Research Article

A New BFGS Algorithm Using the Decomposition Matrix of the Correction Matrix to Obtain the Search Directions

Li Liu

Teaching Department, Jilin City Radio and TV University, Jilin 132002, China

Correspondence should be addressed to Li Liu; liuli2484768@163.com

Received 30 September 2014; Revised 3 January 2015; Accepted 3 January 2015

Academic Editor: Shengwei Mei

Copyright © 2015 Li Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present an improved method for determining the search direction in the BFGS algorithm. Our approach uses the equal inner product decomposition method for positive-definite matrices. The decomposition of an approximated Hessian matrix expresses a correction formula that is independent from the exact line search. This decomposed matrix is used to compute the search direction in a new BFGS algorithm.

1. Introduction

Using the equal inner product decomposition algorithm [1–3] for positive-definite matrices with a vector, a novel quasi-Newton method has been proposed [4, 5]. In this algorithm, the equal inner product decomposition matrix of the correction matrix is utilized to obtain the search directions. That is, for the unrestricted optimization problem

$$\min f(\mathbf{x}), \quad \mathbf{x} \in R^n, \quad (1)$$

where $f: R^n \rightarrow R^1$ is a continuously differentiable function; the equal inner product decomposition correction formula for the approximate Hessian matrix with the objective function gradient vector is used to compute the search directions.

The BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm [6–9] is a typical quasi-Newton method. If we consider $\mathbf{x}^k$ to be an iteration point, and $\mathbf{s}_k = \mathbf{x}^{k+1} - \mathbf{x}^k$, and $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$, $\mathbf{g}_k = \nabla f(\mathbf{x}^k)$, then the correction formulae for the approximate Hessian matrix $\mathbf{B}$ and its inverse $\mathbf{H}$ can be expressed as

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{s}_k \mathbf{s}_k^T \mathbf{B}_k}{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}, \quad (2)$$

$$\mathbf{H}_{k+1} = \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) \mathbf{H}_k \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}. \quad (3)$$

In computing the search direction $\mathbf{d}^{k+1}$ at the iteration point $\mathbf{x}^{k+1}$, we generally solve the equation:

$$\mathbf{B}_{k+1} \mathbf{d}^{k+1} = -\mathbf{g}_{k+1}, \quad (4)$$

but we would rather avoid using (3) to calculate

$$\mathbf{d}^{k+1} = -\mathbf{H}_{k+1} \mathbf{g}_{k+1}. \quad (5)$$

There are clearly more computations involved in (4) than in (5). We take this apparently more complex route because $\mathbf{H}_{k+1}$ may appear to be non-positive-definite or even singular when we calculate (3). Equation (2) can be transformed to

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \left((\mathbf{d}^k)^T \mathbf{g}_k \right)^{-1} \mathbf{g}_k \mathbf{g}_k^T + \left(\alpha_k (\mathbf{d}^k)^T \mathbf{y}_k \right)^{-1} \mathbf{y}_k \mathbf{y}_k^T. \quad (6)$$

Using this form, Gill and Murray [10] developed a method for solving (4) by Cholesky decomposition. First, they found the Cholesky decomposition of $\mathbf{B}_{k+1}$:

$$\mathbf{B}_{k+1} = \mathbf{L}_{k+1} \mathbf{D}_{k+1} \mathbf{L}_{k+1}^T, \quad (7)$$

where $\mathbf{L}_{k+1}$ is a unit lower triangular matrix and $\mathbf{D}_{k+1}$ is a diagonal matrix, and then they successively solved

$$\mathbf{L}_{k+1} \mathbf{v} = -\mathbf{g}_{k+1}, \quad \mathbf{L}_{k+1}^T \mathbf{d}^{k+1} = \mathbf{D}_{k+1}^{-1} \mathbf{v}. \quad (8)$$

Because (6) establishes the iterative formula from $\mathbf{L}_k$ and $\mathbf{D}_k$ to $\mathbf{L}_{k+1}$ and $\mathbf{D}_{k+1}$, the Cholesky decomposition of $\mathbf{B}_k$ need not be repeated on each iteration. The detailed calculation method is described in [10–12].

Unlike the above approach, the algorithm proposed in [4, 5] utilizes the equal inner product decomposition algorithm for a positive-definite matrix with a vector. In this improved technique, the matrix $\mathbf{C}_{k+1}^{-1}$ is computed. This is the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1} = \nabla f(\mathbf{x}^{k+1})$; that is,

$$\mathbf{B}_{k+1} = \mathbf{C}_{k+1} \mathbf{C}_{k+1}^T, \quad \mathbf{C}_{k+1}^{-1} \mathbf{g}_{k+1} = a_{k+1} \mathbf{e}, \quad (9)$$

where $a_{k+1} \in R^1$, $\mathbf{e} = (1, 1, \dots, 1)^T \in R^n$. The search direction at the iteration point $\mathbf{x}^{k+1}$ is then given by

$$\mathbf{d}^{k+1} = -a_{k+1} (d_1^{k+1}, \dots, d_n^{k+1})^T, \quad (10)$$

where d_j^{k+1} ($j = 1, 2, \dots, n$) is the sum of the elements in the j th column of $\mathbf{C}_{k+1}^{-1}$.

Instead of using the correction formula (2), the method of [4, 5] directly establishes the correction formula from $\mathbf{C}_k^{-1}$ to $\mathbf{C}_{k+1}^{-1}$, which are equal inner product decomposition matrixes of $\mathbf{B}$ with the objective function gradient vectors. When the correction matrix $\mathbf{C}_{k+1}^{-1}$ is calculated, the search direction $\mathbf{d}^{k+1}$ is immediately given by (10).

Thus, [4, 5] presented a novel approach to the computation of search directions. However, all the conclusions in [4, 5] relate to the exact line search condition. Therefore, this algorithm is feasible in theory but has restricted applicability.

The BFGS algorithm with the Wolfe line search has global, superlinear convergence [13–15]. Compared with other quasi-Newton methods, BFGS is the most popular and efficient algorithm. In this paper, we focus on the BFGS algorithm and set up a correction formula expressed by the decomposition matrix that is independent of the exact line search. This leads to a new BFGS algorithm that computes the search directions using the equal inner product decomposition matrix of the approximate Hessian matrix.

In Section 3, the procedures of computing search directions in the usual BFGS and in the new BFGS have been compared by a computing example.

2. Correction Formula Expressed by Decomposition Matrix

At the iteration point $\mathbf{x}^k$, we use the method of [3] to directly compute the equal inner product decomposition matrix $\mathbf{C}_k^{-1}$ of $\mathbf{B}_k$ with $\mathbf{g}_k$. However, this decomposition method is computationally expensive. One alternative is to set up a similar correction formula between $\mathbf{C}_k^{-1}$ and $\mathbf{C}_{k+1}^{-1}$ as that between $\mathbf{H}_k$ and $\mathbf{H}_{k+1}$ or between $\mathbf{B}_k$ and $\mathbf{B}_{k+1}$.

We denote $\mathbf{u}_k = \mathbf{C}_k^{-1} \mathbf{y}_k$, $\mathbf{v}_k = \mathbf{C}_k^{-1} \mathbf{g}_{k+1}$, and $\mathbf{w}_k = \mathbf{C}_k^{-1} \mathbf{g}_k$. By Theorem 2.2 in [4], if the equal inner product decomposition matrix $\mathbf{C}_k^{-1}$ of $\mathbf{B}_k$ with $\mathbf{g}_k$ has already been computed in the exact line search condition for the BFGS correction formula

(2), then the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1}$ can be expressed as

$$\mathbf{C}_{k+1}^{-1} = \mathbf{Q}_k \mathbf{C}_k^{-1} \left[\mathbf{I} - \frac{(\mathbf{y}_k - \sqrt{\lambda_k} \mathbf{g}_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right], \quad (11)$$

where $\lambda_k = \mathbf{s}_k^T \mathbf{y}_k / \|\mathbf{w}_k\|^2$, $\mathbf{Q}_k = \mathbf{I} - (2/\|\sigma_k\|^2) \sigma_k \sigma_k^T$, $\sigma_k = \mathbf{w}_k + t_k \mathbf{v}_k$, and $t_k = \|\mathbf{w}_k\| / \|\mathbf{v}_k\|$. By Theorem 2.3 in [4], the equal inner product decomposition matrix of the initial correction matrix $\mathbf{B}_0 = \mathbf{I}$ with $\mathbf{g}_0$ is

$$\mathbf{C}_0^{-1} = \begin{cases} \mathbf{I}, & \mathbf{g}_0 = -\mathbf{e}, \\ \mathbf{I} - \frac{2}{\|\sigma_0\|^2} \sigma_0 \sigma_0^T, & \mathbf{g}_0 \neq -\mathbf{e}, \end{cases} \quad (12)$$

where $\sigma_0 = \sqrt{n} \mathbf{g}_0 + \|\mathbf{g}_0\| \mathbf{e}$.

As defined by (11), $\mathbf{C}_{k+1}^{-1}$ only satisfies $\mathbf{B}_{k+1} = \mathbf{C}_{k+1} \mathbf{C}_{k+1}^T$ under the exact line search condition, that is, when $\mathbf{v}_k^T \mathbf{w}_k = 0$. If the iteration point from $\mathbf{x}^k$ to $\mathbf{x}^{k+1}$ is obtained through an inexact line search, then obviously $\mathbf{C}_{k+1}^{-1}$ in (11) is no longer the decomposition matrix of $\mathbf{B}_{k+1}$ defined by (2). The following conclusions are independent of the exact line search but are dependent on the correction matrix definition for BFGS.

2.1. Main Results

Lemma 1. Suppose $\mathbf{u}, \mathbf{v} \in R^n$. If $\beta \neq 0$, $\gamma \neq 0$, $\mathbf{v}^T \mathbf{u} = \beta^{-1} + \gamma^{-1}$, then the matrix $\mathbf{I} - \beta \mathbf{u} \mathbf{v}^T$ is nonsingular and

$$(\mathbf{I} - \beta \mathbf{u} \mathbf{v}^T)^{-1} = \mathbf{I} - \gamma \mathbf{u} \mathbf{v}^T. \quad (13)$$

The proof is simple and is therefore omitted from this paper.

Lemma 2. For the BFGS correction formula (11), if $\mathbf{C}_k^{-1}$ is the equal inner product decomposition matrix of $\mathbf{B}_k$ with $\mathbf{g}_k$, we denote

$$\mathbf{N}_k = \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T} \mathbf{g}_k} \mathbf{y}_k \mathbf{s}_k^T - \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \mathbf{C}_k. \quad (14)$$

Then, $\mathbf{B}_{k+1} = \mathbf{N}_k \mathbf{N}_k^T$ and

$$\mathbf{N}_k^{-1} = \mathbf{C}_k^{-1} \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{g}_k \mathbf{s}_k^T}{\|\mathbf{w}_k\| \sqrt{\mathbf{s}_k^T \mathbf{y}_k}} \right). \quad (15)$$

Proof. Because $\mathbf{s}_k = \mathbf{x}^{k+1} - \mathbf{x}^k = \alpha_k \mathbf{d}^k$, where α_k is the step length, we have $\mathbf{B}_k \mathbf{s}_k = -\alpha_k \mathbf{g}_k$. By a direct calculation, we have

$$\begin{aligned}
\mathbf{N}_k \mathbf{N}_k^T &= \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k \mathbf{s}_k^T - \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \\
&\quad \cdot \mathbf{B}_k \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{s}_k \mathbf{y}_k^T - \frac{\mathbf{s}_k \mathbf{g}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \\
&= \left(\mathbf{B}_k - \frac{\alpha_k \|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k \mathbf{g}_k^T + \frac{\alpha_k \mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \\
&\quad \cdot \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{s}_k \mathbf{y}_k^T - \frac{\mathbf{s}_k \mathbf{g}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \\
&= \mathbf{B}_k + \frac{\alpha_k \mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} - \frac{\alpha_k \|\mathbf{w}_k\|^2}{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k} \mathbf{y}_k \mathbf{y}_k^T \\
&= \mathbf{B}_k + \frac{\alpha_k \mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \\
&= \mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{s}_k \mathbf{s}_k^T \mathbf{B}_k}{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}.
\end{aligned} \tag{16}$$

Therefore, $\mathbf{B}_{k+1} = \mathbf{N}_k \mathbf{N}_k^T$ holds. Since

$$\begin{aligned}
\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k \mathbf{s}_k^T - \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \\
= \mathbf{I} - \left(-\frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k + \frac{1}{\mathbf{s}_k^T \mathbf{g}_k} \mathbf{g}_k \right) \mathbf{s}_k^T,
\end{aligned} \tag{17}$$

Lemma 1 implies

$$\begin{aligned}
\left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k \mathbf{s}_k^T - \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right)^{-1} \\
= \mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{g}_k \mathbf{s}_k^T}{\|\mathbf{w}_k\| \sqrt{\mathbf{s}_k^T \mathbf{y}_k}}.
\end{aligned} \tag{18}$$

Thus, we have $\mathbf{N}_k^{-1} = \mathbf{C}_k^{-1} (\mathbf{I} - (\mathbf{y}_k \mathbf{s}_k^T / \mathbf{s}_k^T \mathbf{y}_k) + (\mathbf{g}_k \mathbf{s}_k^T / \|\mathbf{w}_k\| \sqrt{\mathbf{s}_k^T \mathbf{y}_k}))$. This completes the proof. $\square$

Lemma 3. Writing $\mathbf{Q}_k = \mathbf{I} - (2\sigma_k \sigma_k^T / \|\sigma_k\|^2)$, where $\sigma_k = \delta_k + (\|\delta_k\| / \|\mathbf{w}_k\|) \mathbf{w}_k$, if $\mathbf{0} \neq \delta_k \in \mathbb{R}^n$, one has $\mathbf{Q}_k \delta_k = -(\|\delta_k\| / \|\mathbf{w}_k\|) \mathbf{w}_k$.

Proof. A direct computation gives $2\sigma_k^T \delta_k = \|\sigma_k\|^2$. Thus, we have

$$\mathbf{Q}_k \delta_k = \delta_k - \frac{2\sigma_k^T \delta_k}{\|\sigma_k\|^2} \sigma_k = \delta_k - \sigma_k = -\frac{\|\delta_k\|}{\|\mathbf{w}_k\|} \mathbf{w}_k. \tag{19}$$

This completes the proof. $\square$

Writing

$$\delta_k = \mathbf{N}_k^{-1} \mathbf{g}_{k+1} = p_k \mathbf{v}_k + q_k \mathbf{w}_k, \tag{20}$$

where $p_k = -\mathbf{s}_k^T \mathbf{g}_k / \mathbf{s}_k^T \mathbf{y}_k$, $q_k = (\|\mathbf{w}_k\| + \sqrt{\mathbf{s}_k^T \mathbf{y}_k}) \mathbf{s}_k^T \mathbf{g}_{k+1} / \|\mathbf{w}_k\| \mathbf{s}_k^T \mathbf{y}_k$, we can state the following results.

Theorem 4. For the BFGS correction formula (2), if the equal inner product decomposition matrix $\mathbf{C}_k^{-1}$ of $\mathbf{B}_k$ with $\mathbf{g}_k$ has already been obtained, then the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1}$ is

$$\mathbf{C}_{k+1}^{-1} = \mathbf{Q}_k \mathbf{C}_k^{-1} \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{g}_k \mathbf{s}_k^T}{\|\mathbf{w}_k\| \sqrt{\mathbf{s}_k^T \mathbf{y}_k}} \right), \tag{21}$$

where $\mathbf{Q}_k = \mathbf{I} - (2/\|\sigma_k\|^2) \sigma_k \sigma_k^T$ and $\sigma_k = \delta_k + (\|\delta_k\| / \|\mathbf{w}_k\|) \mathbf{w}_k$.

Proof. By Lemma 2 and (21), we have

$$\mathbf{C}_{k+1} = \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k \mathbf{s}_k^T - \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \mathbf{C}_k \mathbf{Q}_k^{-1}. \tag{22}$$

Because $\mathbf{Q}_k$ is the Householder matrix, $\mathbf{Q}_k \mathbf{Q}_k^T = \mathbf{I}$. Thus, we have

$$\begin{aligned}
\mathbf{C}_{k+1} \mathbf{C}_{k+1}^T &= \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{y}_k \mathbf{s}_k^T - \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \\
&\quad \cdot \mathbf{B}_k \left(\mathbf{I} + \frac{\|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k}} \mathbf{s}_k \mathbf{y}_k^T - \frac{\mathbf{s}_k \mathbf{g}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right).
\end{aligned} \tag{23}$$

Therefore, Lemma 2 gives $\mathbf{B}_{k+1} = \mathbf{C}_{k+1} \mathbf{C}_{k+1}^T$. By Lemma 3, we get

$$\mathbf{C}_{k+1}^{-1} \mathbf{g}_{k+1} = \mathbf{Q}_k \delta_k = -\frac{\|\delta_k\|}{\|\mathbf{w}_k\|} \mathbf{w}_k = a_{k+1} \mathbf{e}, \tag{24}$$

where $a_{k+1} = -(\|\delta_k\| / \|\mathbf{w}_k\|) a_k$. Thus, $\mathbf{C}_{k+1}^{-1}$ defined by (21) is the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1}$. This completes the proof. $\square$

From the above proof, it can be found that $\mathbf{C}_{k+1}^{-1}$ in (21) is the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ defined by (2), which is independent of the exact line search.

Corollary 5. Under the exact line search condition, (21) is equivalent to (11); that is, (11) is a special case for (21).

Proof. Under the exact line search condition, we have $\mathbf{s}_k^T \mathbf{g}_{k+1} = 0$ and $\mathbf{s}_k^T \mathbf{g}_k = -\mathbf{s}_k^T \mathbf{y}_k$.

From (20), we have

$$\delta_k = \mathbf{N}_k^{-1} \mathbf{g}_{k+1} = \mathbf{v}_k. \tag{25}$$

Thus, $\sigma_k = \delta_k + (\|\delta_k\|/\|\mathbf{w}_k\|)\mathbf{w}_k = (\|\mathbf{v}_k\|/\|\mathbf{w}_k\|)(\mathbf{w}_k + t_k\mathbf{v}_k) = (\|\mathbf{v}_k\|/\|\mathbf{w}_k\|)\tilde{\sigma}_k$, where $\tilde{\sigma}_k = \mathbf{w}_k + t_k\mathbf{v}_k$.

Therefore, $\mathbf{Q}_k = \mathbf{I} - (2\sigma_k\sigma_k^T/\|\sigma_k\|^2) = \mathbf{I} - (2\tilde{\sigma}_k\tilde{\sigma}_k^T/\|\tilde{\sigma}_k\|^2)$.

Under the exact line search condition, $\|\mathbf{w}_k\|/\sqrt{\mathbf{s}_k^T\mathbf{y}_k} = (1/\sqrt{\lambda_k})\mathbf{s}_k^T\mathbf{y}_k$. From the above results, (11) can be directly deduced from (21). This completes the proof. $\square$

2.2. Extension. We can extend the conclusion of Theorem 4 to the following BFGS correction formula using self-scaling [16–18]:

$$\mathbf{B}_{k+1} = \eta_k \left(\mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{s}_k \mathbf{s}_k^T \mathbf{B}_k}{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k} \right) + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}, \quad (26)$$

where $\eta_k = \theta(\mathbf{s}_k^T \mathbf{y}_k / \mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k) + (1 - \theta)(\mathbf{y}_k^T \mathbf{B}_k^{-1} \mathbf{y}_k / \mathbf{s}_k^T \mathbf{y}_k)$ and $\theta \in [0, 1]$. When $\theta = 1$, $\eta_k = \mathbf{s}_k^T \mathbf{y}_k / \mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k$, so (26) can be written as

$$\mathbf{B}_{k+1} = \frac{\mathbf{s}_k^T \mathbf{y}_k}{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k} \left(\mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{s}_k \mathbf{s}_k^T \mathbf{B}_k}{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k} \right) + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}. \quad (27)$$

Equation (27) is well known for its practical applicability. Here, we have only transformed (27) into the form expressed by the decomposition correction matrix. When $\theta \neq 1$, similar methods can be used.

Applying the Sherman-Morrison formula [13] to (27), we can deduce the correction matrix for $\mathbf{H}$:

$$\mathbf{H}_{k+1} = \frac{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{y}_k} \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) \mathbf{H}_k \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}. \quad (28)$$

Writing $\tilde{\delta}_k = \tilde{p}_k \mathbf{v}_k + \tilde{q}_k \mathbf{w}_k$, where $\tilde{p}_k = -\mathbf{s}_k^T \mathbf{g}_k / \mathbf{s}_k^T \mathbf{y}_k$ and $\tilde{q}_k = (\mathbf{s}_k^T \mathbf{g}_{k+1})^2 / \mathbf{s}_k^T \mathbf{y}_k \mathbf{s}_k^T \mathbf{g}_k$, we can state the following for the inexact line search condition.

Theorem 6. For the self-scaling BFGS correction formula (27), if the equal inner product decomposition matrix $\mathbf{C}_k^{-1}$ of $\mathbf{B}_k$ with $\mathbf{g}_k$ is already known, then the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1}$ is

$$\mathbf{C}_{k+1}^{-1} = \frac{\alpha_k \|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k}} \mathbf{Q}_k \mathbf{C}_k^{-1} \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right), \quad (29)$$

where $\mathbf{Q}_k = \mathbf{I} - (2/\|\sigma_k\|^2)\sigma_k\sigma_k^T$, $\sigma_k = \tilde{\delta}_k + (\|\tilde{\delta}_k\|/\|\mathbf{w}_k\|)\mathbf{w}_k$, and $a_{k+1} = -(\alpha_k \|\tilde{\delta}_k\|/\sqrt{\mathbf{s}_k^T \mathbf{y}_k})a_k$ for step length α_k .

Proof. The proof is similar to that for Theorem 4.

Using (29), we have

$$\begin{aligned} (\mathbf{C}_{k+1}^{-1})^T \mathbf{C}_{k+1}^{-1} &= \frac{\alpha_k^2 \|\mathbf{w}_k\|^2}{\mathbf{s}_k^T \mathbf{y}_k} \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{s}_k \mathbf{g}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right) \\ &\quad \cdot \mathbf{H}_k \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{g}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right). \end{aligned} \quad (30)$$

Because of

$$\begin{aligned} \alpha_k^2 \|\mathbf{w}_k\|^2 &= \alpha_k^2 (\mathbf{C}_k^{-1} \mathbf{g}_k)^T (\mathbf{C}_k^{-1} \mathbf{g}_k) = \alpha_k^2 \mathbf{g}_k^T \mathbf{H}_k \mathbf{g}_k \\ &= (-\alpha_k \mathbf{g}_k^T) (-\alpha_k \mathbf{H}_k \mathbf{g}_k) = \mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k, \\ \frac{\mathbf{s}_k \mathbf{g}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \mathbf{H}_k \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) &= \frac{-\mathbf{s}_k \mathbf{s}_k^T}{\alpha_k \mathbf{s}_k^T \mathbf{g}_k} \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) = 0, \\ \frac{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k \mathbf{g}_k^T \mathbf{H}_k \mathbf{g}_k}{(\mathbf{s}_k^T \mathbf{g}_k)^2} &= \frac{\alpha_k^2 (\mathbf{g}_k^T \mathbf{H}_k \mathbf{g}_k)^2}{(\mathbf{s}_k^T \mathbf{g}_k)^2} = 1, \end{aligned} \quad (31)$$

(30) gives

$$\begin{aligned} (\mathbf{C}_{k+1}^{-1})^T \mathbf{C}_{k+1}^{-1} &= \frac{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{y}_k} \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) \mathbf{H}_k \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \\ &= \mathbf{H}_{k+1}. \end{aligned} \quad (32)$$

Therefore, $\mathbf{B}_{k+1} = \mathbf{C}_{k+1} \mathbf{C}_{k+1}^T$ and we have

$$\begin{aligned} \mathbf{C}_{k+1}^{-1} \mathbf{g}_{k+1} &= \frac{\alpha_k \|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k}} \mathbf{Q}_k \tilde{\delta}_k = \frac{\alpha_k \|\mathbf{w}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k}} \left(-\frac{\|\tilde{\delta}_k\|}{\|\mathbf{w}_k\|} \mathbf{w}_k \right) \\ &= -\frac{\alpha_k \|\tilde{\delta}_k\|}{\sqrt{\mathbf{s}_k^T \mathbf{y}_k}} a_k \mathbf{e}, \end{aligned} \quad (33)$$

that is, $\mathbf{C}_{k+1}^{-1} \mathbf{g}_{k+1} = a_{k+1} \mathbf{e}$, where $a_{k+1} = -(\alpha_k \|\tilde{\delta}_k\|/\sqrt{\mathbf{s}_k^T \mathbf{y}_k})a_k$. This completes the proof. $\square$

Theorem 6 admits the following corollary.

Corollary 7. Under the exact line search condition, if the equal inner product decomposition matrix $\mathbf{C}_k^{-1}$ of $\mathbf{B}_k$ with $\mathbf{g}_k$ for the self-scaling BFGS correction formula (27) is already known, then the equal inner product decomposition matrix of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1}$ is

$$\mathbf{C}_{k+1}^{-1} = \sqrt{\alpha_k} \mathbf{Q}_k \mathbf{C}_k^{-1} \left(\mathbf{I} + \frac{\mathbf{g}_{k+1} \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{g}_k} \right), \quad (34)$$

where $\mathbf{Q}_k = \mathbf{I} - (2/\|\sigma_k\|^2)\sigma_k\sigma_k^T$, $\sigma_k = \mathbf{v}_k + (\|\mathbf{v}_k\|/\|\mathbf{w}_k\|)\mathbf{w}_k$, and $a_{k+1} = -\sqrt{\alpha_k}(\|\mathbf{v}_k\|/\|\mathbf{w}_k\|)a_k$ for step length α_k .

Proof. The proof is simple and is therefore omitted. $\square$

3. A New BFGS Algorithm

3.1. Algorithm. Replacing the corresponding steps of the BFGS algorithm by (12), (21), and (10) establishes a new BFGS algorithm that uses the decomposition correction matrix to determine the search direction.

The New BFGS Algorithm

Step 1. Choose a start point $\mathbf{x}^0$, $\varepsilon \geq 0$, and compute $\mathbf{g}_0 = g(\mathbf{x}^0)$. If $\|\mathbf{g}_0\| \leq \varepsilon$, terminate the algorithm; otherwise, go to Step 2.

Step 2. Let $k = 0$. For the initial matrix $\mathbf{B}_0 = \mathbf{I}$, form the equal inner product decomposition with $\mathbf{g}_0$ to give the decomposition matrix $\mathbf{C}_0^{-1}$ and $a_0 = -\|\mathbf{g}_0\|/\sqrt{n}$.

Step 3. Compute the search direction $\mathbf{d}^k$ as

$$\mathbf{d}^k = -a_k (\mathbf{d}_1^k, \dots, \mathbf{d}_n^k)^T, \quad (35)$$

where $\mathbf{d}_j^k$, ($j = 1, 2, \dots, n$) is the sum of the elements in the j th column of $\mathbf{C}_k^{-1}$.

Step 4. Perform a one-dimensional line search (exact line search or Wolfe line search), and compute the step length factor α_k .

Step 5. Let $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha_k \mathbf{d}^k$. Compute $\mathbf{g}_{k+1} = g(\mathbf{x}^{k+1})$. If $\|\mathbf{g}_{k+1}\| \leq \varepsilon$, terminate the algorithm; otherwise, go to Step 6.

Step 6. Using (21), calculate the equal inner product decomposition matrix $\mathbf{C}_{k+1}^{-1}$ of $\mathbf{B}_{k+1}$ with $\mathbf{g}_{k+1}$ and the coefficient $a_{k+1} = -(\|\delta_k\|/\|\mathbf{w}_k\|)a_k$.

Step 7. Let $k = k + 1$ and go to Step 3.

3.2. Example. Here, by an example, we compare the two methods of computing the search directions between the usual BFGS algorithm [10–12] and the new BFGS algorithm. Considering the optimization problem [19],

$$\min_{\mathbf{x} \in \mathbb{R}^2} f(\mathbf{x}) = 2x_1^2 + x_2^2 - 4x_1 + 2, \quad (36)$$

when solving the problems with BFGS algorithm, from the second iterative computation, all the following iteration procedures are similar, so the difference of the two algorithms can be clearly seen by one iterative computing procedure. Here we assume that the first iterative computing procedure is finished and at the first iterative computing, we choose the initial point $\mathbf{x}^0 = (2, 1)^T$, the initial matrix $\mathbf{B}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Noting $g(\mathbf{x}) = \nabla f(\mathbf{x}) = \begin{pmatrix} 4x_1 - 4 \\ 2x_2 \end{pmatrix}$, the initial searching direction is

$$\mathbf{d}^0 = -\mathbf{g}_0 = -g(\mathbf{x}^0) = (-4, -2)^T. \quad (37)$$

At the point $\mathbf{x}^0$, we have done the Wolfe-Powell search along the $\mathbf{d}^0$ and obtained the step length $\alpha_0 = 1/3$ which satisfies

the Wolfe-Powell conditions [13]. Thus we get

$$\begin{aligned} \mathbf{x}^1 &= \mathbf{x}^0 + \alpha_0 \mathbf{d}^0 = \left(\frac{2}{3}, \frac{1}{3}\right)^T, \\ \mathbf{g}_1 &= g(\mathbf{x}^1) = \left(-\frac{4}{3}, \frac{2}{3}\right)^T, \\ \mathbf{s}_0 &= \mathbf{x}^1 - \mathbf{x}^0 = \left(-\frac{4}{3}, -\frac{2}{3}\right)^T, \\ \mathbf{y}_0 &= \mathbf{g}_1 - \mathbf{g}_0 = \left(-\frac{16}{3}, -\frac{4}{3}\right)^T. \end{aligned} \quad (38)$$

Now we compare the two different algorithms of computing search directions by the second iterative computation.

We first compute the search direction $\mathbf{d}^1$ with the usual BFGS algorithm.

Step 1. Compute Cholesky decomposition of $\mathbf{B}_1$.

Obviously, the Cholesky decomposition of $\mathbf{B}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is

$$\begin{aligned} \mathbf{B}_0 &= \mathbf{L}_0 \mathbf{D}_0 \mathbf{L}_0^T = \begin{pmatrix} 1 & 0 \\ l_{21}^{(0)} & 1 \end{pmatrix} \begin{pmatrix} d_{11}^{(0)} & 0 \\ 0 & d_{22}^{(0)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ l_{21}^{(0)} & 1 \end{pmatrix}^T \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \quad (39)$$

Let $\bar{\mathbf{B}}_1 = \mathbf{B}_0 + (1/\mathbf{s}_0^T \mathbf{y}_0) \mathbf{y}_0 \mathbf{y}_0^T$, definite $\gamma_1 = 1/\mathbf{s}_0^T \mathbf{y}_0 = 1/8$, and $\omega^{(1)} = \mathbf{y}_0 = (-16/3, -4/3)^T$.

Compute the following in turn [13]:

$$\begin{aligned} l_1 &= \omega_1^{(1)} = -\frac{16}{3}, \\ \bar{d}_{11} &= d_{11}^{(0)} + \gamma_1 l_1^2 = 1 + \frac{1}{8} \times \left(-\frac{16}{3}\right)^2 = \frac{41}{9}, \\ \beta_1 &= \frac{l_1 \gamma_1}{\bar{d}_{11}} = -\frac{6}{41}, \\ \gamma_2 &= \frac{d_{11}^{(0)} \gamma_1}{\bar{d}_{11}} = \frac{9}{328}, \end{aligned} \quad (40)$$

$$\omega_2^{(2)} = \omega_2^{(1)} - l_1 l_{21}^{(0)} = \left(-\frac{4}{3}\right) - \left(-\frac{16}{3}\right) \times 0 = -\frac{4}{3},$$

$$\bar{l}_{21} = l_{21}^{(0)} + \beta_1 \omega_2^{(2)} = 0 + \left(-\frac{6}{41}\right) \left(-\frac{4}{3}\right) = \frac{8}{41},$$

$$\bar{d}_{22} = d_{22}^{(0)} + \gamma_2 l_2^2 = 1 + \frac{9}{328} \left(-\frac{4}{3}\right)^2 = \frac{43}{41}.$$

Hence,

$$\bar{\mathbf{L}}_1 = \begin{pmatrix} 1 & 0 \\ \bar{l}_{21} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{8}{41} & 1 \end{pmatrix},$$

$$\bar{\mathbf{D}}_1 = \begin{pmatrix} \bar{d}_{11} & 0 \\ 0 & \bar{d}_{22} \end{pmatrix} = \begin{pmatrix} \frac{41}{9} & 0 \\ 0 & \frac{43}{41} \end{pmatrix},$$

$$\bar{\mathbf{B}}_1 = \bar{\mathbf{L}}_1 \bar{\mathbf{D}}_1 \bar{\mathbf{L}}_1^T = \begin{pmatrix} 1 & 0 \\ 8 & 1 \\ 41 & 1 \end{pmatrix} \begin{pmatrix} 41 & 0 \\ 9 & 43 \\ 0 & 41 \end{pmatrix} \begin{pmatrix} 1 & 8 \\ 0 & 41 \\ 0 & 1 \end{pmatrix}. \quad (41)$$

Then let $\mathbf{B}_1 = \bar{\mathbf{B}}_1 - (\mathbf{B}_0 \mathbf{s}_0 \mathbf{s}_0^T \mathbf{B}_0 / \mathbf{s}_0^T \mathbf{B}_0 \mathbf{s}_0) = \bar{\mathbf{B}}_1 - (1/\|\mathbf{s}_0\|^2) \mathbf{s}_0 \mathbf{s}_0^T$. Computing the equation

$$\bar{\mathbf{L}}_1 \mathbf{z} = \frac{1}{\|\mathbf{s}_0\|} \mathbf{s}_0, \quad (42)$$

we get $\mathbf{z} = (z_1, z_2)^T = (-2/\sqrt{5}, -5\sqrt{5}/41)^T$.

Compute the following in turn [13]:

$$\begin{aligned} t_3 &= 1 - \mathbf{z}^T \bar{\mathbf{D}}_1^{-1} \mathbf{z} = 1 - \left(-\frac{2}{\sqrt{5}}, -\frac{25}{41\sqrt{5}} \right) \\ &\quad \cdot \begin{pmatrix} \frac{9}{41} & 0 \\ 0 & 41 \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ -\frac{25}{41\sqrt{5}} \end{pmatrix} = \frac{162}{215}, \\ t_2 &= t_3 + \frac{z_2^2}{\bar{d}_{22}} = \frac{162}{215} + \frac{(25/41\sqrt{5})^2}{43/41} = \frac{169}{205}, \\ t_1 &= t_2 + \frac{z_1^2}{\bar{d}_{11}} = \frac{169}{205} + \frac{(-2/\sqrt{5})^2}{41/9} = 1, \\ d_{11}^{(1)} &= \frac{\bar{d}_{11} t_2}{t_1} = \frac{41}{9} \times \frac{169}{205} = \frac{169}{45}, \\ d_{22}^{(1)} &= \frac{\bar{d}_{22} t_3}{t_2} = \frac{162}{169}, \\ \beta_1 &= -\frac{z_1}{\bar{d}_{11} t_2} = -\frac{-2/\sqrt{5}}{(41/9) \times (169/205)} = \frac{18\sqrt{5}}{169}, \\ \omega_2^{(2)} &= z_2 = -\frac{5\sqrt{5}}{41}, \\ l_{21}^{(1)} &= \bar{l}_{21} + \beta_1 \omega_2^{(2)} = \frac{8}{41} + \frac{18\sqrt{5}}{169} \times \left(-\frac{5\sqrt{5}}{41} \right) = \frac{22}{169}. \end{aligned} \quad (43)$$

Thus we obtain

$$\begin{aligned} \mathbf{L}_1 &= \begin{pmatrix} 1 & 0 \\ l_{21}^{(1)} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 22 & 1 \\ 169 & 1 \end{pmatrix}, \\ \mathbf{D}_1 &= \begin{pmatrix} d_{11}^{(1)} & 0 \\ 0 & d_{22}^{(1)} \end{pmatrix} = \begin{pmatrix} 169 & 0 \\ 45 & 162 \\ 0 & 169 \end{pmatrix}. \end{aligned} \quad (44)$$

Therefore, we get the *Cholesky* decomposition of $\mathbf{B}_1$:

$$\mathbf{B}_1 = \mathbf{L}_1 \mathbf{D}_1 \mathbf{L}_1^T = \begin{pmatrix} 1 & 0 \\ 22 & 1 \\ 169 & 1 \end{pmatrix} \begin{pmatrix} 169 & 0 \\ 45 & 162 \\ 0 & 169 \end{pmatrix} \begin{pmatrix} 1 & 22 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \quad (45)$$

Step 2. Solve the following equations:

$$\mathbf{L}_1 \mathbf{v} = -\mathbf{g}_1, \quad \mathbf{L}_1^T \mathbf{d}^1 = \mathbf{D}_1^{-1} \mathbf{v}. \quad (46)$$

From $\mathbf{L}_1 \mathbf{v} = -\mathbf{g}_1$, there is

$$\begin{aligned} v_1 &= \frac{4}{3}, \\ \frac{22}{169} v_1 + v_2 &= -\frac{2}{3}. \end{aligned} \quad (47)$$

Solving the equations, we get $\mathbf{v} = (v_1, v_2)^T = (4/3, -142/169)^T$. Then from $\mathbf{L}_1^T \mathbf{d}^1 = \mathbf{D}_1^{-1} \mathbf{v}$, there is

$$\begin{aligned} d_1^1 + \frac{22}{169} d_2^1 &= \frac{60}{169}, \\ d_2^1 &= -\frac{71}{81}. \end{aligned} \quad (48)$$

Solving the equations, we get the search direction

$$\mathbf{d}^1 = (d_1^1, d_2^1)^T = \frac{1}{81} (38, -71)^T. \quad (49)$$

We then use the new BFGS algorithm to compute the search direction $\mathbf{d}^1$.

Step 1. First compute the initial correction matrix $\mathbf{C}_0^{-1}$ and the coefficient a_0 .

For $\boldsymbol{\sigma} = \sqrt{2} \mathbf{g}_0 + \|\mathbf{g}_0\| \mathbf{e} = 2 \begin{pmatrix} 2\sqrt{2} + \sqrt{5} \\ \sqrt{2} + \sqrt{5} \end{pmatrix}$, we get

$$\begin{aligned} \mathbf{C}_0^{-1} &= \mathbf{I} - \frac{2}{\|\boldsymbol{\sigma}\|^2} \boldsymbol{\sigma} \boldsymbol{\sigma}^T = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 & -3 \\ -3 & 1 \end{pmatrix}, \\ a_0 &= -\frac{\|\mathbf{g}_0\|}{\sqrt{2}} = -\sqrt{10}. \end{aligned} \quad (50)$$

Step 2. By $\mathbf{C}_0^{-1}$, we compute $\mathbf{C}_1^{-1}$.

Make use of known data, and we get

$$\begin{aligned} \mathbf{s}_0^T \mathbf{g}_0 &= -\frac{20}{3}, \quad \mathbf{s}_0^T \mathbf{g}_1 = \frac{4}{3}, \\ \mathbf{s}_0^T \mathbf{y}_0 &= \mathbf{s}_0^T \mathbf{g}_1 - \mathbf{s}_0^T \mathbf{g}_0 = 8, \\ \mathbf{w}_0 &= \mathbf{C}_0^{-1} \mathbf{g}_0 = -\sqrt{10} (1, 1)^T, \quad \|\mathbf{w}_0\| = 2\sqrt{5}, \\ \mathbf{v}_0 &= \mathbf{C}_0^{-1} \mathbf{g}_1 = \frac{2}{3\sqrt{10}} (-1, 7)^T. \end{aligned} \quad (51)$$

Thus there is

$$\begin{aligned} p_0 &= -\frac{\mathbf{s}_0^T \mathbf{g}_0}{\mathbf{s}_0^T \mathbf{y}_0} = \frac{5}{6}, \\ q_0 &= \frac{\left(\|\mathbf{w}_0\| + \sqrt{\mathbf{s}_0^T \mathbf{y}_0}\right) \mathbf{s}_0^T \mathbf{g}_1}{\|\mathbf{w}_0\| \mathbf{s}_0^T \mathbf{y}_0} = \frac{5 + \sqrt{10}}{30}. \end{aligned} \quad (52)$$

Therefore we have

$$\begin{aligned} \delta_0 &= p_0 \mathbf{v}_0 + q_0 \mathbf{w}_0 = -\frac{1}{9} (3 + 2\sqrt{10}, 3 - 2\sqrt{10})^T, \\ \|\delta_0\| &= \frac{7}{9} \sqrt{2}. \end{aligned} \quad (53)$$

Because

$$\begin{aligned} \sigma_0 &= \delta_0 + \frac{\|\delta_0\|}{\|\mathbf{w}_0\|} \mathbf{w}_0 = -\frac{1}{9} (10 + 2\sqrt{10}, 10 - 2\sqrt{10})^T, \\ \|\sigma_0\|^2 &= \frac{280}{81}, \end{aligned} \quad (54)$$

we obtain

$$\mathbf{Q}_0 = \mathbf{I} - \frac{2}{\|\sigma_0\|^2} \sigma_0 \sigma_0^T = \frac{1}{7} \begin{pmatrix} -2\sqrt{10} & -3 \\ -3 & 2\sqrt{10} \end{pmatrix}. \quad (55)$$

And $\mathbf{y}_0 \mathbf{s}_0^T = (8/9) \begin{pmatrix} 8 & 4 \\ 2 & 1 \end{pmatrix}$, $\mathbf{g}_0 \mathbf{s}_0^T = -(4/3) \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$, so

$$\begin{aligned} \mathbf{C}_1^{-1} &= \mathbf{Q}_0 \mathbf{C}_0^{-1} \left(\mathbf{I} - \frac{\mathbf{y}_0 \mathbf{s}_0^T}{\mathbf{s}_0^T \mathbf{y}_0} + \frac{\mathbf{g}_0 \mathbf{s}_0^T}{\|\mathbf{w}_0\| \sqrt{\mathbf{s}_0^T \mathbf{y}_0}} \right) \\ &= \frac{1}{126} \begin{pmatrix} -9\sqrt{10} - 38 & -18\sqrt{10} + 71 \\ 9\sqrt{10} - 38 & 18\sqrt{10} + 71 \end{pmatrix}. \end{aligned} \quad (56)$$

Thus we can obtain $\mathbf{d}^1 = a_1 (d_1^1, d_2^1)^T$, in which

$$\begin{aligned} d_1^1 &= \frac{1}{126} [(-9\sqrt{10} - 38) + (9\sqrt{10} - 38)] = -\frac{38}{63}, \\ d_2^1 &= \frac{1}{126} [(-18\sqrt{10} + 71) + (18\sqrt{10} + 71)] = \frac{71}{63}, \\ a_1 &= -\frac{\|\delta_0\|}{\|\mathbf{w}_0\|} a_0 = \frac{7}{9}. \end{aligned} \quad (57)$$

Therefore, the search direction is

$$\mathbf{d}^1 = (d_1^1, d_2^1)^T = \frac{1}{81} (38, -71)^T. \quad (58)$$

3.3. Analysis. The above example clearly shows the difference between the two algorithms in computing the search directions. In the new algorithm, correction matrixes are directly used to compute search directions. There is no procedure of *Cholesky* decomposition and no procedure of solving equations. Thus it simplifies the computing procedure of the search directions.

To compute the decomposition correction matrix $\mathbf{C}_{k+1}^{-1}$ at point $\mathbf{x}^{k+1}$ using the new BFGS algorithm, we first need to compute a Householder matrix $\mathbf{Q}_k$. This increases the computational complexity. However, the search direction $\mathbf{d}^{k+1}$ can be directly computed by (10) at the same time. Consequently, the two steps of the correction matrix calculation and search direction calculation are performed concurrently. In practice, the computation of (21) can be simplified using the properties of the Householder matrix.

Let the n row vectors of $\mathbf{N}_k^{-1}$ be denoted as $\alpha_{k1}^T, \alpha_{k2}^T, \dots, \alpha_{kn}^T$; that is,

$$\mathbf{N}_k^{-1} = \mathbf{C}_k^{-1} \left(\mathbf{I} - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \frac{\mathbf{g}_k \mathbf{s}_k^T}{\|\mathbf{w}_k\| \sqrt{\mathbf{s}_k^T \mathbf{y}_k}} \right) = \begin{bmatrix} \alpha_{k1}^T \\ \alpha_{k2}^T \\ \vdots \\ \alpha_{kn}^T \end{bmatrix}. \quad (59)$$

Thus, we can write

$$\sigma_k^T \mathbf{N}_k^{-1} = \sigma_{k1} \alpha_{k1}^T + \dots + \sigma_{kn} \alpha_{kn}^T = \mathbf{r}_k^T, \quad (60)$$

where $\sigma_{k1}, \dots, \sigma_{kn}$ are the n components of σ_k . Therefore, we have

$$\beta_k \sigma_k \sigma_k^T \mathbf{N}_k^{-1} = \begin{bmatrix} \beta_k \sigma_{k1} \mathbf{r}_k^T \\ \beta_k \sigma_{k2} \mathbf{r}_k^T \\ \vdots \\ \beta_k \sigma_{kn} \mathbf{r}_k^T \end{bmatrix}, \quad (61)$$

where $\beta_k = 2/\|\sigma_k\|^2$. Hence,

$$\begin{aligned} \mathbf{C}_{k+1}^{-1} &= \mathbf{Q}_k \mathbf{N}_k^{-1} = (\mathbf{I} - \beta_k \sigma_k \sigma_k^T) \mathbf{N}_k^{-1} = \mathbf{N}_k^{-1} - \beta_k \sigma_k \sigma_k^T \mathbf{N}_k^{-1} \\ &= \begin{bmatrix} \alpha_{k1}^T - \beta_k \sigma_{k1} \mathbf{r}_k^T \\ \alpha_{k2}^T - \beta_k \sigma_{k2} \mathbf{r}_k^T \\ \vdots \\ \alpha_{kn}^T - \beta_k \sigma_{kn} \mathbf{r}_k^T \end{bmatrix}. \end{aligned} \quad (62)$$

As a result, after $\mathbf{N}_k^{-1}$ and σ_k have been computed, the decomposition correction matrix $\mathbf{C}_{k+1}^{-1}$ can be directly obtained from (62), removing the need to compute the Householder matrix $\mathbf{Q}_k$.

We can also compare the condition numbers of $\mathbf{H}_k$ and $\mathbf{C}_k^{-1}$ in the correction formulae (3) and (21). The condition number is directly related to the possible magnification in relative error of the matrix calculations. Thus, a smaller condition number implies better numerical stability.

Because the correction matrix $\mathbf{H}_k$ is symmetric and positive-definite, if the eigenvalues μ_{ki} , ($1 \leq i \leq n$) of $\mathbf{H}_k$ have size order

$$0 < \mu_{k1} \leq \mu_{k2} \leq \dots \leq \mu_{kn}, \quad (63)$$

then the condition number of $\mathbf{H}_k$ in terms of the spectral norm [20] is

$$\text{cond}(\mathbf{H}_k) = \|\mathbf{H}_k\| \|\mathbf{H}_k^{-1}\| = \frac{\mu_{kn}}{\mu_{k1}}. \quad (64)$$

Because $(\mathbf{C}_k^{-1})^T \mathbf{C}_k^{-1} = \mathbf{H}_k$,

$$\text{cond}(\mathbf{C}_k^{-1}) = \frac{\sqrt{\mu_{kn}}}{\sqrt{\mu_{k1}}}, \quad (65)$$

and therefore

$$\text{cond}(\mathbf{C}_k^{-1}) = \sqrt{\text{cond}(\mathbf{H}_k)} \leq \text{cond}(\mathbf{H}_k). \quad (66)$$

Clearly, $\text{cond}(\mathbf{C}_k^{-1}) = \text{cond}(\mathbf{H}_k)$ only when $\mathbf{H}_k$ is a diagonal matrix whose diagonal elements are all the same. Otherwise, $\text{cond}(\mathbf{C}_k^{-1}) < \text{cond}(\mathbf{H}_k)$. Furthermore, $\text{cond}(\mathbf{Q}_k) = 1$ and $\text{cond}(\mathbf{Q}_k \mathbf{C}_k^{-1}) = \text{cond}(\mathbf{C}_k^{-1})$, which is beneficial to the calculation stability of (21). Note that (66) holds for all k , so the numerical stability of correction formula (21) should be much better than that of (3).

In addition, the iterative calculation of the decomposition correction matrix involves only nonsingular matrices. Thus, the calculation accuracy requirements should be lower than in the original BFGS algorithm, which considers symmetric positive-definite matrices.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research was financially supported by Twelfth Five-Year-Plan Science and Technology Research Project of Educational Department, Jilin Province (2014598).

References

- [1] L. Liu and Y. Liu, "Matrix decomposition from the hypothesis testing," *Journal of Beihua University (Natural Science)*, vol. 2, no. 6, pp. 469–471, 2001.
- [2] Y. Liu and L. Liu, "Improvement on the algorithm of equal line sum decomposition for sample correlation coefficient matrix," *Journal of Wuhan University of Technology*, vol. 31, no. 12, pp. 165–168, 2009.
- [3] L. Liu, "Extension and application of a matrix decomposition algorithm," *Mathematics in Practice and Theory*, vol. 41, no. 21, pp. 239–243, 2011.
- [4] L. Liu, "Quasi-Newton methods using equal inner product decomposition matrix of the correction matrix to obtain the search direction," *Mathematics in Practice and Theory*, vol. 43, no. 10, pp. 214–219, 2013.
- [5] L. Liu and Y. Liu, "Correction formula of Broyden's convex class expressed with decomposition matrix," *Journal of Beihua University (Natural Science)*, vol. 15, no. 2, pp. 179–183, 2014.
- [6] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms. II. The new algorithm," *Journal of the Institute of Mathematics and Its Applications*, vol. 6, pp. 222–231, 1970.
- [7] R. Fletcher, "A new approach to variable metric algorithms," *Computer Journal*, vol. 13, no. 3, pp. 317–322, 1970.
- [8] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Mathematics of Computation*, vol. 24, pp. 23–26, 1970.
- [9] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Mathematics of Computation*, vol. 24, pp. 647–656, 1970.
- [10] P. E. Gill and W. Murray, "Quasi-Newton methods for unconstrained optimization," *Journal of the Institute of Mathematics and its Applications*, vol. 9, pp. 91–108, 1972.
- [11] S. L. Xi, *Nonlinear Optimization Method*, Higher Education Press, Beijing, China, 1992.
- [12] N. Y. Deng, M. F. Zhu, D. F. Liu, Z. H. Yang, L. Wang, and Y. Tang, *Unconstrained Optimization Calculation Method*, Science Press, Beijing, China, 1982.
- [13] Y. X. Yuan and W. Y. Sun, *Optimization Theory and Method*, Science Press, Beijing, China, 1997.
- [14] M. J. D. Powell, "Some global convergence properties of a variable metric algorithm for minimization without exact line searches," in *Nonlinear Programming*, R. W. Cottle and C. E. Lemke, Eds., American Mathematical Society, Providence, RI, USA, 1976.
- [15] R. H. Byrd, J. Nocedal, and Y. X. Yuan, "Global convergence of a class of quasi-Newton methods on convex problems," *SIAM Journal on Numerical Analysis*, vol. 24, no. 5, pp. 1171–1190, 1987.
- [16] S. S. Oren and D. G. Luenberger, "Self-scaling variable metric (SSVM) algorithms. I. Criteria and sufficient conditions for scaling a class of algorithms," *Management Science*, vol. 20, pp. 845–862, 1974.
- [17] S. S. Oren, "Self-scaling variable metric (SSVM) algorithms. II. Implementation and experiments," *Management Science*, vol. 20, pp. 863–874, 1974.
- [18] S. S. Oren, "On the selection of parameters in self scaling variable metric algorithms," *Mathematical Programming*, vol. 7, pp. 351–367, 1974.
- [19] H. X. Huang and J. Y. Han, *Mathematical Programming*, Tsinghua University Press, Beijing, China, 2006.
- [20] Q. S. Liu, X. D. Zhang, and B. T. Wang, *Numerical Calculation Method Cause*, Metallurgical Industry Press, Beijing, China, 2005.