A GLOBALY CONVERGENT BFGS METHOD FOR
SYMMETrIC NONLINEAR EQUATIONS

WEIJUN ZHOU∗
Department of Mathematics and Statistics
Changsha University of Science and Technology, Changsha 410114, China
(Communicated by Vladimir Shikhman)

ABSTRACT. A BFGS type method is presented to solve symmetric nonlinear equations, which is shown to be globally convergent under suitable conditions. Compared with some existing Gauss-Newton-based BFGS methods whose iterative matrix approximates the Gauss-Newton matrix, an important feature of the proposed method lies in that the iterative matrix is an approximation of the Jacobian, which greatly reduces condition number of the iterative matrix. Numerical results are reported to support the theory.

1. Introduction. BFGS type quasi-Newton methods are very efficient for optimization [1, 2, 4, 11, 13, 15, 16, 22]. However, the study on global convergence of BFGS type methods for nonlinear equations is relatively rare. This paper is devoted to BFGS type methods for solving symmetric nonlinear equations

\[ F(x) = 0, \tag{1} \]

where \( F : \mathbb{R}^n \to \mathbb{R}^n \) is a continuously differentiable mapping and the symmetry means that the Jacobian \( J(x) = F'(x) \) is symmetric, i.e., \( J(x) = J(x)^T \). Some practical problems such as the KKT system of equality constrained optimization take the form of (1) with symmetric Jacobian [7, 18, 23, 24, 25, 27].

Let \( \{s_k\} \) and \( \{\delta_k\} \) be two given sequences. The BFGS update formula is given by

\[ B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{\delta_k \delta_k^T}{\delta_k^T s_k}, \tag{2} \]

It is well-known that \( B_{k+1} \) is symmetric and positive definite when \( \delta_k^T s_k > 0 \) and \( B_k \) is symmetric and positive definite. Moreover, Byrd and Nocedal [2] revealed the following intrinsic property of the BFGS update formula (2), which is independent of the algorithmic context of the update and very convenient for global analysis of BFGS type methods.

Lemma 1.1. [2, Theorem 2.1] Let \( \{s_k\} \) and \( \{\delta_k\} \) be two given sequences and the sequence \( \{B_k\} \) be generated by the BFGS update formula (2), where \( B_0 \) is symmetric
and positive definite and $\delta_k^T s_k > 0$ for all $k$. If there exist two positive constants $m_1$ and $m_2$ such that for any $k \geq 0$,

$$\frac{\delta_k^T s_k}{s_k^T s_k} \geq m_1, \quad \frac{\delta_k^T \delta_k}{s_k^T s_k} \leq m_2,$$

then for any $p \in (0, 1)$ and $k > 1$, there are positive constants $\beta_i, i = 1, 2, 3, 4$ such that

$$\beta_1 \|s_j\| \leq \|B_j s_j\| \leq \beta_2 \|s_j\|, \quad \beta_3 \|s_j\|^2 \leq s_j^T B_j s_j \leq \beta_4 \|s_j\|^2$$

(3) hold for at least $\lfloor pk \rfloor$ values of $j \in [1, k]$, where $\lfloor t \rfloor$ is the largest integer which is less than or equal to $t$.

BFGS type methods for solving (1) are iterative methods of the form

$$x_{k+1} = x_k + \alpha_k d_k,$$

(4) where $\alpha_k > 0$ is a stepsize given by some line search and $d_k$ is a search direction. Throughout the paper, we set

$$s_k = x_{k+1} - x_k, \quad y_k = F_{k+1} - F_k,$$

(5) where $F_k = F(x_k)$. Different $d_k$ and $\delta_k$ lead to different BFGS type methods.

The classical BFGS method produces the search direction by

$$d_k = -B_k^{-1} F_k, \quad \delta_k = y_k.$$

(6) It is clear that the iterative matrix $B_{k+1}$ generated by (2) approximates the Jacobian $J_{k+1} = J(x_{k+1})$. Under mild conditions, the classical BFGS method possesses locally superlinear convergence property [4]. However, it may fail to converge globally [5, 10] or its global convergence requires very strong assumptions such as the condition $\| (B_k - J_k) d_k \| \leq \epsilon \| F_k \|$ for some very small constant $\epsilon$ used in [12, 14, 17]. Moreover, this method does not utilize the special structure when it is applied to solve the symmetric problem (1).

Li and Fukushima [7] proposed a Gauss-Newton-based BFGS method, that is,

$$d_k = -B_k^{-1} g_k, \quad \delta_k = F(x_k + y_k) - F_k,$$

(7) where

$$g_k = \frac{F(x_k + \alpha_{k-1} F_k) - F(x_k)}{\alpha_{k-1}}.$$  

(8) Under suitable conditions, using some nonmonotone line search, Li and Fukushima [7] proved that the Gauss-Newton-based BFGS method converges globally. To our knowledge, this is the first global convergence result of BFGS type methods for nonlinear equations. It follows from (2), (7) and the symmetry of Jacobian that

$$B_{k+1} s_k = \delta_k \approx J_{k+1}^T J_{k+1} s_k = J_{k+1} J_{k+1}^T s_k,$$

(9) which shows that the iterative matrix $B_{k+1}$ approximates the Gauss-Newton iterative matrix $J_{k+1} J_{k+1}^T$ along direction $s_k$. However, this possibly leads to very large condition number of $B_{k+1}$ even if the condition number of $J_{k+1}$ is not large. Thus computing $d_k = -B_k^{-1} g_k$ may be numerically unstable. Some varieties of the Gauss-Newton-based BFGS method have been proposed such as [3, 6, 20, 21], but these methods still retain the relation (9).

The object of this paper, under the same assumptions as those of the Gauss-Newton-based BFGS method described above, is to design a globally convergent BFGS type method for solving the symmetric problem (1). We will introduce a new term $\delta_k$ by slightly modifying $y_k = F_{k+1} - F_k$, which makes the iterative
matrix $B_k$ approximate the Jacobian $J_k$. Moreover, we will sufficiently utilize the approximation gradient $g_k$ and the residual $F_k$ to construct search direction and adopt the nonmonotone line search proposed by Li and Fukushima \[7\] to compute $\alpha_k$.

The paper is organized as follows. In Section 2, we present the algorithm in detail and study its global convergence property. In Section 3, we report some numerical results. Throughout the paper, $\| \cdot \|$ stands for the 2-norm.

2. Algorithm and global convergence. In this section, we present a new BFGS method for the symmetric problem (1), and give global convergence analysis for the method. We first illustrate our approach which is mainly based on the following consideration.

To suitably modify the term $\delta_k$ in the BFGS update formula (2), we adopt the technique introduced by Zhang and Tang \[19\], where they presented a BFGS method for unconstrained optimization, which is a combination of the MBFGS method and the CBFGS method proposed by Li and Fukushima \[8, 9\]. This hybrid scheme sufficiently utilizes advantages of both methods, that is, it not only reduces to the standard BFGS method for locally strongly convex functions but also regularizes nonconvex functions. We replace the terms $\nabla f(x_{k+1})$ and $\nabla f(x_k)$ of the BFGS update formula in \[19\] by $F_{k+1}$ and $F_k$ respectively. Then we obtain a new term $\delta_k$, which is given by

$$
\delta_k = \begin{cases}
y_k, & \text{if } \frac{y_k^T s_k}{s_k^T s_k} \geq \mu_1 \|F_k\|, \\
y_k + \left( \max \left\{ 0, -\frac{y_k^T s_k}{s_k^T s_k} \right\} + \mu_2 \|F_k\| \right) s_k, & \text{otherwise},
\end{cases}
$$

(10)

where $y_k$ is determined by (5), $\mu_1$ and $\mu_2$ are two given positive constants. It follows from (10) that

$$
\delta_k^T s_k \geq \min \{ \mu_1, \mu_2 \} \|F_k\| s_k^T s_k,
$$

(11)

which ensures the positive definite property of the sequence $\{B_k\}$. Moreover, the new term $\delta_k$ regularizes nonmonotone function and becomes $y_k$ for locally strongly monotone function when $\|F_k\| \to 0$. Note that $d_k = -B_k^{-1} F_k$ possesses local acceleration property and, by (8), $d_k = -B_k^{-1} y_k$ has global feature and utilizes the symmetric structure of the underlying function. Therefore, we use them to produce search direction according to some convex combination. Since the search direction may not be a descent direction of $f(x) = \frac{1}{2} \|F(x)\|^2$, we use the nonmonotone line search proposed by Li and Fukushima \[7\] to compute $\alpha_k$, that is, the stepsize $\alpha_k = \max \{ 1, r, \cdots, r^t, \cdots \}$ satisfying

$$
\|F(x_k + \alpha d_k)\|^2 \leq \|F_k\|^2 - \sigma_1 \|\alpha F_k\|^2 - \sigma_2 \|\alpha d_k\|^2 + \eta_k \|F_k\|^2,
$$

(12)

where $r \in (0, 1)$, $\sigma_1$ and $\sigma_2$ are given positive constants, and the given positive sequence $\{\eta_k\}$ satisfies

$$
\sum_{k=0}^{\infty} \eta_k \leq \eta < \infty.
$$

(13)

Based on the above discussion, we now can present the BFGS method for solving the symmetric problem (1) as follows.

**Algorithm 2.1**

**Step 1.** Select a positive sequence $\{\eta_k\}$ satisfying (13). Choose a starting point $x_0 \in \mathbb{R}^n$, a symmetric and positive definite matrix $B_0 \in \mathbb{R}^{n \times n}$, several constants $\sigma_1, \sigma_2, \mu_1, \mu_2, \alpha^{-1} > 0$, and $r, \rho \in (0, 1)$. Let $k := 0$. 

Step 2. If \( \|F_k\| = 0 \), then stop. Otherwise, go to Step 3.

Step 3. Compute \( d_k \) by solving the following linear equations
\[
B_k d = -\alpha_{k-1} F_k - (1 - \alpha_{k-1}) g_k,
\]
where \( g_k \) is determined by (8).

Step 4. If
\[
\|F(x_k + d_k)\| \leq \rho \|F_k\|,
\]
then set
\[
\alpha_k = 1,
\]
and go to Step 5. Otherwise, compute \( \alpha_k \) by the line search (12).

Step 5. Let \( x_{k+1} = x_k + \alpha_k d_k \).

Step 6. Update \( B_k \) by the BFGS formula (2) and (10).

Step 7. Let \( k := k + 1 \) and go to Step 2.

In this section we show that, under some conditions, Algorithm 2.1 possesses global convergence property. Let the level set be defined by
\[
\Omega = \{ x | \|F(x)\| \leq e^{\eta/2} \|F_0\| \},
\]
where \( \eta \) is a positive constant satisfying (13). To discuss convergence properties of Algorithm 2.1, we use the following lemma and assumptions.

Lemma 2.1. [4, Lemma 3.3] Let \( \{a_k\} \) and \( \{r_k\} \) be positive sequences satisfying
\[
a_{k+1} \leq (1 + r_k) a_k + r_k \quad \text{and} \quad \sum_{k=0}^{\infty} r_k < \infty.
\]
Then \( \{a_k\} \) converges.

Assumption 2.1
(a) \( F \) is continuously differentiable on open convex set \( \Omega_1 \) containing \( \Omega \).
(b) The Jacobian \( J(x) \) is symmetric and bounded on \( \Omega_1 \), that is, there exists a positive constant \( M \) such that
\[
\|J(x)\| \leq M, \quad \forall x \in \Omega_1.
\]
(c) \( J(x) \) is uniformly nonsingular on \( \Omega_1 \), i.e., there is a constant \( m > 0 \) such that
\[
m \|d\| \leq \|J(x)d\|, \quad \forall x \in \Omega_1, d \in \mathbb{R}^n.
\]

Assumption 2.1 implies that
\[
m \|d\| \leq \|J(x)d\| \leq M \|d\|, \quad \forall x \in \Omega_1, d \in \mathbb{R}^n,
\]
\[
\frac{1}{M} \|d\| \leq \|J(x)^{-1}d\| \leq \frac{1}{m} \|d\|, \quad \forall x \in \Omega_1, d \in \mathbb{R}^n,
\]
and
\[
m \|x - y\| \leq \|F(x) - F(y)\| \leq M \|x - y\|, \quad \forall x, y \in \Omega_1.
\]
In particular, for all \( x \in \Omega_1 \), we have
\[
m \|x - x^*\| \leq \|F(x)\| \leq M \|x - x^*\|, \quad \forall x, y \in \Omega_1,
\]
where \( x^* \) stands for the unique solution of (1) in \( \Omega_1 \). Moreover, the level set \( \Omega \) is bounded.

By (15), (12) and (13), we obtain
\[
\|F(x_{k+1})\|^2 \leq (1 + \eta_k) \|F_k\|^2 \leq e^{\eta_k} \|F_k\|^2 \leq e^{\sum_{i=0}^{k} \eta_i} \|F_0\|^2 \leq e^{\eta} \|F_0\|^2.
\]
This together with Lemma 2.1 yields the following result.

Lemma 2.2. Let the sequence \( \{x_k\} \) be generated by Algorithm 2.1, then the sequence \( \{\|F_k\|\} \) converges and \( x_k \in \Omega \) for all \( k \geq 0 \).
The following result gives a bound of \( \alpha_k \) from below, whose proof is similar to that of Lemma 3.2 in [7].

**Lemma 2.3.** Let Assumption 2.1 hold and \( \{x_k\} \) be generated by Algorithm 2.1. If \( \alpha_k \neq 1 \), then we have
\[
\alpha_k \geq \frac{2(d_k^T B_k d_k - \alpha_{k-1}d_k^T(g_k - F_k) - t_k\|d_k\|\|F_k\|)r}{\sigma_1\|F_k\|^2 + (\sigma_2 + M^2)\|d_k\|^2}, \tag{21}
\]
where
\[
t_k = \int_0^1 \|J(x_k + \tau\alpha_k^{-1}d_k) - J(x_k + \tau\alpha_{k-1}F_k)\|d\tau. \tag{22}
\]

**Proof.** If \( \alpha_k \neq 1 \), by (12), we obtain
\[
\|F(x_k + \alpha_k d_k)\|^2 > \|F_k\|^2 - \sigma_1\|\alpha_k'F_k\|^2 - \sigma_2\|\alpha_k'\|d_k\|^2 + \eta_k\|F_k\|^2
\]
where \( \alpha_k' = \alpha_k/r \). It follows from (19) that
\[
\|F(x_k + \alpha_k d_k)\|^2 - \|F_k\|^2 = 2F_k^T(F(x_k + \alpha_k' d_k) - F_k) + \|F(x_k + \alpha_k d_k) - F_k\|^2
\]
\[
\leq 2F_k^T(F(x_k + \alpha_k' d_k) - F_k) + M^2\|\alpha_k' d_k\|^2. \tag{23}
\]
By (8), we know
\[
g_k = G_k F_k, \tag{25}
\]
where
\[
G_k = \int_0^1 J(x_k + \tau\alpha_{k-1}F_k)d\tau.
\]
Moreover, from (25) and (14), we have
\[
F_k^T(F(x_k + \alpha_k' d_k) - F_k) = \alpha_k'F_k^T \int_0^1 J(x_k + \tau\alpha_{k-1}d_k)d_kd\tau
\]
\[
= \alpha_k'F_k^T \int_0^1 J(x_k + \tau\alpha_{k-1}F_k)d_kd\tau + \alpha_k'F_k^T \int_0^1 (J(x_k + \tau\alpha_k' d_k) - J(x_k + \tau\alpha_{k-1}F_k))d_kd\tau
\]
\[
\leq \alpha_k'F_k^T \int_0^1 J(x_k + \tau\alpha_{k-1}F_k)F_k d\tau + \alpha_k'F_k^T \int_0^1 J(x_k + \tau\alpha_{k-1}F_k)F_k d\tau + \alpha_k'F_k^T \int_0^1 \|F_k\|\|d_k\|d\tau
\]
\[
= \alpha_k'F_k^T g_k + \alpha_k't_k\|F_k\|\|d_k\|
\]
\[
= -\alpha_k'F_k^T B_kd_k + \alpha_k'\alpha_{k-1}d_k^T(g_k - F_k) + \alpha_k't_k\|F_k\|\|d_k\|
\]
which together with (23) and (24) yields (21). \( \square \)

The following theorem shows that Algorithm 2.1 converges globally.

**Theorem 2.4.** Let Assumption 2.1 hold and the sequence \( \{x_k\} \) be generated by Algorithm 2.1, then \( \lim_{k \to \infty} \|F_k\| = 0 \) and \( \{x_k\} \) converges to the unique solution \( x^* \) of (1).

**Proof.** Let us denote the index sets
\[
H_j = \{k \leq j\} \ (15) \text{ holds}, \quad G_j = \{0, 1, \cdots, j\}\setminus H_j, \quad j = 1, 2, \cdots.
\]
follows from Lemma 1.1 that, for \( k > 6 \) we have
\[
\|F_{k+1}\|^2 \leq \prod_{i \in G_k} (1 + \eta_i) \prod_{i \in H_k} \rho^2 \|F_0\|^2
\]
\[
= \prod_{i \in G_k} (1 + \eta_i) \rho^2 |H_k| \|F_0\|^2 \leq e^3 \rho^2 |H_k| \|F_0\|^2 \to 0, \quad \text{as } k \to \infty.
\]
This together with (20) shows that \( \{ x_k \} \) converges to the unique solution \( x^* \) of (1).

(ii) If (15) holds only for finite \( k \), in this case, we prove this theorem by contradiction. Suppose that there exists a constant \( \tau_1 > 0 \) such that
\[
\|F_k\| \geq \tau_1. \tag{26}
\]
This together with (11) yields
\[
\delta_k^T s_k \geq \tau_1 \min\{\mu_1, \mu_2\} s_k^T s_k. \tag{27}
\]
Since \( \{\|F_k\|\} \) converges, there exists a positive constant \( M_1 \) such that
\[
\|F_k\| \leq M_1. \tag{28}
\]
Moreover, by (10), (5), (19) and (28), we obtain
\[
\|\delta_k\| \leq 2M \|s_k\| + \mu_2 M_1 \|s_k\| = (2M + \mu_2 M_1) \|s_k\|. \tag{29}
\]
This and (27) yield
\[
\frac{\delta_k^T \delta_k}{\delta_k^T s_k} \leq \frac{(2M + \mu_2 M_1)^2}{\tau_1 \min\{\mu_1, \mu_2\}}. \tag{30}
\]
Inequalities (27) and (30) satisfy the conditions of Lemma 1.1. Since \( s_k = \alpha_k d_k \), it follows from Lemma 1.1 that, for \( k > 1 \), the inequalities
\[
\beta_1 \|d_j\| \leq \|B_j d_j\| \leq \beta_2 \|d_j\|, \quad \beta_3 \|d_j\|^2 \leq d_j^T B_j d_j \leq \beta_4 \|d_j\|^2, \tag{31}
\]
hold for at least \( \lfloor \rho k \rfloor \) values of \( j \in [1, k] \). Now we define the index set
\[
K = \{ j \mid (31) \text{ holds} \}. \tag{32}
\]
By (12), we obtain
\[
\sum_{k=0}^{\infty} \|s_k\|^2 = \sum_{k=0}^{\infty} \|\alpha_k d_k\|^2 < \infty, \quad \sum_{k=0}^{\infty} \|\alpha_k F_k\|^2 < \infty,
\]
which implies
\[
\lim_{k \to \infty} \alpha_k \|d_k\| = 0, \quad \lim_{k \to \infty} \alpha_k \|F_k\| = 0.
\]
This together with (26) yields
\[
\lim_{k \to \infty} \alpha_k = 0. \tag{33}
\]
By (14) and (31)-(32), we have
\[
\beta_1 \|d_k\| \leq \|B_k d_k\| = \|g_k + \alpha_{k-1} (F_k - g_k)\| \leq \beta_2 \|d_k\|, \quad \forall k \in K. \tag{34}
\]
By (25) and (17)-(18), there exist two positive constants \( C_1 \) and \( C_2 \) such that
\[
C_1 \|F_k\| \leq \|g_k\| \leq C_2 \|F_k\|. \tag{35}
\]
Inequalities (33)-(35) indicate that there exist positive constants \( C_3 \) and \( C_4 \) such that
\[
C_3 \|d_k\| \leq \|F_k\| \leq C_4 \|d_k\|, \quad \forall k \in K. \tag{36}
\]

It follows from (21)-(22), (31)-(33) and (35)-(36) that, for large $k \in K$, there exists a constant $C_5 > 0$ such that 

$$\alpha_k \geq C_5.$$ 

This leads to a contradiction with (33). The proof is then completed.

3. Numerical results. In this section, we compare the performance of the following two methods for solving nonlinear equations (1).

- GN-BFGS: the Gauss-Newton-based BFGS method in [7]. We set $B_0 = I$, $\sigma_1 = \sigma_2 = 0.01$, $r = 0.5$, $\rho = 0.95$, $\alpha_{-1} = 0.01$ and $\eta_k = \frac{1}{(k+1)^2}$.
- Algorithm 2.1: we set the same parameters as those in the GN-BFGS method and $\mu_1 = \mu_2 = 0.01$.

The codes were written in Matlab 7.4. We tested both methods on the following two test problems.

Problem 1. The discretized two-point boundary value problem [7]:

$$F(x) = \begin{pmatrix} 2 & -1 & \cdots & -1 \\ -1 & 2 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ -1 & \cdots & \cdots & 2 \end{pmatrix} x + \frac{1}{(n+1)^2} \left( \sin x_1 - 1, \cdots, \sin x_n - 1 \right)^T.$$ 

Problem 2. The gradient function of the Engval function [7, 26]:

$$F_1(x) = x_1(x_1^2 + x_2^2) - 1,$$

$$F_i(x) = x_i(x_{i-1}^2 + 2x_i^2 + x_{i+1}^2) - 1, \quad i = 2, 3, \cdots, n - 1,$$

$$F_n(x) = x_n(x_{n-1}^2 + x_n^2).$$

We stopped the iteration if $k \geq 10^3$ or $\|F_k\| \leq 10^{-6}$. We chose initial points of the form $x_0 = \beta \hat{x}$ with $\hat{x} = (1, 1/2, \cdots, 1/n)^T$. Numerical results are listed in Table 3.1 and Table 3.2, where $N_{iter}$ and $N_F$ mean the total number of iterations and function calculations, $C_{B_k}$ and $\|F_k\|$ are the condition number of $B_k$ with matrix 1-norm and the norm of $F_k$ at the stopping point respectively.

From both Tables, we can see that Algorithm 2.1 outperforms the GN-BFGS method, which requires less iterations and function calculations. Moreover, the GN-BFGS method is numerically unstable since the condition number of $B_k$ becomes very large for some cases. However, we note that the condition number of $B_k$ generated by Algorithm 2.1 is much smaller than that of the GN-BFGS method at the stopping point, which numerically supports the previous discussion.

4. Conclusions. We have presented a BFGS method for solving symmetric nonlinear equations and established its global convergence under mild conditions. The proposed method sufficiently utilizes the symmetric structure of the underlying function and produces iterative matrix whose condition number is much smaller than that of existing Gauss-Newton-based BFGS methods, which makes it more stable in numerical computation.

Acknowledgments. We would like to thank the editor and the referees whose very helpful suggestions led to much improvement of this paper.
Weijun Zhou

Table 1. Test results on Problem 1 with initial point $x_0 = \beta \hat{x}$.

| $\beta$ | $n$ | $N_{iter}$ | $N_F$ | $\|F_k\|$ | $C_{B_k}$ | $N_{iter}$ | $N_F$ | $\|F_k\|$ | $C_{B_k}$ |
|--------|-----|------------|-------|----------|----------|------------|-------|----------|----------|
| 0.01   | 9   | 22         | 63    | 1.9e-007 | 1578     | 17         | 33    | 1.27e-008| 46       |
|        | 49  | 288        | 2041  | 7.85e-007| 1117816  | 135        | 626   | 9.47e-007| 1419     |
|        | 99  | 1000       | 8307  | 3.2e-006 | 18348874 | 1000       | 8072  | 0.000741 | 8672     |
| 0.1    | 9   | 22         | 64    | 4.36e-007| 1663     | 17         | 36    | 5.04e-007| 42       |
|        | 49  | 298        | 2130  | 8.31e-007| 1137498  | 118        | 487   | 8.26e-007| 969      |
|        | 99  | 1000       | 8144  | 0.000850 | 1642370  | 1000       | 7800  | 0.000692 | 10814    |
| 1      | 9   | 23         | 67    | 4.76e-007| 1787     | 19         | 40    | 2.31e-007| 45       |
|        | 49  | 191        | 947   | 1.13e-007| 1061079  | 160        | 697   | 4.13e-007| 1255     |
|        | 99  | 1000       | 9113  | 0.000756 | 6266525  | 1000       | 7474  | 5.27e-005| 10730    |
| 10     | 9   | 24         | 69    | 7.03e-007| 1781     | 20         | 39    | 1.73e-007| 45       |
|        | 49  | 181        | 861   | 2.21e-007| 1110234  | 148        | 609   | 4.41e-007| 1292     |
|        | 99  | 1000       | 8591  | 0.000122 | 1664830  | 1000       | 719   | 5.27e-005| 5989     |
| 50     | 9   | 25         | 71    | 2.51e-007| 1713     | 20         | 39    | 4.5e-007 | 43       |
|        | 49  | 185        | 883   | 1.11e-007| 1098144  | 133        | 531   | 8.71e-007| 1217     |
|        | 99  | 823        | 6041  | 4.55e-008| 18835000 | 548        | 3361  | 8.72e-007| 4638     |

Table 2. Test results on Problem 2 with initial point $x_0 = \beta \hat{x}$.

| $\beta$ | $n$ | $N_{iter}$ | $N_F$ | $\|F_k\|$ | $C_{B_k}$ | $N_{iter}$ | $N_F$ | $\|F_k\|$ | $C_{B_k}$ |
|--------|-----|------------|-------|----------|----------|------------|-------|----------|----------|
| 0.01   | 50  | 4          | 37    | NaN      | NaN      | 60         | 235   | 8.9382e-007 | Inf      |
|        | 100 | 5          | 53    | NaN      | NaN      | 79         | 321   | 8.7896e-007 | Inf      |
|        | 200 | 5          | 53    | NaN      | NaN      | 91         | 362   | 8.363e-007   | 29       |
| 0.1    | 50  | 67         | 306   | 7.6196e-007 | 72    | 59         | 234   | 5.3211e-007 | 13       |
|        | 100 | 132        | 623   | 6.1482e-007 | 133   | 83         | 341   | 6.9712e-007 | 29       |
|        | 200 | 170        | 879   | 8.2061e-007 | 356   | 88         | 358   | 8.8805e-007 | 51       |
| 1      | 50  | 69         | 312   | 9.7115e-007 | 618   | 59         | 225   | 7.0356e-007 | 13       |
|        | 100 | 134        | 619   | 9.3146e-007 | 458   | 79         | 322   | 9.6653e-007 | 33       |
|        | 200 | 186        | 956   | 8.5496e-007 | 595   | 101        | 401   | 9.152e-007  | 43       |
| 10     | 50  | 18         | 241   | NaN      | NaN      | 65         | 265   | 7.7764e-007 | 15       |
|        | 100 | 15         | 194   | NaN      | NaN      | 88         | 385   | 9.7307e-007 | 49       |
|        | 200 | 15         | 188   | NaN      | NaN      | 111        | 461   | 7.7243e-007 | 50       |
| -0.1   | 50  | 77         | 341   | 4.4278e-007 | 59    | 60         | 236   | 9.0422e-007 | 26       |
|        | 100 | 120        | 585   | 9.5748e-007 | 136   | 73         | 305   | 9.8125e-007 | 39       |
|        | 200 | 221        | 1011  | 8.6068e-007 | 456   | 90         | 366   | 8.6948e-007 | 40       |
|        | 500 | 296        | 1684  | 9.1214e-007 | 1095  | 88         | 359   | 9.9204e-007 | 79       |

REFERENCES

[1] S. Bojari and M. R. Eslahchi, Global convergence of a family of modified BFGS methods under a modified weak Wolfe-Powell line search for nonconvex functions, 4OR, 18 (2020), 219–244.
[2] R. H. Byrd and J. Nocedal, A tool for the analysis of quasi-Newton methods with application to unconstrained minimization, SIAM J. Numer. Anal., 26 (1989), 727–739.

[3] H. Cao and D. Li, Adjoint Broyden methods for symmetric nonlinear equations, Pac. J. Optim., 13 (2017), 645–663.

[4] J. E. Dennis and J. J. Moré, A characterization of superlinear convergence and its applications to quasi-Newton methods, Math. Comput., 28 (1974), 549–560.

[5] Y.-H. Dai, Convergence properties of the BFGS algorithm, SIAM J. Optim., 13 (2002), 693–701.

[6] G. Gu, D. Li, L. Qi and S. Zhou, Descent directions of quasi-Newton method for symmetric nonlinear equations, SIAM J. Numer. Anal., 40 (2002), 1763–1774.

[7] D. 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.

[8] D. Li and M. Fukushima, A modified BFGS method and its global convergence in nonconvex minimization, J. Comput. Appl. Math., 129 (2001), 15–35.

[9] D. Li and M. Fukushima, On the global convergence of the BFGS method for nonconvex unconstrained optimization problems, SIAM J. Optim., 11 (2001), 1054–1064.

[10] W. F. Mascarenhas, The BFGS method with exact line searches fails for nonconvex objective functions, Math. Program., 99 (2004), 49–61.

[11] W. Sun and Y. Yuan, Optimization Theory and Methods, Springer Science and Business Media, LLC, New York, 2006.

[12] Z. Wang, Y. Chen, S. Huang and D. Feng, A modified nonmonotone BFGS algorithm for solving smooth nonlinear equations, Optim. Lett., 8 (2014), 1845–1860.

[13] Z. Wan, K. Teo, X. Chen and C. Hu, New BFGS method for unconstrained optimization problem based on modified Armijo line search, Optimization, 63 (2014), 285–304.

[14] G. Yuan and X. Lu, A new backtracking inexact BFGS method for symmetric nonlinear equations, Comput. Math. Appl., 55 (2008), 116–129.

[15] G. Yuan, Z. Sheng, B. Wang, W. Hu and C. Li, The global convergence of a modified BFGS method for nonconvex functions, J. Comput. Appl. Math., 327 (2018), 274–294.

[16] G. Yuan, Z. Wei and X. Lu, Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search, Appl. Math. Model., 47 (2017), 811–825.

[17] G. Yuan and S. Yao, A BFGS algorithm for solving symmetric nonlinear equations, Optimization, 62 (2013), 85–99.

[18] L. Zhang, A derivative-free conjugate residual method using secant condition for general large-scale nonlinear equations, Numer. Algo., 83 (2020), 1277–1293.

[19] L. Zhang and H. Tang, A hybrid MBFGS and CBFGS method for nonconvex minimization with a global complexity bound, Pac. J. Optim., 14 (2018), 693–702.

[20] W. Zhou, A Gauss-Newton-based BFGS method for nonconvex minimization problems, Pac. J. Optim., 9 (2013), 373–389.

[21] W. Zhou, A modified BFGS type quasi-Newton method with line search for symmetric nonlinear equations problems, J. Comput. Appl. Math., 367 (2020), 112454, 8 pp.

[22] W. Zhou and X. Chen, Global convergence of a new hybrid Gauss-Newton structured BFGS methods for nonlinear least squares problems, J. Comput. Optim. Theor. Appl., 164 (2015), 277–289.

[23] W. Zhou and D. Shen, Convergence properties of an iterative method for solving symmetric nonlinear equations, J. Optim. Theory Appl., 14 (2018), 723–737.

[24] W. Zhou and Z. Wang, An inexact PRP conjugate gradient method for symmetric nonlinear equations, Numer. Funct. Anal. Optim., 35 (2014), 370–388.

[25] W. Zhou and F. Wang, A PRP-based residual method for large-scale monotone nonlinear equations, Appl. Math. Comput., 261 (2015), 1–7.

[26] W. Zhou and L. Zhang, A modified Broyden-like quasi-Newton method for nonlinear equations, Pac. J. Optim. Theor. Appl., 12 (2020), 112744, 10 pp.