## SCIENCE \& TECHNOLOGY

PERTANIKA

# Preconditioned Subspace Quasi-Newton Method for Large Scale Optimization 

Hong Seng Sim ${ }^{1 *}$, Wah June Leong ${ }^{2}$, Malik Abu Hassan ${ }^{2}$ and Fudziah Ismail ${ }^{2}$<br>${ }^{1}$ Institute for Mathematical Research, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia<br>${ }^{2}$ Department of Mathematics, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia


#### Abstract

Subspace quasi-Newton (SQN) method has been widely used in large scale unconstrained optimization problem. Its popularity is due to the fact that the method can construct subproblems in low dimensions so that storage requirement as well as the computation cost can be minimized. However, the main drawback of the SQN method is that it can be very slow on certain types of non-linear problem such as ill-conditioned problems. Hence, we proposed a preconditioned SQN method, which is generally more effective than the SQN method. In order to achieve this, we proposed that a diagonal updating matrix that was derived based on the weak secant relation be used instead of the identity matrix to approximate the initial inverse Hessian. Our numerical results show that the proposed preconditioned SQN method performs better than the SQN method which is without preconditioning.


Keywords: Preconditioned, subspace method, limited memory quasi-Newton methods, large scale, unconstrained optimization

## INTRODUCTION

Subspace quasi-Newton (SQN) method is generally used to solve large scale non-linear systems of equations and non-linear least square problems. This method is popular

[^0]because it has the characteristic to force the next iteration in a low dimensional subspace. At each iteration, we searched for a minimum of the objective function over a subspace spanned by the current gradient and by direction of few previous steps.

The main advantage of this method is that it constructs subproblems in low dimensions so that computation cost can be reduced. It also offers a possible way to handle large scale unconstrained optimization problems. Besides, this method can be implemented extremely fast. This happens when the
objective function is a combination of expensive linear mappings with computationally cheap nonlinear functions (Yuan, 2007).

One of the famous subspace algorithms for non-linear optimization is the unbalance property shared by most line search algorithms. Any line search method is considered to have the following form:

$$
\begin{equation*}
x_{k+1}=x_{k}+\alpha_{k} d_{k}, \tag{1}
\end{equation*}
$$

Where, $d_{k}$ is the search direction and $\alpha_{k} \geq 0$ is the step-length that is computed by certain line search technique. Generally, the search direction $d_{k}$ is computed by solving a subproblem which is an approximation to the original non-linear optimization problem. Therefore, there are two parts combined in each iteration of a line search algorithms; the first part is to find $d_{k}$ in the whole $n$ dimensional space, while the other part is to search for a suitable step-length in a fixed one dimensional space spanned by the computed $d_{k}$. As a result, the overall algorithm swings between the $n$ dimensional search and one dimensional search alternately. Some variants of these methods can be found in [3], [4], [5], [9], [10], [11], [12], [17] and [18].

Furthermore, many well-known existing algorithms essentially have certain subspace features. For example, the conjugate gradient method uses a search direction in a two dimensional subspace spanned by the steepest descent direction and the previous step, the dogleg method computes a step that is a convex combination of the steepest descent direction and the Newton's direction, and the limited memory quasi-Newton algorithms will also produce search directions that are spanned in a lower dimensional space to speed up the convergency and lower the computation cost.

## SUBSPACE METHOD APPROACHES

The well-known nonlinear conjugate gradient methods use a linear combination of the steepest descent direction $-g_{k}$ and the previous search direction $d_{k-1}$ to form the new search direction, as follows:

$$
d_{k}=-g_{k}+\beta_{k} d_{k-1}
$$

Hence, one of the important tasks is how to determine the suitable $\beta_{k}$ based on certain conjugate gradient principles. Instead of the conjugate property, Stoer and Yuan (1995) suggested to look at the conjugate gradient method from the subspace point of view. In the conjugate gradient method, $\beta_{k}$ is used to define search direction $d_{k}$ and the stepsize $\alpha_{k}$ to set $x_{k+1}=x_{k}+\alpha_{k} d_{k}$; thus, no matter whatever $\beta_{k}$ and $\alpha_{k}$ are used, the increment in the iterative point will be a linear combination of $-g_{k}$ and $d_{k-1}$. They consider a model subproblem as follows:

$$
\min _{d \in \operatorname{span}\left\{-g_{k}, s_{k-1}\right\}} Q_{k}(d) \approx f\left(x_{k}+d\right)
$$

Let $d_{k}$ be the solution of the above 2 -dimensional subproblem and a successive 2-dimensional search algorithm is presented, which is an example of algorithms that using subspace methods (Stoer \& Yuan, 1995).

Limited memory quasi-Newton method also has the subspace nature. The Quasi-Newton updates have the following form:

$$
B_{k}=U\left(B_{k-1}, s_{k-1}, y_{k-1}\right)
$$

which satisfies

$$
B_{k} s_{k-1}=y_{k-1}
$$

where, $s_{k-1}=x_{k}-x_{k-1}$ and $y_{k-1}=g_{k}-g_{k-1}$. A famous example is the BFGS method.

$$
B_{k}=B_{k-1}-\frac{B_{k-1} s_{k-1} s_{k-1}^{T} B_{k-1}}{s_{k-1}^{T} B_{k-1} s_{k-1}}+\frac{y_{k-1} y_{k-1}^{T}}{s_{k-1}^{T} y_{k-1}}
$$

The limited memory quasi-Newton updates the approximate Hessian repeatedly:

$$
B_{k}^{(i)}=U\left(B_{k}^{(i-1)}, s_{k-m-1+i}, y_{k-m-1+i}\right) \quad i=1,2, \ldots, m
$$

with $B_{k}^{(0)}=\sigma_{k} I$ (Liu \& Nocedal, 1989). There are various formulae for $\sigma_{k}$, with one choice being $\frac{s_{k-1}^{T} y_{k-1}}{y_{k-1}^{T} y_{k-1}}$. The limited memory quasi-Newton matrix can be written as follows:

$$
B_{k}=B_{k}^{(m)}=\sigma_{k} I+\left[\begin{array}{ll}
S_{k} & Y_{k}
\end{array}\right] T_{k}\left[\begin{array}{c}
S_{k}^{T} \\
Y_{k}^{T}
\end{array}\right]
$$

where, $T_{k}$ is a $2 m \times 2 m$ matrix, and

$$
\left[\begin{array}{ll}
S_{k} & Y_{k}
\end{array}\right]=\left[s_{k-1}, s_{k-2}, \ldots, s_{k-m}, y_{k-1}, y_{k-2}, \ldots, y_{k-m}\right] \in \mathfrak{R}^{n \times 2 m}
$$

We have $s_{k}=\alpha_{k} d_{k}=-\alpha_{k} B_{k}^{-1} g_{k}$ in the line search type method, while for a trust region type algorithm, we have $s_{k}=-\left(B_{k}+\lambda_{k} I\right)^{-1} g_{k}$. As a result,

$$
\begin{aligned}
& s_{k}=-\left(\rho_{k} I+\left[\begin{array}{ll}
S_{k} & Y_{k}
\end{array}\right] T_{k}\left[\begin{array}{c}
S_{k}^{T} \\
Y_{k}^{T}
\end{array}\right]\right)^{-1} g_{k} \\
& \in \operatorname{span}\left\{g_{k}, s_{k-1}, \ldots, s_{k-m}, y_{k-1}, \ldots, y_{k-m}\right\} .
\end{aligned}
$$

It has been shown that no matter what, the limited memory quasi-Newton algorithm with line search or trust region will always produce a step in the subspace $\operatorname{span}\left\{g_{k}, s_{k-1}, \ldots, s_{k-m}, y_{k-1}, \ldots, y_{k-m}\right\}$ (Wang et al., 2004).

A model subspace algorithm for unconstrained optimization is suggested, which is a slight modification of the standard trust region algorithm for unconstrained optimization.

## Algorithm 2.1 (A Model Subspace Algorithm for Unconstrained Optimization)

Step 1: Given $x_{1}$, define $S_{1}, \varepsilon>0, k:=1$.
Step 2: Solve a subspace subproblem:

$$
\begin{equation*}
\min _{d \in S_{k}} Q_{k}(d)=g_{k}^{T} d+\frac{1}{2} d^{T} B_{k} d \tag{2}
\end{equation*}
$$

to obtain $s_{k}$. If $\| s_{k} \mid \leq \varepsilon$, then stop.
Step 3: Define

$$
\begin{array}{ll}
x_{k+1}=x_{k}+s_{k} & \text { if } f\left(x_{k}+s_{k}\right)<f\left(x_{k}\right) \\
x_{k+1}=x_{k}, \text { otherwise. }
\end{array}
$$

Step 4: Generate $S_{k+1}$ and $Q_{k+1}(d)$.
Step 5: Set $k:=k+1$, go to step 2 .

The main difference between the above algorithm and the standard whole space algorithm is the constraint for the step $S_{k}$ to be in the subspace $S_{k}$. Thus, the key issue here is how to choose the subspace $S_{k}$. Stoer and Yuan (1995) suggested that the choice for the subspace $S_{k}$ is a generalization of the 2-dimensional subspace, namely, $S_{k}=\operatorname{span}\left\{-g_{k}, s_{k-1}, \ldots, s_{k-m}\right\}$, since all the points in $S_{k}$ can be expressed by:

$$
\begin{equation*}
d=-\sigma g_{k}+\sum_{i=1}^{m} \beta_{i} s_{k-i} \tag{3}
\end{equation*}
$$

using the following approximations,

$$
s_{k-i}^{T} \nabla^{2} f\left(x_{k}\right) s_{k-j} \approx s_{k-i}^{T} y_{k-j}, \quad s_{k-i}^{T} \nabla^{2} f\left(x_{k}\right) g_{k} \approx y_{k-i}^{T} g_{k}
$$

However, the performance of a CG-like search direction can be very slow on certain types of non-linear problem such as ill-conditioned problems. Hence, the main aim of the study is to propose some preconditioners for the search direction (3), namely:

$$
\begin{equation*}
d_{k}=-D_{k}^{-1} g_{k}+\sum_{i=1}^{m} \beta_{i} s_{k-i} \tag{4}
\end{equation*}
$$

where $D_{k}$ is the preconditioner in diagonal matrix form, and it is supposed to have some properties of the Hessian matrix, or a good approximation to the Hessian matrix in some sense.

## DERIVATION OF THE DIAGONAL PRECONDITIONER

In this section, we develop a preconditioner for subspace quasi-Newton algorithm in order to overcome the deficiency of the standard subspace algorithm when solving ill-conditioned optimization problems.

We shall choose a diagonal matrix $D_{k}$ that satisfies the weak-quasi-Newton relation, as below:

$$
\begin{equation*}
y_{k}^{T} D_{k+1} s_{k}=y_{k}^{T} y_{k} \tag{5}
\end{equation*}
$$

where, $y_{k}=g_{k+1}-g_{k}$, and $s_{k}=x_{k+1}-x_{k}$.
Suppose that the Hessian matrix $A$ of an objective function $f(x)=\frac{1}{2} x^{T} A x-b^{T} x$ is positive definite. We let $D_{k}$ be a diagonal matrix to approximate the Hessian matrix. Hence, we form our approximation as follows:

$$
\begin{equation*}
D_{k+1}=D_{k}+\Delta_{k} \tag{6}
\end{equation*}
$$

Our purpose is to construct a $D_{k+1}$ in such a way that it is a good approximation to the actual Hessian matrix.

## Theorem 3.1

Assume that $D_{k}>0$ is a positive definite diagonal matrix and $D_{k+1}$ is the updated version of $D_{k}$, which is also diagonal. Suppose that $s_{k} \neq 0$, the optimal solution of the following minimization problem will then be:

$$
\begin{align*}
& \operatorname{minimize} \frac{1}{2}\left\|\Delta_{k}\right\|_{F}^{2} \\
& \text { subject to } y_{k}^{T} D_{k+1} s_{k}=y_{k}^{T} y_{k} \tag{7}
\end{align*}
$$

and is given by:

$$
\begin{equation*}
D_{k+1}=D_{k}+\frac{\omega_{k}-\mu_{k}}{\gamma_{k}} G_{k} \tag{8}
\end{equation*}
$$

where, $\left\|\Delta_{k}\right\|_{F}=\sqrt{\operatorname{tr}\left(\Delta_{k}^{T} \Delta_{k}\right)}$ is the Frobenius norm and $\operatorname{tr}$ is the trace operator, $\omega=y_{k}^{T} y_{k}$, $\mu=y_{k}^{T} D_{k} s_{k}, \gamma=\sum_{i=1}^{n}\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}$ and $G_{k}=\operatorname{diag}\left(\left(s_{k}^{(1)} y_{k}^{(1)}\right), \ldots,\left(s_{k}^{(i)} y_{k}^{(i)}\right)\right)$ with $y_{k}^{(i)}$ and $s_{k}^{(i)}$ being the $i=t h$ component of the $y_{k}$ and $s_{k}$ respectively.

## Proof

Let $\Delta_{k}=\left(\begin{array}{ccc}a_{k}^{(1)} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & a_{k}^{(k)}\end{array}\right), s_{k}=\left(\begin{array}{c}s_{k}^{(1)} \\ \vdots \\ s_{k}^{(n)}\end{array}\right)$ and $y_{k}=\left(\begin{array}{c}y_{k}^{(1)} \\ \vdots \\ y_{k}^{(k)}\end{array}\right)$ for $i=1,2,3, \ldots, n$.
From equation (7), we have:

$$
\begin{align*}
& \left\|\Delta_{k}\right\|^{2}=\left(\sqrt{\operatorname{tr}\left(\Delta_{k}\right)^{T}\left(\Delta_{k}\right)}\right)^{2} \\
& =\left(\left(a_{k}^{(1)}\right)^{2}+\ldots+\left(a_{k}^{(i)}\right)^{2}+\ldots+\left(a_{k}^{(n)}\right)^{2}\right) . \tag{9}
\end{align*}
$$

Thus, the minimization equation will become:
$\operatorname{minimize} \frac{1}{2}\left(\left(a_{k}^{(1)}\right)^{2}+\ldots+\left(a_{k}^{(i)}\right)^{2}+\ldots+\left(a_{k}^{(n)}\right)^{2}\right)$.
By substituting (6) into (7), we obtain:

$$
\begin{equation*}
y_{k}^{T}\left(D_{k}+\Delta_{k}\right) s_{k}=y_{k}^{T} y_{k} \tag{11}
\end{equation*}
$$

We expand (11) to get the following expression:

$$
y_{k}^{T} D_{k} s_{k}+y_{k}^{T} \Delta_{k} s_{k}=y_{k}^{T} y_{k}
$$

Rearranging the equation, we get:

$$
\begin{equation*}
\mu-\omega+\sum_{i=1}^{n} y_{k}^{(i)} s_{k}^{(i)} a_{k}^{(i)}=0 \tag{12}
\end{equation*}
$$

where $\mu=y_{k}^{T} D_{k} s_{k}$ and $\omega=y_{k}^{T} y_{k}$.
From (12), we have:

$$
\begin{equation*}
\sum_{i=1}^{n} y_{k}^{(i)} s_{k}^{(i)} a_{k}^{(i)}=\omega-\mu \tag{13}
\end{equation*}
$$

Finally, we wish to solve the following:

$$
\begin{align*}
& \operatorname{minimize} \frac{1}{2}\left(\left(a_{k}^{(1)}\right)^{2}+\ldots+\left(a_{k}^{(i)}\right)^{2}+\ldots+\left(a_{k}^{(n)}\right)^{2}\right) \\
& \text { subject to } \mu-\omega+\sum_{i=1}^{n} y_{k}^{(i)} s_{k}^{(i)} a_{k}^{(i)}=0 \tag{14}
\end{align*}
$$

Since the objective function in (14) is convex, there exists a unique solution and its Lagrange function will be:

$$
\begin{equation*}
L=\frac{1}{2}\left(\left(a_{k}^{(1)}\right)^{2}+\ldots+\left(a_{k}^{(n)}\right)^{2}\right)+\lambda\left(\mu-\omega+\sum_{i=1}^{n} y_{k}^{(i)} s_{k}^{(i)} a_{k}^{(i)}\right) \tag{15}
\end{equation*}
$$

where, $\lambda$ is the Lagrange multiplier associated with the constant. We differentiate (15) with respect to $a_{k}^{(i)}$, and setting the result to zero, we obtain,

$$
\begin{equation*}
\frac{\partial L}{\partial a_{k}^{(i)}}=a_{k}^{(i)}+\lambda y_{k}^{(i)} s_{k}^{(i)}=0 \tag{16}
\end{equation*}
$$

From (16), it is clear that,

$$
\begin{equation*}
\lambda y_{k}^{(i)} s_{k}^{(i)}=-a_{k}^{(i)} \tag{17}
\end{equation*}
$$

Multiplying (17) with $s_{k}^{(i)} y_{k}^{(i)}$ for $i=1,2,3, \ldots, n$, respectively, we shall obtain

$$
\begin{equation*}
\lambda\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}=-y_{k}^{(i)} s_{k}^{(i)} a_{k}^{(i)} \tag{18}
\end{equation*}
$$

Summing all of the equations in (18) yields:

$$
\begin{equation*}
\lambda \sum_{i=1}^{n}\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}=-\sum_{i=1}^{n} y_{k}^{(i)} s_{k}^{(i)} a_{k}^{(i)} \tag{19}
\end{equation*}
$$

By equation (13), (19) becomes

$$
\begin{equation*}
\lambda \sum_{i=1}^{n}\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}=\mu-\omega \tag{20}
\end{equation*}
$$

Finally, we get

$$
\begin{equation*}
\lambda=\frac{\mu-\omega}{\gamma} \tag{21}
\end{equation*}
$$

where $\gamma=\sum_{i=1}^{n}\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}$.
Once again, from (17), we get

$$
\begin{equation*}
a_{k}^{(i)}=-\lambda y_{k}^{(i)} s_{k}^{(i)} \tag{22}
\end{equation*}
$$

We substitute (21) into (22), the equation becomes,

$$
\begin{equation*}
a_{k}^{(i)}=\frac{\omega_{k}-\mu_{k}}{\gamma_{k}} y_{k}^{(i)} s_{k}^{(i)} \tag{23}
\end{equation*}
$$

Expression (23) is in the form of each component of $\Delta$. By substituting (23) into the formula of $\Delta_{k}$, we will get the approximation of $D_{k+1}$ as follows:

$$
\begin{equation*}
D_{k+1}=D_{k}+\frac{\omega_{k}-\mu_{k}}{\gamma_{k}} G_{k} \tag{24}
\end{equation*}
$$

where $\omega=y_{k}^{T} y_{k}, \mu=y_{k}^{T} D_{k} s_{k}, \gamma=\sum_{i=1}^{n}\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}$ and $G_{k}=\operatorname{diag}\left(\left(s_{k}^{(1)} y_{k}^{(1)}\right), \ldots,\left(s_{k}^{(i)} y_{k}^{(i)}\right)\right)$ with $y_{k}^{(i)}$ and $s_{k}^{(i)}$ being the $i$-th component of the $y_{k}$ and $s_{k}$ respectively, and the proof is completed.

Now, we give our algorithm for solving large-scale unconstrained optimization, which is called the preconditioned subspace quasi-Newton algorithm.

## Algorithm 3.1 SQN Algorithm

Step 1 : Set $k=0$; select the initial point $x_{0}$ and $\varepsilon$ as a stopping condition.
We also set $D_{0}=I$, where $I$ is $n \times n$ identity matrix.
Step 2 : For $k \geq 0$, compute $g_{k}=A x_{k}-b$. If $\left\|g_{k}\right\| \leq \varepsilon$, stop, else compute $D_{k}$, where $D$ is a specific diagonal preconditioner.
Step 3 : Compute $d_{k+1}=-D_{k+1} g_{k+1}+\sum_{i=1}^{m} \beta_{i} s_{k+1-i}$, where $\beta_{i}=\frac{g_{i+1}^{T} A d_{i}}{d_{i}^{T} A d_{i}}$, $i \leq \min \{k, m\}$.

Step 4 : Compute $\alpha_{k}=-\frac{g_{k}^{T} d_{k}}{d_{k}^{T} A d_{k}}$.

Step 5 : Hence, $x_{k+1}=x_{k}+\alpha_{k} d_{k}$.

Step 6 : Set $k:=k+1$; go to step 2.
The SQN method is tested where in Step 2, $D$ is chosen from theorem 3.1.

## CONVERGENCE ANALYSIS

In this section, we shall look at the convergence properties of the subspace quasi-Newton method. Note that all the Hessian approximations are obtained by updating a bounded matrix using the proposed preconditioned subspace quasi-Newton method. We will prove the convergence properties of our proposed methods based upon the convergence assumptions given by Liu and Nocedal (1989), since it is valid for our preconditioning formulae whose matrices are diagonal and positive definite.

## Assumption 4.1:

(1) The objective function $f$ is twice continuously differentiable.
(2) The level set $D=\left\{x \in \mathfrak{R}^{n}: f(x) \leq f\left(x_{0}\right)\right\}$ is convex.
(3) There exist positive constants $M_{1}$ and $M_{2}$ such that

$$
\begin{equation*}
M_{1}\|z\|^{2} \leq z^{T} G(x) z \leq M_{2}\|z\|^{2} \tag{25}
\end{equation*}
$$

for $\forall z \in \mathfrak{R}^{n}$ and $\forall z \in D$. This implies that the objective function $f$ has a unique minimize $x^{*}$ in $D$.

From (25), we can have another similar inequality, as below:

$$
\begin{equation*}
N_{1}\|z\|^{2} \leq z^{T} G(x)^{-1} z \leq N_{2}\|z\|^{2} \tag{26}
\end{equation*}
$$

where $N_{1}=\frac{1}{M_{2}}$ and $N_{2}=\frac{1}{M_{1}}$ are the constants.

## Lemma 4.1

Let $x_{0}$ be a starting point for which $f$ satisfies Assumptions 4.1, and we take $D_{0}=I$, where $I$ is the $n \times n$ identity matrix. Assume that the matrices $D_{k}^{0}$ are chosen so that $\left\{\left\|D_{k}^{(0)}\right\|\right\}$ and $\left\{\left\|D_{k}^{(0)-1}\right\|\right\}$ are bounded. Then, $\left\{D_{k+1}\right\}$ and $\left\{\left\|D_{k+1}^{-1}\right\|\right\}$ are also bounded, where,

$$
\begin{equation*}
D_{k+1}=D_{k}+\frac{\omega_{k}-\mu_{k}}{\gamma_{k}} G_{k} \tag{27}
\end{equation*}
$$

where $\omega=y_{k}^{T} y_{k}, \mu=y_{k}^{T} D_{k} s_{k}, \gamma=\sum_{i=1}^{n}\left(y_{k}^{(i)} s_{k}^{(i)}\right)^{2}$ and $G_{k}=\operatorname{diag}\left(\left(s_{k}^{(1)} y_{k}^{(1)}\right), \ldots,\left(s_{k}^{(i)} y_{k}^{(i)}\right)\right)$ with $y_{k}^{(i)}$ and $s_{k}^{(i)}$ being the $i-t h$ component of the $y_{k}$ and $s_{k}$, respectively.

## Proof

Without the loss of generality, we shall assume that $D_{0}=I$, where $I$ is the $n \times n$ identity matrix. It is clear that $D_{0}$ is bounded, as follows:

$$
\mu_{0} \leq\left|D_{0}\right|_{F} \leq \omega_{0}
$$

Now, we need to prove $D_{1}$ is bounded.

Let $\nabla^{2} f(\bar{x})$ be defined as:

$$
\nabla^{2} f(\bar{x})=\int_{0}^{1} \nabla^{2} f\left(x_{k}+\tau s_{k}\right) d \tau
$$

Then, we have,

$$
\begin{equation*}
y_{k}=\nabla^{2} f(\bar{x}) s_{k} \tag{28}
\end{equation*}
$$

From (26) and (28), we get

$$
\begin{equation*}
N_{1}\left\|y_{k}\right\|^{2} \leq y_{k}^{T} s_{k} \leq N_{2}\left\|y_{k}\right\|^{2} \tag{29}
\end{equation*}
$$

Where, $N_{1}=\frac{1}{M_{2}}$ and $N_{2}=\frac{1}{M_{1}}$ are the constants.
From (27), we have

$$
\begin{align*}
& \left\|D_{1}\right\|_{F}=\left\|D_{0}+\frac{\omega_{0}-\mu_{0}}{\gamma_{0}} G_{0}\right\|_{F} \\
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+\left\|\frac{\omega_{0}-\mu_{0}}{\gamma_{0}} G_{0}\right\|_{F} \\
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+\frac{\left(\omega_{0}-\mu_{0}\right)}{\gamma_{0}}\left\|G_{0}\right\|_{F} \tag{30}
\end{align*}
$$

where $\|\cdot\|_{F}^{2}$ is the square of Frobenius norm and let $t r$ be the trace operator.

Note that

$$
\begin{align*}
& \left\|y_{0}\right\|^{2}=y_{0}^{(1) 2}+y_{0}^{(2) 2}+\ldots+y_{0}^{(n) 2} \\
& \leq n\left(y_{0}^{M}\right)^{2} \tag{31}
\end{align*}
$$

where $\left(y_{0}^{M}\right)^{2}=\max \left\{y_{0}^{(1) 2}, y_{0}^{(2) 2}, \ldots, y_{0}^{(n) 2}\right\}$.

From (29), (30) and (31), we will get

$$
\begin{align*}
\left\|G_{0}\right\|_{F}^{2} & =\operatorname{tr}\left(G_{0}^{T} G_{0}\right) \\
& =s_{0}^{(1) 2} y_{0}^{(1) 2}+s_{0}^{(2) 2} y_{0}^{(2) 2}+\ldots+s_{0}^{(n) 2} y_{0}^{(n) 2} \\
& \leq\left(y_{0}^{T} s_{0}\right)^{2} \\
\| G_{0} \mid & \leq N_{2} n\left(y_{0}^{M}\right)^{2} . \tag{32}
\end{align*}
$$

From (27), we have

$$
\begin{align*}
\omega_{0} & =y_{0}^{T} y_{0} \\
& =\left\|y_{0}\right\|^{2} \\
& \leq n\left(y_{0}^{M}\right)^{2} \tag{33}
\end{align*}
$$

and

$$
\begin{align*}
& \mu_{0}=y_{0}^{T} D_{0} s_{0} \\
& \mu_{0} \leq M_{3} y_{0}^{T} s_{0} \\
& \mu_{0} \leq M_{3} N_{2}\left\|y_{0}\right\|^{2} \\
& \mu_{0} \leq M_{3} N_{2}\left(y_{0}^{M}\right)^{2} \tag{34}
\end{align*}
$$

where, $M_{3}=1$.
Hence, from (30), (32), (33) and (34), we shall have

$$
\begin{aligned}
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+\frac{\left(n\left(y_{0}^{M}\right)^{2}-N_{2} n\left(y_{0}^{M}\right)^{2}\right)}{\sum_{i=1}^{n}\left(s_{0}^{(i)} y_{0}^{(i)}\right)^{2}} N_{2} n\left(y_{0}^{M}\right)^{2} \\
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+\frac{\left(n-N_{2} n\right)\left(y_{0}^{M}\right)^{2}}{\sum_{i=1}^{n}\left(s_{0}^{(i)} y_{0}^{(i)}\right)^{2}} N_{2} n\left(y_{0}^{M}\right)^{2} \\
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+\frac{\left(1-N_{2}\right) N_{2} n^{2}\left(y_{0}^{M}\right)^{4}}{\sum_{i=1}^{n}\left(s_{0}^{(i)} y_{0}^{(i)}\right)^{2}}
\end{aligned}
$$

$$
\begin{align*}
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+\frac{k N_{2} n^{2}\left(y_{0}^{M}\right)^{4}}{\sum_{i=1}^{n}\left(s_{0}^{(i)} y_{0}^{(i)}\right)^{2}} \\
& \left\|D_{1}\right\|_{F} \leq\left\|D_{0}\right\|_{F}+M_{4} \\
& \left\|D_{1}\right\|_{F} \leq n+M_{4} \tag{35}
\end{align*}
$$

Where, $M_{4}=k N_{2} n^{2}$ and $k=\max \left\{\left(1-N_{2}\right),\left(1+N_{2}\right)\right\}$, and

$$
\begin{equation*}
\frac{\left(y_{0}^{M}\right)^{4}}{\sum_{i=1}^{n}\left(s_{0}^{(i)} y_{0}^{(i)}\right)^{2}} \leq 1 \tag{36}
\end{equation*}
$$

From (35), we can conclude that $\left\|D_{1}\right\|_{F}$ is bounded since $\left\|D_{0}\right\|_{F}$ is also bounded. Now, we assume that $D_{k}$ is bounded, and then, we need to prove that $D_{k+1}$ is also bounded.

From the above, we shall get a similar equation and inequality, as follows:

$$
\begin{align*}
& \left\|G_{k}\right\|_{F} \leq N_{2} n\left(y^{M}\right)^{2},  \tag{37}\\
& \omega_{k} \leq\left\|y_{k}\right\|^{2}  \tag{38}\\
& \mu_{k} \leq M_{3} N_{2}\left\|y_{k}\right\|^{2}  \tag{39}\\
& \left\|y_{k}\right\|^{2} \leq n\left(y_{k}^{M}\right)^{2} \tag{40}
\end{align*}
$$

From (27) and (37)-(40), we obtain

$$
\begin{equation*}
\left\|D_{k+1}\right\|_{F} \leq\left\|D_{k}\right\|_{F}+M_{4} \tag{41}
\end{equation*}
$$

Where, $M_{4}=k N_{2} n^{2}$ and $k=\max \left\{\left(1-M_{3} N_{2}\right),\left(1+M_{3} N_{2}\right)\right\}$.
From the fact that $\left\|D_{k}\right\|_{F}$ is bounded, i.e. $\left\|D_{k}\right\|_{F} \leq M_{5}$, and from (41),

$$
\begin{aligned}
& \left\|D_{k+1}\right\|_{F} \leq M_{5}+M_{6} \\
& \left\|D_{k+1}\right\|_{F} \leq M_{6}
\end{aligned}
$$

Where, $M_{6}=M_{5}+M_{4}$ and it is a constant. Finally, we have shown that $\left\|D_{k+1}\right\|_{F}$ is bounded and the proof is completed.

In this section, we have shown that the proposed preconditioned subspace quasi-Newton methods are convergent on uniformly convex problems and the rate is $R$-linear. This $R$ -linear convergence results obtained are based upon the assumption by Liu and Nocedal (1989).

## COMPUTATIONAL RESULTS AND DISCUSSION

The computational results and discussion on the performance of preconditioner subspace quasiNewton (SQN) method are given in this section. All the algorithms are written in MATLAB 7.0. The total number of tested problems is 4 . All the runs were terminated when

$$
\left\|g_{k}\right\| \leq 10^{-4}
$$

Where, $\|\cdot\|$ denotes the Euclidean norm. Furthermore, we also consider the number of function evaluation and gradient calls. We set our upper bound for the number of function evaluation and gradient call to be 1000 .

The computational results are compared through the number of iterations, gradient evaluations as well as function evaluations. In order to test the efficiency of the proposed preconditioned methods, the number of subspaces that is considered is $m=2$ and $m=3$.

Meanwhile, the SQN method was tested using the following preconditioners:

1. $\mathrm{SQN}(0)-\mathrm{SQN}$ method without preconditioning.
2. SQN(D1)-SQN method with diagonal preconditioner $D$, where $D$ is given by Theorem 3.1.

In order to compare the efficiency of the proposed preconditioned SQN methods with the standard SQN method, the following quadratic test problem is considered:

$$
\begin{equation*}
f(x)=\frac{1}{2} x^{T} A x-b^{T} x \tag{42}
\end{equation*}
$$

where, $A$ is positive definite diagonal matrix and $b=[1,1,1,1,1, \ldots, 1]$.
For all the methods, the initial point is $x_{0}=[0,0,0,0, \ldots, 0]$. A set of unconstrained minimization quadratic problems consisting of 4 test problems were used. We now describe the 4 different quadratic test problems (43) with $n$-dimensional cases.

1. QF 1 , where $A=\operatorname{diag}\left[a_{i i}\right], a_{i i}=i^{2}(\bmod 5), b=[1, \ldots, 1]$.
2. QF2, where $A=\operatorname{diag}\left[a_{i i}\right], a_{i i}=i^{3}(\bmod 5), b=[1, \ldots, 1]$.
3. QF3, where $A=\operatorname{diag}\left[a_{i i}\right], a_{i i}=i^{3}+i(\bmod 5), b=[1, \ldots, 1]$.
4. QF4, where $A=\operatorname{diag}\left[a_{i i}\right], a_{i i}=a_{i-2, i-2}+a_{i-1, i-1}, i \geq 3$ and $a_{11}=1, a_{22}=1, b=[1, \ldots, 1]$.

We tested the above problems by using $m=2$ and $m=3$. In each table, the symbol Ite, $\left\|g_{k}\right\|$, and Fva mean the number of iterations, norm of the gradient and function evaluation, respectively.

## TABLE 1

A comparison of the Method of $m=2$ in solving QF1

| $\mathrm{SQN}(0)$ |  |  |  |  | $\mathrm{SQN}(\mathrm{D} 1)$ |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |
| 10 | 106 | $9.2 \mathrm{e}-5$ | -1.4636 | 40 | $6.2 \mathrm{e}-5$ | -1.4636 |
| 20 | 109 | $9.8 \mathrm{e}-5$ | -2.9272 | 40 | $8.8 \mathrm{e}-5$ | -2.9272 |
| 40 | 113 | $9.6 \mathrm{e}-5$ | -5.8544 | 41 | $6.6 \mathrm{e}-5$ | -5.8544 |
| 80 | 117 | $9.3 \mathrm{e}-5$ | $-1.1709 \mathrm{e}+1$ | 41 | $9.3 \mathrm{e}-5$ | $-1.1709 \mathrm{e}+1$ |
| 100 | 118 | $9.5 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+1$ | 45 | $7.6 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+1$ |
| 200 | 122 | $9.3 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+1$ | 46 | $9.2 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+1$ |
| 500 | 127 | $9.2 \mathrm{e}-5$ | $-7.3181 \mathrm{e}+1$ | 52 | $8.7 \mathrm{e}-5$ | $-7.3181 \mathrm{e}+1$ |
| 1000 | 130 | $9.9 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+2$ | 53 | $8.9 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+2$ |
| 1500 | 133 | $9.1 \mathrm{e}-5$ | $-2.1954 \mathrm{e}+2$ | 56 | $4.6 \mathrm{e}-5$ | $-2.1954 \mathrm{e}+2$ |
| 2000 | 134 | $9.6 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+2$ | 56 | $5.3 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+2$ |

TABLE 2
A comparison of the Method of $m=2$ in solving QF2

| $\mathrm{SQN}(0)$ |  |  |  |  | SQN(D1) |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |
| 10 | 598 | $9.9 \mathrm{e}-5$ | -1.1857 | 66 | $9.8 \mathrm{e}-5$ | -1.1857 |
| 20 | 619 | $9.9 \mathrm{e}-5$ | -2.3713 | 78 | $4.7 \mathrm{e}-5$ | -2.3713 |
| 40 | 640 | $9.9 \mathrm{e}-5$ | -4.7426 | 78 | $6.6 \mathrm{e}-5$ | -4.7426 |
| 80 | 661 | $9.9 \mathrm{e}-5$ | -9.4853 | 79 | $2.5 \mathrm{e}-5$ | -9.4853 |
| 100 | 668 | $9.9 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+1$ | 79 | $3.5 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+1$ |
| 200 | 689 | $9.9 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+1$ | 79 | $5.5 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+1$ |
| 500 | 716 | $1.0 \mathrm{e}-4$ | $-5.9283 \mathrm{e}+1$ | 79 | $6.2 \mathrm{e}-5$ | $-5.9283 \mathrm{e}+1$ |
| 1000 | 737 | $1.0 \mathrm{e}-4$ | $-1.1857 \mathrm{e}+2$ | 91 | $1.0 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+2$ |
| 1500 | 750 | $9.9 \mathrm{e}-5$ | $-1.7785 \mathrm{e}+2$ | 102 | $7.0 \mathrm{e}-5$ | $-1.7785 \mathrm{e}+2$ |
| 2000 | 758 | $1.0 \mathrm{e}-4$ | $-2.3713 \mathrm{e}+2$ | 116 | $3.1 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+2$ |

TABLE 3
A comparison of the Method of $m=2$ in solving QF3

|  | SQN(0) |  |  |  | SQN(D1) |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |
| 10 | 311 | $9.7 \mathrm{e}-5$ | $-6.5573 \mathrm{e}-1$ | 48 | $9.7 \mathrm{e}-5$ | $-6.5573 \mathrm{e}-1$ |
| 20 | 322 | $9.7 \mathrm{e}-5$ | -1.3115 | 54 | $8.1 \mathrm{e}-5$ | -1.3115 |
| 40 | 332 | $1.0 \mathrm{e}-4$ | -2.6229 | 58 | $7.4 \mathrm{e}-5$ | -2.6229 |
| 80 | 343 | $1.0 \mathrm{e}-4$ | -5.2459 | 64 | $9.4 \mathrm{e}-5$ | -5.2459 |
| 100 | 347 | $9.8 \mathrm{e}-5$ | -6.5573 | 66 | $7.2 \mathrm{e}-5$ | -6.5573 |
| 200 | 358 | $9.8 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+1$ | 67 | $5.3 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+1$ |
| 500 | 372 | $9.9 \mathrm{e}-5$ | $-3.2787 \mathrm{e}+1$ | 67 | $8.4 \mathrm{e}-5$ | $-3.2787 \mathrm{e}+1$ |
| 1000 | 383 | $9.9 \mathrm{e}-5$ | $-6.5573 \mathrm{e}+1$ | 68 | $5.3 \mathrm{e}-5$ | $-6.5573 \mathrm{e}+1$ |
| 1500 | 390 | $9.7 \mathrm{e}-5$ | $-9.8360 \mathrm{e}+1$ | 68 | $6.4 \mathrm{e}-5$ | $-9.8360 \mathrm{e}+1$ |
| 2000 | 394 | $9.9 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+2$ | 68 | $7.4 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+2$ |

TABLE 4
A comparison of the Method of $m=2$ in solving QF4

| SQN(0) |  |  |  |  |  | SQN(D1) |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |  |
| 10 | 252 | $9.7 \mathrm{e}-5$ | -1.6652 | 74 | $9.6 \mathrm{e}-5$ | -1.6652 |  |
| 20 | 261 | $9.7 \mathrm{e}-5$ | -3.3305 | 77 | $9.1 \mathrm{e}-5$ | -3.3305 |  |
| 40 | 270 | $9.6 \mathrm{e}-5$ | -6.6609 | 85 | $8.8 \mathrm{e}-5$ | -6.6609 |  |
| 80 | 278 | $1.0 \mathrm{e}-4$ | $-1.3322 \mathrm{e}+1$ | 86 | $8.1 \mathrm{e}-5$ | $-1.3322 \mathrm{e}+1$ |  |
| 100 | 281 | $9.9 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ | 86 | $9.1 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ |  |
| 200 | 290 | $9.9 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+1$ | 92 | $8.2 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+1$ |  |
| 500 | 301 | $9.8 \mathrm{e}-5$ | $-8.3262 \mathrm{e}+1$ | 100 | $1.0 \mathrm{e}-5$ | $-8.3262 \mathrm{e}+1$ |  |
| 1000 | 311 | $9.7 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ | 103 | $7.4 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ |  |
| 1500 | 316 | $9.8 \mathrm{e}-5$ | $-2.4979 \mathrm{e}+1$ | 103 | $9.5 \mathrm{e}-5$ | $-2.4979 \mathrm{e}+1$ |  |
| 2000 | 320 | $9.7 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+2$ | 106 | $9.4 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+2$ |  |

TABLE 5
Comparison of the Method of $m=3$ in solving QF1

|  | $\mathrm{SQN}(0)$ |  |  |  | $\mathrm{SQN}(\mathrm{D} 1)$ |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |
| 10 | 81 | $9.8 \mathrm{e}-5$ | -1.4636 | 36 | $8.6 \mathrm{e}-5$ | -1.4636 |
| 20 | 84 | $9.6 \mathrm{e}-5$ | -2.9272 | 38 | $8.1 \mathrm{e}-5$ | -2.9272 |
| 40 | 87 | $9.4 \mathrm{e}-5$ | -5.8544 | 39 | $7.7 \mathrm{e}-5$ | -5.8544 |
| 80 | 90 | $9.3 \mathrm{e}-5$ | $-1.1709 \mathrm{e}+1$ | 40 | $5.9 \mathrm{e}-5$ | $-1.1709 \mathrm{e}+1$ |
| 100 | 91 | $9.2 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+1$ | 40 | $6.6 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+1$ |

TABLE 5 (continue)

|  | $\operatorname{SQN}(0)$ |  |  | SQN(D1) |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |
| 200 | 94 | $9.0 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+1$ | 40 | $9.4 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+1$ |
| 500 | 97 | $9.9 \mathrm{e}-5$ | $-7.3181 \mathrm{e}+1$ | 47 | $5.6 \mathrm{e}-5$ | $-7.3181 \mathrm{e}+1$ |
| 1000 | 100 | $9.7 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+2$ | 47 | $8.0 \mathrm{e}-5$ | $-1.4636 \mathrm{e}+2$ |
| 1500 | 102 | $9.4 \mathrm{e}-5$ | $-2.1954 \mathrm{e}+2$ | 47 | $9.8 \mathrm{e}-5$ | $-2.1954 \mathrm{e}+2$ |
| 2000 | 103 | $9.6 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+2$ | 48 | $9.5 \mathrm{e}-5$ | $-2.9272 \mathrm{e}+2$ |

TABLE 6
A comparison of the Method of $m=3$ in solving QF2

| $\operatorname{SQN}(0)$ |  |  |  |  |  | SQN(D1) |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | $\left\\|g_{k}\right\\|$ | Fva |  |  |
| 10 | 577 | $1.0 \mathrm{e}-4$ | -1.1857 | 74 | $8.9 \mathrm{e}-5$ | -1.1857 |  |
| 20 | 598 | $9.8 \mathrm{e}-5$ | -2.3713 | 80 | $6.1 \mathrm{e}-5$ | -2.3713 |  |
| 40 | 618 | $9.9 \mathrm{e}-5$ | -4.7426 | 80 | $8.6 \mathrm{e}-5$ | -4.7426 |  |
| 80 | 638 | $9.9 \mathrm{e}-5$ | -9.4853 | 87 | $1.5 \mathrm{e}-5$ | -9.4853 |  |
| 100 | 645 | $9.9 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+1$ | 87 | $6.2 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+1$ |  |
| 200 | 665 | $9.9 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+1$ | 89 | $5.3 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+1$ |  |
| 500 | 692 | $9.9 \mathrm{e}-5$ | $-5.9283 \mathrm{e}+1$ | 89 | $8.8 \mathrm{e}-5$ | $-5.9283 \mathrm{e}+1$ |  |
| 1000 | 712 | $9.9 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+2$ | 90 | $7.9 \mathrm{e}-5$ | $-1.1857 \mathrm{e}+2$ |  |
| 1500 | 724 | $9.9 \mathrm{e}-5$ | $-1.7785 \mathrm{e}+2$ | 92 | $7.6 \mathrm{e}-5$ | $-1.7785 \mathrm{e}+2$ |  |
| 2000 | 732 | $9.9 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+2$ | 96 | $3.2 \mathrm{e}-5$ | $-2.3713 \mathrm{e}+2$ |  |

TABLE 7
A comparison of the Method of $m=3$ in solving QF3

| $\mathrm{SQN}(0)$ |  |  |  |  | $\mathrm{SQN}(\mathrm{D} 1)$ |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | Ite | $\left\\|g_{k}\right\\|$ | Fva |
| 10 | 300 | $9.8 \mathrm{e}-5$ | $-6.5573 \mathrm{e}-1$ | 44 | $8.2 \mathrm{e}-5$ | $-6.5573 \mathrm{e}-1$ |
| 20 | 310 | $1.0 \mathrm{e}-4$ | -1.3115 | 58 | $7.0 \mathrm{e}-5$ | -1.3115 |
| 40 | 321 | $9.8 \mathrm{e}-5$ | -2.6229 | 58 | $1.0 \mathrm{e}-4$ | -2.6229 |
| 80 | 332 | $9.7 \mathrm{e}-5$ | -5.2459 | 60 | $4.6 \mathrm{e}-5$ | -5.2459 |
| 100 | 335 | $9.8 \mathrm{e}-5$ | -6.5573 | 60 | $5.2 \mathrm{e}-5$ | -6.5573 |
| 200 | 346 | $9.7 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+1$ | 60 | $7.3 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+1$ |
| 500 | 359 | $1.0 \mathrm{e}-4$ | $-3.2787 \mathrm{e}+1$ | 62 | $9.0 \mathrm{e}-5$ | $-3.2787 \mathrm{e}+1$ |
| 1000 | 370 | $9.9 \mathrm{e}-5$ | $-6.5573 \mathrm{e}+1$ | 63 | $6.1 \mathrm{e}-5$ | $-6.5573 \mathrm{e}+1$ |
| 1500 | 376 | $9.9 \mathrm{e}-5$ | $-9.8360 \mathrm{e}+1$ | 63 | $7.5 \mathrm{e}-5$ | $-9.8360 \mathrm{e}+1$ |
| 2000 | 381 | $9.7 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+2$ | 63 | $8.7 \mathrm{e}-5$ | $-1.3115 \mathrm{e}+2$ |

TABLE 8
A comparison of the Method of $m=3$ in solving QF4

| SQN(0) |  |  |  |  | SQN(D1) |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| N | Ite | $\left\\|g_{k}\right\\|$ | Fva | $\left\\|g_{k}\right\\|$ | Fva |  |
| 10 | 230 | $9.8 \mathrm{e}-5$ | -1.6652 | 62 | $7.9 \mathrm{e}-5$ | -1.6652 |
| 20 | 238 | $9.8 \mathrm{e}-5$ | -3.3305 | 63 | $5.4 \mathrm{e}-5$ | -3.3305 |
| 40 | 246 | $9.9 \mathrm{e}-5$ | -6.6609 | 63 | $7.7 \mathrm{e}-5$ | -6.6609 |
| 80 | 254 | $9.9 \mathrm{e}-5$ | $-1.3322 \mathrm{e}+1$ | 64 | $8.3 \mathrm{e}-5$ | $-1.3322 \mathrm{e}+1$ |
| 100 | 257 | $9.8 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ | 64 | $9.3 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ |
| 200 | 265 | $9.8 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+1$ | 68 | $3.5 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+1$ |
| 500 | 276 | $9.7 \mathrm{e}-5$ | $-8.3262 \mathrm{e}+1$ | 68 | $9.5 \mathrm{e}-5$ | $-8.3262 \mathrm{e}+1$ |
| 1000 | 284 | $9.8 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ | 73 | $8.2 \mathrm{e}-5$ | $-1.6652 \mathrm{e}+1$ |
| 1500 | 289 | $9.7 \mathrm{e}-5$ | $-2.4979 \mathrm{e}+1$ | 73 | $9.6 \mathrm{e}-5$ | $-2.4979 \mathrm{e}+1$ |
| 2000 | 292 | $9.8 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+2$ | 74 | $8.8 \mathrm{e}-5$ | $-3.3305 \mathrm{e}+2$ |

The number of iterations is the success index in a computational method. In this study, the number of iterations was compared between the standard SQN method and the proposed SQN method.

Tables 1-4 show the comparison results between the proposed preconditioned SQN methods and the standard SQN method for $m=2$. Generally, the computational results show that the proposed methods performed better when compared to that of the standard SQN method. As shown in the Tables, the proposed methods require less number of iterations than the standard method. Although all the methods show the same values of function evaluation, the norms of the gradient for the proposed methods are less than the norms of the gradient of the standard method. Once again, this shows that the proposed SQN methods are promising alternatives as compared to the standard SQN method.

Tables 5-8 show the comparison results between the proposed preconditioned SQN methods and the standard SQN method for $m=3$. Once again, the results reveal that the proposed methods clearly outperform the standard method. The number of iterations and the norms of the gradient are the best evidences to show that the proposed methods generally perform better than the standard SQN method.

## CONCLUSION

The preconditioner for SQN method that is based upon variational technique and weak secant relation is proposed in this paper. The numerical results obtained suggest that the preconditioned SQN method is a good alternative for large-scale unconstrained optimization. Moreover, the preconditioned SQN method is preferred for reasons including simple implementation and it requires only function and gradient values.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledged the financial support of Graduate Research Fellowship (GRF) from Universiti Putra Malaysia and the Ministry of Higher Education Malaysia.

## REFERENCES

Farid, M., Hassan, M. A., \& Leong, W. J. (2011). Improved Hessian approximation with modified secant equations for symmetric rank-one method. J. Comput. Appl. Math., 235, 2423-2431.

Farid, M., Leong, W. J., \& Hassan, M. A. (2010a). A new two-step gradient-type method for large scale unconstrained optimization. Computers and Mathematics with Applications, 59, 3301-3307.

Farid, M., Leong, W. J., \& Hassan, M. A. (2010b). An improved multi-step gradient-type method for large scale optimization. Computers and Mathematics with Applications, 61, 3312-3318.

Leong, W. J., \& Hassan, M. A. (2009). A restarting approach for the symmetric rank one update for unconstrained optimization. Computational Optimization and applications, 43, 327-334.

Leong, W. J., \& Hassan, M. A. (2011). A new gradient method via Least change secant update. International Journal of Computer Mathematics, 88, 816-828.

Leong, W. J., Hassan, M. A., \& Farid, M. (2010a). A monotone Gradient method via weak secant equations for unconstrained optimization. Taiwanese J. Math., 14(2), 413-423.

Leong, W. J., Farid, M., \& Hassan, M. A. (2010b). Improved Hessian approximation with modified quasi-Cauchy relation for gradient-type method. Advanced Modeling and Optimization, 12, 37-44.

Liu, D. C., \& Nocedal, J. (1989). On the limited memory BFGS Method for large scale optimization. Mathematical Programming, 45, 503-528.
Ni, Q., \& Yuan, Y. (1997). A subspace limited memory quasi-Newton algorithm for large scale nonlinear bound constrained optimization Mathematics of Computation, 66, 1509-1520.

Stoer, J., \& Yuan, Y. (1995). A subspace study on conjugate gradient algorithms. ZAMMZ. Angew. Math. Mech., 75, 69-77.
Wang, Z. H. Wen, Z. W., \& Yuan, Y. (2004). A subspace trust region method for large scale unconstrained ooptimization. Numerical Linear Algebra and Optimization, 265-274.
Waziri, M. Y., Leong, W. J., Hassan, M. A., \& Monsi, M. (2010a). A new Newton's method with diagonal Jacobian approximation for systems of nonlinear equations. J. Mathematics and Statistics, 6, 246-252.

Waziri, M. Y., Leong, W. J., Hassan, M. A., \& Monsi, M. (2010b). Jacobian computation-free Newton method for systems of non-linear equations. Journal of numerical Mathematics and Stochastic, 2(1), 54-63.

Yuan, Y. (2007). Subspace techniques for nonlinear optimization. Some Topics in Industrial and Applied Mathematics (Series in Contemporary Applied Mathematics CAM), 8, 206-218.


[^0]:    Article history:
    Received: 22 September 2011
    Accepted: 11 November 2011

    ## E-mail addresses:

    hongseng0505@hotmail.com (Hong Seng Sim),
    leongwj@upm.edu.my (Wah June Leong),
    malik@science.upm.edu.my (Malik Abu Hassan),
    fudziah@upm.edu.my (Fudziah Ismail)
    *Corresponding Author

