A Two-Phase Quasi-Newton Method for Optimization Problem

Suvra Kanti Chakraborty · Geetanjali Panda

Received: date / Accepted: date

Abstract In this paper, a two-phase quasi-Newton scheme is proposed for solving an unconstrained optimization problem. The global convergence property of the scheme is provided under mild assumptions. The super linear rate of the scheme is also proved in the vicinity of the solution. The advantages of the proposed scheme over the traditional scheme are justified with numerical table and graphical illustrations.

Keywords quasi-Newton scheme · Global convergence · Super linear convergence

Mathematics Subject Classification (2000) 90C53 · 90C30

1 Introduction

The basic idea behind the quasi-Newton methods for unconstrained optimization problem lies on updating the hessian approximation in a computationally cheap way, that ensures the secant condition. DFP and BFGS methods are popular quasi-Newton schemes. DFP method is sensitive to inaccurate line searches, whereas BFGS scheme is widely used for its robustness and self correcting properties. quasi-Newton methods have super linear convergence property under exact or inexact line searches. One may see ( [3], [4], [7], [9], [10], [15], [16], [17]) for the recent advances in this direction. Consider a general minimization problem

\[(P) \quad \text{Min}_{x \in \mathbb{R}^n} \ f(x),\]
where \( f : \mathbb{R}^n \to \mathbb{R} \) is a twice continuously differentiable function. Solution of (P) satisfies the system of equations \( \nabla f(x) = 0 \). Several advanced iterative processes exist so far to solve a system of nonlinear equations \( F(x), \ F : \mathbb{R}^n \to \mathbb{R}^m \) \((11), (13), (14)\). The two-phase iterative scheme for the same, proposed by \((2)\) takes the following form.

\[
x_{k+1} = x_k - \frac{1}{2} (J(x_k) + J(z_k))^{-1} F(x_k),
\]

\[
z_k = x_k - J(x_k)^{-1} F(x_k),
\]

where \( J(x_k) \) is the Jacobian of \( F(x) \), at the \( k \)-th iteration of the iterative scheme. These two-phase schemes have been further studied in unconstrained and constrained optimization problems\((5), (6)\).

In this paper, we propose a variant of quasi-Newton scheme for unconstrained optimization problem, which involves two phases and based on the basic idea of above iterative process. The new scheme guarantees the super linear global convergence behavior. In numerical illustrations, it is observed that the number of iterations of the proposed scheme are significantly less than the standard BFGS scheme in comparable execution time.

This concept is organized in following sections. Two-phase quasi-Newton scheme is proposed in Section 2, convergence analysis of the scheme is discussed in Section 3, followed by Numerical illustrations in Section 4.

### 2 Proposing two-phase quasi-Newton scheme

In a general line search method for (P), a descent sequence \( \{x_k\} \) is generated as \( x_{k+1} = x_k + \alpha_k p_k \), where \( p_k \) is the descent direction at \( x_k \) and the positive scalar \( \alpha_k \) is the step length at \( x_k \) along the direction \( p_k \). The condition \( p_k^T \nabla f_k < 0 \) guarantees that the function \( f \) can be reduced along the direction \( p_k \). The search direction may take the form \( p_k = -B_k^{-1} \nabla f_k \), where \( B_k \) is a symmetric positive definite matrix of order \( n \times n \). The most popular quasi-Newton (BFGS) scheme generates a sequence of positive definite hessian approximation \( B_k \) at the current iterate \( x_k \) as

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

where \( s_k = x_{k+1} - x_k \), \( y_k = \nabla f_{k+1} - \nabla f_k \), \( x_{k+1} = x_k - \alpha_k B_k^{-1} \nabla f_k \), \( f_k = f(x_k) \) and \( \alpha_k \) is computed from a line search procedure. The sequence \( \{x_k\} \) shows global super linear convergence behavior under some standard assumptions. The proposed modified scheme is as follows.

Consider a positive definite matrix \( B_0 \) at initial stage \( k = 0 \) and

\[
B_k = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}.
\]
A Two-Phase Quasi-Newton Method for Optimization Problem

$B_k$ is upgraded as

$$B_{k+1} = \lambda B_k + (1 - \lambda)B_k,$$

for $\lambda \in (0, 1)$,

where $s_k = x_k - x$, $y_k = \nabla f_k - \nabla f_k$, $x_k = x_k - \alpha_kB_k^{-1}\nabla f_k$. Upgrade $x_k$ as

$$x_{k+1} = x_k - \alpha_kB_k^{-1}\nabla f_k,$$  \hspace{1cm} (1)

where $\alpha_k$ and $\alpha_k$ are computed by inexact Wolfe search.

In this proposed method we modify the search direction in two phases as follows. The step length $\alpha_k$ and $\alpha_k$ are chosen by Wolfe search method.

$$x_k = x_k + \alpha_k p_k,$$  \hspace{1cm} (2)

and

$$x_{k+1} = x_k + \alpha_k p_k,$$  \hspace{1cm} (3)

If $B_k$ is positive definite then this new matrix $B_{k+1}$ is positive definite being the arithmetic mean of $B_k$ and the standard BFGS updation of it. $B_{k+1}$ is used at two consecutive stages: at the $k$th stage to determine $x_{k+1}$, as well as at $k+1$th stage to determine $x_{k+1}$. In the vicinity of the solution of $(P)$, the parameters $\alpha_k$ and $\alpha_k$ are chosen as unit length for $k$ greater than sufficiently large $k_0 \in \mathbb{N}$. This fact will be used in next section to prove the convergence of the scheme.

3 Convergence of the proposed scheme

**Theorem 1** Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable and $x_k$ is updated by (2) and (3). Assume that the sequence $\{x_k\}$ generated by (1) satisfies $\{x_k\} \to x^*$, $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) > 0$. Then the sequence $\{x_k\}$ converges to $x^*$ super linearly if

$$\lim_{k \to \infty} \frac{\|(B_k - \nabla^2 f^*)p_k\|}{\|p_k\|} = 0.$$  \hspace{1cm} (4)

**Proof** Denote $f(x^*) = f^*$. Newton direction at $x_k$ is $p_k^N = -\nabla^2 f_k^{-1}\nabla f_k$. Then from (2),

$$p_k - p_k^N = \nabla^2 f_k^{-1}(\nabla^2 f_k p_k + \nabla f_k) = \nabla^2 f_k^{-1}(\nabla^2 f_k - B_k)p_k.$$  \hspace{1cm} (5)

Since the limiting hessian $\nabla^2 f^*$ is positive definite, so $\|\nabla^2 f_k^{-1}\|$ is bounded above for $x_k$ sufficiently close to $x^*$, and hence from (2),

$$\|p_k - p_k^N\| \leq M (\|\nabla^2 f_k - B_k\|p_k\| + \|\nabla^2 f_k - \nabla^2 f^*\|\|p_k\|).$$
Dividing both side by $\|p_k\|$, letting $k \to \infty$, $x_k \to x^*$ and using (4), we have

$$\frac{\|p_k - p_k^N\|}{\|p_k\|} \to 0. \quad (6)$$

Note that for sufficiently large $k \in \mathbb{N}$, $\|p_k - p_k^N\| \to 0$, since

$$\|p_k - p_k^N\| = \|x_{k+1} - x_k^N\| \leq \|x_{k+1} - x^*\| + \|x_k^N - x^*\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \quad (7)$$

where $\epsilon$ is arbitrarily small real number. For some $N_2 \in \mathbb{N}$,

$$\|p_k\| = \|x_k - x_k^N\| \leq \|x_k - x^*\| + \|x_k^N - x^*\| \leq 2\|x_k - x^*\| \text{ for all } k > N_2,$$

which implies $\|p_k\| = O(\|x_k - x^*\|)$. Using (6) and (7), we have

$$\|x_k + p_k - x^*\| \leq \|x_k + p_k^N - x^*\| + \|p_k - p_k^N\| + \|p_k^N - p_k^N\|$$

$$\leq O(\|x_k - x^*\|^2) + o(\|p_k\|) \leq o(\|x_k - x^*\|)$$

This shows super linear convergence behavior of the sequence $\{x_k\}$.

### 3.1 Convergence analysis

By Taylor series expansion, $\nabla f_k = \nabla f_k + \int_0^1 \nabla^2 f(x + ts_k)s_k \, dt$. Denote $G_k$ as the average hessian $\int_0^1 \nabla^2 f(x + ts_k) \, dt$. Then

$$y_k = \nabla f_k - \nabla f_k = G_k s_k. \quad (8)$$

**Theorem 2** The sequence $\{x_k\}$ generated by (7) possesses global convergence property under the following assumptions.

1. For some $x_0 \in \mathbb{R}^n$, the level set $\mathcal{L} = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$ is convex and there exist positive constants $m$ and $M$ such that $m\|x\|^2 \leq x^T G_k x \leq M\|x\|^2 \quad \forall x \in \mathbb{R}^n$.

2. $(p_k - p_k^N)^T \nabla f(x_k) < 0$, where $G_k = \nabla^2 f(x_k)$.

3. Step length is found by Wolfe’s inexact line search method.

**Proof** Denote $m_k = \frac{y_k^T s_k}{s_k^T s_k}, \quad M_k = \frac{y_k^T y_k}{y_k^T s_k}$. From Assumption 1,

$$m_k = \frac{y_k^T s_k}{s_k^T s_k} = \frac{s_k^T G_k s_k}{s_k^T s_k} \geq m \text{ and } M_k = \frac{y_k^T y_k}{y_k^T s_k} = \frac{s_k^T G_k^2 s_k}{s_k^T G_k s_k} = \frac{s_k^T G_k z_k}{z_k^T z_k} \leq M. \quad (9)$$

We have

$$B_{k+1} = \lambda B_k + (1 - \lambda) B_k = B_k - (1 - \lambda) \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + (1 - \lambda) \frac{y_k y_k^T}{y_k^T s_k}. \quad (10)$$
Then trace($B_{k+1}$) = trace($B_k$) - (1 - $\lambda$)$\frac{\|B_k s_k\|^2}{s_k^T B_k s_k}$ + (1 - $\lambda$)$\frac{\|y_k\|^2}{y_k^T s_k}$. Let $\theta_k$ be the angle between $s_k$ and $B_k s_k$. Then trace($B_{k+1}$) = trace($B_k$) + (1 - $\lambda$)$\frac{s_k^T B_k s_k}{\|s_k\|^2}$. Hence taking $\psi$ a positive definite matrix. Since the eigenvalues $\lambda_i$’s of $B$ are positive so $\det(B_{k+1}) = \det(B_k)(1 - \lambda)^{\sum_{i=1}^{n} (\lambda_i - \ln(\lambda_i))} > 0$.

Define a function $\psi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ as $\psi(B) = \text{trace}(B) - \ln(\det(B))$ where $B$ is a positive definite matrix. Since the eigenvalues $\lambda_i$’s of $B$ are positive so $\psi(B) = \text{trace}(B) - \ln(\det(B)) = \sum_{i=1}^{n} (\lambda_i - \ln(\lambda_i)) > 0$.

From (11) and (12),

$$\psi(B_{k+1}) = \psi(B_k) + (1 - \lambda)M_k - \frac{q_k}{\cos^2 \theta_k} - \ln \left( \det(B_{k+1}) \right)$$

$$< \psi(B_k) + (1 - \lambda)M_k - \frac{q_k}{\cos^2 \theta_k} - \ln \left( \det(B_k) \right) - \ln(m_k) + \ln(q_k) - \ln (1 - \lambda)^3$$

$$= \psi(B_k) + \left( (1 - \lambda)M_k - \ln(m_k) - (1 - \lambda)^3 \right) + \left[ 1 - \frac{q_k}{\cos^2 \theta_k} + \ln \left( \frac{q_k}{\cos^2 \theta_k} \right) \right] - \ln(\cos^2 \theta_k).$$

(13)

Since the real valued function $h(t) = 1 - t - \ln(t)$ is always non positive for all $t > 0$, so $1 - \frac{q_k}{\cos^2 \theta_k} + \ln \left( \frac{q_k}{\cos^2 \theta_k} \right)$ is non positive. From (9)

$$0 < \psi(B_{k+1}) < \psi(B_0) + c(k + 1) + \sum_{j=0}^{k} \ln \left( \cos^2 \theta_j \right),$$

(14)
where \( c = (1 - \lambda)M_k - \ln(m_k) - \ln(1 - \lambda)^3 \) can be assumed to be positive with out loss of generality.

Assumption 1 implies that \( \nabla f \) Lipschitz continuous. Hence there exists a constant \( L_1 > 0 \) such that \( \| \nabla f(x) - \nabla f(y) \| \leq L_1 \| x - y \| \) for all \( x, y \in \text{int}(\mathcal{L}) \).

Here we employ Wolfe inexact line search to determine the step length \( \alpha_k \) at the iterating point \( x_k \) for which

\[
f(x_k + \alpha_k s_k) \leq f(x_k) + c_1 \alpha_k \nabla f^T s_k,
\]

\[
\nabla f(x_k + \alpha_k s_k)^T s_k \geq c_2 \nabla f^T s_k,
\]

with \( 0 < c_1 < c_2 < 1 \). From second inequality

\[
(\nabla f_k - \nabla f_k)^T s_k \geq (c_2 - 1) \nabla f_k^T s_k,
\]

while the Lipschitz continuity of \( \nabla f \) implies that

\[
(\nabla f_k - \nabla f_k)^T s_k \leq \| \nabla f_k - \nabla f_k \| \| s_k \|
\]

\[
\leq L_1 \| \alpha_k s_k \| \| s_k \| = \alpha_k L_1 \| s_k \|^2.
\]

Combining (15) and (16), we obtain \( \alpha_k \geq \frac{c_2 - 1}{L_1} \frac{\| s_k \|^2}{\| \nabla f_k \|^2} \). Substituting this value of \( \alpha_k \) in the first inequality of Wolfe’s condition we can get

\[
f_k \leq f_k - c_1 \frac{1 - c_2}{L_1} \frac{(\nabla f_k^T s_k)^2}{\| s_k \|^2} \leq f_k - \xi \cos^2 \theta_k \| \nabla f_k \|^2,
\]

where \( \xi = c_1 \frac{1 - c_2}{L_1} \). For each \( k \), Assumption 2 implies \( f_{k+1} \leq f_k \), which gives \( f_{k+1} \leq f_k - \xi \cos^2 \theta_k \| \nabla f_k \|^2 \).

Repeated application of this inequality provides \( f_{k+1} \leq f_0 - \xi \sum_{j=0}^{k} \cos^2 \theta_j \| \nabla f_j \|^2 \).

That is,

\[
\sum_{j=0}^{k} \cos^2 \theta_j \| \nabla f_j \|^2 \leq \frac{1}{\xi} (f_0 - f_{k+1}).
\]

Since \( f \) is bounded below, \( f_0 - f_{k+1} \) is less than some positive constant, for all \( k \). Taking the limit \( k \to \infty \) we can get \( \sum_{k=0}^{\infty} \cos^2 \theta_k \| \nabla f \|^2 < \infty \).

Hence \( \cos^2 \theta_k \| \nabla f \|^2 \to 0 \). Then \( \| \nabla f \| \to 0 \) if \( \cos \theta_k > \frac{1}{M'} \), for some positive number \( M' \). Then \( \cos \theta_j \to 0 \) is not true. Otherwise if \( \cos \theta_j \to 0 \) then there exists a natural number \( k_2 > 0 \) such that for all \( j > k_2 \) we have \( \ln(\cos^2 \theta_j) < -2c \), where \( c \) is the constant defined in (14). Then from (14) we get

\[
0 < \psi(B_0) + c(k + 1) + \sum_{j=0}^{k_1} \ln(\cos^2 \theta_j) + \sum_{j=k_1+1}^{k} ( -2c).
\]

R.H.S. of this inequality is negative for large \( k \), which is a contradiction. Therefore there exists a subsequence of indices \( \{j_k\}_{k=1,2,...} \) such that \( \cos \theta_{j_k} \geq \delta > 0 \). Hence \( \lim \| \nabla f_k \| \to 0 \). Since the problem is strongly convex, it proves that \( x_k \to x^* \).
Theorem 3 Suppose the hessian $G(x)$ is Lipschitz continuous. Then the sequence $\{x_k\}$, generated by (I) converges to minimizer of $(P)$ at super linear order.

Proof Suppose the sequence $\{x_k\}$, generated by (I) converges to $x^*$, which is a solution of $P$. Denote

$$G_* = G(x^*), \quad s_k = G_*^{\frac{1}{2}} s_k, \quad \tilde{y}_k = G_*^{\frac{-1}{2}} y_k, \quad \tilde{B}_k = G_*^{\frac{-1}{2}} B_k G_*^{\frac{1}{2}},$$

If $\tilde{\theta}_k$ is the angle between $\tilde{s}_k$ and $\tilde{B}_k \tilde{s}_k$ then $\cos \tilde{\theta}_k = \frac{\tilde{s}_k \tilde{B}_k \tilde{s}_k}{\|\tilde{s}_k\| \|\tilde{B}_k \tilde{s}_k\|}$. Denote

$$\tilde{q}_k = \frac{\tilde{s}_k \tilde{B}_k \tilde{s}_k}{2 \|\tilde{s}_k\|^2}, \quad \tilde{M}_k = \frac{\|\tilde{y}_k\|^2}{\tilde{y}_k^T \tilde{s}_k} \text{ and } \tilde{m}_k = \frac{\tilde{y}_k^T \tilde{s}_k}{\|\tilde{s}_k\|^2}.$$

By pre and post multiplying two-phase quasi-Newton update (10) by $G_*^{-\frac{1}{2}}$ and grouping we get $\tilde{B}_{k+1} = \tilde{B}_k - (1 - \lambda) \frac{\tilde{B}_k \tilde{s}_k \tilde{B}_k^T}{\tilde{y}_k^T \tilde{s}_k} \tilde{s}_k + (1 - \lambda) \frac{\tilde{q}_k \tilde{q}_k^T}{\tilde{y}_k^T \tilde{s}_k}$. Following the similar logic of (13) from Subsection (3.1), we get

$$\psi(\tilde{B}_{k+1}) < \psi(\tilde{B}_k) + \left( (1 - \lambda) \tilde{M}_k \ln(\tilde{m}_k) - \ln(1 - \lambda)^3 \right) + \left[ 1 - \frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} + \ln \left( \frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} \right) \right].$$

From (5.3), $y_k = G_k s_k$, so $y_k - G_* s_k = (G_k - G_*) s_k$ and $\tilde{y}_k - \tilde{s}_k = G_*^{-\frac{1}{2}} (G_k - G_*) G_*^{\frac{1}{2}} s_k$. Using the Lipschitz continuity property of hessian matrix, we have

$$\|\tilde{y}_k - \tilde{s}_k\| \leq \|G_*^{-\frac{1}{2}}\| \|\tilde{s}_k\| \|G_k - G_*\| \leq \|G_*^{-\frac{1}{2}}\| \|\tilde{s}_k\| L \epsilon_k,$$

where $\epsilon_k = \epsilon_k = \max\{\|x_k - x^*\|, \|x_k - x^*\|\}$ and $L$ is Lipschitz constant. From above inequality,

$$\frac{\|\tilde{y}_k - \tilde{s}_k\|}{\|\tilde{s}_k\|} \leq \bar{c} \epsilon_k, \quad \text{where } \bar{c} = \|G_*^{-\frac{1}{2}}\| L \text{ is a real constant}. \quad (17)$$

Replacing $s_k$ by $p_k$ in Theorem 6.6 (Chapter 6 [12]) and using (17) we conclude that $\lim_{k \to \infty} \frac{\|B_k - G_k\| p_k}{\|p_k\|} = 0$. This is the condition for super linear convergence of the sequence in Theorem 1 with unit step length $\alpha_k = 1$ in the vicinity of the solution. Hence the result.

3.2 Algorithm

In the previous section a two-phase quasi-Newton sequence (I) is developed, $B_k$ is updated at each iteration using intermediate update $B_k$. Since solving matrix system has expensive computational cost, so we frame the algorithm for the proposed scheme such that it includes matrix-vector multiplication
Denote the inverse $B_k$ at $x_k$ as $H_k$. As per the proposed scheme $B_{k+1} = \lambda B_k + (1 - \lambda) B_k^{-1}$, which implies,
\[ H_{k+1} = B_{k+1}^{-1} = \left( \lambda B_k + (1 - \lambda) B_k^{-1} \right)^{-1} = \left( \lambda H_k^{-1} + (1 - \lambda) H_k^{-1} \right)^{-1}. \] (18)

To frame the algorithm we use the standard inverse Hessian update as follows.
\[ H_k = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T, \quad \text{where} \quad \rho_k = \frac{1}{y_k s_k}. \] (19)

Algorithm 1: Two Phase quasi-Newton Algorithm

\begin{algorithm}
\begin{algorithmic}
\State Input: $x_0$, $\epsilon > 0$
\State $H_0 \leftarrow I$
\State $k \leftarrow 0$
\While {$(||\nabla f_k|| > \epsilon)$}
\State $p_k \leftarrow -H_k \nabla f_k$
\State $x_k \leftarrow x_k + \alpha_k p_k$; $\alpha_k$ is the step-length by Wolfe condition.
\State Compute $H_k$ by (19)
\State Compute $H_{k+1}$ by (18)
\State $p_k \leftarrow -H_{k+1} \nabla f_k$
\State $x_{k+1} \leftarrow x_k + \alpha_k p_k$; $\alpha_k$ is the step-length by Wolfe condition.
\State $k \leftarrow k + 1$
\EndWhile
\State Output: $x_{k+1}$
\end{algorithmic}
\end{algorithm}

4 Numerical results and discussions

We fix the set up for numerical computation on MATLAB R2015a platform with 4 GB RAM in a 64 bit computer. We fix $\lambda = 0.5$ for proposed two-phase quasi-Newton scheme. The tolerance limit is taken to be $10^{-6}$. The Armijo parameter, Wolfe parameter, backtracking factors are taken to be $10^{-4}$, 0.9 and 0.5 respectively. We choose 30 test functions of 10 dimensions from [1] and run the traditional BFGS and the proposed algorithm (Two-phase QN) in this platform with the initial points specified in [1]. For each function, both the algorithms are executed for 5 runs. Note the number of iterations and average execution time in Table 1. It presents that two-phase quasi-Newton scheme takes less number of iterations to reach the solution point whereas execution time remains almost same.

Furthermore, we provide a graphical illustration (Figure 1) of performance profile for number of iteration and average execution time for both of the algorithms on this test set. The performance ratio defined by Dolan and More [8] is $\rho_{(p,s)} = \min_{1 \leq r \leq n_s} \frac{r_{(p,s)}}{r_{(p,s)}^{(p,s)}}$, where $r_{(p,s)}$ refers to the iteration (or, execution time) for solver $s$ spent on problem $p$ and $n_s$ refers to the number of
problems in the model test set. In order to obtain an overall assessment of a solver on the given model test set, the cumulative distribution function $P_s(\tau)$ is $P_s(\tau) = \frac{1}{n_p} \text{size}\{p \in \mathcal{T} : \rho(p,s) \leq \tau\}$, where $P_s(\tau)$ is the probability that a performance ratio $\rho(p,s)$ is within a factor of $\tau$ of the best possible ratio.

Table 1: Comparison between BFGS and Two-phase QN

| Sl. | Function                          | BFGS   | Two-Phase QN |
|-----|-----------------------------------|--------|-------------|
| 1   | Almost Perturbed Quadratic        | 18     | 15          |
| 2   | ARWHEAD                           | 7      | 7           |
| 3   | BIGGSB1                           | 12     | 11          |
| 4   | Diagonal 1                        | 17     | 13          |
| 5   | Diagonal 2                        | 23     | 22          |
| 6   | Diagonal 3                        | 19     | 14          |
| 7   | Diagonal 7                        | 5      | 7           |
| 8   | Diagonal 9                        | 22     | 14          |
| 9   | DIXMANNIA DIXMAANL                | 12     | 11          |
| 10  | DQDRTIC                           | 13     | 20          |
| 11  | EDENSCH                           | 23     | 18          |
| 12  | ENGVAL1                           | 30     | 25          |
| 13  | Extended Beale                    | 22     | 20          |
| 14  | Extended DENSCHNB                 | 7      | 7           |
| 15  | Extended Freudenstein and Roth    | 10     | 9           |
| 16  | Extended PSC1                     | 13     | 12          |
| 17  | Extended Tridiagonal 1            | 21     | 20          |
| 18  | Extended Tridiagonal 2            | 11     | 9           |
| 19  | Fletcher                          | 28     | 25          |
| 20  | Generalized PSC1                  | 23     | 14          |
| 21  | Hager                             | 17     | 8           |
| 22  | HIMMELH                           | 7      | 6           |
| 23  | Partial Perturbed Quadratic       | 16     | 16          |
| 24  | Perturbed Quadratic Diagonal      | 10     | 5           |
| 25  | Perturbed Tridiagonal Quadratic   | 18     | 14          |
| 26  | Quadratic QF1                     | 16     | 11          |
| 27  | Quadratic QF2                     | 23     | 17          |
| 28  | Raydan1                           | 18     | 16          |
| 29  | Raydan2                           | 7      | 5           |
| 30  | Tridia                            | 15     | 16          |
5 Conclusion

In this paper we have proposed a two-phase quasi-Newton scheme for optimization problem, which shows super linear convergence property under some suitable assumptions. The algorithm is executed for some test functions and the results show the advantage of the proposed scheme in terms of iterations in comparable time over the traditional schemes, which is provided in the performance profiles.

References

1. Andrei, N.: An unconstrained optimization test functions collection. Adv. Model. Optim 10(1), 147–161 (2008)
2. Babajee, D., Dauhoo, M.: An analysis of the properties of the variants of Newton’s method with third order convergence. Applied Mathematics and Computation 183(1), 659–684 (2006)
3. Biglari, F., Ebadian, A.: Limited memory BFGS method based on a high-order tensor model. Computational Optimization and Applications 60(2), 413–422 (2015)
4. Byrd, R.H., Hansen, S.L., Nocedal, J., Singer, Y.: A stochastic quasi-Newton method for large-scale optimization. SIAM Journal on Optimization 26(2), 1008–1031 (2016)
5. Chakraborty, S.K., Panda, G.: A higher order iterative algorithm for multivariate optimization problem. Journal of applied mathematics & informatics 32(5–6), 747–760 (2014)
6. Chakraborty, S.K., Panda, G.: Two-phase-SQP method with higher-order convergence property. Journal of the Operations Research Society of China 4(5), 385–396 (2016)
7. Chen, Z., Cheng, W., Li, X.: A global convergent quasi-Newton method for systems of monotone equations. Journal of Applied Mathematics and Computing 44(1-2), 455–465 (2014)
8. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles. Mathematical programming 91(2), 201–213 (2002)
9. Jensen, T.L., Diehl, M.: An approach for analyzing the global rate of convergence of quasi-Newton and truncated-Newton methods. Journal of Optimization Theory and Applications pp. 1–16 (2016)
10. Jie, S., Li-Ping, P.: A quasi-Newton bundle method based on approximate subgradients. Journal of Applied Mathematics and Computing 29(1-2), 361–367 (2007)
11. Lukšan, L., Vlček, J.: New quasi-Newton method for solving systems of nonlinear equations. Applications of Mathematics pp. 1–14 (2017)
12. Nocedal, J., Wright, S.J.: Numerical Optimization (2nd Ed.). Springer Series in Operations Research and Financial Engineering. Springer, NY (2006)
13. Sharma, J.R., Guha, R.K., Sharma, R.: An efficient fourth order weighted-newton method for systems of nonlinear equations. Numerical Algorithms 62(2), 307–323 (2013)
14. Soleymani, F., Lotfi, T., Bakhtiari, P.: A multi-step class of iterative methods for nonlinear systems. Optimization Letters 8(3), 1001–1015 (2014)
15. Tapia, R.: On averaging and representation properties of the BFGS and related secant updates. Mathematical Programming 153(2), 363–380 (2015)
16. Xiao, Y.h., Li, T.f., Wei, Z.x.: Global convergence of a modified limited memory BFGS method for non-convex minimization. Acta Mathematicae Applicatae Sinica, English Series 29(3), 555–566 (2013)
17. Yuan, G., Yao, S.: A BFGS algorithm for solving symmetric nonlinear equations. Optimization 62(1), 85–99 (2013)