稿日期: 2025年6月24日; 录用日期: 2025年7月18日; 发布日期: 2025年7月25日

摘 要

本文主要研究区间约束优化问题的求解算法。基于广义Hukuhara导数的Karush-Kuhn-Tucker条件, 结合互补结构, 利用Fischer-Burmeister函数, 将原区间约束优化问题转化为无约束优化问题, 并给出一种谱梯度法求解转化后的问题。最后, 数值实验验证了算法的有效性。

关键词

区间优化, 广义Hukuhara导数, 谱梯度法, Karush-Kuhn-Tucker条件

A Kind of Spectral Gradient Method for Interval-Valued Optimization Problems

Erwei Guo

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Received: Jun. 24th, 2025; accepted: Jul. 18th, 2025; published: Jul. 25th, 2025

Abstract

This paper concentrates on designing an algorithm for solving interval constrained optimization problems. Based on the Karush-Kuhn-Tucker conditions utilizing generalized Hukuhara derivatives, and incorporating complementary structure, the Fischer-Burmeister function is employed to transform the original interval constrained optimization problem into an unconstrained one. A spectral gradient method for solving the transformed problem is presented. Numerical experiments verify the effectiveness of this algorithm.

Keywords

Interval Optimization, Generalized Hukuhara Differentiation, Spectral Gradient Method, Karush-Kuhn-Tucker Conditions



1. 引言

在数学优化领域,学者们对目标函数为区间的优化问题兴趣浓厚,而传统确定性数值估计易致误差,区间方法能有效描述不确定性数据,简化误差处理,而区间优化问题作为集值优化特例,对处理数值精度不足的数据至关重要。在应用方面,区间优化在工程领域具有重要意义,在航空航天、多属性决策、电荷输运等问题中具有广泛应用[1]-[3]。

众多学者深入研究区间优化问题。该问题的求解方法一般分为直接求解和转化法求解两类方法。直接求解法中,一般是定义新的区间序关系,并在此基础上进行直接求解。Inuiguchi 等提出基于最大最小遗憾准则的目标区间函数系数规划方法[4],并通过重复运用单纯性法得到解。Gong 等人[5]基于区间可能度定义区间数的优劣关系,定义了基于区间数的拥挤距离,并基于排序关系和拥挤距离选择最优解。在这类方法中,对区间数排序的比较策略依赖性强,不同的比较策略会影响进化算法搜索最优解的性能。而转化法则是通过不同的转化策略将不确定优化问题转化为确定性优化问题,再用现有成熟的确定性方法进行求解。Ishibuchi 等引入了区间偏序关系,将不确定区间优化问题转化为多目标确定性优化问题[6]。在文献[7]-[9]中,Wu 给出了区间函数及其 Pareto 最优解定义,并研究多目标区间优化问题的 Karush-Kuhn-Tucker 最优性条件。Sun 等定义区间规划的“LU”最优解,给出 Fritz-John 和 Kuhn-Tucker 最优性条件[10]。Villanueva 等人研究目标函数为区间值且约束函数为实值的优化问题,提出基于广义 Hukuhara 次梯度和广义 Hukuhara 方向导数的最优性条件[11]。

谱梯度法因其几乎对所有初始值都成立以及不需要求解目标函数的 Hessian 矩阵的优点被广泛使用。谱梯度法的理论雏形最早可回溯至 1988 年,由 Barzilai 和 Borwein 在研究严格凸二次规划问题时首次提出[12]。其创新性核心思想在于:利用原问题矩阵的 Rayleigh 商与单位矩阵乘积所形成的标量矩阵,构造出一种近似 Hessian 矩阵的有效策略,进而推导出两种固定形式的步长计算公式,并且证明了对于二维凸二次函数,Barzilai 和 Borwein 法超线性收敛于最优解。数值结果表明 Barzilai 和 Borwein 法优于普通的最速下降法,为后续优化算法的发展奠定了关键基础。在 1993 年,Raydan 证明,目标函数为严格凸二次函数时,Barzilai 和 Borwein 法是全局收敛的[13],而在同样前提下,Dai 和 Liao 于 2002 年证明了 Barzilai 和 Borwein 法是 R-线性收敛的[14]。Xiao 等人通过考虑一些修正的拟牛顿方程对这种步长选择给出了两个注释,并且给出了另外两种步长选择[15]。

由于问题的复杂性,对区间优化问题有效的求解算法一直较少,并且还没有学者将谱梯度法应用于区间优化问题。本文针对区间优化问题,旨在为区间优化问题的求解提供一种新思路,利用 Villanueva 等人提出的 Karush-Kuhn-Tucker 条件,引入线性互补问题概念,利用 Fischer-Burmeister 函数,将原问题转化为无约束优化问题,结合无约束优化问题的特征,提出一种谱梯度法求解,数值实验验证了转化技术与算法的有效性。

2. 预备知识

令 $\mathcal{K}_C(\mathbb{R})$ 表示所有实数紧区间的集合。若 $C \in \mathcal{K}_C(\mathbb{R})$, 则 C 可表示为 $[\underline{c}, \bar{c}]$, 其中 $\underline{c}$ 和 $\bar{c}$ 分别表示 C 的下界和上界。令 d_H 表示 Pompeiu-Hausdorff 距离, 定义为:

$$d_H(C, D) = \max \{ |\underline{c} - \underline{d}|, |\bar{c} - \bar{d}| \},$$

其中对任意 $C = [\underline{c}, \bar{c}], D = [\underline{d}, \bar{d}] \in \mathcal{K}_C(\mathbb{R})$ 。

为简化符号, 用 $[d \vee \bar{d}]$ 表示 $[\min\{\underline{d}, \bar{d}\}, \max\{\underline{d}, \bar{d}\}]$, 其中 $\underline{d}, \bar{d} \in \mathbb{R}$ 。

对 $\forall C, D \in \mathcal{K}_C(\mathbb{R})$, 定义如下运算:

- 1) $C + D = \{c + d : c \in C, d \in D\} = [\underline{c} + \underline{d}, \bar{c} + \bar{d}]$;
- 2) $-C = \{-c : c \in C\} = [-\bar{c}, -\underline{c}]$;
- 3) $C - D = \{c - d : c \in C \text{ 且 } d \in D\} = [\underline{c} - \bar{d}, \bar{c} - \underline{d}]$;
- 4) $k + C = [\underline{k} + \underline{c}, \bar{k} + \bar{c}]$ (k 为实数, 可视为区间 $[\underline{k}, \bar{k}]$);
- 5) $kC = [k\underline{c}, k\bar{c}]$ 。

由于区间不能像实数一样比较大小, 于是给出区间的偏序关系的定义[9][16]: 给定区间 $C = [\underline{c}, \bar{c}]$ 和 $D = [\underline{d}, \bar{d}] \in \mathcal{K}_C(\mathbb{R})$, 我们称 $C <_{LU} D$ 当且仅当 $\underline{c} < \underline{d}$ 且 $\bar{c} < \bar{d}$ 。

在区间值优化问题中, 区间可用中心 - 半径形式表示, 即 $A = (a^C, a^R)$, 定义如下: 中心 a^C 为区间 A 的上下界的平均值, 即: $a^C = \frac{\underline{a} + \bar{a}}{2}$, 半径 a^R 为上下界差的一半, 即: $a^R = \frac{\bar{a} - \underline{a}}{2}$ 。因此, 两种表示方法的转换可表示为: $\underline{a} = a^C - a^R$, $\bar{a} = a^C + a^R$ 。于是, 我们可以将区间 A 表示为中心 - 半径形式 $A = (a^C; a^R)$ 。

在此背景下, Stefanini 和 Arana-Jiménez [17]给出了广义 Hukuhara 可微、广义 Hukuhara 导数、广义 Hukuhara 梯度的定义如下:

定义 2.1: 设函数 $F: S \subseteq \mathbb{R}^n \rightarrow \mathcal{K}_C$, 其中:

$$F(x) = (f^C(x); f^R(x)) = [f^C(x) - f^R(x), f^C(x) + f^R(x)],$$

给定 $\delta > 0$, 对 $\forall x_0 \in S$, 满足当 $\forall h = (h_1, \dots, h_n) \in \mathbb{R}^n$, $\|h\| < \delta$ 时, 有 $x_0 + h \in S$ 。若存在向量 $w^C, w^R \in \mathbb{R}^n$, $w^C = (w_1^C, \dots, w_n^C)$, $w^R = (w_1^R, \dots, w_n^R)$, 以及函数 $\varepsilon^C, \varepsilon^R: \mathbb{R}^n \rightarrow \mathbb{R}$ 满足 $\lim_{h \rightarrow 0} \varepsilon^C(h) = \lim_{h \rightarrow 0} \varepsilon^R(h) = 0$, 使得对于所有 $h \neq 0$, 有:

$$\begin{aligned} f^C(x_0 + h) - f^C(x_0) &= \sum_{i=1}^n w_i^C h_i + \|h\| \varepsilon^C(h), \\ |f^R(x_0 + h) - f^R(x_0)| &= \left| \sum_{i=1}^n w_i^R h_i + \|h\| \varepsilon^R(h) \right|. \end{aligned}$$

则称函数 F 在 x_0 处广义 Hukuhara 可微。

定义 2.2: F 在 x_0 处的广义 Hukuhara 导数记为 $D_{gH}F(x_0): \mathbb{R}^n \rightarrow \mathcal{K}_C$, 定义如下:

$$D_{gH}F(x_0)(h) = \left(\sum_{i=1}^n w_i^C h_i, \left| \sum_{i=1}^n w_i^R h_i \right| \right), \quad \forall h = (h_1, \dots, h_n) \in \mathbb{R}^n.$$

定义 2.3: F 在 x_0 处的广义 Hukuhara 梯度记作 $\nabla_{gH}F(x_0)$, 定义为以下 n 元组区间: $\nabla_{gH}F(x_0) = (w_1, \dots, w_n)$, 其中每个分量 w_j 表示为 $w_j = (w_j^C; |w_j^R|)$, 对所有 $j \in \{1, \dots, n\}$ 。

定义 2.4: 对 $\forall x \in \mathbb{R}^n$, 若满足以下条件:

$$x \geq 0, \quad \mathcal{L}(x) \geq 0, \quad x^T \mathcal{L}(x) = 0,$$

我们称 x 和 $\mathcal{L}$ 互补, 其中 $\mathcal{L}(x): \mathbb{R}^n \rightarrow \mathbb{R}^n$ 为连续可微函数。

3. 区间优化问题

本节讨论如下的区间优化问题:

$$\begin{aligned} & \text{Minimize } F(x) = [f(x), \bar{f}(x)] \\ & \text{subject to } g_j(x) \geq 0, j \in \{1, \dots, m\}, x \in S \subseteq \mathbb{R}^n, \end{aligned} \quad (1)$$

其中, 对每个 $j \in \{1, \dots, m\}$, 都有 $g_j: S \rightarrow \mathbb{R}$, 且 S 为开集。我们用集合 $I = \{1, \dots, m\}$ 表示指标集, 并定义所有可行点的集合为:

$$\mathcal{X} = \{x \in S : g_j(x) \geq 0, \forall j \in I\}.$$

对于任意可行点 $x \in \mathcal{X}$, 在 x 处的有效约束集为:

$$I(x) = \{j \in I : g_j(x) = 0\}.$$

给定 $\varepsilon > 0$, $\forall x^* \in \mathbb{R}^n$ 的 ε -邻域表示为:

$$N_\varepsilon(x^*) = \{x \in \mathbb{R}^n : \|x - x^*\| < \varepsilon\}.$$

定义 3.1: 设 $x^* \in \mathcal{X}$, 若 $\exists \varepsilon > 0$, 对 $\forall x \in \mathcal{X} \cap N_\varepsilon(x^*)$, 都满足 $F(x) \geq_{LU} F(x^*)$, 则称 x^* 为问题(1)的局部弱 LU 解。

下面介绍 Karush-Kuhn-Tucker 最优性条件。通常, 推导 Karush-Kuhn-Tucker 条件需要特定的正则性约束条件, 我们采用正线性独立约束条件, 即 Mangasarian-Fromovitz 约束条件[18]:

定义 3.2: 若在可行解 $x^* \in \mathcal{X}$ 处, 不存在全零的非负系数 $\beta_j \geq 0$ ($j \in I(x^*)$)使得下式成立:

$$\sum_{j \in I(x^*)} \beta_j \nabla g_j(x^*) = \mathbf{0}_{\mathbb{R}^n},$$

则称问题(1)的约束条件在 x^* 处满足正线性独立约束条件(PLICQ)。其中, $\nabla g_j(x^*)$ 表示第 j 个约束函数在 x^* 处的梯度。

基于 PLICQ 以及基于广义 Hukuhara 微分给出了区间优化问题的 Karush-Kuhn-Tucker 条件[11]:

定理 3.1: 设 $x^* \in \mathcal{X}$ 为问题(3.1)的弱 LU 解。若问题(1)的约束条件在 x^* 处满足 PLICQ, 则存在 $\mu = (\mu_1, \dots, \mu_m)^T \in \mathbb{R}^m$ 使得:

$$\mathbf{0}_{\mathbb{R}^n} \in \nabla_{gH} F(x^*) \oplus \sum_{j=1}^m \mu_j \nabla g_j(x^*), \quad (2)$$

$$\mu_j g_j(x^*) = 0, \quad \forall j \in I, \quad (3)$$

$$\mu_j \geq 0, \quad \forall j \in I. \quad (4)$$

显然, KKT 条件中的条件(3)和(4)构成一个互补问题:

$$\mu^T g(x^*) = 0, \quad \mu \geq 0, \quad g(x^*) \geq 0, \quad (5)$$

其中, $\mu = (\mu_1, \dots, \mu_m)^T$ 且 $g(x) = (g_1(x), g_2(x), \dots, g_m(x))^T$ 。

针对该问题, 采用 Fischer-Burmeister (FB)函数[19]。函数 $h(a, b)$ 定义如下:

$$h(a, b) = (a + b) - \sqrt{a^2 + b^2}.$$

文献[19]证明了函数 $h(a, b)$ 是全局 Lipschitz 连续的、定向可微的和强半光滑的。于是函数 $h(a, b)$ 的广义梯度 $\partial h(a, b)$ 可表示为所有 (σ_a, τ_b) 组成的集合

$$(\sigma_a, \tau_b) = \begin{cases} \left(1 - \frac{a}{\sqrt{a^2 + b^2}}, 1 - \frac{b}{\sqrt{a^2 + b^2}}\right) & \mu_j \neq 0 \text{ 或 } g_j(x) \neq 0, \\ (1 - \zeta, 1 - \xi) & \mu_j = g_j(x) = 0, \end{cases}$$

其中 (σ_a, τ_b) 是满足 $\sigma_a^2 + \tau_b^2 = 1$ 的任意向量。

于是, 利用以上函数, 将(5)转化为:

$$H(x, \mu) = \begin{pmatrix} (\mu_1 + g_1(x)) - \sqrt{\mu_1^2 + g_1(x)^2} \\ (\mu_2 + g_2(x)) - \sqrt{\mu_2^2 + g_2(x)^2} \\ \vdots \\ (\mu_m + g_m(x)) - \sqrt{\mu_m^2 + g_m(x)^2} \end{pmatrix} = 0.$$

设 $x \in \mathcal{X}$, (5)的解 x^* 满足 $H(x^*, \mu) = 0$ 。于是, $H(x, \mu)$ 的广义梯度 $\partial H(x, \mu)$ 可表示为所有 (σ_μ, τ_g) 组成的集合:

$$(\sigma_\mu, \tau_g) = \begin{cases} \left(\nabla g_j(x) - \frac{g_j(x) \nabla g_j(x)}{\sqrt{(\mu_j)^2 + (g_j(x))^2}}, 1 - \frac{\mu_j}{\sqrt{(\mu_j)^2 + (g_j(x))^2}} \right) & \mu_j \neq 0 \text{ 或 } g_j(x) \neq 0, \\ (1 - \zeta, 1 - \xi) & \mu_j = g_j(x) = 0, \end{cases} \quad (6)$$

其中 (σ_μ, τ_g) 是满足 $\sigma_\mu^2 + \tau_g^2 = 1$ 的任意向量。

经过上述转化过程, 令 $z = \begin{pmatrix} x \\ \mu \end{pmatrix}$, 则求解问题(1)等价于求解 $\phi(z) = 0$, 其中:

$$\phi(z) = \begin{pmatrix} W \\ V \end{pmatrix}.$$

其中, $W \in \nabla_{gH} F(z) \oplus \sum_{j=1}^m \mu_j \nabla g_j(z)$, $V = H(z)$, 令 $\phi(z)$ 的价值函数为 $B(z) = \frac{1}{2} \|\phi(z)\|^2$ 。于是, 求解区间值优化问题转化为求解如下无约束优化问题:

$$\min B(z). \quad (7)$$

$B(z)$ 的梯度由下式给出:

$$\begin{aligned} \nabla B(z) &= \nabla \phi(z)^T \phi(z) \\ &= (w \quad v) \begin{pmatrix} W \\ V \end{pmatrix}, \end{aligned}$$

其中, $w \in \nabla_{gH}^2 f(z) \oplus \sum_{j=1}^m \mu_j \nabla^2 g_j(z)$, $v \in \partial H(z)$, $W \in \nabla_{gH} F(z) \oplus \sum_{j=1}^m \mu_j \nabla g_j(z)$, $V = H(z)$, 且 $\partial H(z)$ 由式(6)给出。

4. 谱梯度方法

谱梯度方法(又名两点步长法)是一种重要的梯度类优化算法, 其本质可归结为以下迭代过程:

$$x_{k+1} = x_k - \alpha_k g_k.$$

在此式中, g_k 表示目标函数 $f(x)$ 在点 x_k 处的梯度方向, α_k 则为沿负梯度方向 $-g_k$ 的步长, 该方法相较于传统共轭梯度法展现出更为显著的数值收敛优势。

以下基于文献[20], 我们给出了求解问题(7)的谱梯度方法:

算法 4.1 (谱梯度方法):

输入: $x_0 \in \mathbb{R}^n, \mu_0 \in \mathbb{R}^m, 0 < \nu < 1, 0 < \varepsilon_1 < \varepsilon_2, 0 < \beta < 1$. 令 $z_0 = \begin{pmatrix} x_0 \\ \mu_0 \end{pmatrix}, \varepsilon = 10^{-5}, k = 0, \gamma_0 = 1$.

步 1: 计算 $\|\nabla B(z_k)\|$, 若 $\|\nabla B(z_k)\| \leq \varepsilon$, 则终止算法, 否则转至步 2。

步 2: 计算下降方向 $d_k = -\frac{1}{\gamma_k} \nabla B(z_k)$ 。

步 3: 计算 $t_k = \beta^l$, 其中 l 是满足以下条件的最小非负整数:

$$B(z_k + t_k d_k) \leq B(z_k) + \nu t_k \nabla B(z_k)^T d_k. \quad (8)$$

步 4: 令 $z_{k+1} = z_k + t_k d_k$, 计算 $B(z_{k+1})$ 。

步 5: 计算公式(10)获得 γ_{k+1} , 若 $\gamma_{k+1} < \varepsilon_1$, 则设 $\gamma_{k+1} = \varepsilon_1$; 若 $\gamma_{k+1} > \varepsilon_2$, 则设 $\gamma_{k+1} = \varepsilon_2$ 。

步 6: 令 $k = k + 1$, 转至步 2。

注 1: 步 5 的目的是调整上升方向并保持序列 $\{\gamma_k\}$ 的一致有界性。事实上, 对于所有 k , 有:

$$0 < \varepsilon_1 \leq \gamma_k \leq \varepsilon_2. \quad (9)$$

注 2: 为了确定步 5 中的参数 γ_k , Zhou [20]通过将 Dennis 和 Wolkowicz [21]以及 Yuan [22]对拟牛顿方程的修改进行线性组合, 得到了参数 γ 的表达式: 设参数 θ 为: $\theta = \frac{\sigma}{\lambda + \sigma}$, 其中 $\lambda + \sigma \neq 0$

$$\gamma_{k+1}(\theta) = \frac{s_k^T y_k + \theta \left[2(B(z_k) - B(z_{k+1})) + (\nabla B(z_k) + \nabla B(z_{k+1}))^T s_k \right]}{s_k^T s_k} = \frac{s_k^T y_k + \theta \Delta_k}{s_k^T s_k}, \quad (10)$$

其中, $\Delta_k = 2(B(z_k) - B(z_{k+1})) + (\nabla B(z_k) + \nabla B(z_{k+1}))^T s_k$, $s_k = z_{k+1} - z_k$, $y_k = \nabla B(z_{k+1}) - \nabla B(z_k)$ 。

下面讨论 $\gamma_{k+1}(\theta)$ 的性质:

定理 4.1: 设 $B(z)$ 充分光滑。若 $\|s_k\|$ 足够小, 则有

$$s_k^T \nabla^2 B(z_{k+1}) s_k - \gamma_{k+1}(\theta) s_k^T s_k = \left(\frac{1}{2} - \frac{\theta}{6} \right) \nabla^3 B(z_{k+1}) \otimes s_k^3 - \left(\frac{1}{6} - \frac{\theta}{12} \right) \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5), \quad (11)$$

其中 $\otimes$ 为张量积, $\nabla^3 B(z_{k+1}) \in \mathbb{R}^{n \times n \times n}$ 和 $\nabla^4 B(z_{k+1}) \in \mathbb{R}^{n \times n \times n \times n}$ 是 $B(z)$ 在 z_{k+1} 处的张量, 满足

$$\nabla^3 B(z_{k+1}) \otimes s_k^3 = \sum_{i,j,l=1}^n \frac{\partial^3 B(z_{k+1})}{\partial z^i \partial z^j \partial z^l} s_k^i s_k^j s_k^l, \quad (12)$$

以及

$$\nabla^4 B(z_{k+1}) \otimes s_k^4 = \sum_{i,j,l,m=1}^n \frac{\partial^4 B(z_{k+1})}{\partial z^i \partial z^j \partial z^l \partial z^m} s_k^i s_k^j s_k^l s_k^m. \quad (13)$$

证明: 利用泰勒公式, 可得

$$B(z_k) = B(z_{k+1}) - \nabla B^T(z_{k+1}) s_k + \frac{1}{2} s_k^T \nabla^2 B(z_{k+1}) s_k - \frac{1}{6} \nabla^3 B(z_{k+1}) \otimes s_k^3 + \frac{1}{24} \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5),$$

以及

$$\nabla B^T(z_k) s_k = \nabla B^T(z_{k+1}) s_k - s_k^T \nabla^2 B(z_{k+1}) s_k + \frac{1}{2} \nabla^3 B(z_{k+1}) \otimes s_k^3 - \frac{1}{6} \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5).$$

由(10)、(12)和(13)式可得

$$\begin{aligned}
& s_k^T \nabla^2 B(z_{k+1}) s_k - \gamma_{k+1}(\theta) s_k^T s_k \\
&= s_k^T \nabla^2 B(z_{k+1}) s_k - s_k^T y_k - \theta \left[2(B(z_k) - B(z_{k+1})) + (\nabla B(z_k) + \nabla B(z_{k+1}))^T s_k \right] \\
&= \left(\frac{1}{2} - \frac{\theta}{6} \right) \nabla^3 B(z_{k+1}) \otimes s_k^3 - \left(\frac{1}{6} - \frac{\theta}{12} \right) \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5).
\end{aligned}$$

证明完毕。

下面考虑 θ 的几种可能的取值的情况：

当 $\theta=0$ 时，有 $\gamma_{k+1}(0) = \frac{s_k^T y_k}{s_k^T s_k}$ ，此时 γ_{k+1} 的定义与 Barzilai 和 Borwein 法的定义相同。由(11)式可得

$$s_k^T \nabla^2 B(z_{k+1}) s_k - \gamma_{k+1}(0) s_k^T s_k = \frac{1}{2} \nabla^3 B(z_{k+1}) \otimes s_k^3 - \frac{1}{6} \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5).$$

令 $\theta=1$ ，此时 $\gamma_{k+1}(1) = \frac{s_k^T y_k + \Delta_k}{s_k^T s_k}$ ，由(11)式可得

$$s_k^T \nabla^2 B(z_{k+1}) s_k - \gamma_{k+1}(1) s_k^T s_k = \frac{1}{3} \nabla^3 B(z_{k+1}) \otimes s_k^3 - \frac{1}{12} \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5).$$

令 $\theta=3$ ，此时 $\gamma_{k+1}(3) = \frac{s_k^T y_k + 3\Delta_k}{s_k^T s_k}$ 。由(11)式可得

$$s_k^T \nabla^2 B(z_{k+1}) s_k - \gamma_{k+1}(3) s_k^T s_k = \frac{1}{12} \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5).$$

求解如下极小化问题：

$$\min_{\theta} \left(\frac{1}{2} - \frac{\theta}{6} \right)^2 + \left(\frac{1}{6} - \frac{\theta}{12} \right)^2,$$

得到 $\theta = \frac{14}{5}$ ，并有 $\gamma_{k+1}\left(\frac{14}{5}\right) = \frac{s_k^T y_k + \frac{14}{5}\Delta_k}{s_k^T s_k}$ 。由(11)式可得

$$s_k^T \nabla^2 B(z_{k+1}) s_k - \gamma_{k+1}\left(\frac{14}{5}\right) s_k^T s_k = \frac{1}{30} \nabla^3 B(z_{k+1}) \otimes s_k^3 + \frac{1}{15} \nabla^4 B(z_{k+1}) \otimes s_k^4 + O(\|s_k\|^5).$$

与文献[14]和[20]类似，我们可以得到算法 4.1 是 R-线性收敛的，下面给出算法 4.1 的收敛性结果，首先给出如下假设。

假设 1: 水平集 $L(z_0) = \{x \in \mathbb{R}^n \mid B(z) \leq B(z_0)\}$ 是有界的。

引理 4.1: 当假设 1 成立时，并且序列 $\{d_k\}$ 由算法 4.1 产生，则 d_k 满足以下条件：

$$\nabla B(z_k)^T d_k \leq -c_1 \|\nabla B(z_k)\|^2, \quad (14)$$

$$\|d_k\| \leq c_2 \|\nabla B(z_k)\|. \quad (15)$$

其中， c_1 和 c_2 是两个正实常数。

证明: 先证(15)式：

对算法 4.1 中的步 2 两端求范数，可得：

$$\|d_k\| = \frac{1}{\gamma_k} \|\nabla B(z_k)\|,$$

结合(9)，有：

$$0 < \varepsilon_1 \leq \frac{\|\nabla B(z_k)\|}{\|d_k\|},$$

两边同时乘以 $\frac{\|d_k\|}{\varepsilon_1}$, 可得:

$$\|d_k\| \leq \frac{1}{\varepsilon_1} \|\nabla B(z_k)\|,$$

令 $c_2 = \frac{1}{\varepsilon_1}$ 即可。

下面证明(14)式:

对算法 4.1 中的步 2 两端左乘 $\nabla B(z_k)^T$, 并在等式两端求范数, 可得:

$$\|\nabla B(z_k)^T d_k\| = -\nabla B(z_k)^T d_k = \frac{1}{\gamma_k} \|\nabla B(z_k)\|^2,$$

结合(9), 有:

$$-\frac{\|\nabla B(z_k)\|^2}{\nabla B(z_k)^T d_k} \leq \varepsilon_2,$$

两端同时乘 $\frac{\nabla B(z_k)^T d_k}{\varepsilon_2}$, 可得

$$\nabla B(z_k)^T d_k \leq -\frac{1}{\varepsilon_2} \|\nabla B(z_k)\|^2,$$

令 $c_1 = \frac{1}{\varepsilon_2}$ 即可。

定理 4.2: 设 $\{z_k\}$ 是由算法 1 生成的序列, 则下列结论之一成立:

存在某个 k 使得 $\|\nabla B(z_k)\| = 0$;

序列 $\{z_k\}$ 的每个极限点都是稳定点。

证明: 若算法在有限次迭代内终止, 即 $\{z_k\}$ 是有限序列, 则存在 k 使得 $\|\nabla B(z_k)\| = 0$, 于是由此可得 z_k 是 $B(z)$ 的稳定点。

若 $\{z_k\}$ 是无限序列, 则存在子列 $\{z_k\}_{K_1} \subset \{z_k\}$ 使得:

$$\lim_{k \in K_1, k \rightarrow \infty} z_k = \tilde{z}.$$

根据(8)可知:

$$\lim_{k \rightarrow \infty} t_k \nabla B(z_k)^T d_k = 0,$$

结合(14)可得:

$$\lim_{k \rightarrow \infty} t_k \|\nabla B(z_k)\|^2 = 0.$$

由此可以得到两种情况:

1) 目标函数的梯度满足:

$$\lim_{k \in K_1, k \rightarrow \infty} \|\nabla B(z_k)\| = 0. \quad (16)$$

2) 存在 $K_2 \subset K_1$ 使得:

$$\lim_{k \in K_2, k \rightarrow \infty} t_k = 0.$$

对于上述两种情况, 第一种情况下由(16)可知 $\nabla B(\tilde{z}_k) = 0$ 。第二种情况下, 根据 t_k 的定义, 存在 $k \geq \tilde{k}$ 且 $k \in K_2$ 满足:

$$\begin{aligned} B\left(z_k + \frac{t_k}{\beta} d_k\right) &> B(z_k) + \mu \frac{t_k}{\beta} \nabla B(z_k)^T d_k \\ &= B\left(z_k - \frac{t_k}{\beta \gamma_k} \nabla B(z_k)\right) \\ &> B(z_k) - \mu \frac{t_k}{\beta \gamma_k} \nabla B(z_k)^T \nabla B(z_k). \end{aligned} \quad (17)$$

其中, $d_k = -\frac{1}{\gamma_k} \nabla B(z_k)$ 。利用泰勒定理展开(17)的左侧并合并同类项, 可得:

$$o(t_k) > (1-\mu) \frac{t_k}{\beta \gamma_k} \nabla B(z_k)^T \nabla_{gH} F_k = (1-\mu) \frac{t_k}{\beta \gamma_k} \|\nabla B(z_k)\|^2. \quad (18)$$

将(18)两边除以 $(1-\mu) \frac{t_k}{\beta \gamma_k}$, 并令 $k \rightarrow \infty, k \in K_2$, 可得:

$$\|\nabla B(\tilde{z}_k)\|^2 \leq 0.$$

由此可知 $\nabla B(\tilde{z}_k) = 0$, 这表明 $\tilde{z}$ 是稳定点。

5. 数值算例

为验证算法 4.1 的有效性, 本节求解文献[11]中的区间优化问题的数值例子。

算法 4.1 的终止条件设为 $\|\nabla B_r(z)\| \leq \varepsilon$ 。计算过程在 MATLAB (R2023a)上实现, 参数设置如下: $\theta = 14/5$, $\beta = 0.5$, $\nu = 0.4$, $\varepsilon = 10^{-3}$, $\varepsilon_1 = 0.01$, $\varepsilon_2 = 100$ 。向量 μ 的初始值的各分量为区间 $(0,1)$ 内的随机数。

在表 1 中, x_0 表示初始点, k 表示迭代次数, x^* 表示最优解, Val 表示 $B(z)$ 的值。

例 5.1:

$$\begin{aligned} \min \quad & [x_1 \vee 2x_2] \\ \text{s.t.} \quad & x_1^2 - 2x_1 - x_2 \leq 0, \\ & x_1^2 - 2x_1 + x_2 \leq 0, \\ & -x_1 + x_2^2 \leq 0, \\ & (x_1, x_2) \in S = \mathbb{R}^2. \end{aligned}$$

数值结果如表 1 及图 1 所示。

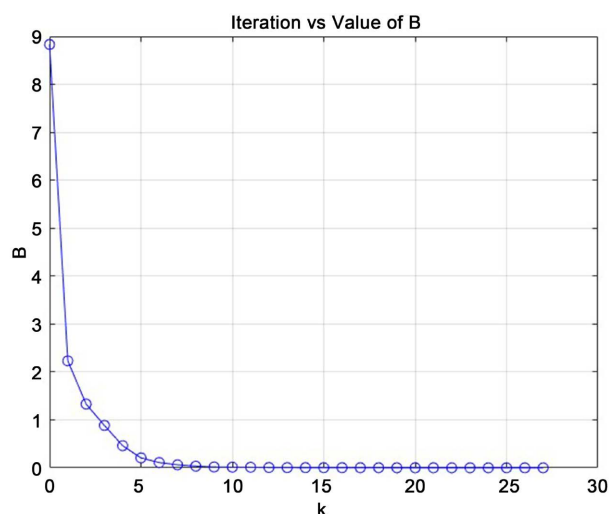
由以上数值算例结果可知, 谱梯度法在求解区间优化问题中有较好的数值表现。

6. 总结与展望

本文对区间优化问题的求解算法进行了研究, 借助区间优化问题的 KKT 条件, 我们引入了互补问题以及次梯度的概念, 利用 Fischer-Burmeister 函数, 将区间优化问题转化为无约束优化问题。根据问题结构我们给出了一类谱梯度算法, 并进行了收敛性分析, 相关的数值实验结果表明算法的有效性。但是

Table 1. The numerical results of example 5.1**表 1.** 例 5.1 的数值结果

x_0	k	x^*	Val
$(0.5383; 0.9961)^T$	46	$(1.9168 \times 10^{-6}; 2.8061 \times 10^{-4})^T$	2.9373×10^{-7}
$(0.8173; 0.8687)^T$	28	$(1.7651 \times 10^{-5}; 2.1348 \times 10^{-4})^T$	3.1067×10^{-7}
$(0.1818; 0.2638)^T$	22	$(4.2169 \times 10^{-5}; 1.7796 \times 10^{-4})^T$	1.8094×10^{-7}

**Figure 1.** Objective function value descent plot of example 5.1**图 1.** 例 5.1 的目标函数值下降图

本文主要考虑区间线性优化问题，对于区间非线性优化问题以及其他类型的区间优化问题还需要结合其他求解技术进一步研究。

参考文献

- [1] Pedroni, N. (2022) Computational Methods for the Robust Optimization of the Design of a Dynamic Aerospace System in the Presence of Aleatory and Epistemic Uncertainties. *Mechanical Systems and Signal Processing*, **164**, Article 108206. <https://doi.org/10.1016/j.ymssp.2021.108206>
- [2] Wu, Q., Liu, X., Qin, J. and Zhou, L. (2021) Multi-Criteria Group Decision-Making for Portfolio Allocation with Consensus Reaching Process under Interval Type-2 Fuzzy Environment. *Information Sciences*, **570**, 668-688. <https://doi.org/10.1016/j.ins.2021.04.096>
- [3] Safi, M.R. and Razmjoo, A. (2013) Solving Fixed Charge Transportation Problem with Interval Parameters. *Applied Mathematical Modelling*, **37**, 8341-8347. <https://doi.org/10.1016/j.apm.2013.03.053>
- [4] Inuiguchi, M. and Sakawa, M. (1995) Minimax Regret Solution to Linear Programming Problems with an Interval Objective Function. *European Journal of Operational Research*, **86**, 526-536. [https://doi.org/10.1016/0377-2217\(94\)00092-q](https://doi.org/10.1016/0377-2217(94)00092-q)
- [5] Gong, D.W., Qin, N.N. and Sun, X.Y. (2010) Evolutionary Algorithms for Multi-Objective Optimization Problems with Interval Parameters. 2010 *IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, Changsha, 23-26 September 2010, 411-420. <https://doi.org/10.1109/bicta.2010.5645160>
- [6] Ishibuchi, H. and Tanaka, H. (1990) Multiobjective Programming in Optimization of the Interval Objective Function. *European Journal of Operational Research*, **48**, 219-225. [https://doi.org/10.1016/0377-2217\(90\)90375-1](https://doi.org/10.1016/0377-2217(90)90375-1)
- [7] Wu, H. (2007) The Karush-Kuhn-Tucker Optimality Conditions in an Optimization Problem with Interval-Valued Objective Function. *European Journal of Operational Research*, **176**, 46-59. <https://doi.org/10.1016/j.ejor.2005.09.007>

- [8] Wu, H.C. (2008) Wolfe Duality for Interval-Valued Optimization. *Journal of Optimization Theory and Applications*, **138**, 497-509. <https://doi.org/10.1007/s10957-008-9396-0>
- [9] Wu, H. (2009) The Karush-Kuhn-Tucker Optimality Conditions in Multiobjective Programming Problems with Interval-Valued Objective Functions. *European Journal of Operational Research*, **196**, 49-60. <https://doi.org/10.1016/j.ejor.2008.03.012>
- [10] Sun, Y.H. and Wang, L.S. (2013) Optimality Conditions and Duality in Nondifferentiable Interval-Valued Programming. *Journal of Industrial & Management Optimization*, **9**, 131-142. <https://doi.org/10.3934/jimo.2013.9.131>
- [11] Villanueva, F.R., de Oliveira, V.A. and Costa, T.M. (2023) Optimality Conditions for Interval Valued Optimization Problems. *Fuzzy Sets and Systems*, **454**, 38-55. <https://doi.org/10.1016/j.fss.2022.06.020>
- [12] Barzilai, J. and Borwein, J.M. (1988) Two-Point Step Size Gradient Methods. *IMA Journal of Numerical Analysis*, **8**, 141-148. <https://doi.org/10.1093/imanum/8.1.141>
- [13] Raydan, M. (1993) On the Barzilai and Borwein Choice of Steplength for the Gradient Method. *IMA Journal of Numerical Analysis*, **13**, 321-326. <https://doi.org/10.1093/imanum/13.3.321>
- [14] Dai, Y.H.J. (2002) R-Linear Convergence of the Barzilai and Borwein Gradient Method. *IMA Journal of Numerical Analysis*, **22**, 1-10. <https://doi.org/10.1093/imanum/22.1.1>
- [15] Xiao, Y., Wang, Q. and Wang, D. (2010) Notes on the Dai-Yuan-Yuan Modified Spectral Gradient Method. *Journal of Computational and Applied Mathematics*, **234**, 2986-2992. <https://doi.org/10.1016/j.cam.2010.04.012>
- [16] Wu, H. (2007) The Karush-Kuhn-Tucker Optimality Conditions for the Optimization Problem with Fuzzy-Valued Objective Function. *Mathematical Methods of Operations Research*, **66**, 203-224. <https://doi.org/10.1007/s00186-007-0156-y>
- [17] Stefanini, L. and Arana-Jiménez, M. (2019) Karush-Kuhn-Tucker Conditions for Interval and Fuzzy Optimization in Several Variables under Total and Directional Generalized Differentiability. *Fuzzy Sets and Systems*, **362**, 1-34. <https://doi.org/10.1016/j.fss.2018.04.009>
- [18] Mangasarian, O.L. and Fromovitz, S. (1967) The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints. *Journal of Mathematical Analysis and Applications*, **17**, 37-47. [https://doi.org/10.1016/0022-247x\(67\)90163-1](https://doi.org/10.1016/0022-247x(67)90163-1)
- [19] Fischer, A. (1992) A Special Newton-Type Optimization Method. *Optimization*, **24**, 269-284. <https://doi.org/10.1080/02331939208843795>
- [20] 周群艳. 解无约束优化问题的一种新的谱梯度方法(英文) [J]. 应用数学, 2020, 33(4): 857-868.
- [21] Dennis, Jr. and Wolkowicz, H. (1993) Sizing and Least-Change Secant Methods. *SIAM Journal on Numerical Analysis*, **30**, 1291-1314. <https://doi.org/10.1137/0730067>
- [22] Yuan, Y. (1991) A Modified BFGS Algorithm for Unconstrained Optimization. *IMA Journal of Numerical Analysis*, **11**, 325-332.