International Journal of Mathematical Archive-3(10), 2012, 3773-3776
Available online through www.ijma.info ISSN 2229 – 5046

A MODIFIED QUASI-NEWTON METHOD FOR UNCONSTRAINED OPTIMIZATION

Zhongqing Hu[1], Zhijun Luo[1*], Mingjian Huan[2] and Jianxue Luo[2]

[1]Department of Mathematics, Hunan University of Humanities, Science and Technology, Loudi, 417000, P. R. China
[2]Xinhua Shang Mei Middle School, Xinhua 417600, P. R. China

(Received on: 03-10-12; Accepted on: 30-10-12)

ABSTRACT

In this paper, a new modified BFGS Quasi-Newton is presented for unconstrained optimization problems. Under some mild conditions, the global convergence of modified Quasi-Newton method is obtained with the Wolfe line search.

Keywords: *Unconstrained optimization; Quasi-Newton method; Wolfe line search; Global convergence*

Mathematics Subject Classification (2000): *90C30 65K05.*

1. INTRODUCTION

Quasi-Newton method is one of the most effective methods for solving the following unconstrained optimization problem:

$$\min_{x\in R^n} f(x) \tag{1.1}$$

where R^n denotes an n-dimensional Euclidean space and $f: R^n \to R$ is a smooth and nonlinear function.

Throughout the paper, we assume that $\|\cdot\|$ denotes the

Euclidean norm, if x^k is the current iterate, we denote $f(x^k) \triangleq f_k$, $\nabla f(x^k) \triangleq g_k$, $\nabla^2 f(x^k) \triangleq G_k$ and $f(x^*) \triangleq f_*$,

respectively. Solving (1.1) by means of the following iteration methods:

$$x^{k+1} = x^k + \alpha_k d^k \qquad k = 0,1,2\cdots,$$

Where α_k is a stepsize, d^k is a descent direction of $f(x)$ at x^k is a step size.

The first quasi-Newton algorithm was proposed by W.C. Davidon [1] and modified by Fletcher and Powell[2]. Newton's method assumes that the function can be locally approximated as a quadratic in the region around the optimum, and uses the first and second derivatives to find the stationary point. In quasi-Newton methods the Hessian matrix of second derivatives of the function to be minimized does not need to be computed. The Hessian is updated by analyzing successive gradient vectors instead. Since second derivatives are not required, quasi-Newton methods are sometimes more efficient than Newton's method. Because the BFGS algorithm is generally considered to be the most effective iterative method, it is a topic of many active researches [3-7], The search direction of the BFGS algorithm is determined as follows:

Corresponding author: *Zhijun Luo[1*], [1]Department of Mathematics, Hunan University of Humanities, Science and Technology, Loudi, 417000, P. R. China*

$$\begin{cases} d_0 = -B_0 g_0 \\ d_k = -B_k g_k \qquad (k \ge 1), \end{cases}$$

B_0 is any given n × n symmetric positive definite matrix, B_k is update by

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k},$$

where $s_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$. As we know, the BFGS update only exploits the gradient information, while the function values available are neglected. Hence, many efficient attempts have been made to modify the usual quasi-Newton methods using both the gradient and function value information (e.g. [5,6]). Lately, in order to get a higher order accuracy in approximating the Hessian matrix of objective function, Wei, Li, and Qi (see [7]) proposed a modified BFGS-type method for the solution of (1.1).

In this paper, put forwards a kind of more general improved Quasi-Newton equation. It extends the existed results and brings out an improved BFGS method correspondence as follows:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^* y_k^{*T}}{y_k^{*T} s_k}, \tag{1.2}$$

$$y_k^* = \frac{1}{\gamma_k}\left[g_{k+1} - \theta_k g_k + A_k s_k\right], \tag{1.3}$$

$$A_k = \frac{2\left[f(x^k) - f(x^{k+1})\right] + \left[g(x^{k+1}) + g(x^k)\right]^T s_k}{\|s_k\|^2}, \tag{1.4}$$

where γ_k, θ_k are nonnegative scalar, satisfying $s_k^T G_{k+1} s_k = s_k^T y_k^*$ and $(\theta_k - 1) g_k^T s_k = 2(\gamma_k - 1)(f_{k+1} - f_k - g_{k+1}^T s_k)$. The Wolfe line search is proposed for the modified BFGS Quasi-Newton method. Under some mild conditions, the Wolfe line search can guarantee the global convergence of modified Quasi-Newton method.

The organization of this paper is as follows. The algorithm is presented in Section 2. In Sections 3 the global convergence is analyzed.

2. DESCRIPTION OF ALGORITHM

Based on (1.2), we state the modified BFGS algorithm as follows.

Algorithm A:

Step 0: Choose $x^0 \in R^n$, an initial positive matrix B_0, $\forall \varepsilon > 0$. Let $k := 0$.

Step 1: If $\| g_k \| \le \varepsilon$ then STOP else go to Step 2.

Step 2: For given x^k and B_k, calculate the search direction d^k by $B_k d^k = -g_k$.

Step 3: Compute steplength α_k by Wolfe line search

$$f(x^k+\alpha_k d^k)-f(x^k)\le \delta\alpha_k g_k^T d^k, \tag{2.1}$$

$$g(x^k+\alpha_k d^k)^T d^k \ge \sigma g_k^T d^k, \tag{2.2}$$

where $\delta,\sigma\in(0,1)$.

Step 4: Let $x^{k+1}=x^k+\alpha_k d^k$, update B_{k+1} by formula (1.2).

Step 5: Set $k=:k+1$ go to step 1.

3. GLOBAL CONVERGENCE OF ALGORITHM

In this section, we discuss the global convergence behavior of Algorithm A. Firstly, we make the following assumptions.

H 3.1 The objective function $f(x)$ is twice continuously differentiable. on R^n.

H 3.2 The gradient $g(x)$ of $f(x)$ is Lipschitz continuous on an open convex set U that contains the level set

$L_0=\{x\in R^n \mid f(x)\le f(x^0)\}$ with x^0 being given, i.e., there exists $L>0$ such that

$$\|g(x)-g(y)\|\le L\|x-y\|, \forall x, y\in U.$$

H 3.3 The function f in (1.1) is uniformly convex, i.e., there exist two positive constants $m\le M$ such that $m\|z\|^2\le z^T G(x) z\le M\|z\|^2$, $\forall z\in R^n$, where $x\in L_0$, $\nabla^2 f(x)=G(x)$.

Lemma3.1 [5]: The sequence $\{x^k\}$ generated by Algorithm A, There exist constant $m_1>0, m_2>0$ such that

$$\frac{s_k^T y_k^*}{s_k^T s_k}\ge m_1,\quad \frac{\|y_k^*\|^2}{y_k^{T*} s_k}\le m_2. \tag{3.1}$$

Hence, for any $p\in(0,1)$, there exist positive constants $\beta_1,\beta_2,\beta_3\ge 0$, for any $k\ge 1$, such that the following inequality:

$$\beta_2\|s_k\|^2\le s_j^T B_j s_j\le \beta_3\|s_j\|^2,\quad \|B_j s_j\|\le \frac{\beta_3}{\beta_1}\|s_j\|, \tag{3.2}$$

holds for at least $[pk]$ values of $j\in[1,k]$.

Lemma3.2: Suppose that assumptions H3.1–H3.3 hold, the sequence $\{x^k\}$ generated by Algorithm A, then $\sum_{k=0}^{\infty} -g_k^T s_k<+\infty$ is hold.

Proof: From the Assumptions, it is easy obtain

$$\sum_{k=0}^{\infty}[f(x_k)-f(x_{k+1})]<+\infty$$

Together with formula (2.1) $f(x^k+\alpha_k d^k)-f(x^k)\le \delta\alpha_k g_k^T d^k$, the result is obtained.

Consequently, according to the assumptions H3.1–H3.3 and Lemma 3.2, it is easy to see, for every k, that the line search at Step3 is always successful. Thereby, it is known that Algorithm A is well-defined.

Theorem 3.1: If the assumptions H3.1–H3.3 hold, the sequence $\{x^k\}$ generated by Algorithm A, then, $\liminf_{k\to\infty} \|g_k\| = 0.$

Proof: Suppose that the desired conclusion is false, i.e. $\liminf_{k\to\infty} \|g_k\| \geq 0.$ There exists a positive constant $\varepsilon > 0$ such that

$$\|g_k\| \geq \varepsilon > 0, k = 1, 2, 3...$$

From Lemma 3.1 and $B_k s_k = -\alpha_k g_k$ we know that for any integer $k \geq 1$

$$+\infty > \sum_{k=0}^{\infty} -g_k^T s_k \geq \sum_{k\in K} -g_k^T s_k = \sum_{k\in K} g_k^T \cdot \frac{\alpha_k g_k}{B_k} = \sum_{k\in K} \alpha_k \|g_k\|^2 \frac{\|s_k\|^2}{\|B_k s_k\|^2} \frac{s_k^T B_k s_k}{\|s_k\|^2}$$

$$\geq \sum_{k\in K} (a\varepsilon^2 \beta_1^2 \beta_2 \beta_3^{-2}) = +\infty,$$

which contradicts Lemma3.2. Therefore, $\liminf_{k\to\infty} \|g_k\| = 0.$

ACKNOWLEDGEMENT

This work was supported in part by Hunan Provincial Department of education teaching reform project **(NO.[2011]333**) and Educational Reform Research Fund of Hunan Institute of Humanities, Science and Technology(**NO.RKJGY1030**).

REFERENCES

[1] W. C. Davidon, Variable Metric Method for Minimization, A. E. C. Res. and Develop. Report ANL-5990 (Rev. TID-4500, 14th ed.) 1959.

[2] M.J.D. Powell, On the convergence of the variable metric algorithm, J. Inst. Math. Appl. 7 (1971) 21–36.

[3] Y.-X.YUAN. A Modified BFGS Algorithm for Unconstrained Optimization, IMAJ Numer Anal (1991) 11 (3): 325-332.

[4] D.T. Zhu, Nonmonotone backtracking inexact quasi-Newton algorithms for solving smooth nonlinear equations, Appl. Math. Comput. 161 (2005) 875-895.

[5] Z. Wei, G. Yu, G. Yuan, Z. Lian, The superlinear convergence of a modified BFGS-type method for unconstrained optimization, Comput.Optim. Appl. 29 (2004) 315–332.

[6] J.Z. Zhang, N.Y. Deng, L.H. Chen, New quasi-Newton equation and related methods for unconstrained optimization, J. Optim. Theory Appl.102 (1999) 147–167.

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

Source of support: Hunan Provincial Department of education teaching, China,

Conflict of interest: None Declared