

The Combined Schubert/Secant Finite-difference Algorithm  
for Solving Sparse Nonlinear Systems of Equations<sup>1</sup>

by

J.E. Dennis, Jr.<sup>2</sup>

and

Guangye Li<sup>3</sup>

Technical Report 86-11, May 1986  
(revised November 1986)

---

<sup>1</sup>This research was partially supported by contracts and grants: DOE DE-AS05-82ER1-13016 and AFOSR 85-0243, and it was also partially supported by Jilin University, People's Republic of China. This work forms a portion of Dr. Li's Ph.D. dissertation in Mathematical Sciences, Rice University written under the direction of Professor Dennis.

<sup>2</sup>Department of Mathematical Sciences, Rice University, P.O. Box 1892, Houston, Texas 77251.

<sup>3</sup>Computer Science Department, Cornell University, Upson Hall, Ithaca, New York 14853. Permanent address: Computer Center, Jilin University, People's Republic of China.

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>NOV 1986</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-1986 to 00-00-1986</b>	
4. TITLE AND SUBTITLE <b>The Combined Schubert/Secant Finite-difference Algorithm for Solving Sparse Nonlinear Systems of Equations</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Computational and Applied Mathematics Department ,Rice University,6100 Main Street MS 134,Houston,TX,77005-1892</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES <b>24</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

**Abstract:** This paper presents an algorithm, the combined Schubert/secant/finite difference algorithm, for solving sparse nonlinear systems of equations. This algorithm is based on dividing the columns of the Jacobian into two parts, and using different algorithms on each part. This algorithm incorporates advantages of both algorithms by exploiting some special structure of the Jacobian to obtain a good approximation to the Jacobian by using as little effort as possible. Kantorovich-type analysis and a locally  $q$ -superlinear convergence result for this algorithm are given.

**Key words:** finite difference, Jacobian,  $q$ -superlinear convergence, Kantorovich type analysis, sparsity, nonlinear system of equations.



## 1. Introduction.

Consider the nonlinear system of equations

$$F(x) = 0, \quad (1.1)$$

where  $F:R^n \rightarrow R^n$  is continuously differentiable on an open convex set  $D \subset R^n$ , and the Jacobian matrix  $F'(x)$  is sparse. To solve the system, we use the iteration

$$\bar{x} = x - B^{-1}F(x), \quad (1.2)$$

where  $x$  is the current iterate,  $\bar{x}$  is the new iterate, and  $B$  is an approximation to  $F'(x)$ , which has the same sparsity as the Jacobian.

Suppose we have finished the current iteration. Then the information we have is  $x, \bar{x}, F(x), F(\bar{x}), B$ . The purpose of this paper is to find a matrix  $\bar{B}$  which is a good approximation to  $F'(\bar{x})$  but to economize on the number of function evaluations required for this approximation.

In 1970 Schubert [11] gave a sparse modification of Broyden's [1] update. Broyden [2] also gave this algorithm independently. In order to present Schubert's algorithm, we introduce the following notation concerning the sparsity pattern of the Jacobian:

*Definition 1.1.* For  $j=1,2,\dots,n$  define the subspace  $Z_j \subset R^n$  determined by the sparsity pattern of the  $j$ th row of the Jacobian:

$$Z_j \equiv \{v \in R^n: e_i^T v = 0 \text{ for all } i \text{ such that } [F'(x)]_{ji} = 0 \text{ for all } x \in R^n\},$$

where  $e_i$  is the  $i$ th column of the  $n \times n$  identity matrix. Define the set of matrices  $Z$  that preserve the sparsity pattern of the Jacobian:

$$Z \equiv \{A \in L(R^n): A^T e_j \in Z_j \text{ for } j=1,2,\dots,n \}.$$

*Definition 1.2.* For  $j=1,2,\dots,n$ , define the projection operator,  $D_j \in L(\mathbb{R}^n)$ , that maps  $\mathbb{R}^n$  onto  $Z_j$ :

$$D_j \equiv \text{diag} (d_{j1}, d_{j2}, \dots, d_{jn}),$$

where

$$d_{ji} = \begin{cases} 1, & \text{if } e_i \in Z_j, \\ 0, & \text{otherwise.} \end{cases}$$

For a scalar  $\alpha \in \mathbb{R}$ , define the pseudo-inverse:

$$\alpha^+ = \begin{cases} \alpha^{-1}, & \text{if } \alpha \neq 0, \\ 0, & \text{if } \alpha = 0. \end{cases}$$

Now Schubert's update is formulated as follows:

$$\bar{B} = B + \sum_{j=1}^n ([s]_j^T [s]_j)^+ e_j e_j^T (y - Bs) [s]_j^T, \quad (1.3)$$

where  $[s]_j = D_j s$ ,  $s = \bar{x} - x$  and  $y = F(\bar{x}) - F(x)$ .

The advantage of Schubert's algorithm is that at each iteration only one function value is required, and it is  $q$ -superlinearly convergent (see Marwil [8]). However it usually requires more iterations than finite difference algorithms (see Li [7]).

Curtis, Powell, and Reid [4] proposed a finite difference algorithm, called the CPR algorithm, which is based on a partition of the columns of the Jacobian. Coleman and Moré [3] associate the partition problem with a graph coloring problem and gave some partitioning algorithms which can make the number of function evaluations needed to approximate the Jacobian by CPR algorithm optimal or nearly optimal.

Following Coleman and Moré, we give some definitions concerning a partition of the columns of the Jacobian.

*Definition 1.3.* A partition of the columns of a matrix  $B$  is a division of the columns into groups  $c_1, c_2, \dots, c_p$  such that each column belongs to one and only one group.

*Definition 1.4.* A partition of the columns of a matrix  $B$  is consistent with the direct determination of  $B$  if whenever  $b_{ij}$  is a nonzero element of  $B$ , then the group containing column  $j$  has no other column with a nonzero element in row  $i$ .

The CPR algorithm can be formulated as follows: for a given consistent partition of the columns of the Jacobian, which divides the set  $\{1, \dots, n\}$  into  $p$  subsets  $c_1, \dots, c_p$  (for convenience,  $c_i, i=1, 2, \dots, p$ , indicates both the sets of the columns and the sets of the indices of these columns), obtain vectors  $d_1, d_2, \dots, d_p$  such that  $B$  is determined uniquely by the equations

$$Bd_i = F(x+d_i) - F(x) \equiv y_i \quad i=1, 2, \dots, p \quad . \quad (1.4)$$

Notice that for the CPR algorithm, the number of function evaluations at each iteration is  $p+1$ . Since the partition of the columns of the Jacobian plays an important role in the CPR algorithm, we call the CPR algorithm based on Coleman and More's algorithms the CPR-CM algorithm.

The advantage of the CPR algorithm is that it usually requires fewer iterations than Schubert's algorithm. However, it requires more function values at each iteration than Schubert's algorithm (see Li [7]).

In an early paper [7], we proposed an algorithm called the secant/finite difference (SFD) algorithm, which is also based on a consistent partition of the columns of the Jacobian. However, it uses the information we already have at every iterative step more efficiently than the CPR algorithm. Let

$$d_i = \sum_{j \in c_i} s_j e_j, \quad i = 1, 2, \dots, p, \quad (1.5)$$

$$g_i = \sum_{j=1}^i d_j, \quad g_0 = 0, \quad (1.6)$$

and

$$y_i = F(\bar{x} - g_{i-1}) - F(\bar{x} - g_i), \quad i = 1, 2, \dots, p, \quad (1.7)$$

where  $s_i = \bar{x}_i - x_i$  indicate the  $i$ th component of  $s$ . The SFD algorithm can be formulated as follows: If  $s_j \neq 0$ , for some  $j \in c_i$ , then the  $j$ th column of  $\bar{B}$  is determined uniquely by equations

$$\bar{B}d_i = y_i.$$

If  $s_j = 0$ , then the  $j$ th column of  $\bar{B}$  is equal to the  $j$ th column of  $B$ .

Since

$$\begin{aligned} y_1 &= F(\bar{x} - g_0) - F(\bar{x} - g_1) = F(\bar{x}) - F(\bar{x} - g_1), \\ y_p &= F(\bar{x} - g_{p-1}) - F(\bar{x} - g_p) = F(\bar{x} - g_{p-1}) - F(x), \end{aligned} \quad (1.8)$$

the number of function evaluations required by the SFD algorithm at each iteration is one less than that required by CPR-CM algorithm.

Now consider the example

$$\begin{bmatrix} \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & \times & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 & \times & 0 \\ \times & \times & \times & 0 & 0 & 0 & 0 & \times \end{bmatrix}. \quad (1.9)$$

The partition  $c_1 = \{1\}$ ,  $c_2 = \{2\}$ ,  $c_3 = \{3\}$ ,  $c_4 = \{4, 5, 6, 7\}$  is an optimal consistent partition of the columns of the Jacobian. For this problem, the CPR-CM algorithm and the SFD algorithm require 5 and 4 function values at each iteration



respectively.

In this paper, we propose an algorithm called the combined Schubert/secant/finite difference (CSSFD) algorithm, which is a combination of the SFD algorithm and Schubert's algorithm (including Broyden's algorithm). For some problems, this algorithm can reduce the number of function values required at each iteration to fewer than the SFD algorithm by considering special structure of the Jacobian. For example (1.9), the number of function evaluations is 2.

The CSSFD algorithm and its properties are given in Section 2. A Kantorovich-type analysis for this algorithm is given in Section 3. A  $q$ -superlinear convergence result is given in Section 4.

In this paper,  $L(R^n)$  denotes the linear space of all real  $n \times n$  matrices,  $\|\cdot\|_F$  indicates the Frobenius norm of a matrix, and  $\|\cdot\|$  indicates the  $l_2$ -vector norm.

## 2. The CSSFD Algorithm and its Properties.

Consider example (1.9). The first 3 columns of the matrix are denser than the other columns, and this makes  $p$ , the number of the groups in the partition, at least 4. The CSSFD algorithm divides the columns of the Jacobian into two parts, and uses different algorithms on each part.

We say a group of the columns of a matrix has 'good sparsity' if the columns in this group have few nonzeros in the same row position. Otherwise, we say the group of the columns has 'bad sparsity'.

Suppose the columns of the Jacobian can be divided into two groups -- the good sparsity group  $c$  and the bad sparsity group  $c_1$ . For convenience, we use  $c$  and  $c_1$  to indicate both the groups of the columns of a matrix and the sets of the indices

of these columns. Then,

$$c \cup c_1 = \{1, \dots, n\}.$$

For any matrix  $A \in L(R^n)$ , let

$$A_1 = A \sum_{j \in c_1} e_j e_j^T, \quad A_2 = A \sum_{j \in c} e_j e_j^T.$$

Then  $A = A_1 + A_2$ . The main idea of the CSSFD algorithm is to use Schubert's update (including Broyden's update) on  $B_1$  and to use the SFD algorithm on  $B_2$ , where  $B = B_1 + B_2$ .

In practice, there are many ways to choose  $c$  and  $c_1$ . For example, we can first partition the columns by using a CPR-CM procedure. Then, if we can afford  $m$  F-values at each iteration, we can keep the columns of the  $m-1$  largest groups of the partition for  $c$  and put all the remaining columns into  $c_1$ .

*Algorithm 2.1.* Given a consistent partition of  $B_2$ , which divides  $c$  into  $p-1$  subsets  $c_2, c_3, \dots, c_p$ , and given an  $x^0 \in R^n$  and a nonsingular matrix  $B_0$  with the same sparsity as the Jacobian, at each step  $k \geq 0$ :

- (1). Solve  $B_k s_N^k = -F(x^k)$ .
- (2). Choose  $x^{k+1}$  by  $x^{k+1} = x^k + s_N^k$  or by a global strategy such as a trust-region method. Let  $s^k = x^{k+1} - x^k$ .
- (3). Check for convergence.
- (4). Update  $B_{k,1}$  by Schubert's update to get  $B_{k+1,1}$ , and update  $B_{k,2}$  by the SFD algorithm to get  $B_{k+1,2}$ .
- (5). Set

$$B_{k+1} = B_{k+1,1} + B_{k+1,2}.$$

Let

$$d_i = \sum_{j \in c_i} s_j e_j, \quad i = 1, \dots, p, \quad (2.1)$$

$$g_i = \sum_{j=1}^i d_j, \quad i = 1, \dots, p, \quad g_0 = 0, \quad (2.2)$$

$$y_i = F(\bar{x} - g_{i-1}) - F(\bar{x} - g_i), \quad i = 1, 2, \dots, p,$$

and

$$J_i = \int_0^1 F'(\bar{x} - g_i + t(g_i - g_{i-1})) dt, \quad i = 1, 2, \dots, p. \quad (2.3)$$

Then,

$$J_i d_i = y_i, \quad i = 1, 2, \dots, p, \quad (2.4)$$

and the update of Algorithm 2.1 can be formulated as

$$\left\{ \begin{array}{l} \bar{B}_1 = B_1 + \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ e_i e_i^T (y_1 - B_1 d_1) [d_1]_i^T, \\ \bar{B}_2 = B_2 + \sum_{i=2}^p \sum_{j \in c_i} s_j^+ s_j (J_i - B_2) e_j e_j^T, \\ \bar{B} = \bar{B}_1 + \bar{B}_2. \end{array} \right. \quad (2.5)$$

Now we give some of the properties of  $\bar{B}$  obtained from (2.5).

*Lemma 2.1.*  $\bar{B}$  satisfies the secant equations

$$\bar{B} d_i = y_i, \quad i = 1, \dots, p, \quad (2.6)$$

and (2.6) implies that

$$\bar{B} s = F(\bar{x}) - F(x) = y. \quad (2.7)$$

*Lemma 2.2.*  $\bar{B}$  is the unique solution to

$$\min \{ \|\hat{B} - B\|_F : \hat{B} d_i = y_i, \quad i = 1, \dots, p, \quad \text{and } \hat{B} \in Z \}. \quad (2.8)$$

The proof of this lemma is similar to that for Schubert's algorithm given by Reid [10] and Marwil [8].

*Theorem 2.3.* If  $A \in L(R^n)$  has the same sparsity as the Jacobian, then

$$\begin{aligned} \|\bar{B}_1 - A_1\|_F^2 &\leq \|B_1 - A_1\|_F^2 - \frac{1}{\|s\|^2} \|(B_1 - A_1)s\|^2 \\ &\quad + \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ [e_i^T (y_1 - A d_1)]^2. \end{aligned} \quad (2.9)$$

*Proof.* Let  $\bar{E}_1 = \bar{B}_1 - A_1$ , and  $E_1 = B_1 - A_1$ . From (2.5), we have

$$e_i^T \bar{B}_1 = e_i^T B_1 + ([d_1]_i^T [d_1]_i)^+ e_i^T (y_1 - B_1 d_1) [d_1]_i^T. \quad (2.10)$$

Subtracting  $e_i^T A_1$  from both sides of (2.10), and noticing that  $e_i^T B_1 d_1 = e_i^T B_1 [d_1]_i$  and that  $e_i^T A_1 d_1 = e_i^T A_1 [d_1]_i$ , we obtain

$$\begin{aligned} e_i^T \bar{E}_1 &= e_i^T E_1 + ([d_1]_i^T [d_1]_i)^+ e_i^T (y_1 - B_1 d_1) [d_1]_i^T \\ &= e_i^T E_1 (I - ([d_1]_i^T [d_1]_i)^+ [d_1]_i [d_1]_i^T) \\ &\quad + ([d_1]_i^T [d_1]_i)^+ e_i^T (y_1 - A_1 d_1) [d_1]_i^T. \end{aligned} \quad (2.11)$$

Since  $([d_1]_i^T [d_1]_i)^+ e_i^T (y_1 - A_1 d_1)$  is a scalar, the first and second terms on the right of (2.11) are perpendicular to each other, and we have

$$\begin{aligned} \|e_i^T \bar{E}_1\|^2 &= \|e_i^T E_1 (I - ([d_1]_i^T [d_1]_i)^+ [d_1]_i [d_1]_i^T)\|^2 + ([d_1]_i^T [d_1]_i)^+ |e_i^T (y_1 - A_1 d_1)|^2 \\ &= \|e_i^T E_1\|^2 - ([d_1]_i^T [d_1]_i)^+ |e_i^T E_1 [d_1]_i|^2 + ([d_1]_i^T [d_1]_i)^+ |e_i^T (y_1 - A_1 d_1)|^2 \\ &\leq \|e_i^T E_1\|^2 - \frac{1}{\|s\|^2} |e_i^T E_1 d_1|^2 + ([d_1]_i^T [d_1]_i)^+ |e_i^T (y_1 - A_1 d_1)|^2. \end{aligned}$$

Therefore,

$$\begin{aligned}
\|\bar{B}_1 - A_1\|_F^2 &= \sum_{i=1}^n \|e_i^T \bar{E}_1\|^2 \\
&\leq \|B_1 - A_1\|_F^2 - \frac{1}{\|s\|^2} \|(B_1 - A_1)d_1\|^2 \\
&\quad + \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ [e_i^T (y_1 - A_1 d_1)]^2 \\
&= \|B_1 - A_1\|_F^2 - \frac{1}{\|s\|^2} \|(B_1 - A_1)s\|^2 \\
&\quad + \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ [e_i^T (y_1 - A_1 d_1)]^2 .
\end{aligned}$$

*Theorem 2.4.* If  $A \in L(R^n)$  has the same sparsity as the Jacobian, then

$$\begin{aligned}
\|\bar{B}_2 - A_2\|_F^2 &\leq \|B_2 - A_2\|_F^2 - \frac{1}{\|s\|^2} \|(B_2 - A_2)s\|^2 \\
&\quad + \sum_{i=2}^p \sum_{j \in c_i} s_j^+ s_j \|(J_i - A)e_j\|^2 .
\end{aligned} \tag{2.12}$$

*Proof.* Let  $\bar{E}_2 = \bar{B}_2 - A_2$ , and  $E_2 = B_2 - A_2$ . It follows from (2.5) that if  $j \in c_i$ ,  $i = 2, \dots, p$ , then

$$\bar{B}_2 e_j = B_2 e_j + s_j^+ s_j (J_i - B_2) e_j . \tag{2.13}$$

Subtracting  $A_2 e_j$  from both sides of (2.13), we obtain

$$\bar{E}_2 e_j = (1 - s_j^+ s_j) E_2 e_j + s_j^+ s_j (J_i - A_2) e_j .$$

Since  $(1 - s_j^+ s_j) s_j^+ s_j = 0$ , we have

$$\begin{aligned}
\|\bar{E}_2 e_j\|^2 &= (1 - s_j^+ s_j) \|E_2 e_j\|^2 + s_j^+ s_j \|(J_i - A_2) e_j\|^2 \\
&= \|E_2 e_j\|^2 - s_j^+ s_j \|E_2 e_j\|^2 + s_j^+ s_j \|(J_i - A_2) e_j\|^2 .
\end{aligned}$$

Therefore,

$$\|\bar{E}_2\|_F^2 = \sum_{j \in c} \|\bar{E}_2 e_j\|^2 \tag{2.14}$$

$$= \|E_2\|_F^2 - \sum_{j \in c} s_j^+ s_j \|E_2 e_j\|^2 + \sum_{i=2}^p \sum_{j \in c_i} s_j^+ s_j \|(J_i - A)e_j\|^2.$$

In addition,

$$\begin{aligned} \sum_{j \in c} s_j^+ s_j \|E_2 e_j\|^2 &= \|E_2 \sum_{j \in c} s_j^+ s_j e_j e_j^T\|_F^2 \\ &\geq \frac{\|E_2 \sum_{j \in c} s_j^+ s_j e_j e_j^T s\|^2}{\|s\|^2} = \frac{\|E_2 s\|^2}{\|s\|^2}. \end{aligned}$$

Thus, (2.12) follows from (2.14).

### 3. A Kantorovich-type Analysis.

To study the convergence properties of Algorithm 2.1, we assume that  $F'$  satisfies the following Lipschitz condition: For every  $i \in c$ , there exists  $\gamma_i > 0$ , such that

$$\|(F'(x) - F'(y))e_i\| \leq \gamma_i \|x - y\|, \quad \forall x, y \in D, \quad (3.1)$$

and there exists  $\Theta_i > 0$ ,  $i = 1, 2, \dots, n$ , such that

$$\|e_i^T (F'(x)_1 - F'(y)_1)\| \leq \Theta_i \|x - y\|, \quad \forall x, y \in D. \quad (3.2)$$

Let  $\gamma = (\sum_{i \in c} \gamma_i^2)^{\frac{1}{2}}$ ,  $\Theta = (\sum_{i=1}^n \Theta_i^2)^{\frac{1}{2}}$ ,  $\alpha = (\gamma^2 + \Theta^2)^{\frac{1}{2}}$ . If  $F'$  satisfies this Lipschitz condition, then the following are true:

$$\|F'(x)_1 - F'(y)_1\|_F \leq \Theta \|x - y\|, \quad \forall x, y \in D, \quad (3.3)$$

$$\|F'(x)_2 - F'(y)_2\|_F \leq \gamma \|x - y\|, \quad \forall x, y \in D, \quad (3.4)$$

and

$$\|F'(x) - F'(y)\|_F \leq \alpha \|x - y\|, \quad \forall x, y \in D. \quad (3.5)$$

*Lemma 3.1.* Let  $F'$  satisfy (3.1) and (3.2), and let  $\bar{B}$  be generated by Algorithm 2.1. If  $\bar{x} \in D$  and  $\bar{x} - d_1 \subset D$ , then for any  $z \in D$ ,

$$\begin{aligned} \|\bar{B}_1 - F'(z)_1\|_F^2 &\leq \|B_1 - F'(z)_1\|_F^2 - \frac{1}{\|s\|^2} \|(B_1 - F'(z)_1)s\|^2 \\ &\quad + \Theta^2(\|\bar{x} - z\| + \frac{1}{2}\|d_1\|)^2. \end{aligned} \quad (3.6)$$

*Proof.* Substituting  $F'(z)$  for  $A$  in (2.9), we obtain

$$\begin{aligned} \|\bar{B}_1 - F'(z)_1\|_F^2 &\leq \|B_1 - F'(z)_1\|_F^2 - \frac{1}{\|s\|^2} \|(B_1 - F'(z)_1)s\|^2 \\ &\quad + \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ [e_i^T (y_1 - F'(z)d_1)]^2. \end{aligned} \quad (3.7)$$

By (2.3), (2.4), (3.3), and Cauchy-Schwarz inequality we have

$$\begin{aligned} \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ [e_i^T (y_1 - F'(z)d_1)]^2 &= \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ (e_i^T (J_1 - F'(z))_1 [d_1]_i)^2 \\ &\leq \sum_{i=1}^n ([d_1]_i^T [d_1]_i)^+ \|e_i^T (J_1 - F'(z))_1\|^2 \| [d_1]_i \|^2 \leq \sum_{i=1}^n \|e_i^T (J_1 - F'(z))_1\|^2 \\ &= \|(J_1 - F'(z))_1\|_F^2 = \left\| \int_0^1 (F'(\bar{x} - (1-t)d_1) - F'(z))_1 dt \right\|_F^2 \\ &\leq \Theta^2(\|\bar{x} - z\| + \frac{1}{2}\|d_1\|)^2. \end{aligned} \quad (3.8)$$

Then (3.6) follows from (3.7) and (3.8).

*Lemma 3.2.* Let  $F'$  satisfy (3.1) and (3.2), and let  $\bar{B}$  be generated by Algorithm

2.1. If  $\bar{x} \in D$  and  $\{\bar{x} - g_i, i = 2, \dots, p\} \subset D$ , then for any  $z \in D$ ,

$$\begin{aligned} \|\bar{B}_2 - F'(z)_2\|_F^2 &\leq \|B_2 - (F'(z))_2\|_F^2 - \frac{1}{\|s\|^2} \|(B_2 - F'(z)_2)s\|^2 \\ &\quad + \gamma^2(\|\bar{x} - z\| + \|s\|)^2. \end{aligned} \quad (3.9)$$

*Proof.* Substituting  $F'(z)$  for  $A$  in (2.12), we obtain

$$\begin{aligned} \|\bar{B}_2 - F'(z)_2\|_F^2 &\leq \|B_2 - F'(z)_2\|_F^2 - \frac{1}{\|s\|^2} \|(B_2 - F'(z)_2)s\|^2 \\ &\quad + \sum_{i=2}^p \sum_{j \in c_i} s_j^+ s_j \|(J_i - F'(z))e_j\|^2. \end{aligned} \quad (3.10)$$

It follows from (2.3) and (3.1) that

$$\begin{aligned} \sum_{i=2}^p \sum_{j \in c_i} s_j^+ s_j \|(J_i - F'(z))e_j\|^2 &\leq \sum_{i=2}^p \sum_{j \in c_i} \|(J_i - F'(z))e_j\|^2 \\ &= \sum_{i=2}^p \sum_{j \in c_i} \left\| \int_0^1 (F'(\bar{x} - g_i + t(g_i - g_{i-1})) - F'(z)) dt e_j \right\|^2 \\ &\leq \sum_{i=2}^p \sum_{j \in c_i} \gamma_j^2 \left( \int_0^1 (\|\bar{x} - z\| + (1-t)\|g_i\| + t\|g_{i-1}\|) dt \right)^2 \\ &\leq \sum_{i=2}^p \sum_{j \in c_i} \gamma_j^2 (\|\bar{x} - z\| + \|s\|)^2 = \gamma^2 (\|\bar{x} - z\| + \|s\|)^2. \end{aligned} \quad (3.11)$$

Thus, (3.9) follows from (3.10) and (3.11).

Let

$$d_i^k = \sum_{j \in c_i} s_j^k e_j,$$

and

$$g_i^k = \sum_{j=1}^i d_j^k, \quad i = 1, 2, \dots, p, \quad g_0^k = 0.$$

We have the following estimate for  $B_{k+1}$ .

*Theorem 3.3.* Let  $F'$  satisfy (3.1) and (3.2), and let  $\{x^k\}$  and  $\{B_k\}$  be generated by Algorithm 2.1. If  $\{x^j\}_{j=0}^{k+1} \subset D$  and  $\{x^{j+1} - g_i^j, i = 1, 2, \dots, p\}_{j=0}^k \subset D$ , then

$$\|B_{k+1} - F'(x^{k+1})\|_F \leq \|B_0 - F'(x_0)\|_F + 2\alpha \sum_{i=0}^k \|x^{i+1} - x^i\|. \quad (3.12)$$

*Proof.* Substituting  $z$  for  $\bar{x}$  in (3.6) and (3.9), we have



$$\|\bar{B}_1 - F'(\bar{x})_1\|_F^2 \leq \|B_1 - F'(\bar{x})_1\|_F^2 + \left(\frac{\Theta}{2}\|d_1\|\right)^2$$

and

$$\|\bar{B}_2 - F'(\bar{x})_2\|_F^2 \leq \|B_2 - F'(\bar{x})_2\|_F^2 + (\gamma\|s\|)^2.$$

Therefore

$$\begin{aligned} \|\bar{B} - F'(\bar{x})\|_F^2 &= \|\bar{B}_1 - F'(\bar{x})_1\|_F^2 + \|\bar{B}_2 - F'(\bar{x})_2\|_F^2 \\ &\leq \|B - F'(\bar{x})\|_F^2 + (\Theta^2 + \gamma^2)\|s\|^2 \\ &= \|B - F'(\bar{x})\|_F^2 + \alpha^2\|s\|^2. \end{aligned}$$

Then

$$\begin{aligned} \|\bar{B} - F'(\bar{x})\|_F &\leq \|B - F'(\bar{x})\|_F + \alpha\|\bar{x} - x\| \\ &\leq \|B - F'(x)\|_F + 2\alpha\|\bar{x} - x\|. \end{aligned} \tag{3.13}$$

Thus, (3.12) follows (3.13).

From (3.12), we have the following Kantorovich-type theorem for Algorithm

2.1.

*Theorem 3.4.* Assume that  $F'$  satisfies (3.1) and (3.2). Also assume that  $x_0 \in D$  and  $B_0 \in L(R^n)$  satisfy

$$\|B_0 - F'(x^0)\|_F \leq \delta, \quad \|B_0^{-1}\|_F \leq \beta, \quad \|B_0^{-1}F(x^0)\| \leq \eta$$

and

$$h \equiv \frac{\alpha\beta\eta}{(1 - 3\beta\delta)^2} \leq \frac{1}{10}, \quad \beta\delta < \frac{1}{3}.$$

If  $\bar{S}(x^0, 2t^*) \equiv \{x: \|x - x^0\| \leq 2t^*\} \subset D$ , where

$$t^* = \frac{1 - 3\beta\delta}{5\alpha\beta} \left(1 - (1 - 10h)^{\frac{1}{2}}\right),$$

then  $\{x^k\}$ , generated by Algorithm 2.1 without any global strategy, converges to  $x^*$ , which is the unique root of  $F(x)$  in  $\bar{S}(x^0, \bar{t}) \cap D$ , where

$$\bar{t} = \frac{1 - \beta\delta}{\alpha\beta} \left[ 1 + \left[ 1 - \frac{2\alpha\beta\eta}{(1 - \beta\delta)^2} \right]^{\frac{1}{2}} \right].$$

*Proof.* Consider the scalar iteration

$$t_{k+1} - t_k = \frac{2\beta}{2 - \beta\delta} f(t_k), \quad t_0 = 0, \quad k = 1, 2, \dots, \quad (3.14)$$

where

$$f(t) = \frac{5}{2} \alpha t^2 - \frac{1 - 3\beta\delta}{\beta} t + \frac{\eta}{\beta}. \quad (3.15)$$

It is easy to show that  $\{t_k\}$  satisfies the difference equation

$$t_{k+1} - t_k = \frac{\beta}{1 - \varphi} [\alpha(t_k - t_{k-1}) + 2\alpha t_{k-1} + \delta](t_k - t_{k-1}), \quad (3.16)$$

where  $\varphi = \frac{3 + \beta\delta}{5} < \frac{2}{3}$ . From (3.16), we see that  $\{t_k\}$  is a monotonically increasing sequence and that

$$\lim_{k \rightarrow \infty} t_k = t^*,$$

where  $t^*$  is the smallest root of (3.15).

Now, by induction, we will prove that

$$\|x^{k+1} - x^k\| \leq t_{k+1} - t_k, \quad k = 1, 2, \dots, \quad (3.17)$$

$$\{x^k\} \subset \bar{S}(x^0, t^*), \quad (3.18)$$

$$\{x^{k+1} + g_i^k, \quad i = 1, 2, \dots, p\} \subset \bar{S}(x^0, 2t^*), \quad (3.19)$$

and

$$\|B_k^{-1}\| \leq \frac{\beta}{1 - \varphi} \leq 3\beta, \quad k = 1, 2, \dots. \quad (3.20)$$

For  $k = 0$ , we have

$$\|x^1 - x^0\| \leq \eta \leq \frac{2}{2 - \beta\delta} \eta = t_1 - t_0 \leq t^* .$$

Thus,

$$\|x^1 - g_i^0 - x^0\| \leq \|x^1 - x^0\| + \|g_i^0\| \leq 2\|x^1 - x^0\| \leq 2t^* .$$

Suppose (3.17) holds for  $k = 0, 1, \dots, m - 1$ . Then,

$$\|x^m - x^0\| \leq \sum_{i=0}^{m-1} (t_{i+1} - t_i) = t_m \leq t^* .$$

Therefore,  $x^m \in \bar{S}(x^0, t^*)$ , and

$$\{x^m - g_i^{m-1}, i = 1, \dots, p\} \subset \bar{S}(x^0, 2t^*) .$$

By Theorem 3.3,

$$\begin{aligned} & \|B_0^{-1}(B_m - B_0)\| \\ & \leq \|B_0^{-1}\|_F (\|B_m - F'(x^m)\|_F + \|F'(x^m) - F'(x^0)\|_F + \|F'(x^0) - B_0\|_F) \\ & \leq \beta(3\alpha \sum_{i=0}^{m-1} \|x^{i+1} - x^i\| + 2\delta) \leq \beta(3\alpha t^* + 2\delta) \leq \frac{3 + \beta\delta}{5} = \varphi . \end{aligned}$$

Thus, by Theorem 3.1.4 of Dennis and Schnabel [6, p.45],

$$\|B_m^{-1}\| \leq \frac{\beta}{1 - \varphi} \leq 3\beta .$$

Therefore,

$$\begin{aligned} \|x^{m+1} - x^m\| & \leq \|B_m^{-1}\|_F \|F(x^m) - F(x^{m-1}) - B_{m-1}(x^m - x^{m-1})\| \\ & \leq \frac{\beta}{1 - \varphi} \left[ \frac{\alpha}{2} \|x^m - x^{m-1}\| + 2\alpha \sum_{i=0}^{m-2} \|x^{i+1} - x^i\| + \delta \right] \|x^m - x^{m-1}\| \\ & \leq \frac{\beta}{1 - \varphi} [\alpha(t_m - t_{m-1}) + 2\alpha t_{m-1} + \delta](t_m - t_{m-1}) = t_{m+1} - t_m . \end{aligned}$$

This completes the induction step. By (3.17), it is easy to show that there is an  $x^* \in D$  such that

$$\lim_{k \rightarrow \infty} x^k = x^* .$$

The uniqueness of  $x^*$  in  $\bar{S}(x^0, \bar{t}) \cap D$  can be obtained from Theorem 12.6.4 of [9] by setting  $A(x) \equiv B_0$ .

#### 4. Local Convergence properties.

To study the local convergence of our algorithm, we assume that  $F: D \subset R^n \rightarrow R^n$  has the following property:

*There is an  $x^* \in D$ , such that  $F(x^*) = 0$  and  $F'(x^*)$  is nonsingular.* (4.1)

*Theorem 4.1.* Let  $F$  satisfy (4.1), and let  $F'$  satisfy (3.1) and (3.2). Also, let  $\{x^k\}$  be generated by Algorithm 2.1 without any global strategy. Then, there exist  $\epsilon, \delta > 0$ , such that if  $x_0 \in D$  and  $B_0$ , a nonsingular  $n \times n$  matrix, satisfy

$$\|x^0 - x^*\| < \epsilon, \quad \|B_0 - F'(x^*)\|_F \leq \delta ,$$

then  $\{x^k\}$  is well defined and converges  $q$ -superlinearly to  $x^*$ .

*Proof.* Notice that when  $\epsilon$  and  $\delta$  are small enough, we have that  $h \leq \frac{1}{10}$ ,  $\beta\delta < \frac{1}{3}$

and that  $\bar{S}(x^0, 2t^*) \subset D$ , where  $h, \beta$  and  $t^*$  are defined in theorem 3.4. Therefore, by Theorem 3.4,

$$\{x^{k+1} + g_i^k, i = 1, 2, \dots, p\} \subset D .$$

Thus, substituting  $x^*$  for  $z$  in (3.6) and (3.9), we have

$$\begin{aligned} \|\bar{B}_1 - F'(x^*)\|_F^2 &\leq \|B_1 - F'(x^*)\|_F^2 - \frac{1}{\|s\|^2} \|(B_1 - F'(x^*)_1)s\|^2 \\ &\quad + \Theta^2(\|\bar{x} - x^*\| + \|s\|)^2 , \end{aligned} \tag{4.2}$$

and

$$\begin{aligned} \|\bar{B}_2 - F'(x^*)_2\|_F^2 &\leq \|B_2 - F'(x^*)_2\|_F^2 - \frac{1}{\|s\|^2} (B_2 - F'(x^*)_2)s\|^2 \\ &\quad + \gamma^2(\|\bar{x} - x^*\| + \|s\|)^2. \end{aligned} \quad (4.3)$$

Then,

$$\begin{aligned} \|\bar{B} - F'(x^*)\|_F^2 &= \|\bar{B}_1 - F'(x^*)_1\|_F^2 + \|\bar{B}_2 - F'(x^*)_2\|_F^2 \\ &\leq \|\bar{B} - F'(x^*)\|_F^2 + \alpha^2(\|\bar{x} - x^*\| + \|s\|)^2 \\ &\leq \|B - F'(x^*)\|_F^2 + (3\alpha\sigma(x, \bar{x}))^2 \end{aligned}$$

where  $\sigma(x, \bar{x}) = \max\{\|\bar{x} - x^*\|, \|x - x^*\|\}$ . Therefore,

$$\|\bar{B} - F'(x^*)\|_F \leq \|B - F'(x^*)\|_F + 3\alpha\sigma(x, \bar{x}).$$

Thus, by Theorem 5.1 of Dennis and Moré [5],  $\{x^k\}$  converges at least  $q$ -linearly to  $x^*$ .

By Theorem 3.1 of Dennis and Moré [5], to prove  $q$ -superlinear convergence, we need only to prove that

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - F'(x^*))s^k\|}{\|s^k\|} = 0. \quad (4.4)$$

Let  $\bar{E} = \bar{B} - F'(x^*)$  and  $E = B - F'(x^*)$ . Then, it follows from (4.2) and (4.3) that

$$\|\bar{E}_1\|_F \leq (\|E_1\|_F^2 - \frac{\|E_1 s\|^2}{\|s\|^2})^{\frac{1}{2}} + 3\Theta\sigma(x, \bar{x}), \quad (4.5)$$

and that

$$\|\bar{E}_2\|_F \leq (\|E_2\|_F^2 - \frac{\|E_2 s\|^2}{\|s\|^2})^{\frac{1}{2}} + 3\gamma\sigma(x, \bar{x}). \quad (4.6)$$

From (4.5) and (4.6), using the same argument for proving the  $q$ -superlinear convergence property of Broyden's algorithm (see Dennis and Moré [5]), we obtain

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - F'(x^*))_1 s^k\|}{\|s^k\|} = 0, \quad (4.7)$$

and

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - F'(x^*))_2 s^k\|}{\|s^k\|} = 0. \quad (4.8)$$

Notice that

$$\|(B_k - F'(x^*))s^k\| \leq \|(B_k - F'(x^*))_1 s^k\| + \|(B_k - F'(x^*))_2 s^k\|.$$

Thus, (4.4) follows from (4.7) and (4.8).

**Acknowledgement.** We would like to thank Professor Richard Tapia and the referees for helpful suggestions and corrections.

## References

- [1]. Broyden, C.G., A class of methods for solving nonlinear simultaneous equations, *Math. Comp.*, 19 (1965), pp. 577-593.
- [2]. Broyden, C.G., The convergence of an algorithm for solving sparse nonlinear systems, *Math. Comp.*, 25 (1971), pp. 285-294.
- [3]. Coleman, T.F., and J.J. Moré, Estimation of sparse Jacobians and graph coloring problems, *SIAM J. Numer. Anal.*, 20 (1983), pp. 187-209.
- [4]. Curtis, A.R., M.J.D. Powell and J.K. Reid, On the estimation of sparse Jacobian matrices, *J.I.M.A.*, 13 (1974), pp. 117-119.
- [5]. Dennis, J.E., Jr., and J.J. Moré, Quasi-Newton Methods, Motivation and Theory, *SIAM Review*, Vol. 19, No. 1 (1977).
- [6]. Dennis, J.E., Jr., and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, New Jersey (1983).
- [7]. Li, Guangye, The secant/finite difference algorithm for solving sparse nonlinear systems of equations, Technical Report 86-1, Math Sciences Dept., Rice Univ. (1986).
- [8]. Marwil, E., Convergence results for Schubert's method for solving sparse nonlinear equations, *SIAM J. Numer. Anal.*, 16 (1979), pp. 588-604.
- [9]. Ortega, J.M., and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York (1970).
- [10]. Reid, J.K., Least squares solution of sparse systems of non-linear equations

by a modified Marquardt algorithm, Proceedings of the NATO Conf. at Cambridge, July 1972, North Holland, Amsterdam, pp. 437-445.

- [11]. Schubert, L.K., Modification of a quasi-Newton method for nonlinear equations with a sparse Jacobian, Math. Comp., 24 (1970), pp. 27-30.