

一类修正的共轭梯度法及其数值计算效果

陈秀芳, 李 锋*

云南师范大学数学学院, 云南 昆明
Email: 1463407198@qq.com, *lf_2364@126.com

收稿日期: 2020年8月19日; 录用日期: 2020年9月8日; 发布日期: 2020年9月15日

摘 要

给出一个新的 β_k^C 和 δ_k 公式, 新的共轭梯度法继承了FR方法的优点, 在SWP线搜索下, 具备良好的全局收敛性。在WWP线搜索下, 给出谱共轭梯度法相应的分析。数值实验表明, 本文提出的共轭梯度法数值计算性能更有效。

关键词

谱共轭梯度法, 共轭梯度法, SWP线搜索, WWP线搜索, 全局收敛性

A Kind of Modified Conjugate Gradient Method and Its Numerical Calculation Effect

Xiufang Chen, Feng Li*

School of Mathematics, Yunnan Normal University, Kunming Yunnan
Email: 1463407198@qq.com, *lf_2364@126.com

Received: Aug. 19th, 2020; accepted: Sep. 8th, 2020; published: Sep. 15th, 2020

Abstract

A new formula β_k^C and δ_k is given. The new conjugate gradient method inherits the advantages of the FR method. Under the SWP line search, it should have good global convergence. Under the WWP line search, the corresponding analysis of the spectral conjugate gradient method is given. Numerical experiments show that the numerical calculation performance of the conjugate gradient method proposed in this paper is more effective.

*通讯作者。

Keywords

Spectral Conjugate Gradient Method, Conjugate Gradient Method, SWP Line Search, WWP Line Search, Global Convergence

Copyright © 2020 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

考虑无约束优化问题:

$$\min \{f(x) : x \in \mathbb{R}^n\} \quad (1)$$

其中, 函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 的光滑函数, 由于该方法低存储性, 易于计算, 目前被广泛运用于石油勘探、大气模拟、航天航空和工程优化等领域的求解问题上[1]。一般地, 求解问题(1), 需给定初始点 $x_0 \in \mathbb{R}^n$, 共轭梯度法由最初的线性优化推广到非线性优化问题上, 用于求解大规模无约束优化问题, 其早期的迭代公式如下:

$$x_{k+1} = x_k + \alpha_k d_k \quad (2)$$

α_k 为搜索步长, d_k 为搜索方向, d_k 的定义如下:

$$d_k = \begin{cases} -g_k; k=1 \\ -g_k + \beta_k d_{k-1}; k>1 \end{cases} \quad (3)$$

其中, $g_k = \nabla f(x_k)$, 是 f 在 x_k 处的梯度, 由(3)式可知, 共轭梯度法的两个核心问题分别是步长因子 α_k 的计算和搜索方向 d_k 的选取, 需要寻找合适的步长搜索规则, 如弱 Wolfe 线搜索(WWP)、强 Wolfe 线搜索(SWP)和推广的 Wolfe 线搜索, 分别如下:

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ g_{k+1}^T d_k \geq \sigma g_k^T d_k \end{cases}$$

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ |g_{k+1}^T d_k| \leq \sigma |g_k^T d_k| \end{cases}$$

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ \sigma g_k^T d_k \leq g_{k+1}^T d_k \leq -\sigma_1 g_k^T d_k \end{cases}$$

当 $0 < \delta < \sigma < 1$ 和 $\sigma_1 \geq 0$ 是常数, β_k 为共轭参数, 常用的 β_k 公式[2] [3] [4] [5]如下:

$$\begin{aligned} \beta_k^{FR} &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \text{ (Fletcher-Reeves, 1964)} \\ \beta_k^{DY} &= \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \text{ (Dai Yuan, 1999)} \\ \beta_k^{PRP} &= \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2} \text{ (Polak-Ribiere-Polyak, 1969)} \\ \beta_k^{HS} &= -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}} \text{ (Hestenses-Stiefel, 1952)} \end{aligned} \quad (4)$$

其中, $\|\cdot\|$ 为 Euclidean 范数, $y_{k-1} = g_k - g_{k-1}$, 一般而言, 目标函数为严格凸时, PRP 方法与 HS 方法的数值计算效果优于 FR 方法和 DY 方法, 但收敛性却没有后两种方法好。当目标函数非凸时, 即使采用 Curry 原则选取步长因子, 即 α_k 为一维函数 $\Phi_k(\alpha) = \{f(x_k + \alpha d_k) : \alpha > 0\}$ 的第一个极小点, 也无法保证它们全局收敛, 以上方法在收敛性分析和数值计算性能上有所差异。近年来, 共轭梯度法的许多变种被广泛研究, 在一定程度上继承了经典方法的某些特性。2012 年, 在文献[6]的基础上, Dai 等[7]提出一种新的 DPRP 方法, 分母变成线性组合的形式, 其共轭参数 β_k^{DPRP} 定义为

$$\beta_k^{DPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\mu |g_k^T d_{k-1}| + \|g_{k-1}\|^2} \quad (5)$$

其中 $\mu > 1$, Dai 等人首先证明了该方法在不进行任何线搜索的情况下具有充分下降性, 其次证明在 Wolfe 线搜索或 Armijo 线搜索下的全局收敛性。并把这个结果扩展到 HS 方法中[8], 也是进行相应的收敛性分析, 数值实验结果表明 DPRP 方法比文献[6]中方法更有效。

2019 年, 景书杰等[9]提出一类改进的谱共轭梯度法, 新的共轭系数和谱系数公式定义为

$$\beta_k^{SN} = \frac{g_k^T \left(\frac{\|g_{k-1}\|}{\|g_k\|} g_k - g_{k-1} - d_{k-1} \right)}{\|d_{k-1}\|^2 + \mu |g_k^T d_{k-1}|}, \mu > 1 \quad (6)$$

$$\theta_k = 1 + \beta_k^{SN} \frac{g_k^T d_{k-1}}{\|g_k\|^2}$$

根据现有的共轭系数公式, 构造一个新的共轭系数公式和谱系数公式, 从而提出一种新的谱共轭梯度法, 保证算法在每次迭代时不依赖于任何线搜索总能产生充分下降的搜索方向, 并在 Armijo 线搜索和一般假设条件下, 给出算法的全局收敛性证明。将式(5)的分子或者分母推广到其它共轭梯度法上, 或者建立混合共轭梯度法, 也具备良好的数值计算效果。更多的可以参看文献[10][11][12][13][14]。

文章结构如下: 第 2 节两种新算法的提出; 第 3 节共轭梯度法及其全局收敛性; 第 4 节谱共轭梯度法及其全局收敛性; 第 5 节数值实验和第 6 节全文总结。

2. 修改的共轭梯度法和谱共轭梯度法

基于以上文献的思想, 特别是 β_k^{DPRP} 和 β_k^{SN} 的构造, 本文对文献[15]的公式做出如下的修改, 给出一个新的 β_k^C 公式

$$\beta_k^C = \frac{g_k^T \left(\frac{2\|g_{k-1}\| - \|g_k\|}{\|g_k\|} g_k - g_{k-1} - d_{k-1} \right)}{\|g_{k-1}\|^2 + \mu |g_k^T d_{k-1}| + \lambda}, \mu \geq 0, \lambda > 0 \quad (7)$$

基于公式(7), 采用 SWP 线搜索条件来修正共轭梯度算法如下:

算法 1

Step 1: 初始值 $x_0 \in R^n$, $\varepsilon > 0$, 令 $\mu \geq 0$, $\lambda > 0$, $0 < \rho < \sigma < 1$, $k := 1$, $d_1 = -g_1$, 计算 g_k , 若 $\|g_1\| \leq \varepsilon$, 则停止。

Step 2: 用 SWP 线搜索公式计算 α_k ,

Step 3: 令 $x_{k+1} = x_k + \alpha_k d_k$, 若 $\|g_{k+1}\| \leq \varepsilon$, 则停止,

Step 4: 用式(3)和(7)分别计算 d_{k+1}, β_{k+1}^C ,

Step 5: 使得 $k := k + 1$, 转 Step 2。

谱共轭梯度法作为求解问题(1)的另一方式被提出, 通过调整谱系数和共轭参数, 建立合适的算法, 本文降低线搜索条件, 故采用 WWP 线搜索条件设计谱共轭梯度算法如下:

算法 2

Step 1: 初始值 $x_0 \in R^n$, $\varepsilon > 0$, 令 $\mu \geq 0$, $\lambda > 0$, $0 < \rho < \sigma < 1$, $k := 1$, $d_1 = -g_1$, 计算 g_k , 若 $\|g_k\| \leq \varepsilon$, 则停止。

Step 2: 用 WWP 线搜索公式计算 α_k ,

Step 3: 令 $x_{k+1} = x_k + \alpha_k d_k$, 若 $\|g_{k+1}\| \leq \varepsilon$, 则停止,

Step 4: 用(7)式计算 β_{k+1}^C , 由下式计算搜索方向 d_{k+1}

$$d_{k+1} = -\delta_{k+1}g_{k+1} + \beta_{k+1}d_k$$

$$\delta_{k+1} = 1 + \beta_{k+1}^C \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}$$

Step 5: 使得 $k := k + 1$, 转 Step 2。

以下常规性假设在共轭梯度法的研究中经常被用到

假设 H [9] (H1)定义目标函数 $f(x)$ 在水平集 $\Omega = \{x \in \mathbb{R}^n; f(x) \leq f(x_0)\}$ 有下界, 且 x_0 为初始点。

(H2) 目标函数 $f(x)$ 在水平集 Ω 的一个领域 N 内连续可微, 其梯度函数满足 Lipschitz 连续, 即存常数 $L > 0$, 使得 $\|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in N$ 。

3. 共轭梯度法及其全局收敛性

引理 3.1: 若假设 H 成立, 考虑序列 $x_{k+1} = x_k + \alpha_k d_k$, 序列 $\{g_k\}$ 和 $\{d_k\}$ 由算法 1 生成, 如果

1) 对所有 k 均成立 $\|g_{k+1}\| > 3L\alpha_k \|d_k\|$

2) 存在常数 $r > 0$, 使得对任给 $k > 0$, 都有 $\|g_k\| > r$ 成立, 则存在一个充分大的 k_0 , 使得 $k > k_0$ 时,

β_{k+1}^C 满足

$$0 \leq \beta_{k+1}^C \leq \beta_{k+1}^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \tag{8}$$

证明: 因为 β_{k+1}^C 的定义为

$$\beta_{k+1}^C = \frac{g_{k+1}^T \left(\frac{2\|g_k\| - \|g_{k+1}\|}{\|g_{k+1}\|} g_{k+1} - g_k - d_k \right)}{\|g_k\|^2 + \mu |g_{k+1}^T d_k| + \lambda} \tag{9}$$

1) 证明分子 $g_{k+1}^T \left(\frac{2\|g_k\| - \|g_{k+1}\|}{\|g_{k+1}\|} g_{k+1} - g_k - d_k \right) \leq \|g_{k+1}\| 3L\alpha_k \|d_k\|$ 成立

$$g_{k+1}^T \left(\frac{2\|g_k\| - \|g_{k+1}\|}{\|g_{k+1}\|} g_{k+1} - g_k - d_k \right) \leq \|g_{k+1}\| \cdot \left\| \frac{2\|g_k\| - \|g_{k+1}\|}{\|g_{k+1}\|} g_{k+1} - g_{k+1} + g_{k+1} - g_k \right\|$$

$$\leq \|g_{k+1}\| \left(\left\| \frac{2\|g_k\| - \|g_{k+1}\|}{\|g_{k+1}\|} g_{k+1} - g_{k+1} \right\| + \|g_{k+1} - g_k\| \right)$$

$$\leq \left((2\|g_k\| - \|g_{k+1}\|) \|g_{k+1}\| - \|g_{k+1}\|^2 \right) + \|g_{k+1}\| \cdot \|g_{k+1} - g_k\|$$

$$\leq 2\|g_k - g_{k+1}\| \cdot \|g_{k+1}\| + \|g_{k+1}\| \cdot \|g_{k+1} - g_k\|$$

$$\leq 3\|g_{k+1}\| \cdot \|g_{k+1} - g_k\| \leq \|g_{k+1}\| 3L\alpha_k \|d_k\|$$

2) 将以上化简结果代入式(9)中

$$\beta_{k+1}^C = \frac{\mathbf{g}_{k+1}^T \left(\frac{2\|\mathbf{g}_k\| - \|\mathbf{g}_{k+1}\|}{\|\mathbf{g}_{k+1}\|} \mathbf{g}_{k+1} - \mathbf{g}_k - d_k \right)}{\|\mathbf{g}_k\|^2 + \mu |\mathbf{g}_{k+1}^T d_k| + \lambda} \leq \frac{\|\mathbf{g}_{k+1}\| 3L\alpha_k \|d_k\|}{\|\mathbf{g}_k\|^2 + \mu |\mathbf{g}_{k+1}^T d_k| + \lambda} \leq \frac{\|\mathbf{g}_{k+1}\| \cdot 3L\alpha_k \|d_k\|}{\|\mathbf{g}_k\|^2}$$

因为

$$0 \leq \alpha_k \|d_k\| < \frac{\|\mathbf{g}_{k+1}\|}{3L}$$

则引理 3.1 得证。

引理 3.2: 如果参数 $\sigma < \frac{1}{2}, \mu \geq 0, \lambda > 0$, 则存在常数 $c_1 = \frac{1-2\sigma+\sigma^k}{1-\sigma}, c_2 = \frac{1-\sigma^k}{1-\sigma}$ 对所有的 k , 有下式成立:

$$-c_2 \|\mathbf{g}_k\|^2 \leq \mathbf{g}_k^T d_k \leq -c_1 \|\mathbf{g}_k\|^2, \forall k \geq 1 \quad (10)$$

证明: 数学归纳法: 当 $k=1, d_1 = -\mathbf{g}_1, \mathbf{g}_1^T d_1 = -\|\mathbf{g}_1\|^2$, 则(10)式成立。

假设 $k > 1$, 都有 $\mathbf{g}_k^T d_k < 0$ 成立, 令(3)式中 k 为 $k+1$ 时与 \mathbf{g}_{k+1}^T 作内积, 结合(8)式得

$$\frac{-\mathbf{g}_{k+1}^T d_{k+1}}{\|\mathbf{g}_{k+1}\|^2} = 1 - \frac{\mathbf{g}_{k+1}^T d_k}{\|\mathbf{g}_k\|^2}$$

在上式中用

$$\sigma d_k^T \mathbf{g}_k \leq d_k^T \mathbf{g}(x_k + \alpha_k d_k) \leq -\sigma d_k^T \mathbf{g}_k$$

则有

$$1 + \sigma \frac{\mathbf{g}_k^T d_k}{\|\mathbf{g}_k\|^2} \leq \frac{-\mathbf{g}_{k+1}^T d_{k+1}}{\|\mathbf{g}_{k+1}\|^2} \leq 1 - \sigma \frac{\mathbf{g}_k^T d_k}{\|\mathbf{g}_k\|^2} \quad (11)$$

利用(11)式递推可得:

$$c_1 = \frac{1-2\sigma+\sigma^k}{1-\sigma} \leq \frac{-\mathbf{g}_{k+1}^T d_{k+1}}{\|\mathbf{g}_{k+1}\|^2} \leq \frac{1-\sigma^k}{1-\sigma} = c_2$$

由归纳法知, 上述引理成立。当 $\sigma < \frac{1}{2}$ 时, 充分下降条件对正常数 $c = \frac{1-2\sigma}{1-\sigma}$ 也成立。

由引理 3.2 和文献[1]引理 1.4.1, 可得以下引理。

引理 3.3: 若假设 H 成立, 序列 $\{x_k\}$ 由算法 1 生成, d_k 满足 $\mathbf{g}_k^T d_k < 0$, α_k 由 SWP 线搜索得到, 则 Zoutendijk 条件成立:

$$\sum_{k \geq 1} \frac{\|\mathbf{g}_k\|^4}{\|d_k\|^2} < +\infty$$

证明: 由 SWP 线搜索得 $d_k^T y_k \geq (\sigma-1)\mathbf{g}_k^T d_k$ 。

由 Lipschitz 条件得 $d_k^T y_k \leq \alpha_k L \|d_k\|^2$, 把两式结合, 则有

$$\alpha_k \geq \frac{\sigma-1}{L} \cdot \frac{\mathbf{g}_k^T d_k}{\|d_k\|^2} \quad (12)$$

由 SWP 线搜索和(12)式得

$$f_k - f_{k+1} \geq c \frac{(\mathbf{g}_k^\top d_k)^2}{\|d_k\|^2}$$

其中, $c = \frac{\delta(1-\sigma)}{L}$, 对上式从 $k=1, 2, \dots$ 求和, 由假设 H(1) 知, $\sum_{k \geq 1} \frac{(\mathbf{g}_k^\top d_k)^2}{\|d_k\|^2} < +\infty$ 成立, 由充分下降条件 $\mathbf{g}_k^\top d_k \leq -c\|\mathbf{g}_k\|^2$ 成立, 则有

$$\sum_{k \geq 1} \frac{\|\mathbf{g}_k\|^4}{\|d_k\|^2} < +\infty$$

以下是算法 1 的全局收敛性结果。

定理 1: 若假设 H 成立, 序列 $\{\mathbf{g}_k\}$ 和 $\{d_k\}$ 由算法 1 生成, 若 $\sigma < \frac{1}{2}$, 则 $\liminf_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0$ 。

证明: 反证法, 即存在常数 $r > 0$, 有 $\|\mathbf{g}_k\| \geq r, \forall k \geq 1$

$$\begin{aligned} \because d_{k+1} &= -\mathbf{g}_{k+1} + \beta_{k+1}^C d_k \\ \|d_{k+1}\|^2 &= (\beta_{k+1}^C)^2 \|d_k\|^2 - 2\beta_{k+1}^C \mathbf{g}_{k+1}^\top d_k + \|\mathbf{g}_{k+1}\|^2 \end{aligned}$$

两边同除 $\|\mathbf{g}_{k+1}\|^4$

$$\frac{\|d_{k+1}\|^2}{\|\mathbf{g}_{k+1}\|^4} = (\beta_{k+1}^C)^2 \frac{\|d_k\|^2}{\|\mathbf{g}_{k+1}\|^4} - 2\beta_{k+1}^C \frac{\mathbf{g}_{k+1}^\top d_k}{\|\mathbf{g}_{k+1}\|^4} + \frac{1}{\|\mathbf{g}_{k+1}\|^2}$$

由 SWP 线搜索、(8)式和(10)式知

$$\begin{aligned} \frac{|\mathbf{g}_{k+1}^\top d_k|}{\|\mathbf{g}_k\|^2} &\leq \frac{\sigma |\mathbf{g}_k^\top d_k|}{\|\mathbf{g}_k\|^2} \leq \frac{\sigma}{1-\sigma} \\ \frac{\|d_{k+1}\|^2}{\|\mathbf{g}_{k+1}\|^4} &\leq (\beta_{k+1}^{FR})^2 \frac{\|d_k\|^2}{\|\mathbf{g}_{k+1}\|^4} + 2\beta_{k+1}^{FR} \frac{|\mathbf{g}_{k+1}^\top d_k|}{\|\mathbf{g}_{k+1}\|^4} + \frac{1}{\|\mathbf{g}_{k+1}\|^2} \\ &\leq \frac{\|d_k\|^2}{\|\mathbf{g}_k\|^4} + 2\sigma \frac{|\mathbf{g}_k^\top d_k|}{\|\mathbf{g}_{k+1}\|^4 \cdot \|\mathbf{g}_k\|^2} + \frac{1}{\|\mathbf{g}_{k+1}\|^2} \leq \frac{\|d_k\|^2}{\|\mathbf{g}_k\|^4} + \frac{1}{\|\mathbf{g}_{k+1}\|^2} \left(1 + \frac{2\sigma}{1-\sigma}\right) \\ &\leq \frac{1+\sigma}{1-\sigma} \sum_{i=1}^{k+1} \frac{1}{\|\mathbf{g}_i\|^2} \leq \frac{1+\sigma}{1-\sigma} \frac{k+1}{r^2} \end{aligned}$$

由上式得

$$\sum_{k \geq 1} \frac{\|\mathbf{g}_{k+1}\|^4}{\|d_{k+1}\|^2} \geq \frac{1-\sigma}{1+\sigma} r^2 \sum_{k \geq 1} \frac{1}{k+1} = +\infty$$

与引理 3.3 矛盾, 故定理 1 得证。

4. 谱共轭梯度法及其全局收敛性

引理 4.1: 序列 $\{\mathbf{g}_k\}$ 和 $\{d_k\}$ 由算法 2 生成, $\forall k \geq 1$, 都有 $\mathbf{g}_k^\top d_k = -\|\mathbf{g}_k\|^2 < 0$ 。

证: 当 $k=1$ 时, $d_1 = -\mathbf{g}_1, \mathbf{g}_1^\top d_1 = -\|\mathbf{g}_1\|^2 < 0$ 成立,

$$\begin{aligned}
\mathbf{g}_k^T \mathbf{d}_k &= -\delta_k \|\mathbf{g}_k\|^2 + \beta_k^C \mathbf{g}_k^T \mathbf{d}_{k-1} \\
\text{当 } k > 1 \text{ 时,} \quad &= -\left(1 + \beta_k^C \frac{\mathbf{g}_k^T \mathbf{d}_{k-1}}{\|\mathbf{g}_k\|^2}\right) \|\mathbf{g}_k\|^2 + \beta_k^C \mathbf{g}_k^T \mathbf{d}_{k-1} \\
&= -\|\mathbf{g}_k\|^2 - \beta_k^C \mathbf{g}_k^T \mathbf{d}_{k-1} + \beta_k^C \mathbf{g}_k^T \mathbf{d}_{k-1} \\
&= -\|\mathbf{g}_k\|^2 < 0
\end{aligned}$$

故算法 2 中产生的搜索方向 \mathbf{d}_k 满足充分下降性。

引理 4.2: 若假设 H 成立, \mathbf{d}_k 满足 $\mathbf{d}_k^T \mathbf{g}_k < 0$, α_k 满足 WWP 线搜索, 则 Zoutendijk 条件成立:

$$\sum_{k \geq 1} \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\|\mathbf{d}_k\|^2} < +\infty$$

定理 2: 若假设 H 成立, 序列 $\{\mathbf{g}_k\}$ 和 $\{\mathbf{d}_k\}$ 由算法 2 生成, 则 $\liminf_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0$ 。

证明: 假设结论不成立, 则存在一个常数 $r > 0$, 使得 $\|\mathbf{g}_k\| \geq r, \forall k \geq 0$, 因为

$$\delta_{k+1} = \left(1 + \beta_{k+1}^C \frac{\mathbf{g}_{k+1}^T \mathbf{d}_k}{\|\mathbf{g}_{k+1}\|^2}\right)$$

由算法 2 中定义的搜索方向 \mathbf{d}_{k+1} , 两边取范数模平方得

$$\|\mathbf{d}_{k+1}\|^2 = (\beta_{k+1}^C)^2 \|\mathbf{d}_k\|^2 - 2\delta_{k+1} \mathbf{d}_{k+1}^T \mathbf{g}_{k+1} - \delta_{k+1}^2 \|\mathbf{g}_{k+1}\|^2$$

两端同除 $\|\mathbf{g}_{k+1}\|^4$, 又因为 $\mathbf{d}_{k+1}^T \mathbf{g}_{k+1} = -\|\mathbf{g}_{k+1}\|^2$ 。

又因为引理 3.1

$$\begin{aligned}
\frac{\|\mathbf{d}_{k+1}\|^2}{\|\mathbf{g}_{k+1}\|^4} &= (\beta_{k+1}^C)^2 \frac{\|\mathbf{d}_k\|^2}{\|\mathbf{g}_{k+1}\|^4} + \frac{2\delta_{k+1}}{\|\mathbf{g}_{k+1}\|^2} - \frac{\delta_{k+1}^2}{\|\mathbf{g}_{k+1}\|^2} \\
&\leq \left(\frac{\|\mathbf{g}_{k+1}\|^2}{\|\mathbf{g}_k\|^2}\right)^2 \frac{\|\mathbf{d}_k\|^2}{\|\mathbf{g}_{k+1}\|^4} - \frac{(\delta_{k+1}-1)^2}{\|\mathbf{g}_{k+1}\|^2} + \frac{1}{\|\mathbf{g}_{k+1}\|^2} \leq \frac{\|\mathbf{d}_k\|^2}{\|\mathbf{g}_k\|^4} + \frac{1}{\|\mathbf{g}_{k+1}\|^2}
\end{aligned}$$

当 $k=0$ 时, $\mathbf{d}_0 = -\mathbf{g}_0$, 上式写成

$$\leq \frac{1}{\|\mathbf{g}_0\|^2} + \frac{1}{\|\mathbf{g}_{k+1}\|^2} \leq \dots \leq \sum_{i=0}^{k+1} \frac{1}{\|\mathbf{g}_i\|^2} \leq \frac{k+2}{r^2}$$

所以有

$$\sum_{i=0}^k \frac{\|\mathbf{g}_{k+1}\|^4}{\|\mathbf{d}_{k+1}\|^2} \geq r^2 \sum_{i=0}^k \frac{1}{k+1} = +\infty$$

此结论与引理 4.2 的 Zoutendijk 条件矛盾, 故定理 2 成立。

5. 数值实验

本节中, 我们分析所提出新方法的数值性能, 分别与文献[15]的方法作比较, 记为 WSWP, 还有文献[16]的方法作比较, 记为 RMIL+, 本文的共轭梯度法和谱共轭梯度法分别记为 FSWP 和 PCWWP, 无约束测试问题选自文献[17], 算法运行的终止条件为 $\|\mathbf{g}_k\| \leq 10^{-6}$ 或迭代次数超过 10,000, 所有代码实现都在 Matlab2016a 中进行, 数值结果以 NI, NF, NG 的形式表示, 其中, NI 指迭代次数, NF 指函数取值,

NG 指梯度取值。“-”指算法运行失效。各符号和参数说明如下:

RMIL + SWP: RMIL + 公式用 SWP 线搜索, $\sigma = 0.1, \delta = 0.01$;

WSPW: 韦等提出的公式用 SWP 线搜索, $\sigma = 0.1, \delta = 0.01$;

FSWP: 本文提出的共轭梯度法用 SWP 线搜索, $\sigma = 0.2, \delta = 0.01, \mu = 4.5, \lambda = 0.2$;

PCWWP: 本文提出的谱共轭梯度法用 WWP 线搜索, $\sigma = 0.1, \delta = 0.01, \mu = 4.5, \lambda = 0$ 。

Table 1. Numerical results of each method

表 1. 各方法的数值计算结果

问题名称	维数 n	RMIL + SWP NI/NF/NG	WSPW NI/NF/NG	FSWP NI/NF/NG	PCWWP NI/NF/NG
ROSE	2	38/438/85	1624/7729/2158	4/107/8	260/816/372
FROYH	2	21/97/36	4/61/10	3/103/7	356/1079/521
BADSCP	2	-	-	5/155/64	-
BADSCB	2	31/254/48	2/53/4	2/53/4	-
BEALE	2	35/254/48	3/102/5	7/65/11	102/190/148
JENSAM	2	2/101/6	27/154/42	6/256/93	64/166/91
HELIX	3	482/2258/831	-	6/159/57	478/1232/664
BARD	3	-	-	6/158/56	1150/2982/160
GAUSS	3	4/9/5	4/9/5	3/103/4	24/93/34
MEYER	3	-	6/109/55	6/109/55	-
CULF	3	2/52/3	2/52/3	1/2/2	3/58/4
BOX	3	1/51/2	1/51/2	1/51/2	1/51/2
SING	4	-	4/62/11	4/60/10	-
WOOD	4	350/1299/570	7/121/62	6/68/10	234/573/334
KOWOSB	4	1345/3954/213	-	6/61/7	735/1906/1027
BD	4	72/432/197	5/158/59	4/106/8	119/404/251
OSB1	5	1/51/2	1/51/2	1/51/2	1/51/2
BIGGS	6	5/109/9	174/1075/226	5/156/43	54/141/76
OSB2	11	-	-	3/103/4	293/863/410
OSB2	20	-	3/104/5	4/154/54	-
ROSEX	8	47/227/103	193/1159/316	4/107/7	1056/2873/149
ROSEX	50	61/532/113	4/48/7	7/207/105	962/2793/1343
ROSEX	100	50/410/95	4/58/9	4/105/7	578/1441/840
SINGX	4	-	4/62/11	4/60/10	-
PEN1	2	43/590/96	-	5/63/11	1500/3882/210
PEN2	4	-	55/798/103	25/1202/74	-
PEN2	50	230/1260/388	4/111/11	8/72/18	2052/4935/290
VARDIM	2	4/12/9	4/12/9	4/11/8	4/12/9
VARDIM	50	10/52/36	10/52/36	15/55/41	10/52/36
TRIG	3	17/137/29	52/248/100	3/106/8	35/67/51
TRIG	50	54/306/104	178/1434/258	4/158/11	101/388/155

Continued

TRIG	100	46/326/120	207/2034/271	3/108/9	95/373/140
BV	3	17/35/24	88/176/171	6/61/7	33/59/44
BV	10	921/2799/1670	-	3/103/5	313/832/432
IE	3	6/14/8	8/19/12	8/161/91	18/34/23
IE	50	7/15/8	13/30/20	5/156/55	20/36/25
IE	100	7/15/9	14/30/20	5/156/55	22/40/27
IE	200	7/63/9	13/123/22	5/202/96	19/33/24
IE	500	7/15/9	45/186/73	3/103/6	21/41/28
TRID	3	14/80/18	5/110/9	6/158/53	57/254/86
TRID	50	29/109/33	4/106/9	5/109/8	53/149/74
TRID	100	31/69/37	4/108/8	8/258/154	54/99/72
TRID	200	31/68/38	-	6/159/58	56/156/80
BAND	3	8/66/13	3/102/7	4/152/8	29/51/33
BAND	50	21/679/30	4/116/11	5/203/12	21/523/27
BAND	100	21/725/29	4/116/11	4/110/11	25/531/36
BAND	200	21/872/28	4/116/11	4/110/11	34/1174/41
LIN	2	1/3/3	1/3/3	1/3/3	1/3/3
LIN	50	1/3/3	1/3/3	1/3/3	1/3/3
LIN	500	1/3/3	1/3/3	1/3/3	1/3/3
LIN	1000	1/3/3	1/3/3	1/3/3	1/3/3
LIN1	2	1/51/2	1/51/2	1/51/2	1/3/3
LIN1	10	1/3/3	1/3/3	1/3/3	1/3/3

从表 1 中可以看出, 基于新的参数公式, 本文提出的共轭梯度法(FSWP)的数值计算性能优于表中其它方法, 能解决绝大部分测试问题, 相反, 本文提出的谱共轭梯度法收敛性条件减弱, 但数值计算性能(PCWWP)却没有共轭梯度法的好, 如何调整谱共轭梯度法的两个参数, 使之达到更好的计算效果, 仍需进一步研究。

6. 总结

本文借鉴已有文献的思路, 对经典的共轭梯度法中的共轭参数作简单的分解或变形, 而不改变原公式的大致结构, 从而提出类似的 β_k 新参数公式, 并分别在 SWP 线搜索下建立新的共轭梯度法的全局收敛性分析和 WWP 线搜索下建立谱共轭梯度法的全局收敛性分析, 数值实验与经典方法作比较, 本文提出的共轭梯度法更有效。

参考文献

- [1] 戴馥虹, 袁亚湘. 非线性共轭梯度法[M]. 上海: 上海科技出版社, 1999.
- [2] Fletcher, R. and Reeves, C.M. (1964) Function Minimization by Conjugate Gradients. *The Computer Journal*, 7, 149-154. <https://doi.org/10.1093/comjnl/7.2.149>
- [3] Dai, Y.H. and Yuan, Y. (1999) A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property.

- SIAM Journal on Optimization*, **10**, 177-182. <https://doi.org/10.1137/S1052623497318992>
- [4] Polak, E. and Ribière, G. (1969) Note sur la convergence de directions conjuguées. *Rev. Française Informat Recherche Opérationnelle, 3e Année e*, **16**, 35-43. <https://doi.org/10.1051/m2an/196903R100351>
- [5] Hestenes, M.R. and Stiefel, E. (1952) Methods of Conjugate Gradients for Solving Linear Systems. *Journal of Research of the National Bureau of Standards*, **49**, 409-436. <https://doi.org/10.6028/jres.049.044>
- [6] Zhang, L. (2009) An Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method for Optimization Computation. *Applied Mathematics and Computation*, **215**, 2269-2274. <https://doi.org/10.1016/j.amc.2009.08.016>
- [7] Dai, Z. and Wen, F. (2012) Another Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method with Sufficient Descent Property. *Applied Mathematics and Computation*, **218**, 7421-7430. <https://doi.org/10.1016/j.amc.2011.12.091>
- [8] 姜彬. 基于 Armijo 搜索的谱共轭梯度法研究[D]: [硕士学位论文]. 太原: 太原科技大学, 2013.
- [9] 景书杰, 李亚敏, 牛海峰. 一类基于 Armijo 线搜索的新的谱共轭梯度法[J]. 河南理工大学学报(自然科学版), 2019, 38(4): 154-158.
- [10] Du, X., Zhang, P. and Ma, W. (2016) Some Modified Conjugate Gradient Methods for Unconstrained Optimization. *Journal of Computational and Applied Mathematics*, **305**, 92-114. <https://doi.org/10.1016/j.cam.2016.04.004>
- [11] Sellami, B., Belloufi, M. and Chaib, Y. (2017) Globally Convergence of Nonlinear Conjugate Gradient Method for Unconstrained Optimization. *RAIRO—Operations Research*, **51**, 1101-1117. <https://doi.org/10.1051/ro/2017028>
- [12] Zheng, X. and Shi, J. (2018) A Modified Sufficient Descent Polak-Ribière-Polyak Type Conjugate Gradient Method for Unconstrained Optimization Problems. *Algorithms*, **11**, 133. <https://doi.org/10.3390/a11090133>
- [13] (2018) Broyden-Fletcher-Goldfarb-Shanno Conjugate Gradient Methods. *Journal of Optimization Theory and Applications*, **178**, 860-884. <https://doi.org/10.1007/s10957-018-1324-3>
- [14] Mtagulwa, P. and Kaelo, P. (2018) A Convergent Modified HS-DY Hybrid Conjugate Gradient Method for Unconstrained Optimization Problems. *Journal of Information and Optimization Sciences*, **40**, 1-18. <https://doi.org/10.1080/02522667.2018.1424087>
- [15] 韦春妙, 庞建华, 黄李韦, 罗杰明. 新的 Armijo 线搜索下 PRP 共轭梯度法及其收敛性分析[J]. 广西科技大学学报, 2019, 30(2): 107-114.
- [16] Rivaie, M., Mamat, M. and Abashar, A. (2015) A New Class of Nonlinear Conjugate Gradient Coefficients with Exact and Inexact Line Searches. *Applied Mathematics and Computation*, **268**, 1152-1163. <https://doi.org/10.1016/j.amc.2015.07.019>
- [17] Moré, J., Garbow, B.S. and Hillstrome, K.E. (1982) Testing Unconstrained Optimization Software. *ACM Transactions on Mathematical Software*, **7**, 17-41. <https://doi.org/10.1145/355934.355936>