

A GLOBALLY CONVERGENT BFGS METHOD FOR NONLINEAR MONOTONE EQUATIONS WITHOUT ANY MERIT FUNCTIONS

WEI-JUN ZHOU AND DONG-HUI LI

ABSTRACT. Since 1965, there has been significant progress in the theoretical study on quasi-Newton methods for solving nonlinear equations, especially in the local convergence analysis. However, the study on global convergence of quasi-Newton methods is relatively fewer, especially for the BFGS method. To ensure global convergence, some merit function such as the squared norm merit function is typically used. In this paper, we propose an algorithm for solving nonlinear monotone equations, which combines the BFGS method and the hyperplane projection method. We also prove that the proposed BFGS method converges globally if the equation is monotone and Lipschitz continuous without differentiability requirement on the equation, which makes it possible to solve some nonsmooth equations. An attractive property of the proposed method is that its global convergence is independent of any merit function. We also report some numerical results to show efficiency of the proposed method.

1. INTRODUCTION

In this paper, we consider the problem of finding a solution of the nonlinear equation

$$(1.1) \quad F(x) = 0,$$

where $F : R^n \rightarrow R^n$ is continuous and monotone. By monotone, we mean

$$\langle F(x) - F(y), x - y \rangle \geq 0, \quad \forall x, y \in R^n.$$

Nonlinear monotone equations have many practical uses such as ballistic trajectory computation and vibration systems [20, 25], the first-order necessary condition of the unconstrained convex optimization problem and the subproblems in the generalized proximal algorithms with Bregman distances [13]. Some monotone variational inequality problems can also be converted into nonlinear monotone equations by means of fixed point maps or normal maps if the underlying function satisfies some coercive conditions [27].

Among numerous algorithms for solving (1.1), quasi-Newton methods are regarded as one of the most efficient classes of methods. Since the first quasi-Newton method for solving nonlinear equations was proposed by Broyden [1], there has been

Received by the editor March 14, 2006 and, in revised form, March 30, 2007.

2000 *Mathematics Subject Classification.* Primary 90C53.

Key words and phrases. BFGS method, monotone equation, hyperplane projection method, global convergence.

This work was supported by the NSF (10771057 and 10701018) of China.

©2008 American Mathematical Society
Reverts to public domain 28 years from publication

significant progress in the theoretical study on quasi-Newton methods, especially in local convergence analysis [2, 6, 7]. Moreover, much effort has been made to establish global convergence of quasi-Newton methods for unconstrained optimization problems, for example [2, 3, 4, 5, 7, 8, 9, 11, 14, 18, 20, 19, 21, 22, 24, 26]. However, the study of globally convergent quasi-Newton methods for solving nonlinear equations is relatively fewer. The major difficulty is the lack of practical line search strategy. To the authors' knowledge, the earliest global convergence result is due to Griewank [10], where a derivative-free line search is proposed. But Li and Fukushima [15] constructed an example to show that the line search in [10] has a certain difficulty in some special cases. By using a nonmonotone line search process, Li and Fukushima [15, 16] proposed a Broyden's method for solving nonlinear equations and a Gauss-Newton-based BFGS method for solving symmetric nonlinear equations and prove that these methods converge globally. Quite recently, Gu et al. [12] introduced a norm descent line search technique and proposed a norm descent Gauss-Newton-based BFGS method for solving symmetric equations with global convergence.

However, to ensure global convergence of the above mentioned quasi-Newton methods, some merit function such as squared norm merit function is often used. Generally, the quasi-Newton direction is not a descent direction for the merit function, which makes it difficult to globalize the method. In this paper, based on the hyperplane projection method [23], we propose a BFGS method for solving nonlinear monotone equations and prove its global convergence property without use of merit functions. The differentiability of the equation is also not assumed. Compared with the Gauss-Newton-based BFGS method in [16], our method is more natural and simpler, which makes our method more implementable in practice.

The paper is organized as follows. We present a BFGS method for solving monotone nonlinear equations in the next section. In Section 3, we establish the global convergence of the proposed method. In Section 4, we report some numerical results to show the efficiency of the proposed method.

2. ALGORITHM

In this section, we describe the proposed method in detail. First we recall the hyperplane projection method [23] for solving nonlinear monotone equations (1.1). By the monotonicity of F , we have

$$\langle F(z_k), \bar{x} - z_k \rangle \leq 0$$

for all \bar{x} such that $F(\bar{x}) = 0$. Suppose that we have obtained a direction d_k . By performing some kind of line search procedure along the direction d_k , a point $z_k = x_k + \alpha_k d_k$ can be computed such that

$$\langle F(z_k), x_k - z_k \rangle > 0.$$

Thus the hyperplane

$$\mathcal{H}_k = \{x \in R^n \mid \langle F(z_k), x - z_k \rangle = 0\}$$

strictly separates the current iterate x_k from zeros of the equation (1.1). Once the separating hyperplane is obtained, the next iterate x_{k+1} is computed by projecting x_k onto the hyperplane.

Now we state our algorithm as follows.

Algorithm 1 (BFGS method)

Step 0. Given initial point $x_0 \in R^n$ and constants $\beta \in (0, 1)$, $\sigma \in (0, 1)$, $h > 0$, $r \geq 0$. Choose $B_0 = I$ (the identity matrix). Let $k := 0$.

Step 1. Compute d_k by

$$(2.1) \quad B_k d_k = -F(x_k).$$

Stop if $d_k = 0$.

Step 2. Determine steplength $\alpha_k = \beta^{m_k}$ such that m_k is the smallest nonnegative integer m satisfying

$$(2.2) \quad -\langle F(x_k + \beta^m d_k), d_k \rangle \geq \sigma \beta^m \|F(x_k + \beta^m d_k)\| \|d_k\|^2.$$

Let $z_k = x_k + \alpha_k d_k$.

Stop if $\|F(z_k)\| = 0$.

Step 3. Compute

$$(2.3) \quad x_{k+1} = x_k - \frac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|^2} F(z_k).$$

Step 4. Compute B_{k+1} by the following BFGS update process:

$$(2.4) \quad B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k},$$

where

$$(2.5) \quad s_k = z_k - x_k, \quad y_k = F(z_k) - F(x_k) + h \|F(x_k)\|^r s_k.$$

Set $k := k + 1$. Go to Step 1.

Remarks. (i) In the update formula (2.4), we used the modified update rule proposed by Li and Fukushima [18] for nonconvex unconstrained optimization problems, which can also be regarded as a regularized BFGS method in the sense that B_k is positive definite and symmetric for all k .

(ii) If we suppose that F is Lipschitz continuous, i.e., there exists a constant $L > 0$ such that

$$(2.6) \quad \|F(x) - F(y)\| \leq L \|x - y\|, \quad \forall x, y \in R^n,$$

then it follows from the monotonicity and Lipschitz continuous property of F that

$$(2.7) \quad h \|F(x_k)\|^r s_k^T s_k \leq y_k^T s_k \leq (L + h \|F(x_k)\|^r) s_k^T s_k,$$

where T stands for transpose.

(iii) The update formula (2.4) is very different from the Gauss-Newton-based BFGS method proposed by Li and Fukushima [16]. Our method is more natural and simpler.

(iv) The line search (2.2) is a little different from that of [23]. It is not difficult to see from (i) that Algorithm 1 is well defined.

3. CONVERGENCE PROPERTY

In this section, we prove global convergence of Algorithm 1. To this end, we introduce some useful lemmas. The following lemma comes from [3].

Lemma 3.1. *Let B_k be updated by the BFGS formula (2.4). Suppose B_0 is symmetric and positive definite. If there are positive constants $m \leq M$ such that for all $k \geq 0$, y_k and s_k satisfy*

$$(3.1) \quad \frac{y_k^T s_k}{\|s_k\|^2} \geq m, \quad \frac{\|y_k\|^2}{y_k^T s_k} \leq M,$$

then for any $\kappa \in (0, 1)$, there exist positive constants $\beta_1, \beta_2, \beta_3$ and β_4 such that inequalities

$$(3.2) \quad \beta_1 \|s_k\| \leq \|B_k s_k\| \leq \beta_2 \|s_k\|, \quad \beta_3 \|s_k\|^2 \leq s_k^T B_k s_k \leq \beta_4 \|s_k\|^2$$

hold for at least $\lceil \kappa k \rceil$ many $j \leq k$.

We define index sets K_k for each k , and K by

$$(3.3) \quad K_t = \{k \leq t \mid (3.2) \text{ hold}\}, \quad K = \bigcup_{t=0}^{\infty} K_t.$$

Since $s_k = \alpha_k d_k$, it is clear that inequalities (3.2) hold true if s_k is replaced by d_k . Moreover, it follows from (2.1) that

$$(3.4) \quad \beta_1 \|d_k\| \leq \|F(x_k)\| \leq \beta_2 \|d_k\|, \quad \forall k \in K.$$

The following lemma comes from [23].

Lemma 3.2. *Let F be monotone and $x, y \in R^n$ satisfy $\langle F(y), x - y \rangle > 0$. Let*

$$x^+ = x - \frac{\langle F(y), x - y \rangle}{\|F(y)\|^2} F(y).$$

Then for any $\bar{x} \in R^n$ such that $F(\bar{x}) = 0$, it holds that

$$\|x^+ - \bar{x}\|^2 \leq \|x - \bar{x}\|^2 - \|x^+ - x\|^2.$$

Now we establish a convergence theorem for Algorithm 1.

Theorem 3.3. *Let $\{x_k\}$ be generated by Algorithm 1. Suppose that F is monotone and Lipschitz continuous and that the solution set of (1.1) is not empty. Then we have for any \bar{x} satisfying $F(\bar{x}) = 0$,*

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \|x_{k+1} - x_k\|^2.$$

In particular, $\{x_k\}$ is bounded. Furthermore, it holds that either $\{x_k\}$ is finite and the last iterate is a solution, or the sequence is infinite and $\lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0$.

Moreover, $\{x_k\}$ converges to some solution of (1.1).

Proof. We first note that if the algorithm terminates at some iteration k , then $d_k = 0$ or $\|F(z_k)\| = 0$. By the positive definiteness of B_k , we have $F(x_k) = 0$ if $d_k = 0$. This means that x_k or z_k is a solution of (1.1).

Suppose that $d_k \neq 0$ and $\|F(z_k)\| \neq 0$ for all k . Then an infinite sequence $\{x_k\}$ is generated. It follows from (2.2) that

$$(3.5) \quad \langle F(z_k), x_k - z_k \rangle = -\alpha_k \langle F(z_k), d_k \rangle \geq \sigma \|F(z_k)\| \alpha_k^2 \|d_k\|^2 > 0.$$

Let \bar{x} be an arbitrary solution of (1.1). By (2.3), (3.5) and Lemma 3.2, we obtain

$$(3.6) \quad \|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \|x_{k+1} - x_k\|^2.$$

In particular, the sequence $\{\|x_k - \bar{x}\|\}$ is decreasing and hence convergent. Also, the sequence $\{x_k\}$ is bounded, and

$$(3.7) \quad \lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0.$$

We obtain from (2.3) and (3.5) that

$$\|x_{k+1} - x_k\| = \frac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|} \geq \sigma \alpha_k^2 \|d_k\|^2.$$

From the last inequality and (3.7), we get

$$(3.8) \quad \lim_{k \rightarrow \infty} \alpha_k \|d_k\| = 0, \quad \lim_{k \in K, k \rightarrow \infty} \alpha_k \|d_k\| = 0.$$

Now we consider the following two possible cases:

- (i) $\liminf_{k \rightarrow \infty} \|F(x_k)\| = 0.$
- (ii) $\liminf_{k \rightarrow \infty} \|F(x_k)\| = \varepsilon > 0.$

If (i) holds, then by the continuity of F and the boundedness of $\{x_k\}$, it is clear that the sequence $\{x_k\}$ has some accumulation point \hat{x} such that $F(\hat{x}) = 0$. We also have from (3.6) that the sequence $\{\|x_k - \hat{x}\|\}$ converges. Therefore, $\{x_k\}$ converges to \hat{x} .

If (ii) holds, then from the boundedness of $\{x_k\}$ and the continuity of F , there exists a positive constant C such that

$$(3.9) \quad \varepsilon \leq \|F(x_k)\| \leq C, \quad \forall k.$$

From (2.7) and (3.9), we have

$$\frac{y_k^T s_k}{\|s_k\|^2} \geq h\varepsilon^r, \quad \frac{\|y_k\|^2}{y_k^T s_k} \leq \frac{(L + hC^r)^2}{h\varepsilon^r}.$$

It then follows from Lemma 3.1 that inequalities (3.2), (3.4) and (3.9) hold for all $k \in K$. So we get from (3.4) and (3.8) that

$$\lim_{k \in K, k \rightarrow \infty} \alpha_k = 0.$$

By the line search rule, we have for all $k \in K$ sufficiently large, $m_k - 1$ will not satisfy (2.2). This means

$$(3.10) \quad -\langle F(x_k + \beta^{m_k-1} d_k), d_k \rangle < \sigma \beta^{m_k-1} \|F(x_k + \beta^{m_k-1} d_k)\| \|d_k\|^2, \quad \forall k \in K.$$

Since $\{x_k\}_K$ is bounded, it follows from (3.4) and (3.9) that $\{d_k\}_K$ is bounded too. We can choose a subsequences of $\{x_k\}_K$ and $\{d_k\}_K$ converging to \hat{x} and \hat{d} , respectively. Taking the limit in (3.10) for the subsequence, we obtain

$$-\langle F(\hat{x}), \hat{d} \rangle \leq 0.$$

However, it is not difficult to deduce from (2.1) and (3.2) (by further taking subsequence if necessary) that

$$-\langle F(\hat{x}), \hat{d} \rangle > 0.$$

This yields a contradiction. Consequently, $\liminf_{k \rightarrow \infty} \|F(x_k)\| > 0$ is not possible. The proof is then complete. \square

4. NUMERICAL RESULTS

In this section, we report some numerical results with the proposed method. We test the performance of Algorithm 1 on the following three problems with various sizes and different initial points.

Problem 1. The elements of function F are given by

$$F_i(x) = x_i - \sin(x_i), \quad i = 1, 2, \dots, n.$$

Problem 2. The function F is given by

$$F(x) = Ax + g(x)$$

where $g(x) = (e^{x_1} - 1, e^{x_2} - 1, \dots, e^{x_n} - 1)^T$ and

$$A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

Problem 3. An application to the two-point boundary value problem:

$$(4.1) \quad u''(t) = f(t, u(t), u'(t)), \quad 0 \leq t \leq 1, u(0) = \alpha, u(1) = \beta.$$

The problem (4.1) is often involved in ballistic trajectory computation and vibration systems et al. [20, 25]. In order to solve (4.1) approximately, we consider its discrete version. Suppose that

$$t_j = jh, h = \frac{1}{n+1}, j = 0, \dots, n+1$$

is a uniform division of $[0,1]$. We approximate $u''(t_j)$ and $u'(t_j)$ by

$$(4.2) \quad u''(t_j) \approx \frac{1}{h^2}[u(t_{j+1}) - 2u(t_j) + u(t_{j-1})],$$

$$(4.3) \quad u'(t_j) \approx \frac{1}{2h}[u(t_{j+1}) - u(t_{j-1})], \quad j = 1, \dots, n.$$

If we denote $x_i = u(t_i), i = 1, 2, \dots, n$, then x_i satisfy the following equations:

$$x_{i+1} - 2x_i + x_{i-1} = h^2 f(t_i, x_i, \frac{1}{2h}(x_{i+1} - x_{i-1})), \quad x_0 = \alpha, x_{n+1} = \beta, i = 1, 2, \dots, n.$$

We consider the following function f (E1.1-1 (d) in [20]),

$$(4.4) \quad f = \frac{1}{2}u^3(t) + 3u'(t) - \frac{3}{2-t} + \frac{1}{2}.$$

Let $\alpha = 0$ and $\beta = 1$ in (4.1). It is easy to see that

$$u(t) = \frac{t}{2-t}$$

is the analytic solution of the problem.

The discrete form of the problem can be written as

$$F(x) = Bx + h(x) = 0,$$

where $x = (x_1, \dots, x_n)^T = (u_1, \dots, u_n)^T$ and

$$\begin{aligned}
 h(x) &= (h_1(x), \dots, h_n(x))^T, \\
 h_i(x) &= h^2 \left(\frac{1}{2} x_i^3 - \frac{3}{2 - ih} + \frac{1}{2} \right), \quad i = 1, \dots, n - 1, \\
 h_n(x) &= h^2 \left(\frac{1}{2} x_n^3 - \frac{3}{2 - nh} + \frac{1}{2} \right) - 1, \\
 B &= \begin{pmatrix} 2 & -1 + \frac{3}{2}h & & & \\ -1 - \frac{3}{2}h & 2 & -1 + \frac{3}{2}h & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 + \frac{3}{2}h \\ & & & -1 - \frac{3}{2}h & 2 \end{pmatrix}.
 \end{aligned}$$

We will test the problem with different divisions, that is, $h = \frac{1}{10}, \frac{1}{20}, \frac{1}{30}$.

We note that Problems 1 and 2 are symmetric while Problem 3 is nonsymmetric.

We compare the performance of Algorithm 1 (BFGS) with the Gauss-Newton-based BFGS method (GNBFGS) in [16] and the Inexact Newton Method (INM) in [23] on Problems 1–3 with different initial points. The algorithms were coded in MATLAB and run on a personal computer with a 3.0GHZ CPU processor.

The results are listed in Tables 1–3 where $x_1 = (0.1, \dots, 0.1)^T$, $x_2 = (1, \dots, 1)^T$, $x_3 = (1, 1/2, \dots, 1/n)^T$, $x_4 = (-10, \dots, -10)^T$, $x_5 = (-0.1, \dots, -0.1)^T$, $x_6 = (-1, \dots, -1)^T$. The parameters are specified as follows. For Algorithm 1, we set $\beta = 0.6$, $\sigma = 10^{-5}$, $r = 0$, $h = 10^{-4}$. For GNBFGS method in [16], we set $r = 0.1$, $\rho = 0.9$, $\sigma_1 = \sigma_2 = 10^{-5}$, $\lambda_{-1} = 0.01$, $\epsilon_k = 1/k^2$ and $B_0 = I$. For INM method in [23], we set $\mu_k = \|F(x_k)\|$, $\rho_k = 0$, $\beta = 0.6$, $\lambda = 0.01$. We stop the iteration if $\|F(x_k)\| \leq 10^{-5}$ or the number of iterations exceeds 10^4 .

The meanings of the columns in Tables 1–3 are stated as follows:

- “n”: the dimension of the problem;
- “ip”: the initial point;
- “iter”: the total number of iterations;
- “time”: the CPU time in seconds;
- “average”: the average number of iterations or the average CPU time.

The numerical results indicates that Algorithm 1 and the INM method terminated successfully for all initial points while the GNBFGS method failed to solve the nonsymmetric Problem 3 with initial point x_6 and dimension $n = 29$. We also see from the tables that the GNBFGS method seems more sensitive to the initial points. In addition, in most cases the INM method performed best. It is interesting to note that Algorithm 1 performed much better than the GNBFGS method did on Problem 3. This might show an advantage of Algorithm 1 on nonsymmetric problems, compared to the GNBFGS method.

5. FINAL REMARK

We have proposed a projection BFGS method for solving nonlinear monotone equations. We have established a global convergence theorem for the method. Numerical results show that the proposed method performed better than the “GNBFGS” method did. It is interesting to note that the symmetry of the Jacobian of the equation is not assumed in the global convergence theorem.

TABLE 1. Test results for BFGS, GNBFGS and INM methods on Problem 1.

		BFGS		GNBFGS		INM	
ip	n	iter	time	iter	time	iter	time
x_1	10	6	0.03	7	0.01	5	0.02
x_2	10	14	0	18	0.01	14	0
x_3	10	18	0.01	382	0.14	11	0.01
x_4	10	24	0.02	18	0	46	0.01
x_5	10	6	0	7	0	5	0
x_6	10	14	0	18	0.01	14	0
average	10	13.6667	0.01	75	0.028333	15.8333	0.0066667
x_1	100	8	0.05	9	0.03	7	0.02
x_2	100	16	0.07	20	0.09	22	0.02
x_3	100	23	0.09	592	1.873	11	0.01
x_4	100	27	0.13	20	0.07	116	0.091
x_5	100	8	0.03	9	0.03	7	0
x_6	100	16	0.06	20	0.08	22	0.01
average	100	16.3333	0.071667	111.6667	0.36217	30.8333	0.025167
x_1	200	8	0.161	10	0.18	7	0.02
x_2	200	16	0.3	21	0.391	26	0.05
x_3	200	23	0.421	1112	22.282	11	0.02
x_4	200	28	0.52	20	0.401	158	0.3
x_5	200	8	0.151	10	0.19	7	0.02
x_6	200	16	0.3	21	0.391	26	0.05
average	200	16.5	0.30883	199	3.9725	39.1667	0.076667

TABLE 2. Test results for BFGS, GNBFGS and INM methods on Problem 2

		BFGS		GNBFGS		INM	
ip	n	iter	time	iter	time	iter	time
x_1	50	50	0.08	36	0.06	29	0.02
x_2	50	95	0.11	50	0.071	36	0.02
x_3	50	57	0.08	44	0.06	24	0.02
x_4	50	110	0.15	145	0.19	104	0.06
x_5	50	49	0.071	36	0.05	27	0.02
x_6	50	58	0.08	48	0.07	34	0.02
average	50	69.8333	0.095167	59.8333	0.0835	42.3333	0.026667
x_1	100	50	0.21	35	0.141	28	0.03
x_2	100	97	0.38	58	0.261	37	0.04
x_3	100	57	0.21	44	0.18	25	0.03
x_4	100	111	0.441	278	1.242	130	0.12
x_5	100	57	0.25	35	0.15	27	0.031
x_6	100	65	0.28	45	0.19	36	0.04
average	100	72.8333	0.29517	82.5	0.36067	47.1667	0.0485
x_1	200	51	0.981	36	0.711	28	0.08
x_2	200	148	2.844	221	4.427	42	0.12
x_3	200	56	1.102	44	0.861	25	0.09
x_4	200	262	5.017	478	9.964	171	0.481
x_5	200	50	0.911	36	0.691	27	0.081
x_6	200	64	1.201	45	0.872	41	0.11
average	200	105.1667	2.0093	143.3333	2.921	55.6667	0.16033

TABLE 3. Test results for BFGS, GNBFGS and INM methods on Problem 3

		BFGS		GNBFGS		INM	
ip	n	iter	time	iter	time	iter	time
x_1	9	93	0.04	19	0.011	73	0.02
x_2	9	100	0.04	229	0.1	77	0.01
x_3	9	64	0.02	24	0.01	143	0.04
x_4	9	122	0.06	142	0.07	134	0.03
x_5	9	93	0.04	52	0.01	98	0.02
x_6	9	103	0.04	652	0.231	112	0.02
average	9	95.8333	0.04	186.3333	0.072	106.1667	0.023333
x_1	19	228	0.1	44	0.03	108	0.03
x_2	19	270	0.11	207	0.111	134	0.03
x_3	19	197	0.09	68	0.04	79	0.01
x_4	19	319	0.13	2867	1.392	189	0.04
x_5	19	238	0.12	51	0.02	109	0.031
x_6	19	253	0.11	8266	4.326	131	0.03
average	19	250.8333	0.11	1917.1667	0.9865	125	0.0285
x_1	29	424	0.261	74	0.05	37	0.02
x_2	29	448	0.24	453	0.32	44	0.02
x_3	29	341	0.181	242	0.17	43	0.01
x_4	29	529	0.28	7168	4.627	117	0.04
x_5	29	384	0.22	90	0.071	43	0.01
x_6	29	491	0.27	10000(Fail)	6.599	56	0.01
average	29	436.1667	0.242	3004.5	1.9728	56.6667	0.018333

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for their valuable suggestions and comments, which improved this paper greatly.

REFERENCES

- [1] C.G. Broyden, A class of methods for solving nonlinear simultaneous equations, *Math. Comput.*, 19 (1965), 577-593. MR0198670 (33:6825)
- [2] C.G. Broyden, J.E. Dennis and Jr. J.J. Moré, On the local and superlinear convergence of quasi-Newton methods, *J. Inst. Math. Appl.*, 12 (1973), 223-246. MR0341853 (49:6599)
- [3] R. Byrd and J. Nocedal, A tool for analysis of quasi-Newton methods with application to unconstrained minimization, *SIAM J. Numer. Anal.*, 26 (1989), 727-739. MR997665 (90f:90116)
- [4] R. Byrd, J. Nocedal and Y.X. Yuan, Global convergence of a class of quasi-Newton methods on convex problems, *SIAM J. Numer. Anal.*, 24 (1987), 1171-1190. MR909072 (88m:65100)
- [5] Y. Dai, Convergence properties of the BFGS algorithm, *SIAM J. Optim.*, 13 (2002), 693-701. MR1972211 (2004b:90164)
- [6] J.E. Dennis and Jr. J.J. Moré, A characterization of superlinear convergence and its application to quasi-Newton methods, *Math. Comput.*, 28 (1974), 549-560. MR0343581 (49:8322)
- [7] J.E. Dennis and Jr. J.J. Moré, Quasi-Newton method, motivation and theory, *SIAM Rev.*, 19 (1977), 46-89. MR0445812 (56:4146)
- [8] J.E. Dennis and R.B. Schnabel, Numerical methods for unconstrained optimization and nonlinear equations, Prentice-Hall, Englewood Cliffs, N.F., 1983. MR702023 (85j:65001)
- [9] L.C.W. Dixon, Variable metric algorithms: necessary and sufficient conditions for identical behavior on nonquadratic functions, *J. Optim. Theory Appl.*, 10 (1972) 34-40. MR0309305 (46:8415)
- [10] A. Griewank, The "global" convergence of Broyden-like methods with a suitable line search, *J. Austral. Math. Soc., Ser. B*, 28 (1986), 75-92. MR846784 (87k:65064)

- [11] A. Griewank, The global convergence of partitioned BFGS on problems with convex decompositions and Lipschitzian gradients, *Math. Program.*, 50 (1991), 141-175. MR1103931 (92d:90079)
- [12] G.Z. Gu, D.H. Li, L. Qi and S.Z. Zhou, Descent directions of quasi-Newton method for symmetric nonlinear equations, *SIAM J. Numer. Anal.*, 40 (2002), 1763-1774. MR1950621 (2003m:65082)
- [13] A.N. Iusem and M.V. Solodov, Newton-type methods with generalized distances for constrained optimization, *Optimization*, 41 (1997), 257-278. MR1473401 (98h:90078)
- [14] T.G. Kolda, D.P. O'Leary and L. Nazareth, BFGS with update skipping and varying memory, *SIAM J. Optim.*, 8 (1998), 1060-1083. MR1646119 (99e:90092)
- [15] D.H. Li and M. Fukushima, A derivative-free line search and global convergence of Broyden-like method for nonlinear equations, *Optim. Methods and Softw.*, 13 (2000), 181-201. MR1785195 (2001e:90146)
- [16] D.H. Li and M. Fukushima, A globally and superlinearly convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations, *SIAM J. Numer. Anal.*, 37 (1999), 152-172. MR1742754 (2000k:65100)
- [17] D.H. Li and M. Fukushima, On the global convergence of the BFGS method for nonconvex unconstrained optimization problems, *SIAM J. Optim.*, 11 (2001), 1054-1064. MR1855221 (2002f:90168)
- [18] D.H. Li and M. Fukushima, A modified BFGS method and its global convergence in nonconvex minimization, *J. Comput. Appl. Math.*, 129 (2001), 15-35. MR1823208 (2002b:90130)
- [19] J. Nocedal, Updating quasi-Newton matrices with limited storage, *Math. Comput.*, 35 (1980), 773-782. MR572855 (81g:65077)
- [20] J.M. Ortega and W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, 1970. MR0273810 (42:8686)
- [21] M.J.D. Powell, On the convergence of the variable metric algorithm, *J. Inst. Math. Appl.*, 7 (1971), 21-36. MR0279977 (43:5698)
- [22] M.J.D. Powell, Some global convergence properties of a variable metric algorithm for minimization without exact line searches, in: R.W. Cottle, C.E. Lemke (Eds.), *Nonlinear Programming, SIAM-AMS Proceedings*, Vol. IX, SIAM, Philadelphia, PA, (1976) 53-72. MR0426428 (54:14371)
- [23] M.V. Solodov and B.F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, in: M. Fukushima, L. Qi (Eds.), *Reformulation: Nonsmooth, Piecewise smooth, Semismooth and Smoothing Methods*, Kluwer Academic Publishers, (1998) 355-369. MR1682755 (2000d:65089)
- [24] Ph.L. Toint, Global convergence of the partitioned BFGS algorithm for convex partially separable optimization, *Math. Program.*, 36 (1986), 290-306. MR866412 (88a:90160)
- [25] E. Zeidler, Nonlinear functional analysis and its applications, II/B: Nonlinear monotone operators, Springer-Verlag, 1990. MR1033498 (91b:47002)
- [26] Y. Zhang and R.P. Tewarson, Quasi-Newton algorithm with updates from the preconvex part of Broyden's family, *IMA J. Numer. Anal.*, 8 (1988), 487-509. MR975609 (90f:65098)
- [27] Y.B. Zhao and D. Li, Monotonicity of fixed point and normal mapping associated with variational inequality and its application, *SIAM J. Optim.*, 4 (2001), 962-973. MR1855216 (2003a:90070)

COLLEGE OF MATHEMATICS AND COMPUTATIONAL SCIENCE, CHANGSHA UNIVERSITY OF SCIENCE AND TECHNOLOGY, CHANGSHA 410076, CHINA

E-mail address: weijunzhou@126.com

COLLEGE OF MATHEMATICS AND ECONOMETRICS, HUNAN UNIVERSITY, CHANGSHA 410082, CHINA

E-mail address: dhli@hnu.cn