English Series | NUMERICAL MATHEMATICS A Journal of Chinese Universities | Vol.14, No.4 Nov. 2005

# THE RELATIONSHIP BETWEEN THE SPARSE SYMMETRIC BROYDEN METHOD AND THE $M$-TIME SECANT-LIKE MULTI-PROJECTION METHOD*

Lin Zhenghua(林正华)

**Abstract** *In this paper, we discuss the relationship between the sparse symmetric Broyden* (SPSB) *method* [1, 2] *and m-time secant-like multi-projection* (SMP) *method* [3] *and prove that when m goes to infinity, the* SMP *method is corresponding to the* SPSB *method.*
**Key words** *sparse unconstrained minimization problem, sparse symmetric Broyden method, m-time secant-like multi-projection method.*
**AMS(2000)subject classifications** 90C05, 65K05

## 1 Introduction

Consider the unconstrained minimization problem

$$\min_{x\in R^n} f(x), \tag{1.1}$$

where $f : R^n \to R$ is twice continuous differentiable. We assume that some sparse pattern is known for the Hessian $H(x) = \bigtriangledown^2 f(x)$, i.e., some entries of $H(x)$ are known to be zeros for all $x \in R^n$. To solve (1.1), we consider the following Newton-like iterative method:

$$x_{k+1} = x_k - B_k^{-1} g(x_k), \qquad k = 0, 1, 2, \cdots, \tag{1.2}$$

where $g(x) = \bigtriangledown f(x)$ is the gradient of $f$, and $B_k$ is an approximation to $H(x_k)$. We assume that $B_k$ is symmetric and has the same sparsity as the Hessian.

To find a $B_k$ which is a good apprixmation to $H(x_k)$ and which is not expensive has been an active topic . Powell[1] and Toint[2] proposed a secant update called the sparse symmetric Broyden (SPSB) update, which can maintain the sparsity, i.e., it does not change the elements of $B_k$ corresponding to known zero entries in the Hessian of the objective function. The SPSB method is quite successful in practice since it not only can save computer storate but also take

*Received: Jan. 7, 2002.

less number of iterations to get the solution than the Powell symmetric Broyden (PSB) method (see [1,4]) for many sparse problems. However, to get the updated matrix, one has to solve an additional sparse, symmetric and positive definite system of linear equations at each iteration, the computational cost of which is significant comparing to the whole procedure. Moreover, the coding is quite complicate, specially for the degenerate case.

In paper [3], we have presented a secant-like multi-projection update. The basic idea of this update is that at step $k$ we do $m(m \geq 1)$ times PSB update, and at each time, we did not compute the updated elements corresponding to the known zero entries of the matrix. We have shown that the $m$-time secant-like multi-projection (SMP) method was locally $q$-superlinearly convergent for any $M \geq 1$.

In this paper, we will show that this procedure is acturally equivalent to two projections. After repeating this procedure $j$ times, we get a sparse matrix $B_k^j, j = 1, 2, \cdots, m$ and we set $B_{k+1} = B_k^{(m)}$. The updated matrix $B_{k+1}$ does not satisfy the secant equation. However, we will show that $B_k^{(m)}$ converges to $B_{k+1}^{\mathrm{SPSB}}$ with $m$ goes to infinity, where $B_{k+1}^{\mathrm{SPSB}}$ is the updated matrix by the SPSB method.

Now we list some notations which will be used in the following context.

(1) Let $L(R^n)$ be the set of all real $n \times n$ matrices.

(2) Let $M_1$ denote the set of all symmetric matrices which satisfy the secant equation, i.e.,

$$M_1 = \{B \in L(R^n) : B^T = B, Bs = y\}.$$

(3) Let $\Omega$ be the set of the index pairs $(i, j)$ where the entry at the $i$th row and $j$th column of the Hessian is known to be nonzero, i.e.,

$$\Omega = \{(i, j) : (H(x))_{ij} \neq 0, \exists x \in R^n\}.$$

(4) Denote the sparse pattern of the $j$th row $(j = 1, 2, \cdots, n)$ by

$$Z_j \in R^n : Z_j = \{v \in R^n : e_j^T v = 0, \quad \text{for all} \quad i \quad \text{such that} \quad (i, j) \notin \Omega\},$$

(5) Let $M_2$ be the below space of all matrices with the same sparsity as the Hessian, i.e.,

$$M_2 = \{A \in L(R^n) : A^T e_j \in Z_j, j = 1, 2, \cdots, n\}.$$

(6) Let $D_j$ be a projection of $R^n$ onto $Z_j$, i.e.,

$$D_j = \mathrm{diag}(d_{j1}, d_{j2}, \cdots, d_{jn}),$$

where

$$d_{ji} = \begin{cases} 1, & \text{if} e_j \in Z_j, \\ 1, & \text{otherwise.} \end{cases}$$

(7) Let $\|\cdot\|$ denote the $l_2$-norm of a vector, $\|\cdot\|_2$ denote the induced matrix norm by the $l_2$-vector norm and $\|\cdot\|_F$ denote the Frobinius norm of a matrix. Let $\langle\cdot,\cdot\rangle$ denote the inner proudct of two vectors.

(8) Let $e_i, i = 1, \cdots, n$ be the $i$th column of the identify matrix.

(9) Suppose $A \in L(R^n)$, let $\sigma(A)$ denote the set of all eigenvalues of $A$.