The global and superlinear convergence of a new nonmonotone MBFGS algorithm on convex objective functions

Liying Liu\textsuperscript{a,*}, Shengwei Yao\textsuperscript{b}, Zengxin Wei\textsuperscript{b}

\textsuperscript{a}College of Mathematics Science, Liaocheng University, 252059, PR China
\textsuperscript{b}College of Mathematics and Information Science, Guangxi University, 530004, PR China

Received 19 November 2006; received in revised form 27 August 2007

Abstract

In this paper, a new nonmonotone MBFGS algorithm for unconstrained optimization will be proposed. Under some suitable assumptions, the global and superlinear convergence of the new nonmonotone MBFGS algorithm on convex objective functions will be established. Some numerical experiments show that this new nonmonotone MBFGS algorithm is competitive to the MBFGS algorithm and the nonmonotone BFGS algorithm.

© 2007 Elsevier B.V. All rights reserved.

MSC: 90C30; 49M37; 65K05

Keywords: Nonmonotone linesearch; MBFGS algorithm; Global convergence; Superlinear convergence; Unconstrained optimization

1. Introduction

In this paper, we consider the following unconstrained optimization problem:

\[
\min f(x), \quad x \in \mathbb{R}^n, \tag{1.1}
\]

where \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is twice continuously differentiable. Throughout the paper, we assume that \( \| \cdot \| \) denotes the Euclidean norm and we abbreviate \( f(x_k), g(x_k) \), etc., as \( f_k, g_k \), etc., respectively.

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

\[
x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \ldots, \tag{1.2}
\]

where \( x_0 \) is any given starting point, \( \alpha_k \) is a stepsize, \( d_k \) is a search direction. It is well known that the BFGS algorithm is generally considered to be the most effective iterative method. The search direction of the BFGS algorithm is determined as follows:

\[
d_0 = -B_0^{-1}g_0, \quad d_k = -B_k^{-1}g_k, \quad k \geq 1,
\]
where $g_k$ is the gradient of $f$ at the point $x_k$, and

$$y_k = g_{k+1} - g_k, \quad s_k = x_{k+1} - x_k,$$

$B_0$ is any given $n \times n$ 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 $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$. In this paper we use the modified BFGS update formula (see [1])

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

where $y^*_k = g_{k+1} - g_k + A^*_k s_k$, $A^*_k = ((2[f(x_k) - f(x_{k+1})] + (g_{k+1} + g_k)^T s_k)/\|s_k\|^2)$.

It is known that the BFGS algorithm is generally associated with either Wolfe-type linesearch or backtracking-type linesearch, and these two types of linesearches are implemented on convex functions (see [2–4]) and uniformly convex function (see [5–7]), respectively. But for nonmonotone algorithm it is difficult to analyze its global convergence. However, in practice, many numerical experiments show that in some cases, nonmonotone algorithms may be more efficient than monotone one (see [8]). Recently, authors analysed the properties of this algorithm from the views of generalizing linesearch procedures (see [10,11]) and introducing nonmonotone algorithm (see [9,12–14]). So, in this paper, we will introduce a new nonmonotone MBFGS algorithm and analyze its convergence. We will use the following nonmonotone linesearch called GLL linesearch.

GLL linesearch (see [9]): Select steplength $\lambda_k$ satisfying

$$f(x_{k+1}) \leq \max_{0 \leq j \leq M_0} f(x_{k-j}) + \epsilon_1 \lambda_k g_k^T d_k,$$

$$g(x_{k+1})^T d_k \geq \max\{\epsilon_2, 1 - (\lambda_k \|d_k\|)^p\} g_k^T d_k, \quad p \in (-\infty, 1),$$

where $\epsilon_1 \in (0, 1)$, $\epsilon_2 \in (0, \frac{1}{2})$, and $M_0$ is nonnegative integer.

The organization of this paper is as follows. In the next section, the new nonmonotone MBFGS algorithm on convex objective functions will be proposed. In Section 3 and in Section 4, under some suitable conditions, the global and superlinear convergence of the given algorithm will be established, respectively. Furthermore, the numerical results will be given in the last section.

2. New algorithm

Based on (1.4), (1.5) and (1.6), we state the nonmonotone MBFGS algorithm as follows.

**Algorithm 2.1 (The nonmonotone MBFGS algorithm).**

**Step 0:** Choose an initial point $x_0 \in \mathbb{R}^n$ and an initial positive matrix $B_0$. Let $k := 0$.

**Step 1:** If $g_k = 0$, then stop.

**Step 2:** For given $x_k$ and $B_k$, solve $B_k d_k + g_k = 0$ to obtain a search direction $d_k$.

**Step 3:** Find a $\lambda_k$ satisfying (1.5) and (1.6).

**Step 4:** Let $x_{k+1} = x_k + \lambda_k d_k$ and update $B_{k+1}$ by formula (1.4).

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

3. Global convergence

In this section, we study the global convergence behavior of Algorithm 2.1. We firstly make the following assumptions:

Assumption (A). (i) The level set

$$L_0 = \{x | f(x) \leq f(x_0)\}$$

is bounded.
(ii) The function \( f \) in (1.1) is continuously differentiable on \( L_0 \) and there is a constant \( L > 0 \) such that for any \( x, y \in L_0 \),

\[
\| g(x) - g(y) \| \leq L \| x - y \|. \tag{3.1}
\]

(iii) The function \( f \) in (1.1) is uniformly convex, i.e., there exist two positive constants \( h \leq H \) such that

\[
h \| d \|^2 \leq d^T G(x) d \leq H \| d \|^2, \quad d \in \mathbb{R}^n
\]

for all \( x \) in the neighborhood of \( x^* \), where \( G(x) = \nabla^2 f(x), \quad x \in L_0 \).

It is obvious that these assumptions imply that there exists a constant \( M > 0 \), such that

\[
\| G(x) \| \leq M, \quad x \in L_0. \tag{3.3}
\]

**Lemma 3.1.** Suppose Assumption (A) holds. Then there exists a constant \( M_1 > 0 \), such that

\[
\frac{\| y_k^* \|^2}{s_k^T y_k^*} \leq M_1. \tag{3.4}
\]

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

\[
\beta_2 \leq \frac{\| B_j s_j \|}{\| s_j \|} \leq \frac{\beta_3}{\beta_1} \equiv \beta,
\]

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

**Proof.** Following the definition of \( y_k^* \) and the Taylor’s formula, we have

\[
s_k^T y_k^* = s_k^T \left( y_k + \frac{2[f(x_k) - f(x_{k+1})] + (g_{k+1} + g_k)^T s_k}{\| s_k \|^2} \right)
\]

\[
= s_k^T y_k + 2[f(x_k) - f(x_{k+1})] + (g_{k+1} + g_k)^T s_k
\]

\[
= 2[f(x_k) - f(x_{k+1})] + 2g_{k+1}^T s_k
\]

\[
= 2[-g_{k+1}^T s_k + \frac{1}{2} s_k^T G(x_{k+1} + \theta(x_{k+1} - x_k)) s_k] + 2g_{k+1}^T s_k
\]

\[
= s_k^T G(x_{k+1} + \theta(x_{k+1} - x_k)) s_k,
\]

where \( \theta \in (0, 1) \). By using (3.2) we have

\[
h \| s_k \|^2 \leq s_k^T G(x_{k+1} + \theta(x_{k+1} - x_k)) s_k \leq H \| s_k \|^2.
\]

Therefore,

\[
\frac{h}{\| s_k \|^2} \leq s_k^T y_k^* \leq \frac{H}{\| s_k \|^2}.
\]

so we have

\[
\frac{s_k^T y_k^*}{\| s_k \|^2} \geq h, \quad \frac{s_k^T y_k^*}{\| s_k \|^2} \leq H.
\]

**(3.7)**
From the definition of $y_k^*$, we have
\[
y_k^* = \frac{2[f(x_k) - f(x_{k+1})] + (g_{k+1} + g_k)^T s_k}{\|s_k\|^2}s_k
\]
\[
\leq \|y_k\| + \frac{|2[f(x_k) - f(x_{k+1})] + (g_{k+1} + g_k)^T s_k|}{\|s_k\|}
\]
\[
= \|y_k\| + \frac{|-2g_k^T s_k - s_k^T G(x_k + \theta(x_{k+1} - x_k))s_k + (g_{k+1} + g_k)^T s_k|}{\|s_k\|}
\]
\[
\leq 2\|y_k\| + \frac{|s_k^T G(x_k + \theta(x_{k+1} - x_k))s_k|}{\|s_k\|}.
\]
By using (3.1) and the right hand of (3.2), we have
\[
2\|y_k\| + \frac{|s_k^T G(x_k + \theta(x_{k+1} - x_k))s_k|}{\|s_k\|} \leq 2L\|s_k\| + H\|s_k\| = (2L + H)\|s_k\|.
\]
Therefore,
\[
\|y_k^*\| \leq (2L + H)\|s_k\|,
\] (3.8)
so we have
\[
\frac{(2L + H)^2 s_k^Ty_k^*}{\|y_k^*\|^2} \geq \frac{s_k^Ty_k^*}{\|s_k\|^2}.
\] (3.9)
By (3.7) and (3.9) we get
\[
\frac{\|y_k^*\|^2}{s_k^Ty_k^*} \leq M_1,
\]
where $M_1 = (2L + H)^2 / h$. From Theorem 2.1 in [15], we get (3.5).

We defined set $K$
\[
K = \{k | k \text{ satisfies (3.5)}\}.
\] (3.10)
Therefore, from the right hand of (3.5) we have
\[
\|B_k s_k\| \leq \beta\|s_k\|, \quad k \in K.
\] (3.11)
Moreover, for all $k \in K$
\[
\|B_k s_k\| = |z_k|\|B_k d_k\| = |z_k|\|g_k\| \leq \beta\|s_k\| = \beta|z_k|\|d_k\|.
\]
So we have
\[
\|g_k\| \leq \beta\|d_k\|.
\]
By using $y_k^* = B_k s_k$ and (3.7) we get
\[
h\|s_k\|^2 \leq |s_k^T B_k s_k| \leq H\|s_k\|^2.
\] (3.12)

Lemma 3.2. $B_k$ is updated by (1.4), then
\[
\det(B_{k+1}) = \det(B_k) \frac{(y_k^*)^T s_k}{s_k^T B_k s_k},
\]
where $\det(B_k)$ denotes the determinant of $B_k$. 
Proof. By taking the determinant in both sides of (1.4), we have
\[
\begin{align*}
\det(B_{k+1}) &= \det\left( B_k \left( I - \frac{s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{B_k^{-1} y_k^*(y_k^*)^T}{s_k^T y_k^*} \right) \right) \\
&= \det(B_k) \det\left( I - \frac{s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{B_k^{-1} y_k^*(y_k^*)^T}{s_k^T y_k^*} \right) \\
&= \det(B_k) \left( 1 - \frac{s_k^T B_k s_k}{s_k^T B_k s_k} \right) + \left( 1 + \left( B_k^{-1} y_k^* \right)^T \frac{y_k^*}{(y_k^*)^T s_k} \right) - \left( -s_k^T \frac{B_k^{-1} y_k^*}{s_k^T B_k s_k} \right) - \left( (B_k s_k)^T \frac{B_k^{-1} y_k^*}{s_k^T B_k s_k} \right) \\
&= \det(B_k) \frac{(y_k^*)^T s_k}{s_k^T B_k s_k}.
\end{align*}
\]
where the third equality follows the formula in Lemma 7.6 in [16].
Our proof is completed. □

From [9], we can get the following Lemmas 2.3–2.6.

**Lemma 3.3.** Assume that Assumption (A) hold. Then, there exists a positive constant \( \varepsilon_0 \) such that
\[
\|z_k\| \geq \varepsilon_0 \min\{\gamma_k, (\gamma_k)^{1/(1-p)}\},
\]
where \( \gamma_k = -g_k^T d_k / \|d_k\| \).

**Lemma 3.4.** Denote that
\[
f(x_{h(k)}) = \max_{0 \leq j \leq M} f(x_k - j), \quad k - M_0 \leq h(k) \leq k.
\]
If \( f(x_{k+1}) \leq f(x_{h(k)}), \ k = 0, 1, \ldots \), then the sequence \( \{ f(x_{h(k)}) \} \) monotonically decreases, and \( x_k \in L_0 \) for all \( k \geq 0 \).

**Lemma 3.5.** Assume that
\[
f(x_{k+1}) \leq f(x_{h(k)}) - \rho_k, \quad k = 0, 1, \ldots ,
\]
where \( \rho_k \geq 0 \). Then
\[
\sum_{k=0}^{\infty} \min_{0 \leq j \leq M_0} \rho_{k+M_0-j} < +\infty.
\]

**Lemma 3.6.** If the sequence of nonnegative numbers \( m_k (k = 0, 1, \ldots) \) satisfies
\[
\prod_{j=0}^{k} m_j \geq c_1^k, \quad c_1 > 0, \quad k = 1, 2, \ldots.
\]
Then
\[
\limsup_k m_k > 0.
\]

**Lemma 3.7.** Assume that Assumption (A) hold. Suppose that \( x_0 \) is any starting point, \( B_0 \) is any symmetric positive definite matrix, and that the sequence \( \{x_k\} \) is generated by the MBFGS algorithm, in which the steplength \( z_k \) is determined by the GLL linesearch (1.5) and (1.6). We conclude that if
\[
\liminf_{k \to \infty} \| g_k \| > 0,
\]
then, there exists a constant $\varepsilon' > 0$ such that

$$
\prod_{j=1}^{k} \gamma_j \geq (\varepsilon')^k \text{ for all } k \geq 1.
$$

**Proof.** We assume that $\lim \inf_{k \to \infty} \|g_k\| > 0$, i.e., there exists $c_2 > 0$ such that

$$
\|g_k\| \geq c_2, \quad k = 0, 1, \ldots
$$

(3.13)

and from (1.4) and Lemma 3.1, we have

$$
\text{Tr}(B_{k+1}) = \text{Tr}(B_k) - \frac{\|B_k s_k\|^2}{s_k^TB_k s_k} + \frac{\|y_k^s\|^2}{(y_k^s)^T s_k} \leq \text{Tr}(B_k) - \frac{\|g_k\|^2}{g_k^T H_k g_k} + M_1
$$

$$
\leq \ldots \leq \text{Tr}(B_1) - \sum_{j=1}^{k} \frac{c_2^2}{g_j^T H_j g_j} + kM_1,
$$

where $\text{Tr}(B_k)$ denotes the trace of $B_k$. Hence,

$$
\text{Tr}(B_{k+1}) \leq \text{Tr}(B_1) + kM_1
$$

(3.14)

and

$$
\sum_{j=1}^{k} \frac{c_2^2}{g_j^T H_j g_j} \leq \frac{\text{Tr}(B_1) + kM_1}{c_2^2}.
$$

(3.15)

From the geometric–arithmetic mean value formula we have

$$
\prod_{j=1}^{k} \frac{g_j^T H_j g_j}{c_2^2} \geq \left[ \frac{k c_2^2}{\text{Tr}(B_1) + kM_1} \right]^k.
$$

(3.16)

Eq. (1.6) and Lemma 3.2 imply that

$$
\det(B_{k+1}) \geq \det(B_k) \frac{\min\{1 - \varepsilon_2, \|s_k\|^p\}}{\alpha_k}
$$

$$
\geq \ldots
$$

$$
\geq \det(B_1) \prod_{j=1}^{k} \frac{\min\{1 - \varepsilon_2, \|s_j\|^p\}}{\alpha_j},
$$

$$
\prod_{j=1}^{k} \max \left\{ \frac{\alpha_j}{1 - \varepsilon_2}, \frac{\alpha_j}{\|s_j\|^p} \right\} \geq \frac{\det(B_1)}{\det(B_{k+1})}.
$$

(3.17)

Again using the geometric–arithmetic mean value formula, we have

$$
\det(B_{k+1}) \leq \left[ \frac{\text{Tr}(B_{k+1})}{n} \right]^n.
$$

(3.18)
From (3.14) and (3.17), we have
\[
\prod_{j=1}^{k} \max \left\{ \frac{\alpha_j}{1 - \epsilon_2}, \frac{\alpha_j}{\|s_j\|^{p'}} \right\} \geq \frac{\det(B_1)n^n}{[\text{Tr}(B_1) + kM_1]^{n}} \\
\geq \frac{\det(B_1)n^n}{k^n[\text{Tr}(B_1) + M_1]^{n}} \\
\geq \left( \frac{1}{\exp(n)} \right)^k \min \left\{ \frac{\det(B_1)n^n}{[\text{Tr}(B_1) + M_1]^{n}}, 1 \right\} \\
\geq c_3^k,
\]
where \(c_3 \leq (1/\exp(n)) \min\{\det(B_1)n^n/[\text{Tr}(B_1) + M_1]^{n}, 1\}\). Denote
\[
\cos \theta_j = \frac{-g_j^T d_j}{\|g_j\| \|d_j\|}.
\]
Multiplying (3.16) with the above inequality, we have for all \(k \geq 1\)
\[
\prod_{j=1}^{k} \max \left\{ \frac{\|s_j\| \|g_j\| \cos \theta_j}{1 - \epsilon_2}, \frac{\|g_j\| \cos \theta_j}{\|s_j\|^{p-1}} \right\} \geq c_3^k \left[ \frac{k c_2^2}{\text{Tr}(B_1) + kM_1} \right]^k \geq \left[ \frac{c_3 c_2^2}{\text{Tr}(B_1) + M_1} \right]^k.
\]
Since
\[
\prod_{j=1}^{k} \max \left\{ \frac{\|s_j\| \|g_j\| \cos \theta_j}{1 - \epsilon_2}, \frac{\|g_j\| \cos \theta_j}{\|s_j\|^{p-1}} \right\} \leq \left( \frac{1}{1 - \epsilon_2} \right)^k \prod_{j=1}^{k} \max\{\|s_j\|, \|s_j\|^{1-p}\}\|g_j\| \cos \theta_j,
\]
thus
\[
\prod_{j=1}^{k} \max\{\|s_j\|, \|s_j\|^{1-p}\}\|g_j\| \cos \theta_j \geq \left[ \frac{1 - \epsilon_2}{\text{Tr}(B_1) + M_1} \frac{c_3 c_2^2}{\text{Tr}(B_1) + M_1} \right]^k.
\]
By Lemma 3.4 and Assumption (A)(i), we know that there exists \(L' > 0\) such that
\[
\|s_k\| = \|x_{k+1} - x_k\| \leq \|x_{k+1}\| + \|x_k\| \leq 2L',
\]
substituting it into (3.19), and noting that \(\|g_j\| \cos \theta_j = \gamma_j\), we have for all \(k \geq 1\)
\[
\prod_{j=1}^{k} \gamma_j \geq \left[ \frac{(1 - \epsilon_2)c_3 c_2^2}{\text{Tr}(B_1) + M_1} \max\{2L', 1, (2L')^{1-p}\} \right]^k = (\delta')^k.
\]
Our proof is completed. \(\square\)

**Theorem 3.8.** Assume that Assumption (A) hold. Suppose that \(x_0\) is any starting point, \(B_0\) is any symmetric positive definite matrix, and that the sequence \(\{x_k\}\) is generated by the MBFGS algorithm, in which the stepsize \(\alpha_k\) is determined by the GLL linesearch (1.5) and (1.6). Then
\[
\liminf_{k \to \infty} \|g_k\| = 0.
\]
**Proof.** By (1.5) and Lemma 3.3 we have
\[
f(x_{k+1}) \leq f(x_{h(k)}) - \varepsilon_1 \|s_k\| \gamma_k
\leq f(x_{h(k)}) - \varepsilon_1 e_0 \min \{\gamma_k^2, (\gamma_k)^{(2-p)/(1-p)}\}.
\]

Set \( \rho_k = \varepsilon_1 e_0 \min \{\gamma_k^2, (\gamma_k)^{(2-p)/(1-p)}\} \). From Lemma 3.5 we have
\[
\sum_{k=1}^{\infty} \min_{0 \leq j \leq M_0} \min \{\gamma_k^2, (\gamma_k)^{(2-p)/(1-p)}\} < +\infty,
\]
\[
\sum_{q=1}^{\infty} \min_{0 \leq j \leq M_0} \min \{\gamma_k^2, (\gamma_k)^{(2-p)/(1-p)}\} < +\infty.
\]

Denote the sequence \( \{p(q)\} \) as follows:
\[
\min \{\gamma_p(q)^2, (\gamma_p(q))^{(2-p)/(1-p)}\} = \min_{0 \leq j \leq M_0} \min \{T_1 j(q), T_2 j(q)\},
\]
\[
T_1 j(q) = (\gamma_p(q)^2),
\]
\[
T_2 j(q) = (\gamma_p(q)^{2-p}/(1-p)),
\]
\[
q(M_0 + 1) \leq p(q) \leq (q + 1)M_0 + q.
\]

Then
\[
p(1) \leq p(2) \leq \cdots \leq p(q - 1) \leq p(q) \leq \cdots,
\]
\[
\lim_{q \to \infty} \min \{\gamma_p(q)^2, (\gamma_p(q))^{(2-p)/(1-p)}\} = 0,
\]
\[
\lim_{q \to \infty} \gamma_p(q) = 0, \tag{3.21}
\]
which just means that \( \lim_{k \in K} \gamma_k = 0, K \subset N \).

Since \( x_k \in L_0, L_0 \) is bounded, we can assume that there exists \( c_4 > 0 \) such that \( \|g_k\| \leq c_4 \). Thus
\[
\gamma_k = -\frac{g_k^T d_k}{d_k} \leq \|g_k\| \leq c_4. \tag{3.22}
\]

Now, we proceed our proof by contradiction. Assume that \( \lim \inf_k \|g_k\| > 0 \), i.e., there exists a positive constant \( c_2 \) such that
\[
\|g_k\| \geq c_2, \quad k = 0, 1, \ldots.
\]

From Lemma 3.7 we know that there exists \( \varepsilon' > 0 \) such that
\[
\prod_{j=1}^{k} \gamma_j \geq (\varepsilon')^k.
\]
From which and (3.22), we know that for any integer \( k \geq 1 \),
\[
(\epsilon')^{(k+1)M_0+k} \leq \prod_{j=1}^{k} \gamma_j
\]
\[
= \frac{1}{\gamma_0} \prod_{q=0}^{k} \prod_{j=q(M_0+1)}^{(q+1)M_0+q} \gamma_j
\]
\[
= \frac{1}{\gamma_0} \prod_{q=0}^{k} \prod_{0 \leq j \leq M_0} \gamma_q(M_0+1)+M_0-j
\]
\[
\leq \frac{1}{\gamma_0} \prod_{q=0}^{k} \left[ \gamma_p(q)(c_4)^M_0 \right]
\]
\[
= \frac{1}{\gamma_0} (c_4)^k M_0 \prod_{q=0}^{k} \gamma_p(q),
\]
\[
\prod_{q=0}^{k} \gamma_p(q) \geq \gamma_0 (\epsilon')^{M_0} \left[ \frac{(\epsilon')^{M_0+1}}{(c_4)^M_0} \right]^k \geq \left[ \frac{(\epsilon')^{M_0+1}}{(c_4)^M_0} \min\{1, \gamma_0^{(\epsilon')} M_0\} \right]^k.
\]

By Lemma 3.6 we have
\[
\limsup_{q \to \infty} \gamma_p(q) > 0,
\]
which contradicts (3.21). Therefore,
\[
\liminf_{k \to \infty} \|g_k\| = 0.
\]
Our proof is completed. \(\square\)

The above theorem established the global convergence of Algorithm 2.1.
Next we will give the superlinear convergence of Algorithm 2.1.

4. Superlinear convergence

In order to give the superlinear convergence of Algorithm 2.1, in addition to Assumption (A), we also need the following Assumption (B). Let \( x^* \) be the limit of the sequence \( \{x_k\} \).

**Assumption (B).**

(i) \( \{x_k\} \) converges to \( x^* \) where \( g(x^*)=0 \) and \( G(x^*) \) is positive definite, where \( G(x) = \nabla^2 f(x), x \in L_0. \)
(ii) \( G(x) \) is Hölder continuous at \( x^* \), i.e., there exist two constants \( \nu \geq 0 \) and \( M_2 \geq 0 \) such that
\[
\|G(x) - G(x^*)\| \leq M_2 \|x - x^*\|^\nu,
\]
for all \( x \) in the neighborhood of \( x^* \).

**Lemma 4.1.** Assume that Assumption (A) hold. Let the sequence \( \{x_k\} \) be generated by Algorithm 2.1. If we use \( \xi_k \) denote the angle between \( s_k \) and \( B_k s_k \), i.e.,
\[
\cos \xi_k = \frac{s_k^T B_k s_k}{\|s_k\| \|B_k s_k\|} = -\frac{g_k^T s_k}{\|g_k\| \|s_k\|},
\]
(4.2)
then there is a constant $a_1 > 0$ such that
\[ a_1 \|g_k\| \cos \hat{\zeta}_k \leq \|s_k\|, \] (4.3)
and there is a constant $a_2 > 0$ satisfying
\[ \cos \hat{\zeta}_k \geq a_2 \quad \text{for all } k \in K. \] (4.4)

**Proof.** By using Assumption (A) and (1.6), we have
\[ y_k^T s_k \geq c_0 (-g_k^T s_k), \quad c_0 \in (0, 1). \] (4.5)
From the fact that \{x_k\} is bounded, by using Assumption (A)(i), (3.3) and (3.4), we can deduce that there exists $M_2 > 0$, such that for all $k$
\[ \|g(x_k)\| \leq M_2. \] (4.6)
Based on (3.3), (3.8) and (4.2), we have
\[ M \|s_k\|^2 \geq \|G^{(\zeta')} s_k\| \|s_k\| = \|g_{k+1} - g_k\| \|s_k\| \geq y_k^T s_k \geq c_0 (-g_k^T s_k) = c_0 \|g_k\| \|s_k\| \cos \zeta_k, \]
where $\zeta' = x_k + \theta'(x_{k+1} - x_k)$, $\theta' \in (0, 1)$. Then we get (4.3). From (3.10) and (3.11), it is clear that when $k \in k$, we can obtain
\[ \cos \hat{\zeta}_k = \frac{s_k^T B_k s_k}{\|s_k\| \|B_k s_k\|} \geq \frac{1}{\bar{B}}. \]
Our proof is completed. □

**Lemma 4.2.** Assume that Assumptions (A) and (B) hold. Let the sequence \{x_k\} be generated by Algorithm 2.1. Then
\[ \sum_{k=0}^{\infty} \|x_k - x^*\| < \infty. \] (4.7)
Moreover, if denote $\tau_k = \max\{\|x_k - x^*\|, \|x_{k+1} - x^*\|\}$, then
\[ \sum_{k=0}^{\infty} \tau_k < + \infty \] (4.8)
and
\[ \sum_{k=0}^{\infty} \|A_k\| < \infty. \] (4.9)

**Proof.** From Assumption (A) and (B), we can obtain that there is a constant $\lambda_3 \geq 0$ such that
\[ \|g(x)\| = \|g(x) - g(x^*)\| \geq \lambda_3 \|x - x^*\| \] (4.10)
for all $x$ in the neighborhood of $x^*$, and by (4.3), (4.4) we can get for all $k$ large enough
\[ -g_k^T s_k = \|g_k\| \|s_k\| \cos \zeta_k \geq a_1 \|g_k\|^2 \cos^2 \zeta_k \geq a_1 \lambda_3^2 \|x_k - x^*\|^2 \cos^2 \zeta_k \geq a_1 a_2 \lambda_3^2 \|x_k - x^*\|^2. \] (4.11)
By the GLL linesearch (1.5) and Lemma 3.5, we get
\[ \sum_{k=0}^{\infty} \min_{0 \leq j \leq M_0} (-e_1 g_{k+M_0-j}^T s_{k+M_0-j}) < + \infty \]
and by (4.11), we get (4.7). Notice that $\tau_k = \max\{\|x_k - x^*\|, \|x_{k+1} - x^*\|\}$, then (4.8) holds.
From Lemma 3.2 in [1], we get

$$\|A_k^*\| \leq \|G(\xi_{k1}) - G(\xi_{k2})\|,$$

where $\xi_{k1} = x_k + \theta_{k1}(x_{k+1} - x_k)$, $\xi_{k2} = x_k + \theta_{k2}(x_{k+1} - x_k)$, and $\theta_{k1}, \theta_{k2} \in (0, 1)$, and by Assumption (A)(ii), we have

$$\sum_{k=0}^{\infty} \|A_k^*\| \leq \sum_{k=0}^{\infty} \|G(\xi_{k2}) - G(\xi_{k1})\|$$

$$\leq \sum_{k=0}^{\infty} \|G(\xi_{k1}) - G(x*)\| + \|G(\xi_{k2}) - G(x*)\|$$

$$\leq \sum_{k=0}^{\infty} M \|\xi_{k1} - x*\|^T + \sum_{k=0}^{\infty} M \|\xi_{k2} - x*\|^T$$

$$\leq 2M \sum_{k=0}^{\infty} \tau_k.$$

Therefore, (4.9) holds. Our proof is completed. \[\square\]

Lemma 4.3. Assume that Assumptions (A) and (B) hold. Let the sequence $\{x_k\}$ be generated by Algorithm 2.1. Denote $Q = G(x*)^{-1}$ and $H_k = B_k^{-1}$. Then there are nonnegative constants $\alpha_1, \alpha_2, b_i$, $i = 1, 2, \ldots, 7$ and $\eta \in (0, 1)$ such that for all large $k$:

$$\|B_{k+1} - G(x*)\|_{Q,F} \leq (1 + \alpha_1 \tau_k) \|B_k - G(x*)\|_{Q,F} + \alpha_2 \tau_k$$

(4.12)

and

$$\|H_{k+1} - G(x*)\|_{Q^{-1},F} \leq [(1 - p\alpha_2^2)^{1/2} + b_4 \tau_k + b_5 \|A_k^*\|] \|H_k - G(x*)^{-1}\|_{Q^{-1},F} + b_6 \tau_k + b_7 \|A_k^*\|.$$  

(4.13)

where $\|A\|_{Q,F} = \|Q^T A Q\|_F$, $\|\cdot\|_F$ is the Frobenius norm of a matrix and $\alpha_k$ is defined as follows:

$$\alpha_k = \frac{\|Q^{-1}(H_k - G(x*)^{-1})y_k^*\|}{\|H_k^* - G(x*)^{-1}\|_{Q^{-1},F} \|Q y_k^*\|}.$$  

(4.14)

In particular, $\{\|B_k^*\|_F\}$ and $\{\|H_k\|_F\}$ are bounded.

Proof. From the definition of $B_{k+1}$, we have

$$\|B_{k+1} - G(x*)\|_{Q,F} = \left\|B_k - G(x*) - \frac{B_k s_k y_k^T}{s_k^T B_k s_k} B_k + \frac{y_k^* y_k^{**} T}{y_k^T s_k} \right\|_{Q,F}$$

$$\leq \left\|B_k - G(x*) - \frac{B_k s_k y_k^T}{s_k^T B_k s_k} B_k + \frac{y_k^* y_k^{**} T}{y_k^T s_k} \right\|_{Q,F} + \left\|\frac{y_k^* y_k^{**} T}{y_k^T s_k} - \frac{y_k y_k^T}{y_k^T s_k} \right\|_{Q,F}$$

$$\leq (1 + \alpha_1 \tau_k) \|B_k - G(x*)\|_{Q,F} + b_1 \tau_k + \left\|\frac{y_k^* y_k^{**} T}{y_k^T s_k} - \frac{y_k y_k^T}{y_k^T s_k} \right\|_{Q,F}.$$
where the last inequality follows the inequality (49) in [17]. Moreover, by using (3.2) and (3.6), we have

\[
\left\| \frac{y_k^* y_k^T}{y_k^T s_k} - \frac{y_k y_k^T}{y_k^T s_k} \right\|_{Q,F} \\
\leq \left\| \frac{(y_k + A_k^* s_k)(y_k + A_k^* s_k)^T}{(y_k + A_k^* s_k)^T s_k} - \frac{y_k y_k^T}{y_k^T s_k} \right\|_{Q,F} \\
\leq \left\| \frac{y_k^T s_k (y_k + A_k^* s_k)(y_k + A_k^* s_k)^T - (y_k + A_k^* s_k)^T s_k y_k y_k^T}{(y_k + A_k^* s_k)^T s_k (y_k^T s_k)^2} \right\|_{Q,F} \\
\leq \left\| A_k^* \right\| \left\| y_k^T s_k s_k y_k^T \right\|_{Q,F} + \left\| y_k^T s_k y_k^T s_k \right\|_{Q,F} + \left\| y_k^T s_k s_k^T A_k^* \right\|_{Q,F} + \left\| s_k^T s_k y_k y_k^T \right\|_{Q,F} \\
\leq \left\| A_k^* \right\| \frac{(3\|s_k\|^2\|y_k\|^2 + \|A_k^*\||\|s_k\|^3\|y_k\|)\|Q\|^2}{(y_k + A_k^* s_k)^T s_k (g_k + g_k)^T s_k} \\
\leq \left\| A_k^* \right\| \frac{(3\|s_k\|^2\|y_k\|^2 + \|A_k^*\||\|s_k\|^3\|y_k\|)\|Q\|^2}{(y_k + A_k^* s_k)^T s_k (s_k^T G(\xi_k) s_k)} \\
\leq \left\| A_k^* \right\| \frac{(3\|s_k\|^2\|y_k\|^2 + \|A_k^*\||\|s_k\|^3\|y_k\|)\|Q\|^2}{h^2\|s_k\|^4} \\
= \left\| A_k^* \right\| \frac{3\|y_k\|^2 + \|A_k^*\||\|s_k\|^3\|y_k\|)\|Q\|^2}{h^2\|s_k\|^2}.
\]

By using Lemma 4.2 we know \( \lim_{k \to \infty} \|A_k^*\| = 0 \), the definition of \( y_k \) and \( \|y_k\| \leq L\|s_k\| \), we can obtain that there exists a positive constant \( b_2 \) such that

\[
\left\| \frac{y_k^* y_k^T}{y_k^T s_k} - \frac{y_k y_k^T}{y_k^T s_k} \right\|_{Q,F} \leq b_2 \|A_k^*\|.
\]

Also from the definition of \( A_k^* \) and Assumption (B)(ii), we have

\[
\|A_k^*\| = \left\| \frac{2[f(x_k) - f(x_{k+1})] + (g_k + g_k)^T s_k}{\|s_k\|^2} \right\| \\
= \left\| \frac{2g_k s_k - \frac{1}{2} s_k^T G(\xi_k) s_k + (g_k + g_k)^T s_k}{\|s_k\|^2} \right\| \\
= \left\| \frac{-s_k^T G(\xi_k) s_k + g_k + g_k s_k}{\|s_k\|^2} \right\| \\
= \left\| \frac{-s_k^T G(\xi_k) s_k + s_k^T G(\xi_k) s_k}{\|s_k\|^2} \right\| \\
\leq \|G(\xi_k) - G(\xi_k)\| \\
\leq \|G(\xi_k) - G(x^*)\| + \|G(\xi_k) - G(x)\| \\
\leq M_2\|\xi_k - x^*\| + M_2\|\xi_k - x^*\| \\
\leq 2M_2 \tau_k,
\]
Proof. Let the sequence \( \{x_k\} \) be generated by Algorithm 2.1. Then, the following Dennis–Moré condition
\[
\lim_{k \to \infty} \frac{\| (B_k - G(x^*))s_k \|}{\| s_k \|} = 0
\] (4.15)
holds.

Proof. Using \( \tau_k \to 0 \), \( \| A_k^* \| \to 0 \), \( \{\| H_k \|\} \) is bounded and the following inequality:
\[
\sqrt{1 - t} \leq 1 - \frac{1}{2} t \quad \text{for all } t \in (0, 1),
\]
can be deduced that there are positive constants \( M_3 \) and \( M_4 \) such that for all large \( k \)
\[
\| H_{k+1} - G(x^*)^{-1} \|_{Q^{-1}, F} \leq \left( 1 - \frac{1}{2} \eta \omega_k^2 \right) \| H_k - G(x^*)^{-1} \|_{Q^{-1}, F} + M_3 \tau_k + M_4 \| A_k^* \|.
\]
By Lemma 4.3 we know that \( \| A_k^* \| \leq 2M_2 \tau_k \). So we have
\[
\| H_{k+1} - G(x^*)^{-1} \|_{Q^{-1}, F} \leq \left( 1 - \frac{1}{2} \eta \omega_k^2 \right) \| H_k - G(x^*)^{-1} \|_{Q^{-1}, F} + M_5 \tau_k,
\]
where \( M_5 = M_3 + 2M_2 M_4 \), i.e.,
\[
\frac{1}{2} \eta \omega_k^2 \| H_k - G(x^*)^{-1} \|_{Q^{-1}, F} \leq \| H_k - G(x^*)^{-1} \|_{Q^{-1}, F} - \| H_{k+1} - G(x^*)^{-1} \|_{Q^{-1}, F} + M_5 \tau_k.
\]
Summing the above inequality over \( k \), we get
\[
\frac{1}{2} \eta \sum_{k=k_0}^{\infty} \omega_k^2 \| H_k - G(x^*)^{-1} \|_{Q^{-1}, F} < + \infty,
\]
where \( k_0 \) is a sufficiently large index such that (4.13) holds for all \( k \geq k_0 \). In particular, we have \( \lim_{k \to \infty} \omega_k^2 \| H_k - G(x^*)^{-1} \|_{Q^{-1}, F} = 0 \), i.e.,
\[
\lim_{k \to \infty} \frac{\| Q^{-1}(H_k - G(x^*)^{-1})y_k \|}{\| Qy_k \|} = 0.
\] (4.16)
Moreover we have
\[
\| Q^{-1}(H_k - G(x^*)^{-1})y_k \| = \| Q^{-1}H_k(G(x^*) - B_k)G(x^*)^{-1}y_k \|
\geq \| Q^{-1}H_k(G(x^*) - B_k)\| - \| Q^{-1}H_k(G(x^*) - B_k)(s_k - G(x^*)^{-1}y_k) \|.
\]
Using the facts that \( \{ 1, B_k^* \} \) and \( \{ H_k \} \) are bounded, \( G(x) \) is continuous, we get

\[
\| Q^{-1} H_k (G(x^*) - B_k) (s_k - G(x^*)^{-1} y_k) \| \\
= \| Q^{-1} H_k (G(x^*) - B_k) G(x^*)^{-1} (G(x^*) s_k - y_k) \| \\
= \| Q^{-1} H_k (G(x^*) - B_k) G(x^*)^{-1} \left[ (G(x^*) - G(x_k)) s_k + (G(x_k) s_k - y_k) \right] \| \\
\leq \| Q^{-1} H_k (G(x^*) - B_k) G(x^*)^{-1} \| \| G(x^*) - G(x_k) \| \| s_k \| + \| (G(x_k) s_k - y_k) \| \\
= o(\| s_k \|).
\]

Therefore, there exists a positive constant \( \kappa > 0 \) such that

\[
\| Q^{-1} (H_k - G(x^*)^{-1}) y_k \| \geq \kappa \| (G(x^*) - B_k) s_k \| - o(\| s_k \|).
\] (4.17)

On the other hand, from (3.8), we have

\[
\| Q y_k \| \leq \| Q \| \| y_k \| \leq (2L + H) \| Q \| \| s_k \|
\]

and by (4.16) we have

\[
\lim_{k \to \infty} \frac{\| (B_k - G(x^*)) s_k \|}{\| s_k \|} = 0.
\]

Our proof is completed. \( \square \)

**Theorem 4.5.** Assume that Assumptions (A) and (B) hold. Let the sequence \( \{ x_k \} \) be generated by Algorithm 2.1. Then \( \{ x_k \} \) tends to \( x^* \) superlinearly.

**Proof.** We verify that \( \lambda \equiv 1 \) for all \( k \in K \) sufficiently large. Since the sequence \( \| B_k \| \) and \( \| B_k^{-1} \| \) is bounded, we have

\[
\| d_k \| = \| B_k^{-1} g_k \| \leq \| B_k^{-1} \| \| g_k \| \to 0.
\]

By Taylor’s expansion and Lemma 3.4, we get

\[
f(x_k + d_k) - f(x_k) - \epsilon_1 g_k^T d_k + f(x_k) - f(x_h(k)) \\
\leq f(x_k + d_k) - f(x_k) - \epsilon_1 g_k^T d_k \\
= (1 - \epsilon_1) g_k^T d_k + \frac{1}{2} d_k^T G(x_{k+1}) d_k + o(\| d_k \|^2) \\
= - (1 - \epsilon_1) d_k^T B_k d_k + \frac{1}{2} d_k^T G(x_{k+1}) d_k + o(\| d_k \|^2) \\
= - \left( \frac{1}{4} - \epsilon_1 \right) d_k^T G(x^*) d_k + o(\| d_k \|^2),
\]

where \( x_{k+1} = x_k + \theta_{k+1} (x_{k+1} - x_k), \theta_{k+1} \in (0, 1), \epsilon_1 \in (0, \frac{1}{2}) \), and the last equality follows from the Dennis–Moré condition (4.15). Therefore,

\[
f(x_k + d_k) - \max_{0 \leq j \leq M_0} f(x_{k-j}) - \epsilon_1 g_k^T d_k \leq 0
\]

holds for all large \( k \). In other words, \( \sigma_k \equiv 1 \) satisfies the first inequality of the GLL (1.5) for all \( k \) sufficiently large.

In (1.6), we denote \( \delta_3 = \max \{ \sigma_2, 1 - \| d_k \|^2 \} \), then \( \delta_3 \in (0, 1) \). On the other hand,

\[
g(x_k + d_k)^T d_k - \delta_3 g_k^T d_k \\
= (g(x_k + d_k) - g(x_k))^T d_k + (1 - \delta_3) g_k^T d_k \\
= d_k^T G(x_k) d_k + (1 - \delta_3) d_k^T B_k d_k \\
= d_k^T G(x^*) d_k + (1 - \delta_3) d_k^T G(x_k) d_k + o(\| d_k \|^2) \\
= \delta_3 d_k^T G(x^*) d_k + o(\| d_k \|^2),
\]
where \( \theta_k \in (0, 1) \). Then we get
\[
g(x_k + d_k)^T d_k \geq \delta_3 g_k^T d_k,
\] (4.18)
which means that \( x_k \equiv 1 \) satisfies the second inequality of the GLL (1.6) for all \( k \) sufficiently large. Therefore, we ensure that \( x_k \equiv 1 \) for all \( k \) sufficiently large. Consequently, we can deduce that \( x_k \) be convergent superlinearly. Our proof is completed. □

5. Numerical experiments

In this section, we test the algorithm (2.1) given by the paper. And the numerical experiments of the original BFGS formula with nonmonotone linesearch and the formulas (1.5) and (1.6) are also contained in this section for comparing. The problems that we tested are from [18]. For each test problems, the termination condition is
\[
\|g(x_k)\| \leq 10^{-5}.
\]

In order to rank the iterative numerical methods, we compute the total number of function and gradient evaluations by the following formula:
\[
N_{\text{total}} = NF + m \times NG.
\] (5.1)
where NF, NG denote the number of function evaluations and gradient evaluations, respectively, and \( m \) is some integer. According to the results on automatic differentiation [1,4], the value of \( m \) can be set to \( m = 5 \). That is to say, one gradient evaluation is equivalent to \( m \) number of function evaluations if automatic differentiation is used.

Table 1 shows the computation results, where the columns have the following meanings:

<table>
<thead>
<tr>
<th>Problems</th>
<th>the name of the test problems in MATLAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dim</td>
<td>the dimension of iterations</td>
</tr>
<tr>
<td>NI</td>
<td>the number of function evaluations</td>
</tr>
<tr>
<td>NG</td>
<td>the number of gradient evaluations</td>
</tr>
<tr>
<td>NMBFGS</td>
<td>the formula (1.4) with nonmonotone linesearch (1.5) and (1.6) in which ( M_0 = 5 )</td>
</tr>
<tr>
<td>WMBFGS</td>
<td>the formula (1.4) with the standard Wolfe linesearch</td>
</tr>
<tr>
<td>NBFGS</td>
<td>the original BFGS formula with nonmonotone linesearch (1.5) and (1.6) in which ( M_0 = 5 )</td>
</tr>
</tbody>
</table>

In order to rank these methods, we compute the total number of function and gradient evaluations by the following formula:
\[
N_{\text{total}} = NF + 5 \times NG.
\] (5.2)

Therefore, in this part, we compare the WMBFGS and NBFGS methods with NMBFGS method as follows: for each testing example \( i \), compute the total numbers of function evaluations and gradient evaluations required by the evaluated method \( j \) (EM(\( j \))) and NMBFGS by formulas (5.2), and denote them by \( N_{\text{total},i}(\text{EM}(j)) \) and \( N_{\text{total},i}(\text{NMBFGS}) \); then, calculate the ratio
\[
\tau_i(\text{EM}(j)) = \frac{N_{\text{total},i}(\text{EM}(j))}{N_{\text{total},i}(\text{NMBFGS})}. \] (5.3)

If EM(\( j_0 \)) does not work for example \( i_0 \), we replace the \( \tau_{i_0}(\text{EM}(j_0)) \) by a positive constant \( \tau \) which is defined as follows:
\[
\tau = \max\{\tau_l(\text{EM}(j)) : \text{where} (i, j) \notin S_1\}, \] (5.4)
where
\[
S_1 = \{(i, j) : \text{method} j \text{ does not work for example} i\}. \]
<table>
<thead>
<tr>
<th>Problem</th>
<th>Dim</th>
<th>Wolfe—MBFGS NI/NF/NG</th>
<th>Nonmonotone—MBFGS NI/NF/NG</th>
<th>Nonmonotone—BFGS NI/NF/NG</th>
</tr>
</thead>
<tbody>
<tr>
<td>rose</td>
<td>2</td>
<td>29/51/30</td>
<td>29/51/30</td>
<td>30/97/31</td>
</tr>
<tr>
<td>froth</td>
<td>2</td>
<td>10/22/11</td>
<td>9/36/10</td>
<td>9/36/10</td>
</tr>
<tr>
<td>badscb</td>
<td>2</td>
<td>13/138/14</td>
<td>13/138/14</td>
<td>14/181/15</td>
</tr>
<tr>
<td>beale</td>
<td>2</td>
<td>15/25/16</td>
<td>13/53/14</td>
<td>14/53/15</td>
</tr>
<tr>
<td>jensam</td>
<td>2</td>
<td>14/26/15</td>
<td>14/26/15</td>
<td>12/24/13</td>
</tr>
<tr>
<td>helix</td>
<td>3</td>
<td>28/55/29</td>
<td>26/145/27</td>
<td>28/95/29</td>
</tr>
<tr>
<td>bard</td>
<td>3</td>
<td>18/44/19</td>
<td>16/59/17</td>
<td>17/69/18</td>
</tr>
<tr>
<td>gauss</td>
<td>3</td>
<td>4/7/5</td>
<td>4/7/5</td>
<td>4/7/5</td>
</tr>
<tr>
<td>gulf</td>
<td>3</td>
<td>1/4/2</td>
<td>1/4/2</td>
<td>1/4/2</td>
</tr>
<tr>
<td>box</td>
<td>3</td>
<td>2/61/22</td>
<td>22/77/23</td>
<td>27/106/28</td>
</tr>
<tr>
<td>sing</td>
<td>4</td>
<td>23/46/24</td>
<td>23/46/24</td>
<td>29/52/30</td>
</tr>
<tr>
<td>wood</td>
<td>4</td>
<td>53/93/54</td>
<td>52/107/53</td>
<td>52/97/53</td>
</tr>
<tr>
<td>kowosb</td>
<td>4</td>
<td>28/32/29</td>
<td>28/32/29</td>
<td>28/32/29</td>
</tr>
<tr>
<td>bd</td>
<td>4</td>
<td>34/49/11/19</td>
<td>23/49/18</td>
<td>22/102/23</td>
</tr>
<tr>
<td>biggs</td>
<td>6</td>
<td>35/46/36</td>
<td>30/142/31</td>
<td>34/161/35</td>
</tr>
<tr>
<td>osb2</td>
<td>11</td>
<td>56/82/57</td>
<td>54/128/55</td>
<td>56/142/57</td>
</tr>
<tr>
<td>watson</td>
<td>20</td>
<td>55/104/56</td>
<td>55/91/56</td>
<td>55/91/56</td>
</tr>
<tr>
<td>roseed</td>
<td>8</td>
<td>80/140/81</td>
<td>70/175/71</td>
<td>77/178/78</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>256/617/257</td>
<td>244/632/245</td>
<td>241/629/242</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>412/1158/413</td>
<td>339/1093/340</td>
<td>379/1188/380</td>
</tr>
<tr>
<td>singx</td>
<td>4</td>
<td>23/46/24</td>
<td>23/46/24</td>
<td>29/52/30</td>
</tr>
<tr>
<td>pen1</td>
<td>2</td>
<td>165/28/2/166</td>
<td>166/269/167</td>
<td>183/309/184</td>
</tr>
<tr>
<td>pen2</td>
<td>4</td>
<td>311/494/312</td>
<td>324/495/325</td>
<td>383/561/384</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>316/931/317</td>
<td>312/892/313</td>
<td>312/914/313</td>
</tr>
<tr>
<td>vardim</td>
<td>2</td>
<td>5/13/6</td>
<td>5/13/6</td>
<td>6/14/7</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>23/103/24</td>
<td>20/158/21</td>
<td>21/85/22</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>42/43/43</td>
<td>42/43/43</td>
<td>44/48/45</td>
</tr>
<tr>
<td>trig</td>
<td>3</td>
<td>13/27/14</td>
<td>13/27/14</td>
<td>12/27/13</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>42/43/43</td>
<td>42/43/43</td>
<td>44/48/45</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>49/52/50</td>
<td>49/52/50</td>
<td>48/51/49</td>
</tr>
<tr>
<td>bv</td>
<td>3</td>
<td>6/14/7</td>
<td>6/14/7</td>
<td>6/14/7</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>18/39/19</td>
<td>18/39/19</td>
<td>18/39/19</td>
</tr>
<tr>
<td>ie</td>
<td>3</td>
<td>7/11/8</td>
<td>7/11/8</td>
<td>7/11/8</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>12/15/13</td>
<td>12/15/13</td>
<td>12/15/13</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>12/15/13</td>
<td>12/15/13</td>
<td>12/15/13</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td>12/15/13</td>
<td>12/15/13</td>
<td>12/15/13</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>12/15/13</td>
<td>12/15/13</td>
<td>13/16/14</td>
</tr>
<tr>
<td>trid</td>
<td>3</td>
<td>12/29/13</td>
<td>12/29/13</td>
<td>13/45/14</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>67/335/68</td>
<td>67/335/68</td>
<td>63/340/64</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>112/644/113</td>
<td>112/644/113</td>
<td>112/637/113</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td>195/1155/196</td>
<td>195/1155/196</td>
<td>216/1223/217</td>
</tr>
<tr>
<td>band</td>
<td>3</td>
<td>–</td>
<td>20/111/21</td>
<td>–</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>lin</td>
<td>2</td>
<td>1/3/2</td>
<td>1/3/2</td>
<td>1/3/2</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>1/3/2</td>
<td>1/3/2</td>
<td>1/3/2</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>1/3/2</td>
<td>1/3/2</td>
<td>1/3/2</td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td>1/3/2</td>
<td>1/3/2</td>
<td>1/3/2</td>
</tr>
<tr>
<td>lin1</td>
<td>2</td>
<td>2/10/3</td>
<td>2/10/3</td>
<td>2/10/3</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>2/10/3</td>
<td>2/10/3</td>
<td>2/10/3</td>
</tr>
<tr>
<td>lin0</td>
<td>4</td>
<td>2/11/3</td>
<td>2/11/3</td>
<td>2/11/3</td>
</tr>
</tbody>
</table>

The geometric mean of these ratios for method EM\( (j) \) over all the test problems is defined by

\[
r(\text{EM}(j)) = \left( \prod_{i \in S} r_i(\text{EM}(j)) \right)^{1/|S|},
\]
Table 2
Numerical results

<table>
<thead>
<tr>
<th></th>
<th>WMBFGS</th>
<th>NBFGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9788</td>
<td>1.0337</td>
</tr>
</tbody>
</table>

where $S$ denotes the set of the test problems and $|S|$ denotes the number of elements in $S$. According to the above rule, it is clear that $r(\text{NMBFGS}) = 1$. The values of $r(\text{WMBFGS})$ and NBFGS are listed in Table 2.

References