doi:10.16055/j.issn.1672-058X.2015.0005.006

强 Wolfe 条件下一类修正 FR 共轭梯度法

张 鹏
(重庆师范大学 数学学院,重庆 401331)

摘 要:通过适当修正 Fletcher-Reeves(FR)方法,提出了一类修正 FR 共轭梯度法方法(MFR*),并证明了 MFR* 方法在强 Wolfe 线搜索下具有充分下降条件和全局收敛性.

关键词:无约束优化;共轭梯度法;充分下降性;全局收敛性

中图分类号:O224 **文献标识码**:A **文章编号**:1672-058X(2015)05-0020-03

0 引 言

考虑无约束优化问题:

$$\min f(x), x \in \mathbf{R}^n \tag{1}$$

其中 $f:\mathbf{R}^n \to \mathbf{R}$ 是连续可微的函数.共轭梯度法因为迭代简单,存储量小,因此常被用来求解具有问题(1)形式的大规模优化问题.共轭梯度法的一般迭代格式为

$$x_{k+1} = x_k + \alpha_k d_k \tag{2}$$

其中 α_k 是由线搜索确定的步长,搜索方向 d_k 满足

$$d_k = \begin{cases} -g_k, k=0 \\ -g_k + \beta_k d_{k-1}, k \geq 1 \end{cases} \tag{3}$$

β_k 是共轭参数,$g_0 \leq \varepsilon$ 的不同取法对应于不同的非线性共轭梯度法.著名的共轭梯度法有 1964 年 Fletcher 和 Reeves 提出的 FR 方法[1],1969 年 Polar-Ribiere 和 Polyak 分别独立提出的 PRP 方法[2,3],1952 年 Hestenes 和 Stiefel 提出的 HS 方法[4],1999 年 Dai 和 Yuan 提出的 DY 方法[5].β_k 分别由下面公式给出:

$$\beta_k^{\mathrm{FR}} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \beta_k^{\mathrm{PRP}} = \frac{g_k^{\mathrm{T}} y_{k-1}}{\|g_{k-1}\|^2}$$

$$\beta_k^{\mathrm{HS}} = \frac{g_k^{\mathrm{T}} y_{k-1}}{d_{k-1}^{\mathrm{T}} y_{k-1}}, \beta_k^{\mathrm{DY}} = \frac{\|g_k\|^2}{d_{k-1}^{\mathrm{T}} y_{k-1}}$$

其中,$\|\cdot\|$ 为欧几里得范数,$y_{k-1}=g_k-g_{k-1}$.

在共轭梯度法的许多理论分析和数值实现中,常常采用非精确线搜索,如强 Wolfe 线搜索.强 Wolfe 线搜索要求步长 α_k 满足

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^{\mathrm{T}} d_k \tag{4}$$

$$|g(x_k + \alpha_k d_k)^{\mathrm{T}} d_k| \leq \sigma |g_k^{\mathrm{T}} d_k| \tag{5}$$

其中 $0<\delta \leq \sigma<1$.

敖卫斌[6]提出一种修正的 DY 共轭梯度法的全局收敛性,并证明了全局收敛性;黎小林[7]提出一种 Glodstein 条件下的修正 CD 共轭梯度法并证明了全局收敛性.

收稿日期:2014-08-24;**修回日期**:2014-10-08.

作者简介:张鹏(1990-),男,河南信阳人,硕士,从事全局最优化理论与算法研究.

Zhang[8]提出一种修正的 PRP 方法,其中 β_k 被定义为

$$\beta_k^{\mathrm{NPRP}}=\frac{\|g_k\|^2-\dfrac{\|g_k\|}{\|g_{k-1}\|}|g_k^{\mathrm{T}}g_{k-1}|}{\|g_{k-1}\|^2}$$

并证明了 NPRP 在强 Wolfe 线搜索下的充分下降性和全局收敛性.

最近,Jiang[9]提出一种修正的 FR 方法(MFR 方法),其中 β_k 具有如下形式

$$\beta_k^{\mathrm{MFR}}=\frac{\|g_k\|^2}{\max\{\|g_k\|^2,\mu|g_k^{\mathrm{T}}d_{k-1}|\}}(\mu>1)$$

证明了 MFR 不依赖线搜索,总是产生一个下降方向,并且在强 Wolfe 线搜索下全局收敛.

受 Zhang[8]和 Jiang[9]的启发,构造一类如下的修正 FR 方法,共轭梯度参数 β_k 具有如下形式

$$\beta_k^{\mathrm{MFR}^*}=\frac{\|g_k\|^2-\dfrac{\|g_k\|}{\|d_{k-1}\|}|g_k^{\mathrm{T}}d_{k-1}|}{\max\{\|g_{k-1}\|^2,\mu|g_k^{\mathrm{T}}d_{k-1}|\}}\tag{6}$$

其中 $\mu>1$.当目标函数是严格凸二次函数并采用精确线搜索时,MFR* 退化为经典的 FR 方法,可以证明新方法不依赖线搜索具有充分下降性且在强 Wolfe 线搜索下具有全局收敛性.

第 2 部分给出 MFR* 算法;第 3 部分证明了 MFR* 算法的充分下降性和在强 Wolfe 线搜索下的全局收敛性.

1　算法设计

MFR* 算法:

Step0　给定常数 $\sigma\in(0,1)$,$\delta\in(0,\sigma)$,$\varepsilon\geqslant0$,选取初始点 $x_0\in\mathbf{R}^n$,计算 $d_0:=-g(x_0)$,置 $k:=0$;

Step1　如果 $\|g_k\|_\infty\leqslant\varepsilon$,则算法终止;

Step2　计算 $\alpha_k>0$ 满足强 Wolfe 线搜索的式(4)(5);

Step3　令 $x_{k+1}=x_k+\alpha_kd_k$,$g_{k+1}=g(x_{k+1})$,如果 $\|g_{k+1}\|_\infty\leqslant\varepsilon$,则算法终止;

Step4　由式(3)(6)计算 d_{k+1},置 $k:=k+1$,转 Step2.

2　MFR* 算法的收敛性

假设 A　目标函数 $f(x)$ 在水平集 $L=\{x\in\mathbf{R}^n\mid f(x)\leqslant f(x_0)\}$ 上有下界,其中 $x_0\in\mathbf{R}^n$ 为算法初始点.

假设 B　目标函数 $f(x)$ 在水平集 L 的一个领域 N 内连续可微,且其梯度函数 g 满足 Lipschitz 条件,即存在常数 $l>0$,使 $\|g(x)-g(y)\|\leqslant l\|x-y\|$,$\forall x,y\in L$.

引理 1　$\beta_k^{\mathrm{MFR}^*}$ 满足 $0\leqslant\beta_k^{\mathrm{MFR}^*}\leqslant\beta_k^{\mathrm{FR}}$.

证明　由 $\beta_k^{\mathrm{MFR}^*}$ 的表达式(6),知

$$\beta_k^{\mathrm{MFR}^*}=\frac{\|g_k\|^2-\dfrac{\|g_k\|}{\|d_{k-1}\|}|g_k^{\mathrm{T}}d_{k-1}|}{\max\{\|g_{k-1}\|^2,\mu|g_k^{\mathrm{T}}d_{k-1}|\}}\geqslant\frac{\|g_k\|^2-\dfrac{\|g_k\|}{\|d_{k-1}\|}\|g_k\|\|d_{k-1}\|}{\max\{\|g_{k-1}\|^2,\mu|g_k^{\mathrm{T}}d_{k-1}|\}}=0$$

并且有

$$\beta_k^{\mathrm{MFR}^*}\leqslant\frac{\|g_k\|^2}{\|g_{k-1}\|^2}=\beta_k^{\mathrm{FR}}$$

引理 2　考虑具有式(2)(3)格式的迭代方法,其中 $\beta_k=\beta_k^{\mathrm{MFR}^*}$,则对任意的线搜索有

$$g_k^{\mathrm{T}}d_k\leqslant-\left(1-\frac{1}{\mu}\right)\|g_k\|^2,\mu>1$$

证明 由 $\beta_k^{\mathrm{MFR}^*}$ 的表达式(6),有

$$0 \leqslant \beta_k^{\mathrm{MFR}^*} = \frac{\| g_k \|^2 - \dfrac{\| g_k \|}{\| d_{k-1} \|} | g_k^{\mathrm{T}} d_{k-1} |}{\max\{ \| g_{k-1} \|^2, \mu | g_k^{\mathrm{T}} d_{k-1} | \}} \leqslant \frac{\| g_k \|^2}{\mu | g_k^{\mathrm{T}} d_{k-1} |}$$

因此

$$\begin{aligned} g_k^{\mathrm{T}} d_k = & - \| g_k \|^2 + \beta_k^{\mathrm{MFR}^*} g_k^{\mathrm{T}} d_{k-1} \leqslant \\ & - \| g_k \|^2 + | \beta_k^{\mathrm{MFR}^*} | | g_k^{\mathrm{T}} d_{k-1} | \leqslant \\ & - \| g_k \|^2 + \frac{\| g_k \|^2}{\mu | g_k^{\mathrm{T}} d_{k-1} |} | g_k^{\mathrm{T}} d_{k-1} | = \\ & - (1 - \frac{1}{\mu}) \| g_k \|^2 \end{aligned}$$

由引理 1 和文献[10]中的定理 3.2,可以直接得到定理 1.

定理 1 若假设(A)(B)成立,MFR*算法中步长 α_k 满足强 Wolfe 条件式(4)和(5),参数 $0<\delta<\sigma<\frac{1}{2}$,则

$$\lim_{k\to\infty} \inf \| g_k \| = 0$$

参考文献:

[1] FLETCHER R,REEVES C M.Function Minimization by Conjugate Gradients[J].The Computer Journal,1964,7(2):149-154

[2] POLAK E,RIBIERE G.Note Sur La Convergence De Methodes De Directions Conjuguees[J].ESAIM:Mathematical Modelling and Numerical Analysis-Modélisation Mathematique et Analyse Numerique,1969,3(R1):35-43

[3] POLYAK B T.The Conjugate Gradient Method in Extremal Problems[J].USSR Computational Mathematics and Mathematical Physics,1969,9(4):94-112

[4] HESTENES M R,STIEFEL E.Methods of Conjugate Gradients for Solving Linear Systems[M].NBS,1952

[5] DAI Y H,YUAN Y.A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property[J].SIAM Journal on Optimization,1999,10(1):177-182

[6] 敖卫斌.一种修正的 DY 共轭梯度法的全局收敛性[J].重庆工商大学学报:自然科学版,2013,30(10):17-20

[7] 黎小林.Glodstein 条件下的一种修正 CD 共轭梯度法[J].重庆工商大学学报:自然科学版,2014,31(3):24-26

[8] ZHANG L. An Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method for Optimization Computation[J]. Applied Mathematics and Computation,2009(6):2269-2274

[9] JIANG X,JIAN J. A Sufficient Descent Dai-Yuan Type Nonlinear Conjugate Gradient Method for Unconstrained Optimization Problems[J].Nonlinear Dynamics,2013,72(1-2):101-112

[10] GILBERT J C,NOCEDAL J.Global Convergence Properties of Conjugate Gradient Methods for Optimization[J].SIAM Journal on Optimization,1992,2(1):21-42

A Class of Modified FR Method with Strong Wolfe Line Search

ZHANG Peng

(College of Mathematics Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: Abstract: This paper proposes a class of modified FR method (MFR*) by modifying Fletcher-Reeves (FR), and proves MFR* with strong Wolfe line search satisfies sufficient descent condition and global convergence.

Key words: unconstrained optimization; conjugate gradient method; sufficient descent condition; global convergence