

具有充分下降性的改进FR共轭梯度法

马国栋, 江羨珍, 靳文慧

(广西民族大学数学与物理学院, 广西 南宁 530006)

摘要: 本文研究了大规模无约束优化问题, 提出了一个基于改进的FR共轭参数公式的共轭梯度法. 不依赖于任何线搜索准则, 算法所产生的搜索方向总是充分下降的. 在标准Wolfe线搜索准则下, 获得了新算法的全局收敛性. 最后, 对所提出的算法进行了初步数值实验, 其结果表明所改进的方法是有效的.

关键词: 无约束优化; 共轭梯度法; 标准Wolfe线搜索; 全局收敛性

MR(2010)主题分类号: 90C30; 65K05 中图分类号: O221.2

文献标识码: A 文章编号: 0255-7797(2021)03-0212-07

1 引言

共轭梯度法是求解大规模光滑无约束优化问题 $\min\{f(x) | x \in R^n\}$ 的有效方法之一, 其具有算法结构简单, 无需矩阵存储和二次终止性等优点, 特别适用于求解大规模优化问题, 其迭代公式一般形如

$$x_{k+1} = x_k + \alpha_k d_k, \quad (1.1)$$

$$d_k = \begin{cases} -g_k, & k = 1; \\ -g_k + \beta_k d_{k-1}, & k \geq 2, \end{cases} \quad (1.2)$$

其中 $g_k = \nabla f(x_k)$, α_k 为搜索步长, d_k 为搜索方向, β_k 为搜索方向 d_k 的调控参数. 不同的 β_k 对应不同的共轭梯度法, 著名的 β_k 计算公式有:

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T (g_k - g_{k-1})}, \quad \beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2},$$
$$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2}, \quad \beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T (g_k - g_{k-1})}.$$

与上述四个公式相对应的共轭梯度法分别称为HS方法[1]、FR方法[2]、PRP方法[3, 4]和DY方法[5]. 其中, FR方法和DY方法具有良好的收敛性质, 而HS方法和PRP方法则数值性能优良. 因此, 为寻求收敛性和数值效果都理想的方法, 基于上述调控参数所建立的共轭梯度法已被

*收稿日期: 2020-07-05 接收日期: 2020-11-16

基金项目: 广西自然科学基金资助 (2018GXNSFAA281099); 国家自然科学基金资助 (11771383).

作者简介: 马国栋(1983-), 男, 湖南邵阳, 副教授, 主要研究方向: 最优化理论与方法.

通讯作者: 靳文慧.

广泛研究, 且获得很多理论性质和数值效果均优良的改进共轭梯度法[6-14]. 如, 文[9]中的算法改进了DY公式

$$\beta_k^{JMJ} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} |g_k^T d_{k-1}|}{d_{k-1}^T (g_k - g_{k-1})}, \quad (1.3)$$

在标准Wolfe线搜索准则下获得了算法的全局收敛性, 且具有较为理想的数值效果. 基于文[9]的思想, 文[10]修正了FR共轭梯度参数公式

$$\beta_k^N = \frac{g_k^T (g_k - \frac{\|g_k\|}{\|d_{k-1}\|} d_{k-1})}{\|g_{k-1}\|^2}, \quad (1.4)$$

由式(1.4)产生的搜索方向在强Wolfe线搜索准则下满足充分下降条件, 然而由此共轭梯度参数公式所建立的算法未能获得全局收敛性. 鉴于此, 受文[9, 10]思想的启发, 本文将公式(1.4)进一步修正为:

$$\beta_k^{MJJ} = \frac{\|g_k\|^2 - \frac{(g_k^T d_{k-1})^2}{\|d_{k-1}\|^2}}{\|g_{k-1}\|^2 + u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}}, \quad (1.5)$$

其中参数 $u > 1$. 显然, 在精确线搜索条件下, 且目标函数为严格凸二次函数, 则式(1.5)退化为FR共轭梯度参数公式.

本文余下部分组织如下: 基于公式(1.5), 第二部分建立了算法框架, 不依赖于任何线搜索条件, 以新公式为方向调控参数所产生的搜索方向总是充分下降的. 第三部分在标准Wolfe非精确线搜索条件下, 获得了算法的全局收敛性; 最后对所提出的算法进行数值测试, 并与同类算法进行比对.

2 算法框架及充分下降性

基于本文所改进的共轭参数公式(1.5), 建立算法框架A:

初始化 任取初始点 $x_1 \in R^n$, 给定精度 $\varepsilon > 0$, 参数 $u > 1$, $d_1 = -g_1, k = 1$.

步骤1 若 $\|g_k\| < \varepsilon$, 则停止. 否则, 转步骤2.

步骤2 由某个非精确线搜索准则求步长 α_k .

步骤3 令 $x_{k+1} = x_k + \alpha_k d_k$, 由式(1.5)计算 β_{k+1}^{MJJ} . 令 $d_{k+1} = -g_{k+1} + \beta_{k+1}^{MJJ} d_k$, $k := k+1$, 返回步骤1.

首先, 给出共轭参数公式(1.5)和算法框架A所产生的搜索方向具有如下的性质.

引理 2.1 对于任意 $k \geq 1$, 总有 $0 \leq \beta_k^{MJJ} \leq \beta_k^{FR}$ 恒成立.

证 设 θ_k 为 g_k^T 和 d_{k-1} 的夹角, 有

$$|\cos\theta_k|^2 = \frac{|g_k^T d_{k-1}|^2}{\|g_k\|^2 \|d_{k-1}\|^2}.$$

于是

$$\beta_k^{MJJ} = \frac{\|g_k\|^2 [1 - \frac{(g_k^T d_{k-1})^2}{\|g_k\|^2 \|d_{k-1}\|^2}]}{\|g_{k-1}\|^2 + u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}} = \frac{\|g_k\|^2 (1 - |\cos\theta_k|^2)}{\|g_{k-1}\|^2 + u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}},$$

进而, 可得

$$0 \leq \beta_k^{MJJ} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2 + u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \beta_k^{FR}.$$

故引理获证.

引理 2.2 对所有的 $k \geq 1$ 和 $u > 1$, 由算法框架A所产生的搜索方向 d_k 满足

$$g_k^T d_k \leq -\left(1 - \frac{1}{u}\right) \|g_k\|^2,$$

即方向 d_k 是充分下降的.

证 由 β_k^{MJJ} 的定义, 易知

$$0 \leq \beta_k^{MJJ} = \frac{\|g_k\|^2 - \frac{(g_k^T d_{k-1})^2}{\|d_{k-1}\|^2}}{\|g_{k-1}\|^2 + u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}} \leq \frac{\|g_k\|^2}{u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}}. \quad (2.1)$$

将(1.2)两端与 g_k 作内积, 利用(2.1), 得

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{MJJ} d_{k-1}) \leq -\|g_k\|^2 + \beta_k^{MJJ} |g_k^T d_{k-1}| \\ &\leq -\|g_k\|^2 + \frac{\|g_k\|^2}{u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}} |g_k^T d_{k-1}| \\ &\leq -\left(1 - \frac{1}{u}\right) \|g_k\|^2. \end{aligned}$$

因此, 算法框架A所产生的搜索方向是充分下降的. 故引理获证.

由引理2.2可知, 不依赖于任何线搜索条件, 以新公式为方向调控参数所产生的搜索方向总是充分下降的.

为了便于后续的描述, 算法框架A采用如下标准 *Wolfe* 非精确线搜索准则求步长所产生的算法称为算法MJJ, 即步长 α_k 满足:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (2.2)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \quad (2.3)$$

其中参数 $0 < \delta < \sigma < 1$.

3 算法的全局收敛性

为获得算法MJJ的全局收敛性, 首先给出算法所需的两个常规假设条件.

(A1) 目标函数 $f(x)$ 在水平集 $\Lambda = \{x \in R^n | f(x) \leq f(x_1)\}$ 上有下界, 其中 x_1 为算法的初始点.

(A2) 目标函数 $f(x)$ 在水平集 Λ 的某邻域 U 内可微, 其梯度函数 $g(x) = \nabla f(x)$ 满足 Lipschitz 条件, 即存在常数 $L > 0$, 使 $\|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in U$.

下面引理给出著名的 Zoutendijk 条件, 其在共轭梯度法全局收敛性分析中起着非常重要的作用.

引理 3.1 设假设A1和A2成立, 考虑一般迭代方法(1.1)-(1.2), 若搜索方向 d_k 满足 $g_k^T d_k < 0$, 步长因子 α_k 满足标准Wolfe线搜索准则, 则 $\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$ 成立. 特别地, 若方向 d_k 是充分下降的, 有 $\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$ 成立.

由引理2.1, 引理2.2和引理3.1可建立算法MJJ的全局收敛性定理.

定理 3.2 设假设A1和A2成立, 点列 $\{x_k\}$ 由(1.1)和(1.2)产生, 步长 α_k 满足标准Wolfe非精确线搜索准则(2.2)-(2.3), 则 $\liminf_{k \rightarrow \infty} \|g_k\| = 0$, 即算法MJJ是全局收敛的.

证 由反证法. 若定理3.2不成立, 不妨假设存在常数 $\varepsilon > 0$, 使得 $\|g_k\|^2 \geq \varepsilon, \forall k \geq 0$. 将式(1.2)两端取平方, 得

$$\|d_k\|^2 = \|g_k\|^2 - 2\beta_k^{MJJ} g_k^T d_{k-1} + (\beta_k^{MJJ})^2 \|d_{k-1}\|^2. \quad (3.1)$$

由式(2.1)可得

$$-2\beta_k^{MJJ} g_k^T d_{k-1} \leq 2|\beta_k^{MJJ}| |g_k^T d_{k-1}| \leq \frac{2\|g_k\|^2}{u \max\{|g_k^T d_{k-1}|, |g_k^T g_{k-1}|\}} |g_k^T d_{k-1}| \leq \frac{2\|g_k\|^2}{u}.$$

于是, 将上式代入(3.1), 并利用引理2.1, 有

$$\|d_k\|^2 \leq \|g_k\|^2 + \frac{2\|g_k\|^2}{u} + (\beta_k^{FR})^2 \|d_{k-1}\|^2 = \|g_k\|^2 \left(1 + \frac{2}{u}\right) + \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_{k-1}\|^2. \quad (3.2)$$

将式(3.2)两端同除以 $\|g_k\|^4$, 得

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{1}{\|g_k\|^2} \left(1 + \frac{2}{u}\right) + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4}. \quad (3.3)$$

注意到式(1.2)及 $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$, 由(3.3)立得

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \left(1 + \frac{2}{u}\right) \sum_{i=1}^k \frac{1}{\|g_i\|^2}.$$

又由 $\|g_k\|^2 \geq \varepsilon$, 有 $\frac{\|g_k\|^4}{\|d_k\|^2} \geq \frac{\varepsilon}{k} \frac{u}{u+2}$. 进而, 立得 $\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = +\infty$, 这与引理3.1矛盾. 故定理获证.

4 数值试验

为检验本文所提出的算法实际数值效果, 数值试验共测试了43个问题, 所有算例均取自于无约束优化测试问题集[15], 并与文[9]中的MJJ方法, 文[10]中的NJJ方法, 文[2]中的FR方法及文[6]中的PRP+方法进行比较. 所有测试都采用标准Wolfe线搜索准则(2.2)和(2.3)获得步长. 测试的环境为MATLAB R2017b, Windows 10操作系统, 计算机硬件为Inter(R) Core(TM) i5-8250U CPU 1.80 GHz和8 GB RAM. 相关参数选取如下: $\delta = 0.01, \sigma = 0.1, u = 2.5$. 算法的终止准则为以下两种情形之一: (1) $\|g_k\| < 10^{-5}$; (2) 迭代次数 $\text{Itr} > 2000$. 终止准则(2)出现时认为该方法对这个数值例子失效, 并记为“F”.

在试验中, 对所测试算法的迭代次数(Itr), 目标函数函数值计算次数(NF), 梯度计算次数(NG), 计算时间(Tcpu)(单位为秒)及算法终止时目标函数梯度的2-范数($\|g_*\|$)等5个

重要指标进行观测和比较, 数值结果详见表1. 为进一步直观刻划试验效果, 我们还采用Dolan和Moré[16]性能图对试验效果进行比对. 图1-4分别就算法MJJ、算法JMJ、算法NJJ、算法FR及算法PRP+在迭代次数、函数值计算次数、梯度计算次数及计算时间进行比较. 从图1-4可以更直观地看出, 算法MJJ在所考查的4个指标中均明显优于其他4个算法. 因此, 算法MJJ不仅在解决问题的个数上占据优势, 并且鲁棒性最优. 综上所述, 本文所提出的算法是有效的.

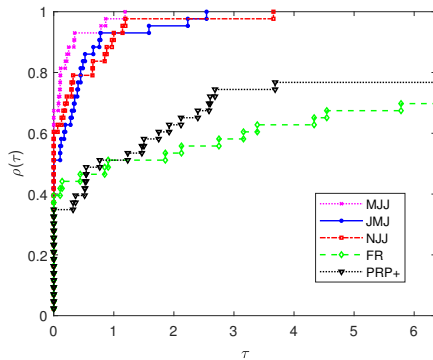


图1 迭代次数比较

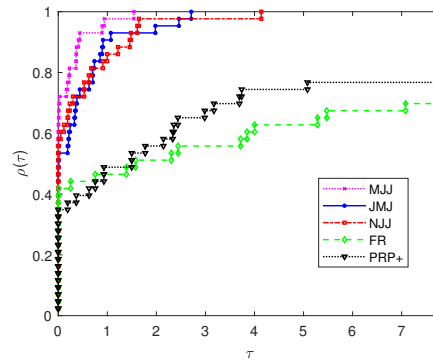


图2 函数值计算次数比较

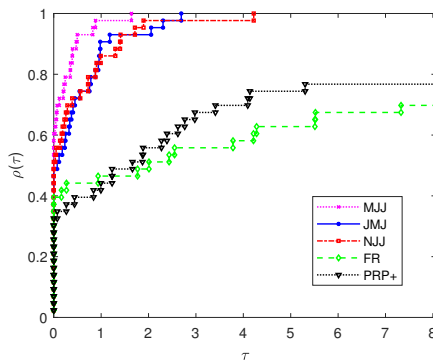


图3 梯度计算次数比较

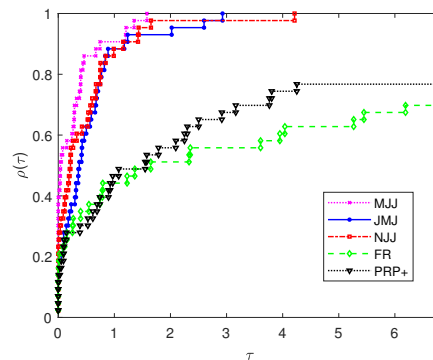


图4 计算时间比较

表1 数值试验报告

	MJJ	JMJ	NJJ	FR	PRP+
问题/n	Iter/NF/NG/Tcpu //g _r	Iter/NF/NG/Tcpu //g _r	Iter/NF/NG/Tcpu //g _r	Iter/NF/NG/Tcpu //g _r	Iter/NF/NG/Tcpu //g _r
bdexp 10	3/1/3/0.000/2.71e-49	3/1/3/0.000/5.63e-49	3/1/3/0.000/1.27e-57	3/1/3/0.000/1.93e-53	3/1/3/0.000/2.50e-48
bdexp 100	3/2/3/0.000/1.18e-82	3/2/3/0.001/1.24e-82	3/2/3/0.000/4.38e-84	3/2/3/0.000/2.19e-83	3/2/3/0.000/1.33e-82
bdexp 1000	3/2/3/0.001/4.44e-107	3/2/3/0.001/4.45e-107	3/2/3/0.001/2.84e-107	3/2/3/0.001/3.52e-107	3/2/3/0.001/4.45e-107
bdexp 10000	3/2/3/0.005/1.14e-109	3/2/3/0.008/1.14e-109	3/2/3/0.007/1.08e-109	3/2/3/0.005/1.11e-109	3/2/3/0.005/1.14e-109
bdexp 20000	3/2/3/0.010/1.07e-109	3/2/3/0.010/1.07e-109	3/2/3/0.010/1.05e-109	3/2/3/0.009/1.06e-109	3/2/3/0.012/1.07e-109
himmelbg 200	3/6/7/0.001/7.11e-29	3/6/7/0.001/7.12e-29	3/6/7/0.001/6.77e-29	3/6/7/0.001/6.94e-29	3/6/7/0.001/7.14e-29
himmelbg 1000	3/6/7/0.002/1.59e-28	3/6/7/0.002/1.59e-28	3/6/7/0.002/1.51e-28	3/6/7/0.002/1.55e-28	3/6/7/0.001/1.60e-28
himmelbg 2000	3/6/7/0.002/2.25e-28	3/6/7/0.002/2.25e-28	3/6/7/0.002/2.14e-28	3/6/7/0.002/2.19e-28	3/6/7/0.001/2.26e-28
himmelbg 5000	3/6/7/0.003/3.56e-28	3/6/7/0.004/3.56e-28	3/6/7/0.005/3.39e-28	3/6/7/0.005/3.47e-28	3/6/7/0.006/3.57e-28
genquartic 1000	15/169/71/0.020/7.28e-06	21/355/161/0.028/5.20e-06	27/528/264/0.034/5.10e-06	825/22847/11405/1.456/9.34e-06	193/5713/2806/0.380/4.52e-06
genquartic 1500	23/356/169/0.027/1.90e-06	33/643/311/0.051/7.09e-07	45/977/449/0.074/2.49e-06	537/15802/7761/1.196/8.55e-06	54/1216/619/0.095/6.27e-06
biggab1 5	38/872/415/0.043/3.33e-06	60/1453/714/0.070/2.67e-06	70/1741/825/0.086/2.44e-06	398/11468/5710/0.526/9.99e-06	165/4458/2173/0.206/1.12e-06
biggab1 10	52/1196/542/0.057/9.94e-07	88/2225/1052/0.102/2.58e-06	115/2783/1334/0.128/6.01e-06	227/6502/3172/0.293/7.70e-06	197/5285/2619/0.245/9.47e-06
fletcbv3 10	1/1/1/0.000/5.97e-06	1/1/1/0.000/5.97e-06	1/1/1/0.000/5.97e-06	1/1/1/0.000/5.97e-06	1/1/1/0.000/5.97e-06
nonscomp 50	47/887/414/0.042/5.24e-06	53/1047/487/0.052/2.35e-06	593/15629/7714/0.773/7.84e-06	F/F/F/F/F	F/F/F/F/F
dixmaana 1500	16/266/117/0.229/3.88e-06	17/260/118/0.222/3.02e-06	14/207/89/0.173/8.73e-06	125/3316/1652/2.850/1.87e-06	47/1121/547/0.983/7.68e-06
dixmaab 1500	11/140/49/0.112/2.90e-06	11/140/53/0.113/2.90e-06	11/140/55/0.115/2.90e-06	74/2015/952/1.736/5.65e-07	66/1811/863/1.526/5.31e-06
dixmaanc 1500	19/327/139/0.274/3.51e-06	21/363/165/0.309/1.82e-06	16/279/121/0.232/3.67e-06	58/1378/654/1.165/8.55e-06	20/360/164/0.302/2.88e-06
dixmaand 1500	14/184/72/0.152/6.33e-06	16/214/90/0.179/4.25e-06	28/562/236/0.479/6.04e-06	F/F/F/F/F	90/2456/1229/2.129/5.37e-06
dqdrtic 1000	173/4509/2021/0.278/3.87e-06	554/15900/7084/1.006/9.16e-06	95/2428/1097/0.166/2.99e-06	F/F/F/F/F	269/6858/3468/0.495/8.88e-06
dqdrtic 3000	238/6667/2993/0.812/6.39e-06	648/19134/8363/2.687/7.49e-06	151/3769/1731/0.473/1.58e-06	152/4145/2037/0.425/8.57e-06	138/3480/1691/0.353/8.54e-06
dqrtc 50	18/230/92/0.016/6.86e-06	17/202/72/0.012/5.09e-06	19/209/86/0.013/1.31e-06	F/F/F/F/F	47/1018/489/0.061/3.53e-06
dqrtc 100	22/296/115/0.024/2.25e-06	30/482/227/0.039/6.40e-06	27/427/190/0.034/7.88e-06	F/F/F/F/F	32/564/270/0.047/7.40e-07
edensch 100	43/896/395/0.070/4.70e-06	53/1147/576/0.094/3.38e-06	45/1054/482/0.082/7.38e-07	F/F/F/F/F	F/F/F/F/F
edensch 200	38/836/379/0.100/2.72e-06	36/831/368/0.097/8.20e-06	41/954/430/0.112/1.75e-06	F/F/F/F/F	F/F/F/F/F
edensch 1000	37/726/311/0.303/4.07e-06	49/1121/525/0.456/8.17e-07	43/893/407/0.364/9.54e-06	F/F/F/F/F	F/F/F/F/F
fletcher 100	145/3951/1768/0.177/5.66e-06	146/4018/1823/0.180/8.84e-06	134/3521/1733/0.159/3.73e-06	F/F/F/F/F	F/F/F/F/F
liarwhd 20	130/3516/1574/0.173/8.75e-07	130/3464/1612/0.164/8.04e-06	161/4205/2002/0.212/4.95e-06	103/2732/1357/0.133/6.80e-06	F/F/F/F/F
penalty1 1000	14/243/93/0.718/2.34e-06	14/243/93/0.717/2.34e-06	14/243/93/0.711/2.34e-06	19/406/178/1.232/9.37e-07	14/241/98/0.725/4.82e-07
penalty1 2000	10/129/39/1.287/5.53e-06	10/129/39/1.284/5.53e-06	10/129/39/1.238/5.53e-06	18/385/157/3.881/3.60e-08	17/364/144/3.624/6.45e-06
quarte 20	16/222/85/0.011/3.86e-06	16/218/81/0.011/2.11e-06	20/347/149/0.017/1.05e-06	30/571/276/0.028/6.25e-06	23/337/145/0.017/4.72e-06
quarte 100	22/296/115/0.024/2.25e-06	30/482/227/0.041/6.40e-06	27/427/190/0.035/7.88e-06	F/F/F/F/F	32/564/270/0.045/7.40e-07
raydan2 1000	11/173/64/0.013/6.93e-06	11/173/59/0.013/6.97e-06	11/173/59/0.013/6.97e-06	11/173/66/0.014/6.95e-06	F/F/F/F/F
raydan2 7000	14/237/100/0.078/6.45e-06	14/237/91/0.115/4.29e-07	13/235/82/0.113/8.25e-06	14/237/72/0.103/2.29e-06	F/F/F/F/F
raydan2 10000	14/266/117/0.168/3.35e-06	14/238/94/0.177/5.55e-07	13/235/91/0.190/9.14e-06	13/206/83/0.177/2.74e-07	F/F/F/F/F
diagonal1 12	105/3134/1543/0.159/6.45e-06	105/3134/1481/0.151/6.46e-06	105/3134/1533/0.151/6.46e-06	F/F/F/F/F	135/3580/1781/0.161/7.50e-06
diagonal2 20	82/2383/1162/0.108/7.31e-07	108/3227/1550/0.146/6.21e-07	36/818/373/0.036/7.72e-06	F/F/F/F/F	F/F/F/F/F
diagonal3 40	57/1355/639/0.064/3.74e-06	49/1019/502/0.051/7.82e-06	112/2845/1324/0.136/6.09e-06	F/F/F/F/F	298/8064/3951/0.383/8.62e-06
bv 1000	1/1/1/0.000/4.99e-06	1/1/1/0.000/4.99e-06	1/1/1/0.000/4.99e-06	1/1/1/0.000/4.99e-06	1/1/1/0.000/4.99e-06
bv 10000	1/1/1/0.000/5.00e-08	1/1/1/0.000/5.00e-08	1/1/1/0.000/5.00e-08	1/1/1/0.000/5.00e-08	1/1/1/0.000/5.00e-08
ie 200	14/225/93/3.573/3.73e-06	14/225/92/3.504/3.06e-06	11/166/71/2.701/5.48e-06	222/6471/3237/103.941/9.77e-06	58/1497/755/24.210/3.22e-06
gauss 3	15/364/162/0.034/5.78e-06	24/589/263/0.045/3.96e-06	22/588/256/0.046/5.33e-06	14/313/137/0.025/3.58e-06	20/528/273/0.042/9.38e-06
lin 500	2/2/2/0.022/9.93e-14	2/2/2/0.022/9.93e-14	2/2/2/0.016/9.93e-14	2/2/2/0.016/9.93e-14	2/2/2/0.018/9.93e-14

参考文献

- [1] Hestenes M R, Stiefel E. Method of conjugate gradient for solving linear equations [J]. J. Res. Natl. Bur. Stand., 1952, 49: 409–436.
- [2] Fletcher R, Reeves C. Function minimization by conjugate gradients [J]. Comput. J., 1964, 7(2): 149–154.
- [3] Polak E, Ribière G. Note sur la convergence de directions conjuguées [J]. Rev. Fr. Inform. Rech. Oper., 1969, 16(3): 35–43.
- [4] Polak B T. The conjugate gradient method in extreme problems [J]. USSR Comput. Math. Math. Phys., 1969, 9(4): 94–112.
- [5] Dai Y H, Yuan Y X. A nonlinear conjugate gradient method with a strong global convergence property [J]. SIAM J. Optim., 1999, 10(1): 177–182.
- [6] Glibert J C, Nocedal J. Global convergence properties of conjugate gradient method for optimization [J]. SIAM J. Optim., 1992, 2(1): 21–42.
- [7] Hager W W, Zhang H C. A new conjugate gradient method with guaranteed descent and an efficient line search [J]. SIAM J. Optim., 2005, 16(1): 170–192.

- [8] Andrei N. Hybrid conjugate gradient algorithm for unconstrained optimization[J]. *J. Optim. Theory Appl.*, 2009, 141(2): 249–264.
- [9] 江羨珍, 马国栋, 简金宝. Wolfe 线搜索下一个新的全局收敛共轭梯度法[J]. *工程数学学报*, 2011, 28(6): 779–786.
- [10] Jiang X Z, Jian J B. Two modified conjugate gradient methods with disturbance factors for unconstrained optimization[J]. *Nonlinear Dynam.*, 2014, 77(1–2): 387–397.
- [11] 马国栋, 简金宝, 江羨珍. 一个具有下降性的改进Fletcher-Reeves共轭梯度法[J]. *应用数学学报*, 2015, 38(1): 89–97.
- [12] 景书杰, 王慧婷, 牛海峰, 陈耀. 精确线搜索下一种新的混合共轭梯度法[J]. *数学杂志*, 2018, 38(3): 520–524.
- [13] Jiang X Z, Jian J B. Improved Fletcher-Reeves and Dai-Yuan conjugate gradient methods with the strong Wolfe line search[J]. *J. Comput. Appl. Math.*, 2019, 348: 525–534.
- [14] 李春梅, 王翠方, 段雪峰. 求解一类矩阵迹极小化问题的非线性共轭梯度法[J]. *数学杂志*, 2020, 40(3):323–331.
- [15] Morè J J, Garbow B S, Hillstrome K E. Testing unconstrained optimization software[J]. *ACM T. Math. Softwar.*, 1981, 7(1): 17–41.
- [16] Dolan E D, Morè J J. Benchmarking optimization software with performance profiles[J]. *Math. Program*, 2002, 91(2): 201–213.

THE IMPROVED FR CONJUGATE GRADIENT METHOD WITH SUFFICIENT DESCENT PROPERTY

MA Guo-dong, JIANG Xian-zhen, JIN Wen-hui

(*College of Mathematics and Physics, Guangxi University for Nationalities, Nanning Guangxi 530006*)

Abstract: In this paper, we consider solving large-scale unconstrained optimization, based on the improved parameter formula of the FR method, a conjugate gradient method that is proposed. Without any line search, we proved that the search direction always satisfied sufficient descent condition at each iteration. The global convergence of the proposed method is proved under the standard Wolfe inexact line search condition. Finally, some elementary numerical experiments are reported, which show that the algorithm is promising.

Keywords: unconstrained optimization; conjugate gradient method; Wolfe line search; global convergence

2010 MR Subject Classification: 90C30; 65K05