Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 8342536, 9 pages
https://doi.org/10.1155/2021/8342536

Research Article

The Global Convergence of a Modified BFGS Method under Inexact Line Search for Nonconvex Functions

Pengyuan Li, Junyu Lu, and Haishan Feng

College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, China

Correspondence should be addressed to Haishan Feng; 516791535@qq.com

Received 21 August 2020; Accepted 22 April 2021; Published 6 May 2021

Academic Editor: Wenjie Liu

Among the quasi-Newton algorithms, the BFGS method is often discussed by related scholars. However, in the case of inexact Wolfe line searches or even exact line search, the global convergence of the BFGS method for nonconvex functions is not still proven. Based on the aforementioned issues, we propose a new quasi-Newton algorithm to obtain a better convergence property; it is designed according to the following essentials: (1) a modified BFGS formula is designed to guarantee that B_{k+1} inherits the positive definiteness of B_k; (2) a modified weak Wolfe–Powell line search is recommended; (3) a parabola, which is considered as the projection plane to avoid using the invalid direction, is proposed, and the next point x_{k+1} is designed by a projection technique; (4) to obtain the global convergence of the proposed algorithm more easily, the projection point is used at all the next iteration points instead of the current modified BFGS update formula; and (5) the global convergence of the given algorithm is established under suitable conditions. Numerical results show that the proposed algorithm is efficient.

1. Introduction

Consider

$$\min\{f(x)|x\in\Re^n\},\tag{1}$$

where $f\colon \Re^n \longrightarrow \Re$ and $f\in C^2$. The multitudinous algorithms for (1) often use the following iterative formula:

$$x_{k+1}=x_k+s_k,\quad k=0,1,2\ldots,\tag{2}$$

where x_k is the current point, $s_k=x_{k+1}-x_k=\alpha_k d_k$, α_k is a step size, and d_k is a search direction at x_k. There exist many algorithms for (1) [1–9]. Davidon [10] pointed out that the quasi-Newton method is one of the most effective methods for solving nonlinear optimization problems. The idea of the quasi-Newton method is to use the first derivative to establish an approximate Hessian matrix in many iterations, and the approximation is updated by a low-rank matrix in each iteration. The primary quasi-Newton equation is as follows:

$$B_{k+1}s_k=y_k, y_k=g_{k+1}-g_k.\tag{3}$$

The search direction d_k of the quasi-Newton method is generated by the following equation:

$$d_0=-H_0 g_0, d_k=-H_k g_k,\quad k\ge 1,\tag{4}$$

where H_0 is any given $n\times n$ symmetric positive-definite matrix, $H_k=B_k^{-1}$, the Hessian approximation matrix B_k is the quasi-Newton update matrix, and $g_k=g(x_k)$ is the gradient of $f(x)$ at x_k. The BFGS (Broyden [11], Fletcher [12], Goldfarb [13], and Shanno [14]) method is one of the quasi-Newton line search methods and has great numerical stability. The famous BFGS update formula is

$$B_{k+1}=B_k-\frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}+\frac{y_k y_k^T}{s_k^T y_k},\tag{5}$$

which is effective for solving (1) [15–18]. Powell [19] first proved that the BFGS method possesses global convergence for convex functions under Wolfe line search. Some global convergence results for the BFGS method for convex minimization problems can be found in [19–26]. However, Dai [16] proposed a counterexample to illustrate that the standard BFGS method may not be applicable to nonconvex functions with Wolfe line search, and Mascarenhas [27] demonstrated the nonconvergence of the standard BFGS method even with exact line search. To verify the global convergence of the BFGS method for general functions, some modified BFGS methods [28–31] have been also presented for nonconvex minimization problems. Aiming to obtain a better approximation of the objective function Hessian matrix, Wei et al. [32] proposed a new BFGS method, whose formula is

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^* y_k^{*T}}{s_k^T y_k^*}, \tag{6}$$

where $y_k^* = y_k + C_k/\|s_k\|^2 s_k$ and $C_k = 2[f(x_k) - f(x_k + \alpha_k d_k)] + (g(x_k + \alpha_k d_k) + g(x_k))^T s_k$, and the corresponding quasi-Newton equation is as follows:

$$B_{k+1} s_k = y_k^*. \tag{7}$$

For convex functions, convergence analysis of the new BFGS algorithm was given for weak Wolfe–Powell line search:

$$\begin{aligned} f(x_k + \alpha_k d_k) &\le f_k + \delta \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k &\ge \sigma g_k^T d_k, \end{aligned} \tag{8}$$

where $\delta \in (0, 1/2)$ and $\sigma \in (\delta, 1)$.

Motivated by the above formula and other observations, Yuan and Wei [33] defined a modified quasi-Newton equation as follows:

$$B_{k+1} s_k = y_k^m, \tag{9}$$

where $y_k^m = y_k + \max\{C_k, 0\}/\|s_k\|^2 s_k$ and

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^m y_k^{mT}}{s_k^T y_k^m}. \tag{10}$$

It is obvious that if $C_k > 0$ holds, then the quasi-Newton method is the method (7); otherwise, it is the standard BFGS method. Therefore, when $C_k > 0$ holds, the modified quasi-Newton method (10) and the quasi-Newton method (6) have the same approximation of the Hessian matrix. Inspired by their views, we will demonstrate the global convergence of the modified BFGS (MBFGS) method (10) for nonconvex functions with the modified weak Wolfe–Powell (MWWP) line search [34], whose form is as follows:

$$f(x_k + \alpha_k d_k) \le f_k + \delta \alpha_k g_k^T d_k + \alpha_k \min\left[-\delta_1 g_k^T d_k, \delta \frac{\alpha_k}{2} \|d_k\|^2\right], \tag{11}$$

$$g(x_k + \alpha_k d_k)^T d_k \ge \sigma g_k^T d_k + \min\left[-\delta_1 g_k^T d_k, \delta \alpha_k \|d_k\|^2\right], \tag{12}$$

where $\delta \in (0, 1/2), \delta_1 \in (\delta/2, \delta)$, and $\sigma \in (\delta, 1)$. The parameter δ_1 is different from that in paper [34].

This article is organized as follows: Section 2 introduces the motivation and states the given technique and algorithm. In Section 3, we prove the global convergence of the modified BFGS method with MWWP line search under some reasonable conditions. Section 4 reports the results of the numerical experiments to show the performance of the algorithms. The last section presents the conclusion. Throughout the article, $f(x_k)$ and $f(x_{k+1})$ are replaced by f_k and f_{k+1}, and $g(x_k)$ and $g(x_{k+1})$ are replaced by g_k and g_{k+1}. $\|\cdot\|$ denotes the Euclidean norm.

2. Motivation and Algorithm

The global convergence of the BFGS algorithm has been established for the uniformly convex functions which have many advantages. It is worth considering whether we can use these properties of uniformly convex functions in the BFGS algorithm to obtain global convergence. This idea motivates us to propose a projection technique to acquire better convergence properties of the BFGS algorithm. Given a new numerical formula for (1):

$$z_k = x_k + \alpha_k d_k, \tag{13}$$

where z_k is the next point generated by the classical BFGS formula. Moreover, a parabolic form is given as follows:

$$\left\{x \mid g(z_k)^T (z_k - x) + \lambda \left\|z_k - x\right\|^2 = 0\right\}, \tag{14}$$

where $\lambda > 2$ is a constant. It is not difficult to see that $Q_k(x) = g(z_k)^T (z_k - x) + \lambda \|z_k - x\|^2$ can be considered as the first two terms of the expansion of a quadratic function at z_k, whose Hessian matrix is a diagonal matrix with eigenvalue -2λ. Therefore, the BFGS method is globally convergent. By projecting x_k onto (14), we obtain the next step x_{k+1}:

$$x_{k+1} = x_k + \frac{g(z_k)^T (z_k - x_k) + \lambda \left\|z_k - x_k\right\|^2}{\left\|g(z_k) - g(x_k)\right\|^2} \left[g(z_k) - g(x_k)\right]. \tag{15}$$

The idea of the projection can also be found in [6, 8, 35]. Based on the above discussions, the modified algorithm is given in Algorithm 1.

Remark 1

(i) x_{k+1} is the defined projection point in Step 6, and vector y_k^τ is the same as vector y_k in Step 5, where the projection point x_{k+1} does not work in (10) but does in the next iteration.

(ii) If $-\delta_1 g_k^T d_k > \delta \alpha_k \|d_k\|^2$ holds in Step 5, then the global convergence of the algorithm can be obtained by the modified weak Wolfe–Powell line search, (11) and (12). If not, we can ensure the global convergence of the algorithm using the projection method (15).

3. Convergence Analysis

In this section, we concentrate on the global convergence of the modified projection BFGS algorithm. The following assumptions are required.

Step 0: given a point $x_0 \in \Re^n$, constants $\varepsilon \in (0,1), \delta \in (0,1/2), \sigma \in (\delta,1), \delta_1 \in (\delta/2,\delta), \lambda \in (2,+\infty)$, and an $n \times n$ symmetric positive-definite matrix B_0, set $k=0$.
Step 1: stop if $\|g_k\| \le \varepsilon$.
Step 2: obtain a search direction d_k by solving
$B_k d_k + g_k = 0$
Step 3: calculate α_k using the inequalities (11) and (12).
Step 4: set $w_k = x_k + \alpha_k d_k$.
Step 5: if $-\delta_1 g_k^T d_k > \delta\alpha_k \|d_k\|^2$ holds, then let $x_{k+1} = w_k$, $y_k = g(x_{k+1}) - g(x_k)$, and go to Step 7; otherwise, go to Step 6.
Step 6: let x_{k+1} be defined by (15), $y_k^\tau = g(w_k) - g(x_k)$, and $y_k := y_k^\tau$.
Step 7: update B_{k+1} using formula (10).
Step 8: let $k = k+1$, and go to Step 1.

Algorithm 1: Algorithm MBFGS.

Assumption 1

(i) The level set $\Gamma = \{x \in \Re^n | f(x) \le f(x_0)\}$ is bounded.

(ii) The function $f(x)$ is twice continuously differentiable and bounded from below, and its gradient function $g(x)$ is Lipschitz continuous, that is,

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad x, y \in \Re^n, \tag{16}$$

holds, where $L > 0$ is a constant.Assumption 1(ii) indicates that the relation

$$\left\|G(\xi)s_k\right\| \le L\left\|s_k\right\|, G(\xi) = \nabla^2 f(\xi), \quad \xi \in \left(x_k, x_{k+1}\right), \tag{17}$$

holds.

Theorem 1. *Suppose that Assumption 1 and $g_k^T d_k \le 0$ hold. Then, there exists a constant α satisfying (11) and (12), where $\alpha \in [p_m, p_n]$, and $p_n > p_m > 0$ are constants.*

Proof. The detailed proof of the rationality of the line search is given in paper [35]. □

Lemma 1. *Let Assumption 1 and $g_k^T d_k \le 0$ hold. If the sequence $\{x_k\}$ is generated by Algorithm 1, then we have*

$$s_k^T y_k^m \ge \beta\left\|\alpha_k d_k\right\|^2, \tag{18}$$

where $\beta > 0$ is a constant.

Proof. According to Lemma 1 of paper [35], the following relations are reasonable:

$$s_k^T y_k \ge \delta\left\|\alpha_k d_k\right\|^2, \quad \text{if } -\delta_1 g_k^T d_k > \delta\alpha_k\left\|d_k\right\|^2, \tag{19}$$

$$s_k^T y_k \ge \beta_1\left\|\alpha_k d_k\right\|^2, \quad \text{if } -\delta_1 g_k^T d_k \le \delta\alpha_k\left\|d_k\right\|^2, \tag{20}$$

□

where parameter $\beta_1 = \lambda - \delta(\sigma - \delta_1)/\delta_1 > 0$. Combining (19) and (20), we obtain the following formula with $\beta = \min\{\delta, \beta_1\}$:

$$s_k^T y_k \ge \beta\left\|\alpha_k d_k\right\|^2. \tag{21}$$

Using the definition of y_k^m, we obtain

$$s_k^T y_k^m = s_k^T y_k + \max\{C_k, 0\} \ge s_k^T y_k. \tag{22}$$

Therefore, $s_k^T y_k^m \ge \beta\|\alpha_k d_k\|^2$. The proof is complete.

Lemma 2. *If the sequence $\{B_k\}$ is generated by Algorithm 1 and Assumption 1 holds, then the matrix B_k is positive definite for all k.*

Proof. According to (18), the relation $s_k^T y_k^m > 0$ is valid. Thus, the proof is complete. □

Lemma 3. *If the sequence $\{\alpha_k, d_k, x_k, g_k\}$ is generated by Algorithm 1 and Assumption 1 holds, then*

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\left\|d_k\right\|^2} < \infty. \tag{23}$$

Proof. According (11) and Assumption 1(ii), the following formula obviously exists

$$\sum_{k=0}^{\infty}\left(-\alpha_k g_k^T d_k\right) < \infty. \tag{24}$$

□

Combining (12) with (16), we obtain

$$\begin{aligned}\alpha_k L\left\|d_k\right\|^2 &\ge \left(g_{k+1} - g_k\right)^T d_k \\ &\ge -(1-\sigma)g_k^T d_k + \min\left[-\delta_1 g_k^T d_k, \delta\alpha_k\left\|d_k\right\|^2\right] \\ &\ge -(1-\sigma)g_k^T d_k.\end{aligned} \tag{25}$$

Thus,

$$\alpha_k \ge \frac{-(1-\sigma)g_k^T d_k}{L\left\|d_k\right\|^2}. \tag{26}$$

Substituting the above inequality into (24), we have (23). The proof is complete.

Lemma 4. *Let Assumption 1 and the inequality $g_{k+1}^T d_k \le -\delta g_k^T d_k$ hold. If there exist constants $\rho_1 \ge \rho_2 > 0$, then the following hold:*

$$\|B_k s_k\| \le \rho_1 \|s_k\|, \tag{27}$$

$$s_k^T B_k s_k \ge \rho_2 \|s_k\|^2, \tag{28}$$

for at least $[t/2]$ values of $k \in [1, t]$ for any positive integer t.

Proof. The proof will be completed using the following two cases:

Case 1. If $-\delta_1 g_k^T d_k > \delta \alpha_k \|d_k\|^2$ is true, then $y_k = g(z_k) - g(x_k) = g(x_{k+1}) - g(x_k)$ and $s_k = z_k - x_k = x_{k+1} - x_k = \alpha_k d_k$. By (16) and Assumption 1,

$$\|y_k\| = \|g(x_{k+1}) - g(x_k)\| \le L\|x_{k+1} - x_k\| = L\|\alpha_k d_k\|. \tag{29}$$

Combining the above inequality with (19), we obtain

$$\frac{s_k^T y_k}{\|s_k\|^2} \ge \delta \text{ and } \frac{\|y_k\|^2}{s_k^T y_k} \le \frac{L^2}{\delta}. \tag{30}$$

Assumption 1(ii) and the definition of C_k imply that

$$|C_k| \le \|y_k\|\|s_k\| + L\|s_k\|^2. \tag{31}$$

By (31), (19), and the definition of y_k^m, we have

$$\begin{aligned} \|y_k^m\| &\le \|y_k\| + \left\| \frac{\max\{C_k, 0\} s_k}{\|s_k\|^2} \right\| \\ &\le 2\|y_k\| + \frac{L}{\delta}\|y_k\| = \left(2 + \frac{L}{\delta}\right)\|y_k\|. \end{aligned} \tag{32}$$

Relations (22), (30), and (32) indicate that

$$\frac{s_k^T y_k^m}{\|s_k\|^2} \ge \delta \text{ and } \frac{\|y_k^m\|^2}{s_k^T y_k^m} \le \left(2 + \frac{L}{\delta}\right)^2 \frac{L^2}{\delta}. \tag{33}$$

Case 2. $-\delta_1 g_k^T d_k \le \delta \alpha_k \|d_k\|^2$ holds. From Step 6 of Algorithm 1, we obtain

$$\|y_k\|^2 = \|y_k^\tau\|^2 = \|g(w_k) - g(x_k)\|^2 \le L^2 \|\alpha_k d_k\|^2. \tag{34}$$

Together with (32), we have

$$\|y_k^m\|^2 \le \left(2L + \frac{L^2}{\delta}\right)^2 \|\alpha_k d_k\|^2. \tag{35}$$

Integrating (35) with (20), we obtain

$$\frac{\|y_k^m\|^2}{s_k^T y_k^m} \le \frac{1}{\beta_1}\left(2L + \frac{L^2}{\delta}\right)^2. \tag{36}$$

Using the definition of s_k, we have

$$\begin{aligned} \|s_k\| &= \|x_{k+1} - x_k\| \\ &= \left\| \frac{g(z_k)^T (z_k - x_k) + \lambda \|z_k - x_k\|^2}{\|g(z_k) - g(x_k)\|^2} [g(z_k) - g(x_k)] \right\| \\ &\le \frac{g(z_k)^T (z_k - x_k) + \lambda \|z_k - x_k\|^2}{\|g(z_k) - g(x_k)\|} \\ &\le \frac{-\delta \alpha_k g_k^T d_k + \lambda \|z_k - x_k\|^2}{\|g(z_k) - g(x_k)\|} \\ &\le \frac{(\delta/\delta_1 + \lambda)\|\alpha_k d_k\|^2}{\|g(z_k) - g(x_k)\|}, \end{aligned} \tag{37}$$

□

where the second inequality follows from $g_{k+1}^T d_k \le -\delta g_k^T d_k$, and the last inequality follows from (16). Combining the above formula with (20), we obtain

$$\frac{s_k^T y_k^m}{\|s_k\|^2} \ge \frac{\beta_1 \delta_1^2 L^2}{(\delta + \lambda \delta_1)^2}. \tag{38}$$

All in all, the following formula always holds:

$$\frac{s_k^T y_k^m}{\|s_k\|^2} \ge \omega_1 \text{ and } \frac{\|y_k^m\|^2}{s_k^T y_k^m} \le \omega_2, \tag{39}$$

where $\omega_1 = \min\{\delta, \beta_1 \delta_1^2 L^2 / (\delta + \lambda \delta_1)^2\}$ and $\omega_2 = \max\{(2 + L/\delta)^2 L^2/\delta, 1/\beta_1 (2L + L^2/\delta)^2\}$. Similar to the proof of Theorem 1 in [36], we obtain (27) and (28). The proof is complete.

Theorem 2. *If the conditions of Lemma 4 hold, then we obtain*

$$\lim_{k \to \infty} \inf \|g_k\| = 0. \tag{40}$$

Proof. By (23), we can obtain

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \longrightarrow 0, \quad k \longrightarrow \infty. \tag{41}$$

□

Then, using Algorithm 1, we have

$$\frac{(d_k^T B_k d_k)^2}{\|d_k\|^2} \longrightarrow 0, \quad k \longrightarrow \infty. \tag{42}$$

The relationship between (27) and (28) indicates that

$$0 \le \rho_2^2 \|d_k\|^2 \le \frac{\left(d_k^T B_k d_k\right)^2}{\|d_k\|^2} \longrightarrow 0, \quad k \longrightarrow \infty, \tag{43}$$

which means that $\|d_k\| \longrightarrow 0$ holds for $k \longrightarrow \infty$. Combining Lemma 4 and $g_k = -B_k d_k$, we obtain

$$\rho_2 \|d_k\| \le \|g_k\| \le \rho_1 \|d_k\|. \tag{44}$$

This implies that (40) holds. The proof is complete.

4. Numerical Results

In this section, we perform some numerical experiments to test Algorithm 1 with the modified weak Wolfe–Powell line search and compare its performance with that of the normal BFGS method. We call Algorithm 1 as MBFGS.

4.1. General Unconstrained Optimization Problems

Tested problems: the problems are obtained from [37, 38]. There are 74 test questions in total, which are listed in Table 1.

Dimensionality: problem instances with 300, 900, and 2700 variables are considered.

Himmelblau stop rule [39]: if $|f(x_k)| > e_1$, then set $\text{Ter}_1 = |f(x_k) - f(x_{k+1})|$, or let $\text{Ter}_1 = |f(x_k) - f(x_{k+1})| / |f(x_k)|$. If $\|g(x)\| < \varepsilon$ or $\text{Ter}_1 < e_2$ holds, then the program is stopped, where $e_1 = e_2 = 10^{-5}$ and $\varepsilon = 10^{-6}$.

Parameters: in Algorithm 1, $\delta = 0.2$, $\delta_1 = 0.15$, $\sigma = 0.85$, $\lambda = 5$, and $B_0 = I$, which is the unit matrix.

Experiment environment: all programs are written in MATLAB R2014a and run on a PC with an Inter(R) Core(TM) i5-4210U CPU @ 1.70 GHz, 8.00 GB of RAM, and the Windows 10 operating system.

Symbol representation: some definitions of the notation used in Tables 1 and 2 are as follows:

No: the test problem number.
CPUTime: the CPUtime in seconds.
NI: the number of iterations.
NFG: the total number of function and gradient evaluations.

Image description: Figures 1–3 show the profiles of CPUTime, NI, and NFG. It is easy to see from these figures that the MBFGS method possesses the best performance since its performance curves of CPUTime, NI, and NFG are better than those of the BFGS method. In addition, numerical results of the total CPUTime, NI, and NFG of the modified BFGS method are lower than those of the BFGS method.

4.2. The Muskingum Model in Engineering Problems. The Muskingum model, whose definition is as follows, is presented in this subsection. The key work is to numerically estimate the model using Algorithm 1.

Table 1: The test problems.

No.	Test problem
1	Extended Freudenstein and Roth function
2	Extended trigonometric function
3	Extended Rosenbrock function
4	Extended White and Holst function
5	Extended Beale function
6	Extended penalty function
7	Perturbed quadratic function
8	Raydan 1 function
9	Raydan 2 function
10	Diagonal 1 function
11	Diagonal 2 function
12	Diagonal 3 function
13	Hager function
14	Generalized tridiagonal-1 function
15	Extended tridiagonal-1 function
16	Extended three exponential terms function
17	Generalized tridiagonal-2 function
18	Diagonal 4 function
19	Diagonal 5 function
20	Extended Himmelblau function
21	Generalized PSC1 function
22	Extended PSC1 function
23	Extended Powell function
24	Extended block diagonal BD1 function
25	Extended Maratos function
26	Extended Cliff function
27	Quadratic diagonal perturbed function
28	Extended wood function
29	Extended Hiebert function
30	Quadratic function QF1 function
31	Extended quadratic penalty QP1 function
32	Extended quadratic penalty QP2 function
33	A quadratic function QF2 function
34	Extended EP1 function
35	Extended tridiagonal-2 function
36	BDQRTIC function (CUTE)
37	TRIDIA function (CUTE)
38	ARWHEAD function (CUTE)
39	NONDIA function (CUTE)
40	NONDQUAR function (CUTE)
41	DQDRTIC function (CUTE)
42	EG2 function (CUTE)
43	DIXMAANA function (CUTE)
44	DIXMAANB function (CUTE)
45	DIXMAANC function (CUTE)
46	DIXMAANE function (CUTE)
47	Partial perturbed quadratic function
48	Broyden tridiagonal function
49	Almost perturbed quadratic function
50	Tridiagonal perturbed quadratic function
51	EDENSCH function (CUTE)
52	VARDIM function (CUTE)
53	STAIRCASE S1 function
54	LIARWHD function (CUTE)
55	DIAGONAL 6 function
56	DIXON3DQ function (CUTE)
57	DIXMAANF function (CUTE)
58	DIXMAANG function (CUTE)
59	DIXMAANH function (CUTE)
60	DIXMAANI function (CUTE)

Table 1: Continued.

No.	Test problem
61	DIXMAANJ function (CUTE)
62	DIXMAANK function (CUTE)
63	DIXMAANL function (CUTE)
64	DIXMAAND function (CUTE)
65	ENGVAL1 function (CUTE)
66	FLETCHCR function (CUTE)
67	COSINE function (CUTE)
68	Extended DENSCHNB function (CUTE)
69	Extended DENSCHNF function (CUTE)
70	SINQUAD function (CUTE)
71	BIGGSB1 function (CUTE)
72	Partial perturbed quadratic PPQ2 function
73	Scaled quadratic SQ1 function
74	Scaled quadratic SQ2 function

Table 2: Numerical results for the total CPUTime, NI, and NFG.

Algorithm	CPUTime	NI	NFG
BFGS	62299.85938	23833	57444
MBFGS	60990.64063	22357	51370

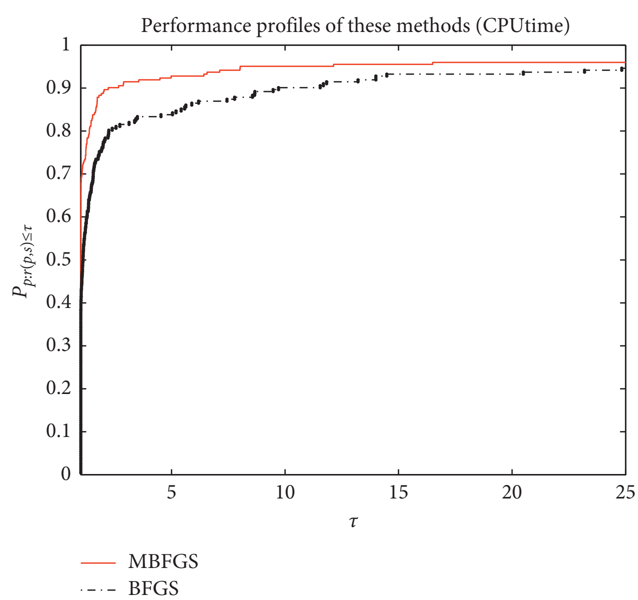

Figure 1: Performance profiles of the methods (CPUTime).

Muskingum Model [40]:

$$\begin{aligned}\min f(x_1,x_2,x_3) = \sum_{i=1}^{n-1}\Bigg(&\left(\left(1-\frac{\Delta t}{6}\right)x_1\left(x_2 I_{i+1}+(1-x_2)Q_{i+1}\right)^{x_3}\right. \\ &-\left(1-\frac{\Delta t}{6}\right)x_1\left(x_2 I_i+(1-x_2)Q_i\right)^{x_3} \\ &\left.-\frac{\Delta t}{2}\left(I_i-Q_i\right)+\frac{\Delta t}{2}\left(1-\frac{\Delta t}{3}\right)\left(I_{i+1}-Q_{i+1}\right)\right)^2,\end{aligned} \tag{45}$$

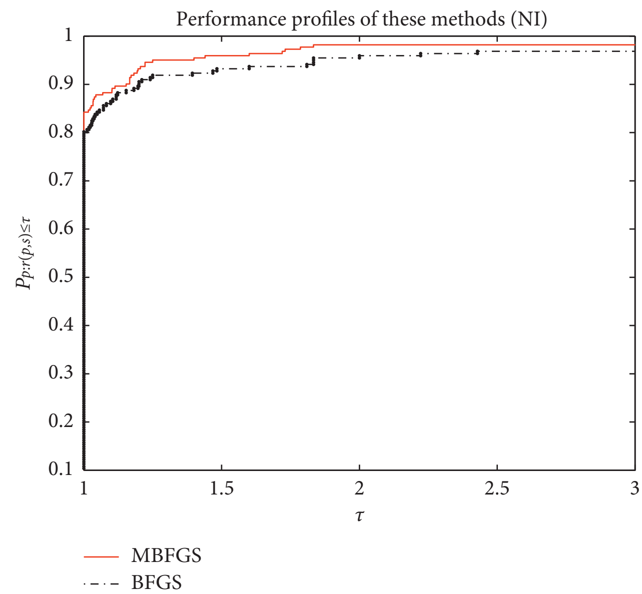

Figure 2: Performance profiles of these methods (NI).

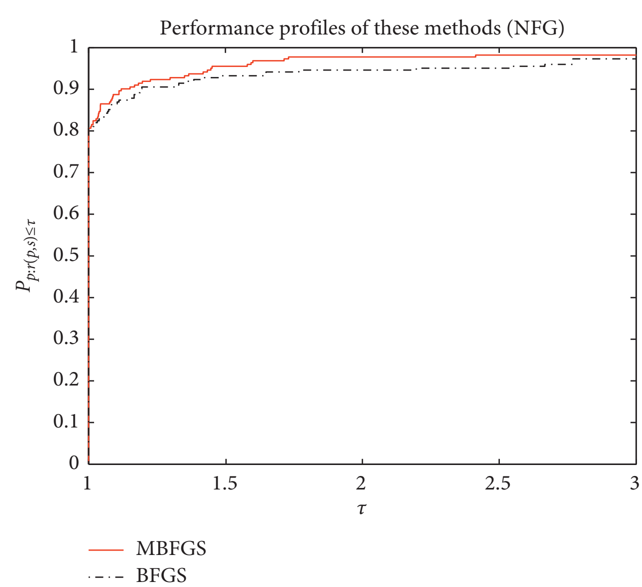

Figure 3: Performance profiles of these methods (NFG).

whose symbolic representation is as follows: x_1 is the storage time constant, x_2 is the weight coefficient, x_3 is an extra parameter, I_i is the observed inflow discharge, Q_i is the observed outflow discharge, Δt is the time step at time t_i $(i = 1, 2, \ldots, n)$, and n is the total time.

The observed data of the experiment are derived from the process of flood runoff from Chenggouwan and Linqing of Nanyunhe in the Haihe Basin, Tianjin, China. To obtain better numerical results, the initial point $x = [0, 1, 1]^T$ and the time step $\Delta t = 12\,(h)$ are selected. The specific values of I_i and Q_i for the years 1960, 1961, and 1964 are stated in article [41]. The test results are listed in Table 3.

Table 3: Results of the three algorithms.

Algorithm	x_1	x_2	x_3
BFGS [42]	10.8156	0.9826	1.0219
HIWO [40]	13.2813	0.8001	0.9933
MBFGS	11.1832	1.0000	0.9996

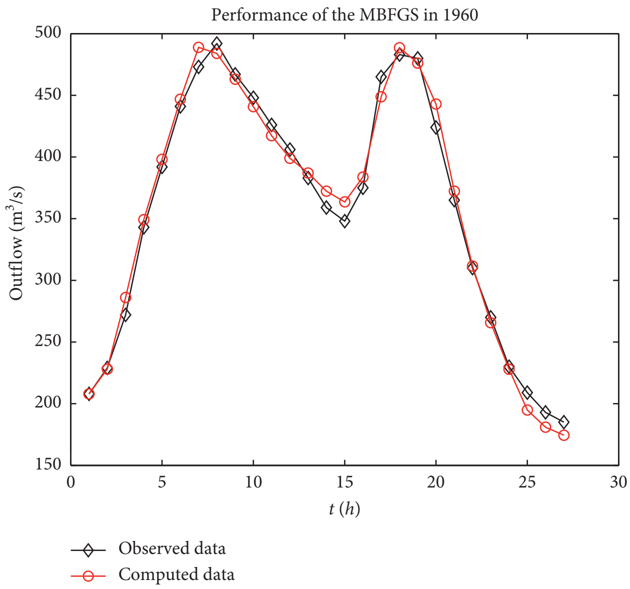

Figure 4: Performance of the MBFGS in 1960.

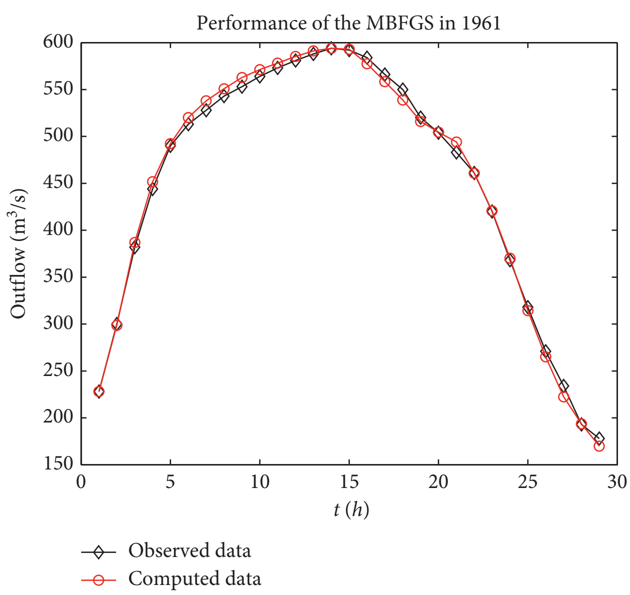

Figure 5: Performance of the MBFGS in 1961.

The following three conclusions are apparent from Figures 4–6 and Table 3: (1) Combined with the Muskingum model, the MBFGS method has great numerical experiment performance, similar to the BFGS method and the HIWO method, and the efficiency of these three algorithms is fascinating. (2) The final points (x_1, x_2, and x_3) of the MBFGS method are emulative of those of other similar methods. (3) Owing to the final points of these three methods being distinct,

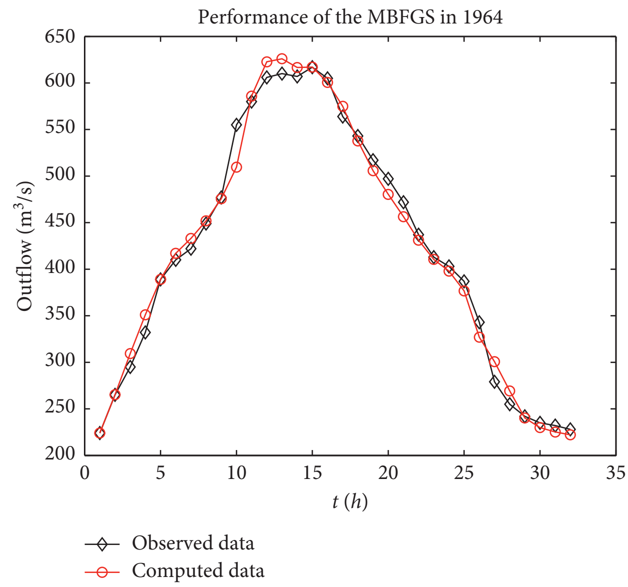

Figure 6: Performance of the MBFGS in 1964.

the Muskingum model may have more approximation optimum points.

5. Conclusion

This paper gives a modified BFGS method and studies its global convergence under an inexact line search for nonconvex functions. A new algorithm is proposed, which has the following properties: (i) The search direction and its associated step size are accepted if a positive condition holds, and the next iterative point is designed; otherwise, a parabola is introduced, which is regarded as the projection surface to avoid using the failed direction, and the next point x_{k+1} is designed by a projection technique. (ii) To easily obtain global convergence of the proposed algorithm, the projection point is used at all the next iteration points instead of the current modified BFGS update formula. The global convergence for nonconvex functions and the numerical results of the proposed algorithm indicated that the given method is competitive with other similar methods. As for future work, we have the following points to consider: (a) Is there a new projection technique suitable for the global convergence of the modified BFGS method? (b) Application of the modified BFGS method (10) to other line search techniques should be discussed. (c) Whether the combination of the projection technique mentioned and a conjugate gradient method, especially the PRP method, has good numerical experimental results is worthy of investigation.

Data Availability

All data supporting the findings are included in the paper.

Conflicts of Interest

There is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 11661009), the High Level Innovation Teams and Excellent Scholars Program in Guangxi Institutions of Higher Education (Grant No. [2019] 52)), the Guangxi Natural Science Key Fund (No. 2017GXNSFDA198046), and the Special Funds for Local Science and Technology Development Guided by the Central Government (No. ZY20198003).

References

[1] Z. Dai, H. Zhou, J. Kang, and F. Wen, "The skewness of oil price returns and equity premium predictability," *Energy Economics*, vol. 94, p. 105069, 2021.
[2] Z. Dai and J. Kang, "Some new efficient mean-variance portfolio selection models," *International Journal of Finance and Economics*, vol. 2021, pp. 1–13, 2021.
[3] Z. Dai and H. Zhu, "A modified Hestenes-Stiefel-type derivative-free method for large-scale nonlinear monotone equations," *Mathematics*, vol. 8, no. 2, p. 168, 2020.
[4] G. Yuan, T. Li, and W. Hu, "A conjugate gradient algorithm for large-scale nonlinear equations and image restoration problems," *Applied Numerical Mathematics*, vol. 147, pp. 129–141, 2020.
[5] G. Yuan, J. Lu, and Z. Wang, "The PRP conjugate gradient algorithm with a modified WWP line search and its application in the image restoration problems," *Applied Numerical Mathematics*, vol. 152, pp. 19–11, 2020.
[6] G. Yuan, X. Wang, and Z. Sheng, "The projection technique for two open problems of unconstrained optimization problems," *Journal of Optimization Theory and Applications*, vol. 186, no. 2, pp. 590–619, 2020.
[7] G. Yuan, X. Wang, and Z. Sheng, "Family weak conjugate gradient algorithms and their convergence analysis for nonconvex functions," *Numerical Algorithms*, vol. 84, no. 3, pp. 935–956, 2020.
[8] G. Yuan, Z. Wei, and Y. Yang, "The global convergence of the Polak-Ribière-Polyak conjugate gradient algorithm under inexact line search for nonconvex functions," *Journal of Computational and Applied Mathematics*, vol. 362, pp. 262–275, 2019.
[9] G. Yuan, Z. Wang, and P. Li, "A modified Broyden family algorithm with global convergence under a weak Wolfe-Powell line search for unconstrained nonconvex problems," *Calcolo*, vol. 57, pp. 1–21, 2020.
[10] W. C. Davidon, "Variable metric method for minimization," *SIAM Journal on Optimization*, vol. 1, no. 1, pp. 1–17, 1991.
[11] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms," *IMA Journal of Applied Mathematics*, vol. 6, no. 3, pp. 222–231, 1970.
[12] R. Fletcher, "A new approach to variable metric algorithms," *The Computer Journal*, vol. 13, no. 3, pp. 317–322, 1970.
[13] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Mathematics of Computation*, vol. 24, no. 109, p. 23, 1970.
[14] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Mathematics of Computation*, vol. 24, no. 111, p. 647, 1970.
[15] J. E. Dennis and R. B. Schnabel, "Numerical methods for unconstrained optimization and nonlinear equations," 1983.
[16] Y. Dai, "Convergence properties of the BFGS algorithm," *SIAM Journal on Optimization*, vol. 13, pp. 693–701, 2006.
[17] R. Fletcher, "Practical methods of optimization," 2013.
[18] J. Han and G. Liu, "Global convergence analysis of a new nonmonotone BFGS algorithm on convex objective functions," *Computational Optimization and Applications*, vol. 7, no. 3, pp. 277–289, 1997.
[19] M. J. D. Powell, "Some global convergence properties of a variable metric algorithm for minimization without exact line searches," 1976.
[20] C. G. Broyden, J. E. Dennis Jr., and J. J. Moré, "On the local and superlinear convergence of quasi-Newton methods," *IMA Journal of Applied Mathematics*, vol. 12, no. 3, pp. 223–245, 1973.
[21] R. H. Byrd, J. Nocedal, and Y.-X. Yuan, "Global convergence of a cass of quasi-Newton methods on convex problems," *SIAM Journal on Numerical Analysis*, vol. 24, no. 5, pp. 1171–1190, 1987.
[22] L. C. W. Dixon, "Variable metric algorithms: necessary and sufficient conditions for identical behavior of nonquadratic functions," *Journal of Optimization Theory and Applications*, vol. 10, no. 1, pp. 34–40, 1972.
[23] J. E. Dennis and J. J. Moré, "Quasi-Newton methods, motivation and theory," *SIAM Review*, vol. 19, pp. 46–89, 1977.
[24] A. Griewank, "The "global" convergence of Broyden-like methods with suitable line search," *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, vol. 28, no. 1, pp. 75–92, 1986.
[25] A. Griewank and P. L. Toint, "Local convergence analysis for partitioned quasi-Newton updates," *Numerische Mathematik*, vol. 39, no. 3, pp. 429–448, 1982.
[26] P. L. Toint, "Global convergence of the partitioned BFGS algorithm for convex partially separable optimization," *Mathematical Programming*, vol. 36, no. 3, pp. 290–306, 1986.
[27] W. F. Mascarenhas, "The BFGS method with exact line searches fails for non-convex objective functions," *Mathematical Programming*, vol. 99, no. 1, pp. 49–61, 2004.
[28] D.-H. Li and M. Fukushima, "A modified BFGS method and its global convergence in nonconvex minimization," *Journal of Computational and Applied Mathematics*, vol. 129, no. 1-2, pp. 15–35, 2001.
[29] D.-H. Li and M. Fukushima, "On the global convergence of the BFGS method for nonconvex unconstrained optimization problems," *SIAM Journal on Optimization*, vol. 11, no. 4, pp. 1054–1064, 2001.
[30] M. J. D. Powell, "A new algorithm for unconstrained optimization," *Nonlinear Programming*, vol. 23, pp. 31–65, 1970.
[31] Z. Wei, L. Qi, and X. Chen, "An SQP-type method and its application in stochastic programs," *Journal of Optimization Theory and Applications*, vol. 116, no. 1, pp. 205–228, 2003.
[32] Z. Wei, G. Yu, G. Yuan, and Z. Lian, "The superlinear convergence of a modified BFGS-type method for unconstrained optimization," *Computational Optimization and Applications*, vol. 29, no. 3, pp. 315–332, 2004.
[33] G. Yuan and Z. Wei, "Convergence analysis of a modified BFGS method on convex minimizations," *Computational Optimization and Applications*, vol. 47, no. 2, pp. 237–255, 2010.
[34] G. Yuan, Z. Wei, and X. Lu, "Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search," *Applied Mathematical Modelling*, vol. 47, pp. 811–825, 2017.
[35] G. Yuan, Z. Sheng, B. Wang, W. Hu, and C. Li, "The global convergence of a modified BFGS method for nonconvex

functions," *Journal of Computational and Applied Mathematics*, vol. 327, pp. 274–294, 2018.

[36] R. H. Byrd and J. Nocedal, "A tool for the analysis of quasi-Newton methods with application to unconstrained minimization," *SIAM Journal on Numerical Analysis*, vol. 26, no. 3, pp. 727–739, 1989.

[37] I. Bongartz, A. R. Conn, N. Gould, and P. L. Toint, "Cute," *ACM Transactions on Mathematical Software*, vol. 21, no. 1, pp. 123–160, 1995.

[38] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "Testing unconstrained optimization software," *ACM Transactions on Mathematical Software*, vol. 7, no. 1, pp. 17–41, 1981.

[39] Y. Yuan and W. Sun, "Theory and methods of optimization," 1999.

[40] A. Ouyang, L.-B. Liu, Z. Sheng, and F. Wu, "A class of parameter estimation methods for nonlinear Muskingum model using hybrid invasive weed optimization algorithm," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–15, 2015.

[41] A. Ouyang, Z. Tang, K. Li, A. Sallam, and E. Sha, "Estimating parameters of Muskingum model using an adaptive hybrid PSO algorithm," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 28, pp. 1–29, 2014.

[42] Z. W. Geem, "Parameter estimation for the nonlinear Muskingum model using the BFGS technique," *Journal of Irrigation and Drainage Engineering*, vol. 132, no. 5, pp. 474–478, 2006.