Sparse quasi-Newton updates with positive definite matrix completion

Nobuo Yamashita*

March 11, 2005
(Revised August 4, 2005)

Abstract

The quasi-Newton method is a powerful method for solving unconstrained minimization problems. However, since the approximate Hessian generated by the usual quasi-Newton update (e.g. BFGS or DFP) becomes dense, the quasi-Newton method cannot be applied for large-scale problems due to lack of memory. To overcome this difficulty, we propose sparse quasi-Newton updates with positive definite matrix completion that exploit the sparsity pattern $E := \{(i,j) \mid (\nabla^2 f(x))_{ij} \neq 0 \text{ for some } x \in R^n\}$ of the Hessian. The proposed method first calculates a partial approximate Hessian $H_{ij}^{QN}, (i,j) \in F$, where $F \supseteq E$, by using an existing quasi-Newton update formula such as BFGS or DFP. Next, we obtain a full matrix H_{k+1}, which is a maximum-determinant positive definite matrix completion of $H_{ij}^{QN}, (i,j) \in F$. If the sparsity pattern E (or its extension F) has a property related to a chordal graph, then the matrix H_{k+1} can be expressed as products of some sparse matrices. Therefore, if the Hessian is sparse, the time and space complexities of the proposed method are far fewer than those of the BFGS or the DFP. In particular, when the Hessian matrix is tridiagonal, the complexities become $O(n)$. We show that the proposed method has superlinear convergence under the usual assumptions.

Key words: quasi-Newton method, large-scale problems, sparsity, positive definite matrix completion.
AMS subject classifications. 90C53, 90C06

1 Introduction

In this paper we consider the following unconstrained minimization problem:

$$\begin{array}{ll} \min & f(x) \\ \text{subject to} & x \in R^n. \end{array} \tag{1}$$

Throughout the paper we assume that f is twice continuously differentiable, n is huge and $\nabla^2 f(x)$ is sparse. For solving the unconstrained minimization problem, there exist several useful methods,

*Department of Applied Mathematics of Physics, Graduate School of Informatics, Kyoto University, 606-8501, Japan. nobuo@i.kyoto-u.ac.jp

including the steepest descent method, the Newton method, the quasi-Newton method, the conjugate gradient method and the trust region method [10]. Among others, the quasi-Newton method is easy to implement and has good convergence properties.

The quasi-Newton method generates a sequence $\{x_k\}$ by $x_{k+1} = x_k - H_k \nabla f(x_k)$ with an approximate inverse Hessian H_k. The approximate inverse Hessian usually satisfies the secant condition:

$$H_{k+1} y_k = s_k, \tag{2}$$

where

$$\begin{aligned} s_k &= x_{k+1} - x_k \\ y_k &= \nabla f(x_{k+1}) - \nabla f(x_k). \end{aligned}$$

The quasi-Newton updates that satisfy the secant condition are BFGS and DFP. The BFGS and DFP update formulae are given by

$$H_{k+1}^{BFGS} = H_k - \frac{H_k y_k s_k^T + s_k (H_k y_k)^T}{s_k^T y_k} + \left(1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k} \tag{3}$$

and

$$H_{k+1}^{DFP} = H_k - \frac{H_k y_k (H_k y_k)^T}{y_k^T H_k y_k} + \frac{s_k s_k^T}{s_k^T y_k}, \tag{4}$$

respectively. It is known that both H_{k+1}^{BFGS} and H_{k+1}^{DFP} are positive definite when $s_k^T y_k > 0$ and H_k is positive definite. Moreover, the update can be calculated within $O(n^2)$ arithmetic operations, whereas the Newton method requires $O(n^3)$ arithmetic operations to solve Newton equations. The quasi-Newton method has superlinear convergence under appropriate conditions [2, 10]. Therefore, the quasi-Newton method is very efficient for small- and medium-scale problems. For large-scale problems, the Hessian $\nabla^2 f(x_k)$ usually becomes sparse. By exploiting the sparsity, the Newton method and the trust region method can be implemented with little memory. Thus, these methods are applicable for such problems. However, since $s_k s_k^T$ in (3) or (4) becomes dense, the updated matrix H_{k+1} (or its inverse B_{k+1}) is also dense even if the Hessian is sparse. Storing the full matrix H_{k+1} requires $O(n^2)$ memory, and thus BFGS and DFP are not applicable for large-scale problems.

In order to overcome this difficulty, several methods have been proposed [4, 9, 12]. The limited-memory BFGS (L-BFGS) [9] is widely used in practice. The L-BFGS stores a few vector pairs $(s^i, y^i), i = k-m+1, \ldots, k-1, k$, and constructs an approximate Hessian by BFGS with the vector pairs. The approximate Hessian satisfies the secant condition and becomes positive definite. The time and space complexities per iteration of the L-BFGS are $\mathrm{O}(mn)$, and it is shown that the L-BFGS converges linearly [8]. However, since L-BFGS does not use much information of the Hessian, it converges very slowly for ill-posed problems.

In this paper we propose quasi-Newton updates that exploit the sparsity of the Hessian. Although Toint [12] and Fletcher [4] have previously proposed updates that exploit the sparsity, these methods involve the solution of a convex programming problem at each iteration in order to obtain approximate Hessians. Moreover, since these methods require the sparsity and secant conditions simultaneously, the approximate Hessian tends to be ill-posed when $(s_k)_i = 0$ for some i [11]. The method proposed herein is based on positive definite matrix completion. For a given set $F \subseteq \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$ and a partial matrix $\bar{X}_{ij}(i, j) \in F$, we assume that $\bar{X}_{ij}, (i, j) \in F$ has a positive definite matrix completion (PDMC) X or that X is a PDMC of $\bar{X}_{ij}, (i, j) \in F$ if X is an $n \times n$ symmetric positive definite matrix and $X_{ij} = \bar{X}_{ij}, \forall (i, j) \in F$. The PDMC has been investigated extensively [7, 6, 5]. Recently, the PDMC has been used for the interior point method for solving the sparse semidefinite programming problem [5]. The results reported in [7, 5] are as follows: (i) If F and $\bar{X}_{ij}, (i, j) \in F$ satisfy some properties related to a chordal graph (for the definition of "chordal", see Section 2), then $\bar{X}_{ij}, (i, j) \in F$ has a PDMC. (ii) If X is the maximum-determinant PDMC, then $(\bar{X})^{-1}_{ij} = 0, (i, j) \in F$. (iii) The maximum-determinant PDMC is expressed as products of sparse matrices. Based on these results, we propose new sparse quasi-Newton updates. The proposed methods first calculate a partial approximate inverse Hessian $H^{QN}_{ij}, (i, j) \in F$, where F is an extension of the sparsity pattern $E = \{(i, j) \mid (\nabla^2 f(x))_{ij} \neq 0 \text{ for some } x \in R^n\}$ of the Hessian, by using the existing quasi-Newton updates, such as BFGS (3) and DFP (4). We then obtain a full matrix H_{k+1}, which is the maximum-determinant PDMC of $H^{QN}_{ij}, (i, j) \in F$. When the Hessian is sparse, the time and space complexities of the proposed method become much fewer than those of BFGS and DFP. Since the updates do not require the sparsity and secant conditions simultaneously, they do not suffer from Sorensen's example [11], i.e., the approximate Hessian does not become ill-posed even if $(s_k)_i = 0$ for some i. Moreover, we will show that the proposed method has local and superlinear convergence under the usual assumptions.

The paper is organized as follows. In Section 2, we introduce some results regarding PDMC. The results are based primarily on [7, 5] and are rearranged slightly for our purpose. In Section 3, we propose the sparse quasi-Newton updates with PDMC and discuss their time and space complexities per iteration. In Section 4, we examine the behavior of the proposed method for Sorensen's example, which indicates that the proposed method is better than existing sparse quasi-Newton updates. We then show that the proposed method with DFP, which is a special case of the proposed methods, has local and superlinear convergence under appropriate conditions in Section 5. Section 6 presents a number of numerical experiments, and we present concluding remarks in Section 7.

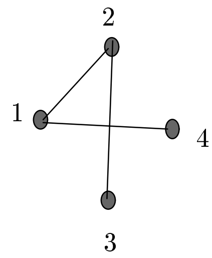
The following notation is used throughout the present paper. We denote V by $\{1, 2, \ldots, n\}$. For a given set $F \subset V \times V$, $F_i = \{j \in V \mid (i, j) \in F\}$ and $|F|$ denotes the number of elements of F. For an $n \times n$ matrix H, $\|H\|$ denotes the Frobenius norm of H and $H \succeq 0$ indicates that H is positive definite. For a vector $z \in R^n$ and a set $S \subseteq V$, z_S denotes the $|S|$-dimensional vector with components $z_i, i \in S$. For an $n \times n$ matrix A and sets $S, U \subseteq V$, A_{SU} denotes the $|S| \times |U|$ matrix with components $A_{ij}, (i, j) \in S \times U$.

2 Positive definite matrix completion

In this section we introduce some results regarding the PDMC, which will be used in subsequent sections. Most of these results are found in [7, 5].

Let $F \subseteq V \times V$. Throughout this section we assume that $(i, i) \in F$ for $i \in V$, and $(i, j) \in F$ if $(j, i) \in F$. For a given $\bar{X}_{ij}, (i, j) \in F$, the problem of finding a PDMC of $\bar{X}_{ij}, (i, j) \in F$ is usually formulated as a semidefinite programming problem, and thus it is not easy to obtain the PDMC. However, if F and $\bar{X}_{ij}, (i, j) \in F$ have certain properties, then the PDMC can be calculated directly. Such properties are related to a graph $G(V, \bar{F})$ induced from F, where $G(V, \bar{F})$ is a graph having a vertex set

$$M = \begin{pmatrix} * & * & 0 & * \\ * & * & * & 0 \\ 0 & * & * & 0 \\ * & 0 & 0 & * \end{pmatrix} \Leftrightarrow$$



$$M = \begin{pmatrix} * & 0 & 0 & 0 \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{pmatrix} \Leftrightarrow$$

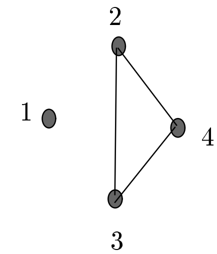
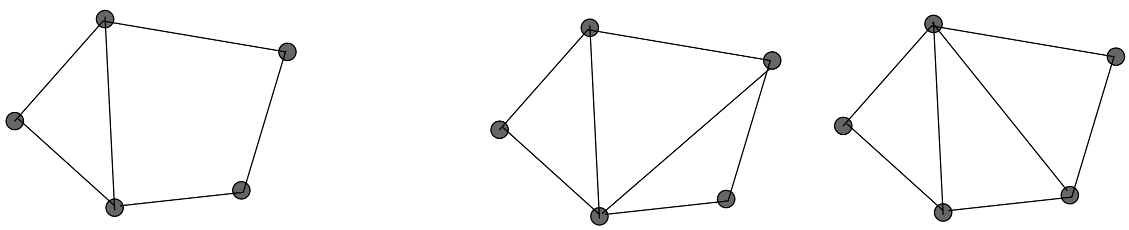


Figure 1: $F = \{(i,j) \in V \times V \mid M_{ij} \neq 0\}$ and its related graph



Non-chordal graph

Chordal graphs

Figure 2: Chordal graph

V and an edge set $\bar{F} := F \backslash \{(i,i) \mid i = 1, \dots, n\}$ (Figure 2).

We recall the following concepts of graph theory, which are related to PDMC.

Definition 1

- *Two vertices $u, v \in V$ are adjacent if $(u,v) \in \bar{F}$. The set of the vertices adjacent to $v \in V$ is denoted by $Adj(v)$.*
- *A graph is complete if every pair of vertices is adjacent.*
- *For a subset V' of V, the induced subgraph on V' is a graph $G(V', \bar{F}')$ with the edge set $\bar{F}' = \bar{F} \cap (V' \times V')$.*
- *A clique of a graph is an induced subgraph that is complete.*
- *A clique is maximal if its vertices do not constitute a proper subset of another clique.*
- *A vertex is simplicial if its adjacent vertices induce a clique.*
- *For a cycle, an edge is a cord of the cycle if it joins two nonconsecutive vertices of the cycle.*
- *A graph is chordal if every cycle of length greater than 3 has a chord (Figure 2).*

When $G(V, \bar{F})$ is a chordal graph, there exists a family $\{C_r \mid r = 1, \dots, l\}$ of maximal cliques of $G(V, \bar{F})$ such that $F = \cup_{i=1}^{l} C_l \times C_l$ [1]. One of the necessary conditions for $\bar{X}_{ij}, (i,j) \in F$ to have a PDMC is that $\bar{X}_{C_r C_r}$ is positive definite for all $r = 1, \dots, l$. We call this condition the clique positive definite condition (CPDC). When $G(V, \bar{F})$ is chordal, it becomes a sufficient condition [7]. Moreover, [7] reported the following properties.

Theorem 1 **(a)** *$G(V, \bar{F})$ is a chordal graph if and only if $\bar{X}_{ij}, (i,j) \in F$ satisfying the CPDC has a PDMC.*

(b) *Suppose that $G(V, \bar{F})$ is a chordal graph and $\bar{X}_{ij}, (i,j) \in F$ satisfies the CPDC. Then a maximum-determinant PDMC of $\bar{X}_{ij}, (i,j) \in F$, i.e., a solution of*

$$\begin{array}{ll} \max & \det(X) \\ \text{subject to} & X_{ij} = \bar{X}_{ij}, \ \forall (i,j) \in F \\ & X = X^T \\ & X \succeq 0 \end{array}$$

is unique and $X^{-1}_{ij} = 0$ for all $(i,j) \notin F$.

Next, we consider how to compute the maximum-determinant PDMC of $\bar{X}_{ij}, (i,j) \in F$. To this end it is important to specify the family $\{C_r \mid r = 1, \ldots, l\}$ of maximal cliques of the chordal graph $G(V, \bar{F})$.

The chordal graph has a simplicial vertex [1]. Let the vertex be v_1. Then the subgraph induced on $V \backslash \{v_1\}$ becomes a chordal graph, and thus it has a simplicial vertex v_2. By repeating this process, we can construct an ordering $(v_1, v_2, \ldots, v_n)$. (We call such an ordering a perfect elimination ordering.) The maximal cliques can be enumerated from the ordering. Note that, since v_1 is simplicial, a maximal clique containing v_1 is given by $\{v_1\} \cup \mathrm{Adj}(v_1)$. Furthermore, a maximal clique not containing v_1 is a maximal clique of the subgraph induced on $\{v_2, v_3, \ldots, v_n\}$. Therefore, we can construct $\{C_r \subseteq V \mid r = 1, \ldots, l\}$ as

$$C_r = \{v_i\} \cup (Adj(v_i) \cap \{v_{i+1}, v_{i+2}, \ldots, v_n\})$$

for $i = \min\{j \mid v_j \in C_r\}$. Thus, maximal cliques $\{C_r \subseteq V \mid r = 1, \ldots, l\}$ can be computed within $O(n+m)$ by the maximum cardinality search [1], where m is the number of edges. Moreover, the maximal cliques can be indexed in such a way that for each $r = 1, 2, \ldots, l-1$, the following holds:

$$\exists s > r \text{ such that } C_r \cap (C_{r+1} \cup C_{r+2} \cdots \cup C_l) \subsetneq C_s.$$

This is called the running intersection property (RIP) and is easily obtained by using the clique tree [1].

Next, we suppose that $\{C_r \mid r = 1, \ldots, l\}$ are indexed as satisfying the RIP. Then, we can define the following families of subsets of $\{C_r\}$.

$$S_r = C_r \backslash (C_{r+1} \cup C_{r+2} \cup \cdots \cup C_l), \quad r = 1, \ldots, l \tag{5}$$

$$U_r = C_r \cap (C_{r+1} \cup C_{r+2} \cup \cdots \cup C_l), \quad r = 1, \ldots, l \tag{6}$$

By definition each element of S_1 is simplicial. Moreover, each element of S_{i+1} is a simplicial node of the graph induced from $S_i \cup S_{i+1} \cup \cdots \cup S_l$. Therefore, we can construct the perfect elimination ordering from $\{S_r \mid r = 1, \ldots, r\}$. Let P be the permutation matrix of this ordering. Then, the maximum-determinant PDMC of $\bar{X}_{ij}, (i,j) \in F$ is given as follows [5, Sparse clique-factorization formula (2.16)]:

$$X = P^T L_1^T L_2^T \cdots L_l^T D L_l L_{l-1} \cdots L_2 L_1 P, \tag{7}$$

where the factors $\{L_r\}$ and D are given by

$$[L_r]_{ij} = \begin{cases} 1 & i = j \\ (\bar{X}^{-1}_{U_r U_r} \bar{X}_{U_r S_r})_{ij} & (i,j) \in U_r \times S_r \\ 0 & \text{otherwise} \end{cases}$$

for $r = 1, \ldots, l-1$, and

$$D = \begin{pmatrix} D_{S_1 S_1} & & & \\ & D_{S_2 S_2} & & \\ & & \ddots & \\ & & & D_{S_l S_l} \end{pmatrix}$$

with

$$D_{S_r S_r} = \begin{cases} \bar{X}_{S_r S_r} - \bar{X}_{S_r U_r} \bar{X}^{-1}_{U_r U_r} \bar{X}_{U_r S_r} & r \leq l-1 \\ \bar{X}_{S_r S_r} & r = l. \end{cases}$$

3 Sparse quasi-Newton updates with positive definite matrix completion

In this section, we propose new sparse quasi-Newton updates.

Fletcher [3] showed that H^{DFP}_{k+1} is the unique solution of the following problem:

$$\begin{array}{ll} \min_H & \psi(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}}) \\ \text{subject to} & H y_k = s_k, H = H^T \\ & H \succeq 0, \end{array} \tag{8}$$

where $\psi : R^{n \times n} \to R$ is a strictly convex function defined by

$$\psi(A) = \text{trace}(A) - \ln \det(A). \tag{9}$$

When A is symmetric positive definite and its eigenvalues are $\lambda_i, i = 1, \ldots, n$, we have $\psi(A) = \sum_{i=1}^{n} (\lambda_i - \ln \lambda_i)$. Therefore, the minimum of ψ on $A \succeq 0$ is attained at $\lambda_i = 1, i = 1, \ldots, n$. This implies that $\psi(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}})$ denotes a kind of distance from H_k to H, and thus the solution H_{k+1} of (8) is the "nearest" positive semidefinite matrix satisfying the secant condition from H_k. On the other hand, B^{BFGS}_{k+1}, the inverse of H^{BFGS}_{k+1}, is a solution of the following problem [3]:

$$\begin{array}{ll} \min_B & \psi(H_k^{\frac{1}{2}} B H_k^{\frac{1}{2}}) \\ \text{subject to} & B s_k = y_k, B = B^T \\ & B \succeq 0. \end{array} \tag{10}$$

The above problems (8) and (10) do not include the information of the sparsity of the Hessian. If we exploit this information, we may construct a new approximate Hessian with less memory. Therefore, rather than (8), we consider the following problem:

$$\begin{array}{ll} \min_H & \psi(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}}) \\ \text{subject to} & H y_k = s_k, H = H^T \\ & (H^{-1})_{ij} = 0, (i,j) \notin F \\ & H \succeq 0, \end{array} \tag{11}$$

where $F \supseteq E = \{(i,j) \mid \nabla^2 f(x)_{i,j} \neq 0 \text{ for some } x \in R^n\}$. We refer to E as the sparsity pattern of the Hessian and F as an extension of E. (Of course, it is favourable to choose $F = E$, but certain properties of F are required, as will be discussed later.) Throughout the paper we assume that $(i,i) \in F$ for all $i \in V$ and that $(i,j) \in F$ if $(j,i) \in F$. Fletcher [4] considered the problem (10) with the sparsity conditions $B_{ij} = 0, (i,j) \notin F$, and proposed the use of its exact solution as B_{k+1}. Since the problem is a nonlinear convex programming problem, a great deal of time is required in order to obtain the exact solution. Moreover, as shown in Section 4, B_{k+1} sometimes becomes unstable due to the simultaneous requirement of the sparsity and secant conditions [11]. In this paper we consider the use of an approximate solution of (11) as H_{k+1} rather than the exact solution. More precisely, we propose the following new updates:

Step 1: Obtain a partial matrix $H_{ij}^{QN}, (i,j) \in F$ by using existing quasi-Newton updates, such as BFGS and DFP.

Step 2: Obtain a solution H_{k+1} of the following problem with $H_{ij}^{QN}, (i,j) \in F$ as given constants.

$$\begin{array}{ll} \min & \psi(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}}) \\ \text{subject to} & H_{ij} = H_{i,j}^{QN}, (i,j) \in F \\ & H = H^T \\ & (H^{-1})_{ij} = 0, (i,j) \notin F \\ & H \succeq 0 \end{array} \tag{12}$$

Remark 1 *If we use DFP in Step 1, then H^{QN} is a solution of problem (11) without the sparsity constraints, i.e., problem (8).*

Remark 2 *The secant condition $Hy_k = s_k$ in problem (11) is replaced with the constraints $H_{ij} = H_{i,j}^{QN}, (i,j) \in F$ in problem (12). Therefore, as shown in Section 4, H_{k+1} is stable even if $(s_k)_i = 0$ for some i.*

Remark 3 *When $F = V \times V$, the proposed updates are reduced to the existing quasi-Newton updates used in Step 1.*

Remark 4 *A matrix satisfying the constraints of (12) may be unique. However, we prefer to use the optimization formulation (12) for the subsequent analysis.*

The proposed method is illustrated in Figure 3.

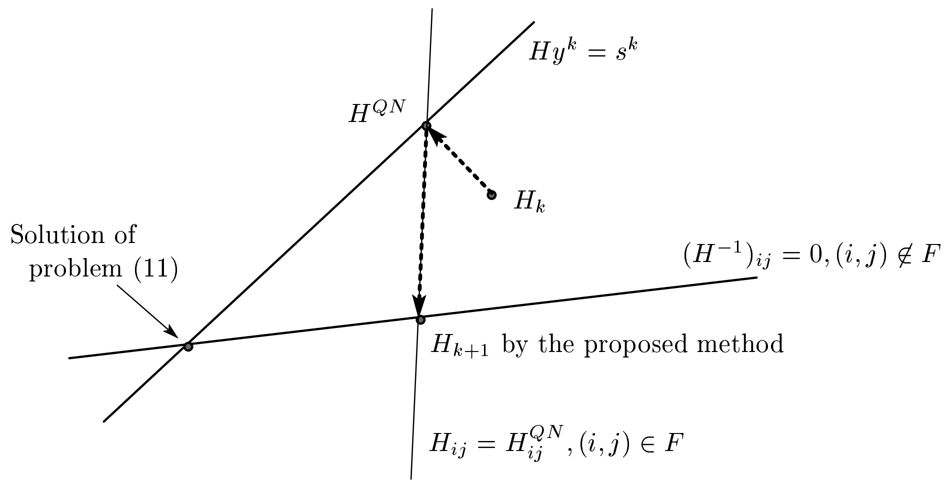


Figure 3: Proposed method

From the above remarks and Figure 3, the updated matrix H_{k+1} is regarded as a kind of approximate solution of (11). However, problem (12) still seems to be difficult. Fortunately, as shown below, if $G(V, \bar{F})$ is chordal, then problem (12) is equivalent to finding a maximum-determinant PDMC of $H_{ij}^{QN}, (i,j) \in F$, i.e.,

$$\begin{array}{ll} \max & \det(H) \\ \text{subject to} & H_{ij} = H_{i,j}^{QN}, (i,j) \in F \\ & H = H^T \\ & H \succeq 0. \end{array} \tag{13}$$

Theorem 2 *Suppose that $s_k^T y_k > 0$, H_k is symmetric positive definite and $(H_k^{-1})_{ij} = 0, \forall (i,j) \notin F$. If $G(V, \bar{F})$ is a chordal graph, then problem (12) is equivalent to problem (13). Moreover, the solution H_{k+1} of the problem (12) forms the sparse clique-factorization formula (7).*

Proof. We first show that problem (12) is equivalent to

$$\begin{array}{ll} \max & \det(H) \\ \text{subject to} & H_{ij} = H_{i,j}^{QN}, (i,j) \in F \\ & H = H^T \\ & (H^{-1})_{ij} = 0, (i,j) \notin F \\ & H \succeq 0. \end{array} \tag{14}$$

Since, based on the assumption, we have $(H_k^{-1})_{ij} = 0, \forall (i,j) \notin F$, and from the constraint of (12) we have $H_{ij} = H^{QN}_{i,j}, (i,j) \in F$, we have

$$\mathrm{trace}(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}}) = \mathrm{trace}(H H_k^{-1}) = \sum_{i=1}^{n}\sum_{j=1}^{n} H_{ij}(H_k^{-1})_{ji} = \sum_{i=1}^{n}\sum_{j \in F_i} H_{ij}(H_k^{-1})_{ij} = \sum_{i=1}^{n}\sum_{j \in F_i} H^{QN}_{ij}(H_k^{-1})_{ij},$$

which shows that $\mathrm{trace}(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}})$ is constant on the feasible set of (12). Moreover, we have

$$\ln\det(H_k^{-\frac{1}{2}} H H_k^{-\frac{1}{2}}) = 2\ln\det(H_k^{-\frac{1}{2}}) + \ln\det(H).$$

Therefore, problem (12) is equivalent to problem (14).

Next, we show that problem (14) is equivalent to problem (13). Suppose that $\{C_r \mid r = 1, \ldots, r\}$ is a family of maximal cliques of $G(V, \bar{F})$. Since $s_k^T y_k > 0$ and H_k is positive definite, H^{QN} is also positive definite. Therefore, the submatrices $H^{QN}_{C_r C_r}, r = 1, \ldots, l$ are positive definite, i.e., $H^{QN}_{ij}, (i,j) \in F$ satisfies the CPDC. The desired relation then follows from Theorem 1 (b). □

Remark 5 *Fletcher [4] showed that the problem (10) with the sparsity conditions $B_{ij} = 0, (i,j) \notin F$ can be efficiently solved by the Newton method if a factorization of B_k has no fill-in, which implies that $G(V, \bar{F})$ is chordal.*

We now describe the proposed method as follows:

Matrix Completion Quasi-Newton method (MCQN)

Step 0: Obtain an extension F of E such that $G(V, \bar{F})$ is chordal. Calculate a family $\{C_r \mid r = 1 \ldots, l\}$ of maximum cliques of $G(V, \bar{F})$, $\{S_r \mid r = 1, \ldots, l\}$ and $\{U_r \mid r = 1, \ldots, l\}$ by (5) and (6). Choose $x_0 \in R^n$ and a positive definite matrix H_0 with $(H_0^{-1})_{ij} = 0, \forall (i,j) \notin F$. Set $k = 0$.

Step 1: If x_k satisfies the termination criterion, then stop.

Step 2: $x_{k+1} = x_k - H_k \nabla f(x_k)$.

Step 3: Obtain $H^{QN}_{ij}, (i,j) \in F$ by the existing quasi-Newton updates.

Step 4: Obtain the sparse clique-factorization formula (7) of H_{k+1} with $\bar{X}_{ij} = H^{QN}_{ij}, (i,j) \in F$.

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

Next, we estimate the time and space complexities per iteration of MCQN. In order to obtain H^{QN}_{ij} in Step 3, we may employ the BFGS or DFP update formula. Let us assume the use of the BFGS. Step 3 is then calculated as follows:

$$H^{QN}_{i,j} = (H_k)_{i,j} + \rho s_i s_j - \frac{(H_k y_k)_i (s_k)_j + (s_k)_j (H_k y_k)_j}{s_k^T y_k} \quad \forall (i,j) \in F, \tag{15}$$

where

$$\rho = \frac{1}{s_k^T y_k} + \frac{(y_k)^T H_k y_k}{(s_k^T y_k)^2}.$$

We first estimate the time complexity per iteration. To compute $(H^{QN}_{U_rU_r})^{-1}$ for each r, we need $O(|C_r|^3)$ arithmetic operations. Therefore, the calculation of $H_k v$ for given $v \in R^n$ requires $O(\sum_{i=1}^{l} |C_r|^3)$ arithmetic operations, and thus the time complexity of Step 2 is $O(\sum_{i=1}^{l} |C_r|^3)$. In Step 3, we first calculate $H_k y_k$, then we compute $H^{QN}_{ij}, (i,j) \in F$. The calculation of $H_k y_k$ is $O(\sum_{i=1}^{l} |C_r|^3)$. Moreover, since $|F| \leq \sum_{r=1}^{l} |C_r|^2$, $O(\sum_{r=1}^{l} |C_r|^2)$ arithmetic operations are required for (15). Consequently, the time complexity of Step 3 is $O(\sum_{r=1}^{l} |C_r|^3)$. Step 4 is a dummy step because we compute the factorization (7) of H^{k+1} whenever we compute $H_k v$ for given v. Consequently, when $\nabla f(x_k)$ is given, the time complexity per iteration of MCQN is $O(\sum_{r=1}^{l} |C_r|^3)$. If we store $((H_k)_{U_rU_r})^{-1}$ for all $r = 1, \ldots, l$, we can reduce the time complexity to $O(\sum_{r=1}^{l} |C_r|^2)$. For clarification, note that

$$(H_{k+1})_{U_rU_r} = H^{QN}_{U_rU_r} = (H_k)_{U_rU_r} + \rho s_{U_r} s_{U_r}^T - \frac{(H_k y_k)_{U_r} s_{U_r}^T + s_{U_r}(H_k y)_{U_r}^T}{s_k^T y_k}.$$

Thus, using the Sherman-Morrison formula, we can compute $((H_{k+1})_{U_rU_r})^{-1}$ from $((H_k)_{U_rU_r})^{-1}$ within $O(|C_r|^2)$ arithmetic operations. By using the stored $((H_k)_{U_rU_r})^{-1}$, the time complexity of the computations of $H_k v$ becomes $O(\sum_{r=1}^{l} |C_r|^2)$.

Next, we estimate the space complexity. When we do not store $((H_{k+1})_{U_rU_r})^{-1}$ for all r, we only need to store $(H_k)_{ij}, (ij) \in F$. Therefore, the space complexity is $O(|F|)$. When we store $((H_{k+1})_{U_rU_r})^{-1}$ for each r, the space complexity becomes $O(\sum_{i=1}^{l} |C_r|^2)$.

When the Hessian is sparse, in general, C_r becomes much less than n. Since $l \leq n$, $\sum_{r=1}^{l} |C_r|^2$ is usually smaller than n^2. For example, as shown below, when the Hessian is tridiagonal, $l = n$ and $|C_r| = 2$ for all $r = 1, \ldots, n$. Then, the time and space complexities become $O(n)$.

In Step 0 of the proposed method, we must obtain the chordal extension $G(V, \bar{F})$ of $G(V, \bar{E})$. The problem of finding a minimum chordal extension of a general graph is NP complete. The minimum chordal extension is obtained via the minimum fill-in Cholesky factorization of a positive definite matrix with sparsity pattern E. Therefore, we may employ various existing heuristic methods, such as minimum degree ordering and nested dissection ordering, for the minimum fill-in Cholesky factorization. On the other hand, when the sparsity pattern E has a special structure, we can easily obtain the minimum chordal extension $G(V, \bar{F})$. The following are practical examples in which F becomes E [5].

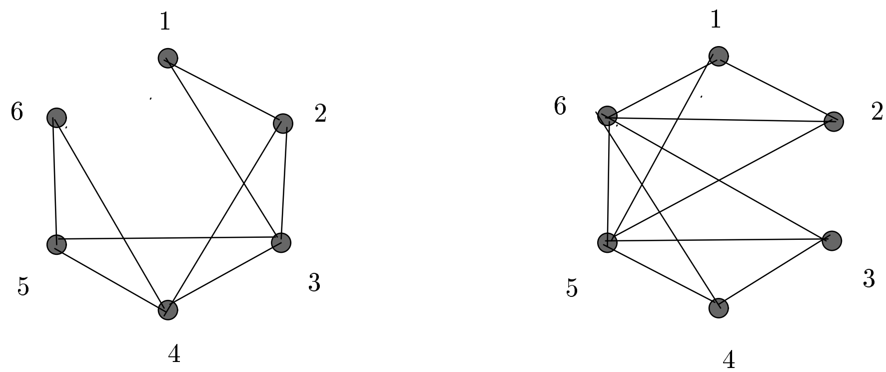
Multidiagonal: Suppose that a sparsity pattern E is given by $E = \{(i,j) \in V \times V \mid |i-j| \leq \beta\}$ with a positive integer β. Let

$$C_r = \{i \in V \mid (r-1)\kappa < i \leq \beta + r\kappa\}, r = 1, \ldots, l$$

with a positive integer κ and the smallest positive integer l satisfying $\beta + l\kappa \geq n$ and $F = \cup_{r=1}^{l} C_r \times C_r$. Then, $G(V, \bar{F})$ is chordal and $\{C_r \mid r = 1, \ldots, l\}$ is a family of its maximum cliques. Figure 3 (a) shows the case for $n = 6$ and $\beta = 2$, and we verify that the graph is chordal.

Note that the integer κ corresponds to $|C_r|$. Moreover, as κ becomes large, l becomes small and $|F|$ becomes large. If $\kappa = 1$, then $l = n - \beta$, $|C_r| = \beta$ and $F = E$.

Now let us consider (7) the case in which the Hessian is tridiagonal, i.e., $\beta = 1$ and $\kappa = 1$. In this case, we have $S_r = \{r\}, r = 1, \ldots, l-1$, $S_l = \{n-1, n\}$, $U_r = \{r+1\}, r = 1, \ldots, l-1$ and $U_l = \emptyset$.



(a) multidiagonal ($\beta = 2$) (b) borederd block-diagonal ($|S_r| = 2$)

Figure 4: Special cases of $G(V, \bar{F})$

Therefore, $L_l = I$ and $L_r, r = 1, \ldots, l-1$ are given by

$$[L_r]_{ij} = \begin{cases} 1 & i = j \\ H^{QN}_{r+1,r}/H^{QN}_{(r+1),(r+1)} & (i,j) = (r+1, r) \\ 0 & \text{otherwise} \end{cases}$$

and $D_{S_r S_r}$ are given by

$$D_{S_r S_r} = \begin{cases} H^{QN}_{r,r} - (H^{QN}_{r,r+1})^2/H^{QN}_{r+1,r+1} & r \le l-1 \\ H^{QN}_{S_r S_r} & r = l. \end{cases}$$

Therefore, we can compute all L_r and $D_{S_r S_r}$ with $O(1)$ arithmetic operations, and the space complexity is $O(1)$. For given v, we can compute Hv with $O(n)$ arithmetic operations.

Bordered block-diagonal Consider the case in which the Hessian has the following form:

$$\begin{pmatrix} B_{S_1 S_1} & 0 & \cdots & 0 & B_{S_1 S_0} \\ 0 & B_{S_2 S_2} & \cdots & 0 & B_{S_2 S_0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{S_l S_l} & B_{S_l S_0} \\ B_{S_0 S_1} & B_{S_0 S_2} & \cdots & B_{S_0 S_l} & B_{S_0 S_0}, \end{pmatrix}.$$

Let $C_r = S_0 \cup S_r$. Then $E = F = \cup_{r=1}^{l} C_r \times C_r$ and $G(\bar{F}, V)$ is a chordal graph. Now suppose that n is even and $|S_r| = 2, r = 0, \ldots, l$. Then we have $l = n/2 - 1$, $S_0 = \{n-1, n\}$, $S_r = \{2r-1, 2r\}$ and $U_r = S_0, r = 1, \ldots, l$ (the case in which $n = 6$ is illustrated in Figure 3 (b)). Therefore, $H^{QN}_{U_r U_r}$ becomes a 2×2 matrix for each r, and thus L_r and D_r can be calculated within $O(1)$ arithmetic operations. Consequently, the time complexity per iteration becomes $O(n)$.

4 Behavior of MCQN for Sorensen's example

In this section, we show the behavior of the proposed method for Sorensen's example [11]:

$$f(x) = \frac{1}{8}(x_1 - 1)^2 (x_1 + 1)^2 x_3^2 + x_2^2 + (x_2 - x_3)^2 \tag{16}$$

with $x_0 = (0, 0, \sqrt{432/55} - \varepsilon)^T, x_1 = (-5/6, 1, \sqrt{432/55})^T, \varepsilon = 10^{-6}$. As shown in [11, p.149], if the secant condition $B_1 s_0 = y_0$ is imposed, then

$$(B_1)_{13} = \frac{1 + 5(B_1)_{11}/6}{\varepsilon},$$

and thus numerical difficulty occurs. Therefore, most existing sparse quasi-Newton updates suffer from this problem.

The Hessian of f has the following form:

$$\begin{pmatrix} * & 0 & * \\ 0 & * & * \\ * & * & * \end{pmatrix}$$

Therefore, its sparsity pattern E is bordered diagonal, and thus $G(V, \bar{E})$ is chordal and its maximum cliques are $C_1 = \{1, 3\}$ and $C_2 = \{2, 3\}$. When B_0 is the identity matrix, the new matrix B_1 updated by MCQN with BFGS becomes

$$B_1 = \begin{pmatrix} 0.3421 & 0 & 0.2373 \\ 0 & 2.0629 & -1.7167 \\ 0.2373 & -1.7167 & 2.5931 \end{pmatrix}$$

This shows that the proposed method does not suffer from Sorensen's problem.

Next, we show the behavior of MCQN with BFGS for the solution of (16) in Table 1. (We employed the Armijo step size rule presented in Section 6 for global convergence.)

After nine iterations, the method obtains an approximate stationary point of f. Moreover, even if the true Hessians are singular (see $k = 7, 8, 9$), the approximate Hessians B_k are still positive definite and stable.

5 Local and superlinear convergence of MCQN with DFP

In this section, we show that the MCQN with DFP in Step 3 has local and superlinear convergence.

This is proven in a manner similar to [10, 8.4 Convergence Analysis], where the superlinear convergence of the BFGS method is demonstrated using the following property of the function ψ defined by (9).

$$0 < \psi(B_{k+1}^{BFGS}) \leq \psi(B_k) + \frac{y_k^T y_k}{y_k^T s_k} - \frac{\|B_k s_k\|^2}{s_k^T B_k s_k} - \ln \frac{y_k^T s_k}{\|s_k\|^2} + \ln \frac{s_k B_k s_k}{\|s_k\|^2}. \tag{17}$$

Iteration	B_k	$\nabla^2 f(x_k)$	$\|\nabla f(x_k)\|$
k=1	$\begin{pmatrix} 0.3421 & 0 & 0.2373 \\ 0 & 2.0629 & -1.7167 \\ 0.2373 & -1.7167 & 2.5931 \end{pmatrix}$	$\begin{pmatrix} 8.5091 & 0 & 0.7136 \\ 0 & 4 & -2 \\ 0.7136 & -2 & 2.0233 \end{pmatrix}$	4.13
k=2	$\begin{pmatrix} 3.6201 & 0 & -3.4288 \\ 0 & 1.5396 & -0.8149 \\ -3.4288 & -0.8149 & 3.5658 \end{pmatrix}$	$\begin{pmatrix} 26.3805 & 0 & -17.7974 \\ 0 & 4 & -2 \\ -17.7974 & -2 & 10.5672 \end{pmatrix}$	8.41
k=3	$\begin{pmatrix} 4.2031 & 0 & -1.3889 \\ 0 & 2.3174 & 0.6037 \\ -1.3889 & 0.6037 & 10.4431 \end{pmatrix}$	$\begin{pmatrix} 9.9069 & 0 & -16.6557 \\ 0 & 4 & -2 \\ -16.6557 & -2 & 22.3021 \end{pmatrix}$	6.07
k=4	$\begin{pmatrix} 2.0461 & 0 & 0.6472 \\ 0 & 2.3485 & 0.1128 \\ 0.6472 & 0.1128 & 10.4969 \end{pmatrix}$	$\begin{pmatrix} 3.3856 & 0 & -2.4441 \\ 0 & 4 & -2 \\ -2.4441 & -2 & 3.1845 \end{pmatrix}$	4.21
k=5	$\begin{pmatrix} 2.0584 & 0 & 1.3355 \\ 0 & 2.3740 & 0.3049 \\ 1.3355 & 0.3049 & 9.4283 \end{pmatrix}$	$\begin{pmatrix} -0.0516 & 0 & 0.0531 \\ 0 & 4 & -2 \\ 0.0532 & -2 & 2.2249 \end{pmatrix}$	1.51
k=6	$\begin{pmatrix} 2.1656 & 0 & 1.1738 \\ 0 & 2.1526 & 1.0876 \\ 1.1738 & 1.0876 & 8.2456 \end{pmatrix}$	$\begin{pmatrix} -0.003 & 0 & 0.0001 \\ 0 & 4 & -2 \\ 0.0001 & -2 & 2.25 \end{pmatrix}$	3.70E-1
k=7	$\begin{pmatrix} 2.1462 & 0 & 1.0924 \\ 0 & 2.0614 & 1.1157 \\ 1.0924 & 1.1157 & 8.2908 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & -2 & 2.25 \end{pmatrix}$	1.58E-2
k=8	$\begin{pmatrix} 2.0837 & 0 & 1.0747 \\ 0 & 1.5396 & 1.0563 \\ 1.0747 & 1.0563 & 8.2706 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & -2 & 2.25 \end{pmatrix}$	7.23E-4
k=9	$\begin{pmatrix} 2.0231 & 0 & 1.0629 \\ 0 & 2.0294 & 1.0483 \\ 1.0629 & 1.0483 & 8.2687 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & -2 & 2.25 \end{pmatrix}$	2.34E-5

Table 1: Behavior of MCQN for Sorensen's example

Here, $B_k = H_k^{-1}$ and $B_{k+1}^{BFGS} = (H_{k+1}^{BFGS})^{-1}$. Since MCQN updates H_k and (17) is the inequality for B_k, we cannot directly apply the proof technique to show the superlinear convergence of MCQN. Moreover, since H_{k+1} is the maximum-determinant PDMC of $H_{ij}^{QN}, (i,j) \in F$, we have $\det(H_{k+1}) \geq \det(H^{QN})$, and thus $\det(B_{k+1}) \leq \det(B^{QN})$ where $B^{QN} = (H^{QN})^{-1}$. Therefore, when we consider MCQN with BFGS in Step 3, i.e., $B^{QN} = B_{k+1}^{BFGS}$, it is difficult to derive inequalities like (17) due to the definition of ψ. Taking these difficulties into account, we consider MCQN with DFP because the update formula (4) of DFP has a form similar to that of B_{k+1}^{BFGS}. We will derive an inequality similar to (17) for H_{k+1} updated by MCQN with DFP.

For our purposes, the following assumptions are necessary:

Assumption 1 *Let x_* be a solution of (1) and let $\mathcal{C} = \{x \in R^n \mid \|x - x_*\| \leq b\}$ with a positive constant b.*

(i) *The objective function f is twice continuously differentiable on $\mathcal{C}$.*

(ii) *There exist positive constants m and M such that*

$$m\|z\|^2 \leq z^T(\nabla^2 f(x))^{-1} z \leq M\|z\|^2 \ \ \forall z \in R^n$$

for all $x \in \mathcal{C}$.

If the second-order sufficient optimality condition holds at the solution x_*, then Assumption 1 (ii) holds. From Assumption 1 (i), $\nabla^2 f(x)$ is Lipschitz continuous on $\mathcal{C}$. Then, from Lemmas 4.1.12 and 4.1.15 in [2], there exist L_1 and L_2 such that

$$\|y_k - \nabla^2 f(x_*)s_k\| \leq L_1\|s_k\|^2 \tag{18}$$

and

$$\|y_k - \nabla^2 f(x_*)s_k\| \leq L_2\varepsilon_k\|s_k\|, \tag{19}$$

where ε_k is defined by

$$\varepsilon_k = \max\{\|x_{k+1} - x_*\|, \|x_k - x_*\|\} \tag{20}$$

Moreover, from Eq. (8.12) of [10] we have

$$y_k = \bar{G}_k s_k, \tag{21}$$

where $\bar{G}_k$ is the average Hessian defined by $\bar{G}_k = \int_0^1 \nabla^2 f(x_k + ts_k)dt$.

For convenience in our analysis, we use the following notations, which are used in [10]:

$$G_* = \nabla^2 f(x_*), H_* = \nabla^2 f(x_*)^{-1},$$

$$\tilde{s}_k = H_*^{-1/2}s_k, \ \ \tilde{y}_k = H_*^{1/2}y_k, \ \ \tilde{H}_k = H_*^{-1/2}H_k H_*^{-1/2}, \ \ \tilde{H}^{QN} = H_*^{-1/2}H^{QN}H_*^{-1/2},$$

$$\cos\tilde{\theta}_k = \frac{\tilde{y}_k^T\tilde{H}_k\tilde{y}_k}{\|\tilde{y}_k\|\|\tilde{H}_k\tilde{y}_k\|}. \ \ \tilde{q}_k = \frac{\tilde{y}_k^T\tilde{H}_k\tilde{y}_k}{\|\tilde{y}_k\|^2},$$

$$\tilde{M}_k = \frac{\|s_k\|^2}{\tilde{y}_k^T\tilde{s}_k}, \ \ \tilde{m}_k = \frac{\tilde{y}_k^T\tilde{s}_k}{\tilde{y}_k^T\tilde{y}_k}.$$

We will make frequent use of the following inequality in our analysis:

$$h(t) := t - \ln t - 1 \geq 0 \quad \forall t > 0. \tag{22}$$

The inequality can be shown from the fact that h is strictly convex on $t > 0$, and its minimum is attained at $t = 1$.

First, we show the following two basic lemmas:

Lemma 1 *Suppose that Assumption 1 holds. Then there exists $c \in (0, \infty)$ and $\gamma \in (0, b)$ such that*

$$\ln\tilde{m}_k \geq -2c\varepsilon_k$$

$$\tilde{M}_k \leq 1 + c\varepsilon_k$$

whenever $\varepsilon_k < \gamma$.

Proof. Since

$$y_k - G_* s_k = (\bar{G}_k - G_*) s_k$$

from (21), we have

$$\begin{aligned} \tilde{y}_k - \tilde{s}_k &= G_*^{-1/2}(y_k - G_* s_k) \\ &= G_*^{-1/2}(\bar{G}_k - G_*) s_k \\ &= G_*^{-1/2}(\bar{G}_k - G_*) G_*^{-1/2} \tilde{s}_k. \end{aligned}$$

Thus, there exists a positive constant $\bar{c}$ such that

$$\|\tilde{y}_k - \tilde{s}_k\| \le \|G_*^{-1/2}\|^2 \|\tilde{s}_k\| \|\bar{G}_k - G_*\| \le \bar{c} \|\tilde{s}_k\| \varepsilon_k, \tag{23}$$

where the first inequality follows from the Cauchy-Schwartz inequality and the second inequality follows from the Lipschitz continuity of $\nabla^2 f$. It follows from the triangle inequalities $\pm\|y_k\| \mp \|s_k\| \le \|y_k - s_k\|$ that

$$(1 - \bar{c}\varepsilon_k)\|\tilde{s}_k\| \le \|\tilde{y}_k\| \le (1 + \bar{c}\varepsilon_k)\|\tilde{s}_k\|. \tag{24}$$

Moreover, squaring both sides of (23) and using (24), we have

$$\tilde{y}_k^T \tilde{s}_k \ge (1 - \bar{c}\varepsilon_k)\|\tilde{s}_k\|^2.$$

It then follows that

$$\tilde{m}_k = \frac{\tilde{y}_k^T \tilde{s}_k}{\|\tilde{y}_k\|^2} \ge \frac{(1 - \bar{c}\varepsilon_k)}{(1 + \bar{c}\varepsilon_k)^2}.$$

Suppose that γ is sufficiently small. Then, since $\varepsilon_k \le \gamma$, there exists a positive constant c_1 such that

$$\tilde{m}_k \ge 1 - c_1 \varepsilon_k. \tag{25}$$

In a similar manner, we can show that there exists a positive constant c_2 such that

$$\tilde{M}_k = \frac{\|\tilde{s}_k\|^2}{\tilde{y}_k^T \tilde{s}_k} \le \frac{1}{1 - \bar{c}\varepsilon_k} \le 1 + c_2 \varepsilon_k.$$

It then follows from (22) that

$$\frac{-c_1\varepsilon_k}{1 - c_1\varepsilon_k} - \ln(1 - c_1\varepsilon_k) = 1 - \frac{1}{1 - c_1\varepsilon_k} + \ln\left(\frac{1}{1 - c_1\varepsilon_k}\right) = h\left(\frac{1}{1 - c_1\varepsilon_k}\right) \le 0,$$

and thus

$$\frac{-c_1\varepsilon_k}{1 - c_1\varepsilon_k} \le \ln(1 - c_1\varepsilon_k) \tag{26}$$

Since we choose γ sufficiently small, we may suppose that $c_1\varepsilon_k < \frac{1}{2}$. Thus, from (26) we have

$$\ln(1 - c_1\varepsilon_k) \ge \frac{-c_1\varepsilon_k}{1 - c_1\varepsilon_k} \ge -2c_1\varepsilon_k.$$

It then follows from (25) that

$$\ln \tilde{m}_k \ge \ln(1 - c_1\varepsilon_k) \ge -2c_1\varepsilon_k.$$

Letting $c = \max\{c_1, c_2\}$, we have the desired inequalities. $\square$

Lemma 2 *Suppose that Assumption 1 holds and* $H^{QN} = H_{k+1}^{DFP}$*. Then we have*

$$\psi(\tilde{H}_{k+1}) \leq \psi(\tilde{H}^{QN}),$$

where ψ *is defined by (9).*

Proof. We investigate the determinant term and the trace term of ψ separately. Since H^{QN} is feasible for problem (13) and H_{k+1} is the unique maximizer of (13), we have $\det(H^{QN}) \leq \det(H_{k+1})$. Moreover, since $H_*^{-1/2}$ is positive definite by Assumption 1, we have

$$\begin{aligned} \det(\tilde{H}^{QN}) &= \det(H_*^{-1/2})\det(H^{QN})\det(H_*^{-1/2}) \\ &\leq \det(H_*^{-1/2})\det(H_{k+1})\det(H_*^{-1/2}) \\ &= \det(\tilde{H}_{k+1}). \end{aligned} \tag{27}$$

Next, we show that $\text{trace}(\tilde{H}^{QN}) = \text{trace}(\tilde{H}_{k+1})$. Note that $H_{ij}^{QN} = (H_{k+1})_{ij}, \forall(i,j) \in F$ and $(G_*)_{ij} = 0, \forall(i,j) \notin F$. Therefore, we have

$$\begin{aligned} \text{trace}(\tilde{H}^{QN}) &= \text{trace}(H_*^{-1/2} H^{QN} H_*^{-1/2}) \\ &= \text{trace}(H^{QN} G_*) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} H_{ij}^{QN}(G_*)_{ji} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} H_{ij}^{QN}(G_*)_{ij} \\ &= \sum_{i=1}^{n}\sum_{j\in F_i} H_{ij}^{QN}(G_*)_{ij} \\ &= \sum_{i=1}^{n}\sum_{j\in F_i} (H_{k+1})_{ij}(G_*)_{ij} \\ &= \text{trace}(H_{k+1}G_*) \\ &= \text{trace}(\tilde{H}_{k+1}). \end{aligned} \tag{28}$$

Combining (27) and (28), we have the desired inequality. □

By using the above lemmas, we show the following key inequality, which corresponds to (17).

Lemma 3 *Suppose that Assumption 1 holds and* $H^{QN} = H_{k+1}^{DFP}$*. Suppose also that* γ *is the constant specified in Lemma 1. If* $\varepsilon_k \leq \gamma$*, then we have*

$$\psi(\tilde{H}_{k+1}) + \ln\frac{1}{\cos^2\tilde{\theta}_k} - \left[1 - \frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k} + \ln\left(\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}\right)\right] \leq \psi(\tilde{H}_k) + 3c\varepsilon_k. \tag{29}$$

Proof. By Assumption 1 (ii), we have

$$\frac{y_k^T s_k}{y_k^T y_k} = \frac{y_k^T \bar{H}_k y_k}{y_k^T y_k} \geq m$$

and

$$\frac{y_k^T y_k}{y_k^T s_k} = \frac{z_k^T \bar{H}_k z_k}{z_k^T z_k} \leq M,$$

where $z_k = \bar{H}_k^{1/2} y_k$ and $\bar{H}_k = \bar{G}_k^{-1}$.

Since H^{QN} is obtained from H_k by the DFP formula (4), we have

$$\begin{aligned}
\tilde{H}^{QN} &= H_*^{-1/2} H^{QN} H_*^{-1/2} \\
&= H_*^{-1/2} H_k H_*^{-1/2} + H_*^{-1/2} \left(-\frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{s_k s_k^T}{y_k^T s_k} \right) H_*^{-1/2} \\
&= \tilde{H}_k - \frac{\tilde{H}_k H_*^{1/2} y_k y_k^T H_*^{1/2} \tilde{H}_k}{y_k^T H_*^{1/2} H_*^{-1/2} H_k H_*^{-1/2} H_*^{1/2} y_k} + \frac{H_*^{-1/2} s_k s_k^T H_*^{-1/2}}{y_k^T H_*^{1/2} H_*^{-1/2} s_k} \\
&= \tilde{H}_k - \frac{\tilde{H}_k \tilde{y}_k \tilde{y}_k^T \tilde{H}_k}{\tilde{y}_k^T \tilde{H}_k \tilde{y}_k} + \frac{\tilde{s}_k \tilde{s}_k^T}{\tilde{y}_k^T \tilde{s}_k}.
\end{aligned} \tag{30}$$

In a manner similar to the use of Eqs. (8.44) and (8.55) in [10], we can show that

$$\operatorname{trace}(\tilde{H}^{QN}) = \operatorname{trace}(\tilde{H}_k) - \frac{\|\tilde{H}_k y_k\|^2}{\tilde{y}_k \tilde{H}_k \tilde{y}_k} + \frac{\|\tilde{s}_k\|^2}{\tilde{y}_k^T \tilde{s}_k} \tag{31}$$

and

$$\det(\tilde{H}^{QN}) = \det(\tilde{H}_k) \frac{\tilde{y}_k^T \tilde{s}_k}{\tilde{y}_k \tilde{H}_k \tilde{y}_k}. \tag{32}$$

Moreover, by simple calculations, we have

$$\frac{\tilde{y}_k^T \tilde{s}_k}{\tilde{y}_k \tilde{H}_k \tilde{y}_k} = \frac{\tilde{y}_k^T \tilde{s}_k}{\|\tilde{y}_k\|^2} \frac{\|\tilde{y}_k\|^2}{\tilde{y}_k \tilde{H}_k \tilde{y}_k} = \frac{\tilde{m}_k}{\tilde{q}_k} \tag{33}$$

and

$$\frac{\|\tilde{H}_k y_k\|^2}{\tilde{y}_k \tilde{H}_k \tilde{y}_k} = \frac{\tilde{y}_k \tilde{H}_k \tilde{y}_k}{\|\tilde{y}_k\|^2} \frac{\|\tilde{H}_k y_k\|^2 \|\tilde{y}_k\|^2}{(\tilde{y}_k \tilde{H}_k \tilde{y}_k)^2} = \frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k}. \tag{34}$$

It then follows from (30)-(34) that

$$\begin{aligned}
\psi(\tilde{H}^{QN}) &= \operatorname{trace}(\tilde{H}^{QN}) - \det(\tilde{H}^{QN}) \\
&= \operatorname{trace}(\tilde{H}_k) + \tilde{M}_k - \frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} - \ln \det(\tilde{H}_k) - \ln \tilde{m}_k + \ln \tilde{q}_k \\
&= \psi(\tilde{H}_k) + \tilde{M}_k - \ln(\tilde{m}_k) - 1 + 1 - \frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} + \ln \left(\frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} \right) + \ln \cos^2 \tilde{\theta}_k.
\end{aligned}$$

Then, from Lemmas 1 and 2, we have

$$\psi(\tilde{H}_{k+1}) \leq \psi(\tilde{H}^{QN}) \leq \psi(\tilde{H}_k) + 3c\varepsilon_k + \ln\cos^2\tilde{\theta}_k + 1 - \frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k} + \ln\left(\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}\right),$$

and thus

$$\psi(\tilde{H}_{k+1}) + \ln\frac{1}{\cos^2\tilde{\theta}_k} - \left[1 - \frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k} + \ln\left(\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}\right)\right] \leq \psi(\tilde{H}_k) + 3c\varepsilon_k,$$

which is the desired inequality. □

By using the inequality (29), we will show the local and superlinear convergence. We first show the local convergence. To this end, we need the following relation between $\psi(\tilde{H}_k)$ and the distance $\|H_k - H_*\|$.

Lemma 4 *Suppose that Assumption 1 holds. Suppose also that $H \in R^{n\times n}$ is symmetric positive definite and $\tilde{H} = H_*^{-\frac{1}{2}} H H_*^{-\frac{1}{2}}$.*

(a) *Let $\mu_i, i = 1, \ldots, n$ be the eigenvalues of H. Then $\psi(H) = \sum_{i=1}^n(\mu_i - \ln\mu_i)$ and $\psi(H) - n \geq 0$.*

(b) *For any $\rho > 0$, there exists δ such that $\psi(\tilde{H}) - n < \delta$ implies $\|H - H_*\| < \rho$.*

(c) *For any $\delta > 0$, there exists ρ such that $\|H - H_*\| < \rho$ implies $\psi(\tilde{H}) - n < \delta$.*

Proof. To show (a), note that $\det(H) = \Pi_{i=1}^n \mu_i$ and $\mathrm{trace}(H) = \sum_{i=1}^n \mu_i$. Thus, we have $\psi(H) = \sum_{i=1}^n(\mu_i - \ln\mu_i)$. It then follows from (22) that $\psi(H) - n = \sum_{i=1}^n(\mu_i - \ln\mu_i - 1) \geq 0$.

Next, we show (b) and (c). Let $\lambda_i, i = 1, \ldots, n$ be the eigenvalues of $\tilde{H}$. We then have

$$\|\tilde{H} - I\| = \sqrt{\sum_{i=1}^n(\lambda_i - 1)^2}. \tag{35}$$

Moreover, since $\|H - H_*\| = \|H_*^{\frac{1}{2}}(\tilde{H} - I)H_*^{\frac{1}{2}}\|$ and H_* is positive definite, there exist positive constants a_1 and a_2 such that

$$a_1\|\tilde{H} - I\| \leq \|H - H_*\| \leq a_2\|\tilde{H} - I\|.$$

It then follows from (35) that

$$a_1\sqrt{\sum_{i=1}^n(\lambda_i - 1)^2} \leq \|H - H_*\| \leq a_2\sqrt{\sum_{i=1}^n(\lambda_i - 1)^2}. \tag{36}$$

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

$$0 \leq \psi(\tilde{H}) - n = \sum_{i=1}^n(\lambda_i - \ln\lambda_i - 1). \tag{37}$$

Since $h(t) = t - \ln t - 1$ is strictly convex and $h(1) = 0$, $\psi(\tilde{H}) - n \to 0$ implies that $\lambda_i \to 1$ for all i. Therefore, we have (b) and (c). □

Theorem 3 *Suppose that Assumption 1 holds and* $H^{QN} = H^{DFP}_{k+1}$. *Then, for any* $\alpha \in (0,1)$, *there exists* τ *such that* $\|x_0 - x_*\| \le \tau$ *and* $\|H_0 - \nabla^2 f(x_*)\| \le \tau$ *imply*

$$\|x_{k+1} - x_*\| \le \alpha \|x_k - x_*\|$$

for all k.

Proof. Suppose that $\alpha \in (0,1)$. We will show that the following inequalities hold for all k.

$$\|x_{k+1} - x_*\| \le \alpha \|x_k - x_*\| \tag{38}$$

$$\|H_k - \nabla^2 f(x_*)^{-1}\| \le \frac{\alpha}{2}. \tag{39}$$

First, note that by choosing τ to be sufficiently small, we have

$$L_1 M \tau < \frac{\alpha}{2}, \quad \tau \le \gamma \tag{40}$$

where L_1, M and γ are the constants specified in (18), Assumption 1 (ii) and Lemma 1, respectively. Moreover, by choosing τ to be sufficiently small, if necessary, from Lemma 4 (b) and (c), there exists δ such that

$$\psi(\tilde{H}_0) - n < \frac{\delta}{2}, \tag{41}$$

$$\psi(\tilde{H}) - n < \delta \implies \|H - \nabla^2 f(x_*)^{-1}\| \le \frac{\alpha}{2}. \tag{42}$$

and

$$\frac{3c\tau}{1-\alpha} \le \frac{\delta}{2}, \tag{43}$$

where H is a symmetric positive definite matrix, $\tilde{H} = H_*^{-\frac{1}{2}} H H_*^{-\frac{1}{2}}$, and c is the constant specified in Lemma 3.

We show the inequalities (38) and (39) by induction. When $k = 0$, the inequality (39) holds from (41) and (42). Moreover, we have

$$\begin{aligned}
\|x_1 - x_*\| &= \|x_0 - H_0 \nabla f(x_0) - x_*\| \\
&\le \|x_0 - x_* - \nabla^2 f(x_*)^{-1} \nabla f(x_0)\| + \|(H_0 - \nabla^2 f(x_*)^{-1})(x_0 - x_*)\| \\
&\le \|\nabla^2 f(x_*)^{-1}(\nabla f(x_*) - \nabla f(x_0) + \nabla^2 f(x_*)(x_0 - x_*))\| + \|H_0 - \nabla^2 f(x_*)^{-1}\| \|x_0 - x_*\| \\
&\le L_1 \|\nabla^2 f(x_*)^{-1}\| \|x_0 - x_*\|^2 + \frac{\alpha}{2} \|x_0 - x_*\| \\
&\le (L_1 M \tau + \frac{\alpha}{2}) \|x_0 - x_*\| \\
&\le \alpha \|x_0 - x_*\|,
\end{aligned}$$

where the third inequality follows from (18) and the last inequality follows from (40).

Next, we suppose that (38) and (39) hold for $k = 0, 1, \ldots, l$ and show the inequalities for $k = l + 1$. Similar to the case in which $k = 0$, we have

$$\begin{aligned}
\|x_{l+1} - x_*\| &= \|x_l - H_l \nabla f(x_l) - x_*\| \\
&\leq \|x_l - x_* - \nabla^2 f(x_*)^{-1} \nabla f(x_l)\| + \|(H_l - \nabla^2 f(x_*)^{-1})(x_l - x_*)\| \\
&\leq \|\nabla^2 f(x_*)^{-1} (\nabla f(x_*) - \nabla f(x_l) + \nabla^2 f(x_*)(x_l - x_*))\| + \|H_l - \nabla^2 f(x_*)^{-1}\| \|x_l - x_*\| \\
&\leq L_1 \|\nabla^2 f(x_*)^{-1}\| \|x_l - x_*\|^2 + \frac{\alpha}{2} \|x_l - x_*\| \\
&\leq (L_1 M \|x_l - x_*\| + \frac{\alpha}{2}) \|x_l - x_*\| \\
&\leq (L_1 M (\alpha)^l \tau + \frac{\alpha}{2}) \|x_l - x_*\| \\
&\leq \alpha \|x_l - x_*\|,
\end{aligned}$$

where the fifth inequality follows from the fact that $\|x_l - x_*\| \leq (\alpha)^l \|x_0 - x_*\|$. This shows (38) for $k = l + 1$. Next, we show (39) by using (29) in Lemma 3. Summing up the inequalities (29) with $k = 0, 1, \ldots, l$, we have

$$\psi(\tilde{H}_{l+1}) + \sum_{k=0}^{l} \left(\ln \frac{1}{\cos^2 \tilde{\theta}_k} - \left[1 - \frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} + \ln \left(\frac{\tilde{q}_k}{\cos^2 \tilde{\theta}_k} \right) \right] \right) \leq \psi(\tilde{H}_0) + 3c \sum_{k=0}^{l} \varepsilon_k.$$

Since $0 < \cos \tilde{\theta}_k \leq 1$ and the term in the square brackets is nonpositive by (22), we have

$$\psi(\tilde{H}_{l+1}) - n \leq \psi(\tilde{H}_0) - n + 3c \sum_{k=0}^{l} \varepsilon_k. \tag{44}$$

From (38), we have

$$\varepsilon_k = \|x_{k+1} - x_*\| \leq (\alpha)^{k+1} \tau$$

for $k = 0, \ldots, l$, and thus

$$\sum_{k=0}^{l} \varepsilon_k \leq \frac{1 - (\alpha)^{l+1}}{1 - \alpha} \tau \leq \frac{\tau}{1 - \alpha}.$$

It then follows from (44), (41) and (43) that

$$\begin{aligned}
\psi(\tilde{H}_{l+1}) - n &\leq \psi(\tilde{H}_0) - n + \frac{3c\tau}{1 - \alpha} \\
&\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.
\end{aligned}$$

From (42) we have $\|H_{l+1} - \nabla^2 f(x_*)^{-1}\| \leq \frac{\alpha}{2}$, which is (39) for $k = l + 1$. □

Next, we show the superlinear convergence. The following are the sufficient conditions for the superlinear convergence of quasi-Newton methods [10].

$$\lim_{k \to \infty} \frac{\|(B_k - G_*) s_k\|}{\|s_k\|} = 0. \tag{45}$$

By using (29) and Theorem 3 we will show that

$$\lim_{k\to\infty} \frac{\|(H_k - H_*)y_k\|}{\|y_k\|} = 0. \tag{46}$$

In order to show the superlinear convergence, the following relation between (46) and the superlinear convergence condition (45) is necessary.

Lemma 5 *Suppose that Assumption 1 holds and that $H^{QN} = H^{DFP}_{k+1}$. Suppose also that $\|x_0 - x_*\| \le \tau$ and $\|H_0 - \nabla^2 f(x_*)\| \le \tau$ with the constant τ specified in Theorem 3 for sufficiently small $\alpha \in (0,1)$. Then (46) implies (45).*

Proof. Let $\lambda_i^k, i = 1,\dots,n$ be the eigenvalues of H_k. Since the inequality (39) holds for sufficiently small α, we may assume that there exists $\lambda_{\min} > 0$ such that $\lambda_i^k \ge \lambda_{\min}$ for all i and k.

From Assumption 1 (i) there exists a positive constant L_3 such that

$$\|y_k\| = \|\nabla f(x_{k+1}) - \nabla f(x_k)\| \le L_3\|s_k\| \quad \text{for all } k. \tag{47}$$

Moreover, since $y_k = G_* s_k + (\bar{G}_k - G_*)s_k$ from (21), we have

$$\begin{aligned}
\|(H_k - H_*)y_k\| &= \|(H_k - H_*)G_* s_k + (H_k - H_*)(\bar{G}_k - G_*)s_k\| \\
&\ge \|H_k(G_* - B_k)s_k\| - \|H_k - H_*\|\|\bar{G}_k - G_*\|\|s_k\| \\
&\ge \lambda_{\min}\|(B_k - G_*)s_k\| - \|H_k - H_*\|\|\bar{G}_k - G_*\|\|s_k\|.
\end{aligned}$$

It then follows from (47) that

$$\frac{\|(H_k - H_*)y_k\|}{\|y_k\|} \ge \frac{\lambda_{\min}\|(B_k - G_*)s_k\|}{L_3\|s_k\|} - \frac{\|H_k - H_*\|\|\bar{G}_k - G_*\|}{L_3}.$$

Since $\bar{G}_k = \int_0^1 \nabla^2 f(x_k + ts_k)dt$ and $x_k \to x_*$ by Theorem 3, the second term of the right-hand side of the inequality converges to 0 as $k \to \infty$. It then follows from (46) that

$$\lim_{k\to\infty} \frac{\|(B_k - G_*)s_k\|}{\|s_k\|} = 0,$$

which is the desired inequality. □

We can now show the main result of this section.

Theorem 4 *Suppose that Assumption 1 holds. Suppose also that $\|x_0 - x_*\| \le \tau$ and $\|H_0 - \nabla^2 f(x_*)^{-1}\| \le \tau$ hold for sufficiently small $\tau > 0$. The sequence $\{x_k\}$ generated by MCQN with DFP then converges to x_* superlinearly.*

Proof. From Lemma 5 it is sufficient to show (46). Summing up the inequalities (29) in Lemma 3, we have

$$\sum_{k=0}^{\infty}\left(\ln\frac{1}{\cos^2\tilde{\theta}_k} - \left[1 - \frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k} + \ln\left(\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}\right)\right]\right) \le \psi(\tilde{H}_0) + 3c\sum_{k=0}^{\infty}\varepsilon_k < \infty,$$

where the first inequality follows from the fact that $\psi(\tilde{H}_k) > 0$ for all k, and the last inequality follows from the local linear convergence of $\{x_k\}$ (Theorem 3). Since $0 < \cos\tilde{\theta}_k \le 1$, $\ln(1/\cos^2\tilde{\theta}_k)$ must be nonnegative. Moreover, the term in square brackets is nonpositive. Therefore, we have

$$\lim_{k\to\infty} \cos\tilde{\theta}_k = 1, \quad \lim_{k\to\infty} \tilde{q}_k = 1. \tag{48}$$

Furthermore, we have

$$\begin{aligned}\frac{\|H_*^{-1/2}(H_k - H_*)y_k\|^2}{\|H_*^{1/2}y_k\|^2} &= \frac{\|(\tilde{H}_k - I)\tilde{y}_k\|^2}{\|\tilde{y}_k\|^2} \\ &= \frac{\|\tilde{H}_k\tilde{y}_k\|^2 - 2\tilde{y}_k^T\tilde{H}_k\tilde{y}_k + \|\tilde{y}_k\|^2}{\|\tilde{y}_k\|^2} \\ &= \frac{\tilde{q}_k^2}{\cos^2\tilde{\theta}_k} - 2\tilde{q}_k + 1,\end{aligned}$$

where the last equality follows from the fact that

$$\frac{\tilde{q}_k^2}{\cos^2\tilde{\theta}_k} = \frac{\left(\tilde{y}_k^T\tilde{H}_k\tilde{y}_k\right)^2}{\|\tilde{y}_k\|^4}\frac{\|\tilde{y}_k\|^2\|\tilde{H}_k\tilde{y}_k\|^2}{\left(\tilde{y}_k^T\tilde{H}_k\tilde{y}_k\right)^2} = \frac{\|\tilde{H}_k\tilde{y}_k\|^2}{\|\tilde{y}_k\|^2}.$$

It then follows from (48) and the positive definiteness of H_* that we have the desired inequality (46). □

As in the proofs, we can show the superlinear convergence under Assumption 1 and the assumptions that (a) $\{H_k\}$ is uniformly positive definite and (b) $\sum_{k=0}^{\infty}\varepsilon_k < \infty$. The assumptions on the initial data, i.e., the assumptions that $\|x_0 - x_*\| \le \tau$ and $\|H_0 - \nabla^2 f(x_*)^{-1}\| \le \tau$ hold for sufficiently small τ, are sufficient conditions for (a) and (b).

6 Numerical Experiments

In this section, we report numerical results for the proposed method MCQN as well as for the existing BFGS and L-BFGS methods.

We solved the following problems:

Problem 1: $f(x) = \sum_{i=1}^{n-1} i(x_{i+1} - x_i)^2$

Problem 2: $f(x) = \sum_{i=1}^{n-1} \sin(x_{i+1} - x_i)$

Problem 3: $f(x) = (n+1)(x_{n-1}^2 + x_n^2) + \sum_{i=1}^{(n-2)/2}(x_{2i-1}^2 + x_{2i-1}x_{2i} + ix_{2i}^2 + x_{2i-1}x_{n-1} + x_{2i}x_n)$

Problem 4: $f(x) = \sum_{i=1}^{(n-2)/2}(\sin(x_{2i-1} - x_{2i}) + (x_{2i-1} - x_{n-1})^2 + (x_{2i} - x_n)^2)$

Problem	n	BFGS	L-BFGS	MCQN with DFP	MCQN with BFGS
Problem 1	10	16.4	76.7	30.1	26.7
	100	115.2	1363.2	123.1	122.9
	1000	1028.5	F	815.1	785.2
Problem 2	10	18.4	33.3	20.8	20.4
	100	115.7	309.5	88.5	85.2
	1000	1021.1	2373.6	533.5	498.3
Problem 3	10	17.5	80.3	27.3	22.3
	100	69.3	234.7	126.2	99.2
	1000	646.3	742.7	1030.9	786.0
Problem 4	10	19.5	38.2	20.3	27.6
	100	27.8	81.7	32.3	37.3
	1000	29.7	F	103.0	108.3

Table 2: MCQN vs. existing methods

Problems 1 and 3 are convex quadratic minimization problems, and Problems 2 and 4 are nonconvex and nonlinear. The conditions numbers of Problems 1 and 3 are very large, and thus they are ill-conditioned. The sparsity patterns of Problems 1 and 2 are tridiagonal, whereas those of Problems 3 and 4 are bordered block-diagonal. Therefore, we can easily obtain the chordal extensions of their sparsity pattern.

We employed the following termination criterion:

$$\|\nabla f(x_k)\| \leq \varepsilon \text{ or } k \geq 5000.$$

The second criterion implies that the method fails to obtain a solution. We set $\varepsilon = 10^{-5}$ for nonconvex problems, Problems 2 and 4, and $\varepsilon = 10^{-6}$ for other problems. Since Wolfe's rule has a high cost, we employed the Armijo rule:

$$f(x_k + (0.5)^{t_k} d_k) - f(x_k) \leq 0.001(0.5)^{2t_k} \nabla f(x_k)^T d_k,$$

to obtain a step size t_k and set $x_{k+1} = x_k - t_k H_k \nabla f(x_k)$. In order to guarantee the positive definiteness of H_k, we set $H_{k+1} = H_k$ if $s_k^T y_k$ is less than 2.2×10^{-15}. The initial points are randomly chosen from $[-10, 10]^n$. We set $m = 5$, which is the number of stored vectors of L-BFGS. The algorithms were implemented in Matlab 6.1.

We solved problems of various dimensions, i.e., $n = 10, 100, 1000$ by MCQN with BFGS, BFGS and L-BFGS. The results are listed in Table 2. The table lists the total number of iterations (average of 10 independent runs), and the symbol" F" denotes that the number is over 5000.

Table 2 shows that the number of iterations of MCQN is almost equal to or less than those of the existing methods. In particular, L-BFGS converged very slowly for the ill-posed problem (Problem 1), whereas MCQN converged after far fewer iterations for this problem.

	BFGS	L-BFGS	MCQN (with DFP)
Positive definiteness	Yes	Yes	Yes
Secant condition	Yes	Yes	No
Sparsity pattern	No	No	Yes
Rate of convergence	superlinear	linear	superlinear

Table 3: Comparison between MCQN and existing updates

7 Concluding remarks

In this paper, we proposed the quasi-Newton methods using the PDMC. We showed that the method with DFP has local and superlinear convergence under the usual assumptions. The method requires fewer space and time complexities than those for existing quasi-Newton methods that have superlinear convergence. In particular, when the Hessian has a special structure, such as multidiagonal or bordered block-diagonal, as discussed in Section 4, the complexities are drastically decreased. The simple numerical results suggest that the proposed method is very promising. The properties of the proposed method, compared with existing methods, are summarized in Table 3.

Investigation of the sparse quasi-Newton update with PDMC is a relatively new area of research. We are considering the following future research topics.

- How do we obtain a sparsity pattern E of the Hessian ?

  It is not easy for non-specialists to obtain the sparsity pattern E and thus it is important to construct an automatic procedure for obtaining E. To this end, automatic differentiation may be useful.

- How do we obtain an appropriate chordal extension ?

  In general, the problem of finding a minimum chordal extension is NP complete. We may use a heuristic method, such as the minimum degree ordering, for this problem. However, minimum degree ordering sometimes adds many edges (fill-in) in order to obtain the chordal extension. We note that, although we need $F \supseteq E$ for superlinear convergence, F is not required be a superset of E in practice. For example, if F is multidiagonal, i.e., $F = \{(i,j) \in V \times V \mid |i-j| \leq \beta\}$ regardless of the value of E, then the sparse clique-factorization (7) can be computed quickly, even though the number of iterations may increase. An investigation of the value of F that is efficient for practical situations would be useful.

- Can we show the superlinear convergence of MCQN with BFGS?

  We have shown that MCQN with DFP has superlinear convergence. Generally, BFGS is faster than DFP. Therefore, it is important to show the superlinear convergence of MCQN with BFGS.

- Can we extend MCQN to the constrained minimization problem ?

  The interior point method and the sequential quadratic programming method are widely used for solving constrained minimization problems. These methods usually exploit approximate Hessians of the Lagrange function, which are generated by the BFGS update formula. We may employ MCQN for such algorithms.

Acknowledgement

The author would like to thank Prof. M. Fukushima of Kyoto University for his constructive comments and suggestions, which helped to improve the readability of our paper.

References

[1] J.R.S. Blair and B. Peyton, An introduction to chordal graphs and clique trees, in Graph Theory and Sparse Matrix Computation, A. George, J.R. Gilbert and J.W.H. Liu, eds., pp. 1–29, Springer-Verlag, New York, 1993.

[2] J.E. Dennis, Jr. and R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Inc., New Jersey, 1983.

[3] R. Fletcher, A new result for quasi-Newton formulae, SIAM Journal on Optimization, 1, pp. 18–21, 1991.

[4] R. Fletcher, An optimal positive definite update for sparse Hessian matrices, SIAM Journal on Optimization, 5, pp. 192–218, 1995.

[5] M. Fukuda, M. Kojima, K. Murota and K. Nakata, Exploiting sparsity in semidefinite programming via matrix completion I: General framework, SIAM Journal on Optimization, 11, pp. 647–674, 2000.

[6] A. George and J.W.H. Liu, Computer solution of large sparse positive definite systems, Prentice-Hall, Englewood Cliff, NJ, 1981.

[7] R. Grone, C.R. Johnson, E. M. Sá and H. Wolkowicz, Positive definite completions of partial Hermitian matrices, Linear Algebra and its Applications., 58, pp. 109–124, 1984.

[8] D.C. Liu and J. Nocedal, On the limited memory BFGS method for large scale optimization, Mathematical Programming, 45, pp. 503-528, 1989.

[9] J. Nocedal, Updating quasi-Newton matrices with limited storage, Mathematics of Computation, 35, pp. 773-782, 1980.

[10] J. Nocedal and S.J. Wright, Numerical Optimization, Springer-Verlag, New York, 1999.

[11] D.C. Sorensen, Collinear scaling and sequential estimation in sparse optimization algorithm, Mathematical Programming Study, 18, pp. 135–159, 1982.

[12] P.L. Toint, On sparse and symmetric matrix updating subject to a linear equation, Mathematics of Computation, 31, pp. 954–961, 1977.