

文章编号:1007-6735(2013)02-0113-03

一种非线性扩展混合共轭梯度算法的全局收敛性

陈元媛^{1,2}, 高 岩¹

(1. 上海理工大学 管理学院, 上海 200093; 2. 青岛大学 数学科学学院, 青岛 266071)

摘要: 描述了非线性 FR 共轭梯度法、非线性 PRP 共轭梯度法、非线性 DY 共轭梯度法等求解大规模无约束优化问题的有效算法. 研究了计算更为有效的适合求解无约束优化问题的一种非线性扩展混合共轭梯度算法; 给出了在 Wolfe 型线搜索下的非线性扩展混合共轭梯度法, 算法产生的方向为下降方向. 在一般的条件下, 给出了算法的全局收敛结果, 且数值实验表明算法十分有效.

关键词: 无约束优化; 非线性扩展混合共轭梯度法; 线搜索; 全局收敛

中图分类号: O 224; N 94 **文献标志码:** A

Global Convergence of A Nonlinear Extended Hybrid Conjugate Gradient Method with Wolfe Type Line Search

CHEN Yuanyuan^{1,2}, GAO yan¹

(1. Business School, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. College of Mathematics, Qingdao University, Qingdao 266071, China)

Abstract: The nonlinear FR conjugate gradient method, nonlinear PRP conjugate gradient method, nonlinear DY conjugate gradient method etc. for large scale unconstrained optimization problems were described. The nonlinear extended hybrid conjugate gradient method, which is more efficient for solving unconstrained optimization problems, was stressed and a new nonlinear extended hybrid conjugate gradient method with Wolfe type line search was proposed. The direction generated by the new method produces sufficiently descentding search direction. Under some conditions, the global convergence of the new method was investigated. Some numerical results show its efficiency.

Key words: unconstrained optimization; nonlinear extended hybrid conjugate gradient method; line search; global convergence

1 问题的提出

共轭梯度法起源于 Hestenes 和 Stiefel 在 1952

年提出的求解线性方程组的方法, 是著名的共轭方向法. 由于解非线性方程组等价于极小化一个正定二次函数, 故 1964 年 Fletcher 和 Reeves 提出了无约束优化的共轭梯度法. 共轭梯度法由于计算过程

收稿日期: 2012-03-12
基金项目: 国家自然科学基金资助项目(11171221, 11101231); 上海市一流学科建设资助项目(XTKX2012); 山东省高校科技计划资助项目(J10LA05)
第一作者: 陈元媛(1978-), 女, 讲师. 研究方向: 系统分析与优化. E-mail: usstchenyuanyuan@163.com
通讯作者: 高 岩(1962-), 男, 教授. 研究方向: 系统分析与最优化、金融数学与金融工程. E-mail: gaoyan@usst.edu.cn

中只要目标函数值和梯度函数值,不要矩阵存储,却比最速下降算法有好的数值效果,一直为一种广泛应用的无约束优化算法^[1-8]. 本文主要对大规模的无约束优化问题给出了一种非线性扩展混合共轭梯度法. 考虑求解的无约束优化问题为

$$\min_{x \in R^n} f(x) \quad (1)$$

式中, $f(x)$ 为 R^n 上的连续可微函数, 梯度 $\nabla f(x)$ 记为 $g(x)$. 算法的一般迭代形式为

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

式中, $\{x_k\}$ 为迭代点列; α_k 为步长因子; d_k 为搜索方向, 即

$$d_k = \begin{cases} -g_k, & k=1 \\ -g_k + \beta_k d_{k-1}, & k \geq 2 \end{cases} \quad (3)$$

式中, g_k 为 $\nabla f(x_k)$; β_k 为标量, 常用式为

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}}, \beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}$$

$$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}}, \beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}$$

关于上述方法的文章可以参见文献[5-12]. 步长因子 α_k 一般由 Wolfe 线搜索、Armijo 线搜索等非精确线搜索得到^[3-6]. 最近, 文献[13]中给出了一种新的 β_k , 即

$$\beta_k^{H_1} = \max\{0, \min\{\beta_k^1, \beta_k^2\}\} \quad (4)$$

$$\beta_k^{H_2} = \max\{-\beta_k^1, \min\{\beta_k^1, \beta_k^2\}\} \quad (5)$$

其中, $\beta_k^1 = \frac{\|g_k\|^2}{\|g_{k-1}\|^2 + (g_k^T d_{k-1})^2}$, $\beta_k^2 = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2 + (g_k^T d_{k-1})^2}$, $\|\cdot\|$ 为 Euclidean 范数.

在文献[13]的基础上, 结合一种新的 Wolfe 型线搜索, 给出了一种新的非线性扩展混合共轭梯度法. 在下文中对这种共轭下降算法的全局收敛性进行理论分析, 在一般假设条件下给出全局收敛性定理, 并且给出了该算法的数值实验结果与讨论.

2 非线性扩展混合共轭梯度法

2.1 假设条件

a. $f(x)$ 在水平集 $L_0 = \{x \in R^n \mid f(x) \leq f(x_0)\}$ 有界.

b. $f(x)$ 在 L_0 的一个邻域 U 内连续可微, 且其导数 $g(x)$ 满足 Lipschitz 条件, 即存在常数 $L > 0$ 使得 $\|g(x) - g(y)\| \leq L\|x - y\|$, $\forall x, y \in U$.

另外, 给出 Wolfe 型线搜索, 此线搜索在文献

[14]中用过. 选取 $\alpha_k > 0$, 满足

$$f(x_k + \alpha_k d_k) - f(x_k) \leq -\rho \alpha_k^2 \|d_k\|^2 \quad (6)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq -2\sigma \alpha_k \|d_k\|^2 \quad (7)$$

其中, $0 < \rho < \sigma < 1$.

2.2 计算步骤

a. 给出 $x_1 \in R^n$, $\varepsilon > 0$, $d_1 = -g_1$, $k := 1$. 若 $\|g_1\| \leq \varepsilon$, 停.

b. 求 $\alpha_k > 0$ 满足 Wolfe 型线搜索式(6)和式(7), 由式(2)得 x_{k+1} , 若 $g_{k+1} \leq \varepsilon$, 则停, 否则转 c.

c. 由式(4)和式(5)计算 β_{k+1} , d_{k+1} 的计算为

$$d_{k+1} = -(1 + \beta_{k+1} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}) g_{k+1} + \beta_{k+1} d_k \quad (8)$$

d. $k := k + 1$, 转 b.

为建立非线性扩展混合共轭梯度算法的全局收敛结果, 给出引理:

引理 1^[14] 若假设条件满足, 则算法中线搜索式(6)和式(7)可行.

引理 2^[13] 若 d_k 由式(8)产生, 则对 $k \geq 1$ 有充分下降条件 $g_k^T d_k = -\|g_k\|^2$.

引理 3 若假设条件满足, 步长 α_k 由式(6)和式(7)得到, 则

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \quad (9)$$

证明 由式(6)和式(7)和假设条件, 得到

$$(2\sigma + L)\alpha_k \|d_k\|^2 \geq -g_k^T d_k$$

则

$$\alpha_k \|d_k\| \geq \frac{1}{2\sigma + L} \left(\frac{-g_k^T d_k}{\|d_k\|} \right)$$

对其两边平方, 得到

$$\alpha_k^2 \|d_k\|^2 \geq \left(\frac{1}{2\sigma + L} \right)^2 \left(\frac{-g_k^T d_k}{\|d_k\|} \right)^2$$

由式(6), 得

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \leq (2\sigma + L)^2 \sum_{k=1}^{\infty} \alpha_k^2 \|d_k\|^2 \leq$$

$$\left(\frac{2\sigma + L}{\rho} \right)^2 \sum_{k=1}^{\infty} \{f(x_k) - f(x_{k+1})\} < +\infty$$

即得到式(9), 结论成立.

引理 4 若假设条件满足, d_k 由式(8)产生, 步长 α_k 由式(6)和式(7)得到, 则

$$\sum_{k \geq 0} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty \quad (10)$$

证明 由引理 2 和引理 3, 得到式(10).

3 全局收敛结果

定理 1 若假设条件满足, $\{x_k\}$ 由算法产生, 则

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0$$

证明 反证法.若结论不成立,则 $\exists \epsilon > 0$,使 $\|g_k\| \geq \epsilon$,由式(8)和引理 2,得

$$\|d_k\|^2 = \beta_k \|d_{k-1}\|^2 - 2\xi_k g_k^T d_k + \xi_k^2 \|g_k\|^2$$

其中, $\xi_k = 1 + \beta_k \frac{g_k^T d_{k-1}}{\|g_k\|^2}$

对上式两边同时除 $\|g_k\|^4$,得到

$$\frac{\|d_k\|^2}{\|g_k\|^4} = \beta_k^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - (\xi_k - 1)^2 \frac{1}{\|g_k\|^2} +$$
$$\frac{1}{\|g_k\|^2} \leq \beta_k^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} + \frac{1}{\|g_k\|^2} \leq$$
$$(\beta_k^{FR})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} + \frac{1}{\|g_k\|^2} \leq \sum_{i=1}^k \frac{1}{\|g_i\|^2} \leq \frac{k}{\epsilon^2}$$

即

$$\sum_{k \geq 1} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \epsilon^2 \sum_{k \geq 1} \frac{1}{k} = +\infty$$

与式(10)矛盾,定理成立.

4 数值实验与讨论

给出了该算法的数值实验与相关的讨论,数值问题取自文献[15], α_k 由式(6)和式(7)得到,其中 $\rho = 0.001, \sigma = 0.01. \|g_k\| \leq 10^{-6}$ 为停止规则. 利用 Matlab7.0 在 CPU1.60GHZ 上对测试函数进行了计算,具体的数值结果见表 1,其中 N_I 表示迭代的次数, N_F 表示计算函数值的次数, N_G 表示计算函数梯度值的次数.

表 1 数值结果
Tab.1 Numerical results

函数名称	维数	N_I	N_F	N_G
GULF	3	6	50	3
VARDIM	2	8	54	6
LIN	50	1	3	3
LIN	1 000	6	9	9
LIN1	4	1	3	3

由非线性扩展混合共轭梯度算法的全局收敛分析知,只要 d_k 满足充分下降条件,在 Wolfe 型线搜索式(6)和式(7)下结合式(3),不需其它条件就可得到该算法的全局收敛结果.如最近文献[16]中

$$\beta_k = -\frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}} + \beta_k^{PRP}$$

其中, $s_{k-1} = x_k - x_{k-1}, y_{k-1} = g_k - g_{k-1}$.

参考文献:

[1] Fletcher R,Reeves C. Function minimization by conjugate

gradients[J]. J Comput,1964,7(2):149-154.

[2] Polyak B T. The conjugate gradient method in extreme problems[J]. USSR Comp Math Physk, 1969, 9(4): 94-112.

[3] Dai Y H, Yuan Y X. A nonlinear conjugate gradient method with nice global convergence properties[J]. SIAM J Optim,1999,10(1):177-182.

[4] Wei Z, Li G, Qi L. New quasi-Newton methods for unconstrained optimization problems[J]. Appl Math Comput,2006,175(2):1156-1188.

[5] Fletcher R. Practical methods of optimization, unconstrained optimization[M]. New York;Wiley,1987.

[6] Liu Y, Storey C. Efficient generalized conjugate gradient algorithms, Part 1: Theory [J]. J Optim Theory Appl,1991,69(1):129-137.

[7] Raydan M. The Barzilai and Borwein gradient method for the large unconstrained minimization problem[J]. SIAM J Optim,1997,7(1):26-33.

[8] Birgin E G, Martinez J M. A spectral conjugate gradient method for unconstrained optimization[J]. Appl Math Optim,2001,43(2):117-128.

[9] Gilbert J C, Nocedal J. Global convergence properties of conjugate gradient methods for optimization[J]. SIAM J Optim,1992,2(1):21-24.

[10] Sun J, Zhang J. Convergence of conjugate gradient methods without line search[J]. Annals Operations Research,2001,103(1):161-173.

[11] Zhang L, Zhou W, Li D. Global convergence of a modified fletcher-reeves conjugate gradient method with Armijo-type line search[J]. Numer Math, 2006, 104(4):561-572.

[12] Hager W W, Zhang H. A new conjugate gradient method with guaranteed descent and an efficient line search[J]. SIAM J Optim,2005,16(1):170-192.

[13] Hui Y,Chen L P. Extended AS-GN hybrid conjugate gradient method[J]. OR Transactions, 2010, 14(3): 122-128.

[14] Wang C, Chen Y, Du S. Further insight into the Shamanskii modification of Newton method[J]. Appl Math Comput,2006,18(1):46-52.

[15] More J J, Garbow B S, Hillstrom K E. Testing unconstrained optimization software[J]. ACM Trans Math Software,1981,7(1):17-41.

[16] Wu C Y. A modified PRP conjugate gradient algorithm for unstrained optimization problems[J]. Mathematica Applicata,2011,24(1):25-29.

(编辑:金 虹)