文章编号：1674-8085(2018)01-0021-03

一类改进的拟牛顿算法

*徐莹莹，张　惠

(郑州工业应用技术学院基础教学部　河南，郑州 451150)

摘　要：在利用拟牛顿算法求解非线性无约束优化问题中，本文在文献[8]提出的拟牛顿方程基础上，通过加权形式构造一类改进拟牛顿方程，产生了修正的 BFGS 校正公式，进而提出改进的拟牛顿算法，在一定条件下证明新算法的全局收敛性。数值实验结果表明，与文献[12]中的拟牛顿算法对比，新算法在迭代次数上更有优势。

关键词：无约束优化；拟牛顿方程；线性搜索准则；全局收敛性

中图分类号：O224　　**文献标识码**：A　　DOI:10.3969/j.issn.1674-8085.2018.01.005

A CLASS OF NEW MODIFIED QUASI-NEWTON ALGORITHM

* XU Ying-ying，ZHANG Hui

(Zhengzhou University of Industrial Technology, Zhengzhou, Henan 451150, China)

Abstract: In solving nonlinear unconstrained optimization problem by using quasi - Newton algorithm. Though the weighted form a class of new quasi-newton algorithm is constructed based on the quasi-newton equation and the method of literature. Furthermore, a new quasi-newton algorithm is proposed combining the modified BFGS correction formula of the new quasi-newton equation. The new global convergence of the algorithm is proved under certain conditions. Finally, through numerical experiments show that this new algorithm has much more advantages in the number of iterations.

Key words: unconstrained optimization; quasi-Newton equation; linear search; global convergence

0　引言

对于求解无约束优化问题 $\min f(x), x\in R^n$，其中 $f(x)$ 是 $R^n \to R$ 连续可微函数。拟牛顿算法是一类最有效、理论上也是最成熟的算法之一[1]。许多学者对拟牛顿算法做了深入而广泛的研究，并取得了丰硕的成果。近年来，基于修改原始拟牛顿方程的新拟牛顿算法[2-5]和修正拟牛顿方程的非拟牛顿算法[1]的研究吸引了不少国内外学者。文献[6-7]提出了一类广义拟牛顿算法，在目标函数一致凸的条件下，结合 wolfe 线搜索准则，证明了其算法的全局收敛性以及局部超线性收敛性。文献[2-3]改进了文献[6-7]中的广义拟牛顿算法，将其扩展到了一般形式，并对改进后的算法给出了全局收敛性证明。本文基于文献[8]提出的拟牛顿方程，结合文献[6]中的方法，通过加权形式构造了一类改进拟牛顿方程，提出了一类改进拟牛顿算法，并在一定条件下证明了新算法的全局收敛性。

1　新拟牛顿方程的提出

2006 年，Wei 在文献[8]中利用目标函数的泰勒展式提出了一种新的拟牛顿方程

$$B_{k+1}s_k = y_k^* \tag{1}$$

其中

$$y_k^* = y_k + \frac{\theta_k}{\|s_k\|^2}s_k$$

收稿日期：2017-04-11；修改日期：2017-09-20

作者简介：*徐莹莹(1988-)，女，河南焦作人，硕士生，主要从事最优化理论与算法研究(E-mail: xy010007@126.com);
张　惠(1990-)，女，河南新郑人，硕士生，主要从事计算数学研究(E-mail: piaoyao3652@126.com).

$$\theta_k = 2\left(f_k - f_{k+1}\right) + \left(g_k + g_{k+1}\right)^T s_k \qquad (2)$$

另一方面，当$\|s_k\|$充分小时，由 Taylor 公式知$2R_k \approx s_k^T G_{k+1} s_k$，其中定义

$$R_k = f_{k+1} - f_k - g_k^T s_k \qquad (3)$$

结合式（2）、式（3），考虑它们的加权形式，有：

$$s_k^T G_{k+1} s_k = t(s_k^T y_k + \theta_k) + 2(1-t)R_k$$

进而构造出一类新的拟牛顿方程

$$\tilde{y}_k = y_k + \frac{(2t-1)\theta_k}{s_k^T u_k} u_k, t \in [0,1] \qquad (4)$$

显然，当$t=1, u_k = s_k$时，文献[3]中的拟牛顿方程为（4）式中特例；当$t = \frac{1}{2}$时，（4）式退化为一般拟牛顿方程。

由（4）式给出如下校正公式

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{\tilde{y}_k \tilde{y}_k^T}{s_k^T \tilde{y}_k}$$

由于 wolfe 搜索准则不能保证$s_k^T \tilde{y}_k > 0$，我们给出如下校正公式

$$B_{k+1} = \begin{cases} B_k - \dfrac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \dfrac{\tilde{y}_k \tilde{y}_k^T}{s_k^T \tilde{y}_k}, & k \in K \\ B_k, & 否则 \end{cases} \qquad (5)$$

式中$K = \left\{ k \middle| \dfrac{s_k^T \tilde{y}_k}{\|s_k\|^2} \ge \beta \|g_k\|^\gamma \right\}$，其中$\beta$为正常数，$\gamma \in [\mu_1, \mu_2], 0 < \mu_1 < \mu_2$。

2 新拟牛顿算法的提出

新算法采用 wolfe 搜索准则，其算法步骤如下：

Step1 给出初始点$x_0 \in R^n$和初始正定矩阵$B_0 \in R^{n \times n}$，令$k := 0$；

Step2 若$\|g_k\| < \varepsilon$,则停止。否则计算$B_k d_k = -g_k$得搜索方向d_k；

Step3 利用 wolfe 线性搜索准则

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta_1 \alpha_k g_k^T d_k$$

$$g(x_k + \alpha_k d_k) \ge \delta_2 g_k^T d_k, 0 < \delta_1 < \delta_2 < 1$$

求得步长α_k；

Step4 令$x_{k+1} = x_k + \alpha_k d_k$得到下一个迭代点。若$\|g_{k+1}\| < \varepsilon$,则迭代停止，否则转步 5；

Step5 利用公式（5）校正B_k得B_{k+1}，令$k := 0$，转步 2。

3 算法收敛性分析

我们做如下假设：

假设(P)：目标函数$f(x)$在R^n上二阶连续可微，水平集$L = \{x | f(x) \le f(x_1)\}$有界，$x^*$是$f(x)$的局部最优解。

定理 1 设B_k正定，则由校正公式（5）可保证B_{k+1}的正定性。

证明： 将B_k分解为$B_k = H_k H_k^T$，对任意$z \in R^n, z \ne 0$，令$a_k = H_k^T z, b_k = H_k^T s_k$，则

$$z^T (B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}) z = a_k^T a_k - \frac{(a_k^T b_k)^2}{b_k^T b_k} \ge 0 \qquad (6)$$

$$z^T \frac{\tilde{y}_k \tilde{y}_k^T}{s_k^T \tilde{y}_k} z = \frac{(z^T \tilde{y}_k)^2}{s_k^T \tilde{y}_k} \ge 0 \qquad (7)$$

由于不等式（6）、（7）中等号不同时成立，因此有$z^T B_{k+1} z > 0$，即B_{k+1}正定。

定理 2 若假设(P)成立，当k取某个值时，有$g_k = 0$，或者当$k \to +\infty$，有$\liminf_{k \to \infty} \|g_k\| = 0$。

证明： 记下标集$\tilde{K} = \{1, 2, \cdots\}$，若对所有的$k$有$g_k \ne 0$，下证$\liminf_{k \to \infty} \|g_k\| = 0$成立即可。

用反证法。假设（8）式不成立，则存在$\varepsilon > 0$，使得对所有k，总有$\|g_k\| \ge \varepsilon$，由B_k正定性可得$g_k^T s_k < 0$成立，故当$k \to +\infty$时，有$f_{k+1} - f_k \to 0$。

设L^*为包含水平集L的最小闭集。由于新算法是下降算法，有极小化序列$\{x_k\} \subset L^*$，故存在收敛子列$\{x_k\} (k \in K_1 \subset K)$，当$k \to +\infty$时，有$\|x_{k+1} - x_k\| \to 0$，又由$f(x)$的连续性可得，对函数子列$\{f_k\}_{K_1}$，有$f_{k+1} - f_k \to 0\ (k \to \infty, k \in K_1 \subset K)$。

另外有$g_{k+1}^T s_k = g_k^T s_k + o(\|s_k\|)$，当$k \to +\infty$时，有$\dfrac{g_{k+1}^T s_k}{g_k^T s_k} \to 1$，这与$\dfrac{g_{k+1}^T s_k}{g_k^T s_k} \le \delta_2 \le 1$矛盾。因而

$\lim\limits_{k\to\infty}\inf\|g_k\|=0$。

由定理 2 可知，$\{x_k\}$ 有子列收敛于（1）的稳定点，当 $f(x)$ 为凸函数时，$\{x_k\}$ 的每个聚点都是原问题（1）的全局最优解。当 $f(x)$ 为非凸函数时，该结论不成立。但当 $k\to\infty, s_k\to\infty$，存在 $\{x_k\}$ 的一个聚点 x^* 满足 $g(x^*)=0$，$G(x^*)$ 正定，则 $\{x_k\}$ 收敛于 x^*。

4 数值试验

为检验新算法的实验效果，本文将此算法与文献[12]算法进行了比较，列举了下表测试函数，数值实验通过 Windows 2007 操作系统运行在 MATLAB 2010 上，进行数值实验时的参数设置为：$t=0.75$；$\mu_k=y_k$. 终止条件为：$\|g(x)\|\le 10^{-5}$。k_0 表示文献[12]中的算法迭代次数，k_1 表示新算法迭代次数。

例 1：$min\, f(x)=5x_1^2+7.5x_2^2+10x_3^2+2x_1x_2+4x_2x_3+x_1x_3+e^{x_1+x_2+x_3}$

例 2：$min\, f(x)=10x_1^4+25x_2^4+12x_3^4+18x_1^2+13x_2^2+10x_3^2+2x_1x_2+2x_2x_3-5x_1-3x_2-x_3$

计算结果见表 1：

表 1 数值实验结果

Table 1 Numerical experiments results

算例	x_0	k_0	k_1	x_k
1	(-1,1.5,-0.5)	13	7	(-0.0754,-0.0391,-0.0316)
2	(-0.4,3.2,0.15)	10	8	(0.1309,0.0986,0.0400)

k_0 表示文献[14]中的算法对算例 1、算例 2 的迭代次数，k_1 表示新算法迭代次数。

5 结语

本文通过加权形式构造了一类改进的拟牛顿方程，提出了一类改进的拟牛顿算法，在一定条件下证明了新算法的全局收敛性。并通过数值试验表明新算法是有效的，相对文献[14]在迭代次数上有一定提高。并希望通过新算法，结合文献[16]的研究方向，求解非线性方程组。

参考文献

[1] 袁亚湘,孙文瑜. 最优化理论与方法[M].北京:科学出版社,1997(7):183-218.

[2] 刘辉．解非线性方程的牛顿迭代法及其应用[J]．重庆工学院学报:自然科学版,2007,21(8):5-8．

[3] 楚添定,马柏林. 基于新的拟牛顿方程的 Broyden-Fletcher-Goldfarb-Shanno 算法[J].应用数学与计算数学学报,2012,26(4):360-367.

[4] 王海滨. 基于新拟牛顿方程的一类超线性收敛的改进 BFGS 算法[J].兰州理工大学学报,2007,33(4):150-152.

[5] 郑发美,刘辉辉. 一类新拟牛顿算法的全局收敛性与数值试验[J].河南师范大学学报:自然科学版,2010,38(2):35-38.

[6] 陈兰平,焦宗聪. 一般无约束优化问题的广义拟牛顿法[J].数学进展,2007,36(1):81-85．

[7] 焦宝聪. 一类超线性收敛的广义拟牛顿算法[J].高等学校计算数学学报,1999,21(6):178-188.

[8] Wei Zengxin, Li Guoyin, Qi Liqun. New quasi-Newton methods for unconstrained optimization problems[J]. Applied Mathematics and Computation 2006,175: 1156-1188.

[9] 时平平,王希云. 基于新拟牛顿方程的拟牛顿法对一般目标函数的全局收敛性[J].太原科技大学学报,2008,29(3):1673-2057.

[10] Li Donghui,Fukushima．On the global convergence of the BFGS method for nonconvex unconstrained optimization problems[J]．SIMA J．Optim．Anal,2001(4)：1054-1064．

[11] Xiao Wei,Sun Fengjian．On quasi-Newton methods with modified quasi-Newton equation[C]．In Proceeding of 2008 International Pre—Olympic CongressonComputer Science,Volume II:Information Science and Engineering，edited by Yong Jiang and JianLiang Li. World-AcademicPress.2008:359-363.

[12] Yuan Gonglin, Wei Zengxin. The superlinear convergence analysis of a non-monotone BFGS algorithm on convex objectivefunctions[J]. Acta Mathmatic Sinica English-Series,2008,19(1):35-42.

[13] 赵云彬,段虞荣. 伪 Newton-δ 族算法对一般目标函数的收敛性[J].数值计算与计算机应用,1996(1):36-47．

[14] 赵云彬,易正俊. 伪 Newton-δ 族的导出和全局收敛性[J].数值计算与计算机应用,1995(1):53-62．

[15] 宫恩龙,段立宁,高苗苗,等. 基于修正拟牛顿方程的两阶段非单调稀疏对角变尺度梯度投影算法[J].数学的实践与认识,2017,47(6):233-242．

[16] 徐林．基于改进拟牛顿法求解非线性方程组[J]．济宁学院学报,2016,37(6):54-57.