

The Secant/Finite Difference Algorithm for
Solving Sparse Nonlinear Systems of Equations¹

by

Guangye Li²

Technical Report 86-1, March 1986.
(Revised May 1986)

¹This research was partially supported by Jilin University, P.R. of China. It also was partially supported by contracts and grants: DOE DE-AS05-82ER1-13016 and AFOSR 85-0243. This work was done while the author was a visiting scholar in Mathematical Sciences Department, Rice University.

²Computer 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 MAY 1986		2. REPORT TYPE		3. DATES COVERED 00-00-1986 to 00-00-1986	
4. TITLE AND SUBTITLE The Secant/Finite Difference Algorithm for Solving Sparse Nonlinear Systems of Equations				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) Computational and Applied Mathematics Department ,Rice University,6100 Main Street MS 134,Houston,TX,77005-1892				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 Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 22	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Abstract

This paper presents an algorithm, the secant/finite difference algorithm, for solving sparse nonlinear systems of equations. This algorithm is a combination of a finite difference method and a secant method. A q -superlinear convergence result and an r -convergence rate estimate show that this algorithm has good local convergence properties. The numerical results indicate that this algorithm is probably more efficient than some currently used algorithms.

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 consider 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 the Jacobian $F'(x)$ with the same sparsity as the Jacobian. After we finish this step, we have the information: x , $\bar{x}$, and B . Our purpose is to get a $\bar{B}$, a good approximation to $F'(\bar{x})$, by using as little effort as possible.

Curtis, Powell and Reid [3] gave an efficient algorithm for sparse problems called the CPR Algorithm, which is a finite difference algorithm, but which can take the advantage of the sparsity to make the number of the function evaluations small. To describe the CPR Algorithm, Coleman and Moré [2] gave some definitions concerning a partition of the columns of a matrix B .

Definition 1.1. A partition of the columns of 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 1.2. A partition of the columns such that columns in a given group do not have a nonzero element in the same row position is *consistent* with the direct determination of B .

In order to have a good partition of the columns of B , Coleman and Moré [2] associated the partition problem with a certain graph coloring problem and gave some partitioning algorithms which can make p optimal or nearly optimal. For convenience, we call the CPR algorithm based on Coleman and Moré's algorithms the CPR-CM algorithm.

The CPR Algorithm can be formulated as follows: For a given consistent partition of the columns of B , obtain vectors $d_1, d_2, \dots, d_p$ such that $\bar{B}$ is determined uniquely by the equations

$$\bar{B}d_i = F(\bar{x} + d_i) - F(\bar{x}) \equiv y_i, \quad i = 1, \dots, p.$$

Notice that at each iterative step we need only to compute $p + 1$ function values rather than the $n + 1$ values required by a straightforward column-by-column finite-difference algorithm.

As an example we consider the matrix with a tridiagonal structure:

$$\begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}. \quad (1.3)$$

A consistent partition of the columns of the matrix is $c_1 = \{1, 4\}$, $c_2 = \{2, 5\}$, and $c_3 = \{3, 6\}$.

If we take

$$d_1 = (h, 0, 0, h, 0, 0)^T,$$

$$d_2 = (0, h, 0, 0, h, 0)^T,$$

$$d_3 = (0, 0, h, 0, 0, h)^T,$$

then $\bar{B}$ is determined uniquely, and the number of function evaluations required at each iteration is 4.

In this paper, we propose an algorithm called the secant/finite difference (SFD) algorithm for solving sparse nonlinear systems of equations. This algorithm 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. The secant equation is also satisfied by the SFD algorithm. Therefore, this algorithm can be seen as a combination of the CPR-CM algorithm and a secant algorithm. The SFD algorithm reduces the number of function evaluations required by the CPR-CM algorithm by one, and it has good local convergence properties. Our numerical results show that the SFD algorithm is probably more efficient than the CPR-CM algorithm. The SFD algorithm and some of its properties are given in Section 2. A Kantorovich-type analysis for the SFD algorithm is given in Section 3. A q -superlinear convergence result and an r -convergence order estimate of the SFD algorithm are given in Section 4. Some numerical results are given in Section 5.

In this paper, $||.||_F$ indicates the Frobenius norm of a matrix, and $||.||$ indicates the l_2 vector norm. We use ' $\setminus$ ' to denote the subtraction of two sets; that is,

$$A \setminus B = \{v : v \in A \text{ and } v \notin B\}.$$

For a sparse matrix B , we use M to denote the set of pairs of indices (i, j) , where b_{ij} is a structurally nonzero element of B , i.e.

$$M = \{(i, j) : b_{ij} \neq 0\}.$$

2. The SFD Algorithm and its Properties.

Given a 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), also given $x, \bar{x} \in R^n$, let

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

and

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

If $s_j \neq 0, j = 1, \dots, n$, then $\bar{B}$ is determined uniquely by the equations

$$\begin{aligned} \bar{B}d_1 &= \bar{B}(\bar{x} - (\bar{x} - g_1)) = F(\bar{x}) - F(\bar{x} - g_1) \equiv y_1, \\ &\dots \\ \bar{B}d_i &= \bar{B}(\bar{x} - g_{i-1} - (\bar{x} - g_i)) = F(\bar{x} - g_{i-1}) - F(\bar{x} - g_i) \equiv y_i, \\ &\dots \\ \bar{B}d_p &= \bar{B}(\bar{x} - g_{p-1} - x) = F(\bar{x} - g_{p-1}) - F(x) \equiv y_p. \end{aligned} \tag{2.1}$$

Let $\bar{B} = [\bar{b}_{lm}]$. By (2.1), if $(l, m) \in M$, then

$$\bar{b}_{lm} = \frac{e_l^T y_i}{s_m}, \tag{2.2}$$

where $m \in c_i, \quad i = 1, 2, \dots, p$.

Notice that by (2.1), to get $\bar{B}$, we need only to compute p function values since we already have the value $F(x)$ at the current step. The number of function evaluations at each iteration is

one less than the CPR-CM algorithm. For example (1.3), we take

$$d_1 = (s_1, 0, 0, s_4, 0, 0)^T,$$

$$d_2 = (0, s_2, 0, 0, s_5, 0)^T,$$

$$d_3 = (0, 0, s_3, 0, 0, s_6)^T,$$

and

$$g_1 = (s_1, 0, 0, s_4, 0, 0)^T,$$

$$g_2 = (s_1, s_2, 0, s_4, s_5, 0)^T.$$

Then, we need only to compute the values of $F(\bar{x})$, $F(\bar{x}-g_1)$, and $F(\bar{x}-g_2)$, and the number of function evaluations is 3 per iteration instead of 4 required by the CPR-CM algorithm.

Let

$$J_i = \int_0^1 F'(\bar{x} - g_i + td_i) dt, \quad i = 1, \dots, p. \quad (2.3)$$

Then

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

Let $J_i = [J_{lm}^i]$. Since J_i has the same sparsity as the Jacobian, by (2.4), we have that if $(l, m) \in M$, then

$$J_{lm}^i = \frac{e_l^T y_i}{s_m}, \quad (2.5)$$

where $m \in c_i$, $i = 1, 2, \dots, p$. Comparing (2.5) with (2.2), we have

$$J_i e_j = \bar{B} e_j, \quad (2.6)$$

where $j \in c_i$, $i = 1, \dots, p$. Therefore, $\bar{B}$ can be written as

$$\bar{B} = \sum_{i=1}^p \sum_{j \in c_i} J_i e_j e_j^T. \quad (2.7)$$

To study the properties of the SFD algorithm, we assume that $F'(x)$ satisfies the Lipschitz condition, i.e. that there exist $\alpha_i > 0$, $i = 1, \dots, n$ such that

$$\|(F'(x) - F'(y))e_i\| \leq \alpha_i \|x - y\|, \quad i = 1, 2, \dots, n, \quad x, y \in D. \quad (2.8)$$

Let $\alpha = \left(\sum_{i=1}^n \alpha_i^2\right)^{\frac{1}{2}}$, then

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

Now we have the following estimate for $\bar{B}$:

Theorem 2.1. Suppose $F'(x)$ satisfies Lipschitz condition (2.8), and $\bar{B}$ is determined by (2.1). If

$\bar{x} \in D$, $\bar{x} - g_i \in D$, $i = 1, \dots, p$, and $s_i \neq 0$, $i = 1, 2, \dots, n$, then

$$\|F'(\bar{x}) - \bar{B}\|_F \leq \alpha \|\bar{x} - x\|. \quad (2.9)$$

Proof. By (2.6) and (2.7),

$$\begin{aligned} \|F'(\bar{x}) - \bar{B}\|_F^2 &= \sum_{i=1}^n \|(F'(\bar{x}) - \bar{B})e_i\|^2 \\ &= \sum_{i=1}^p \sum_{j \in c_i} \|(F'(\bar{x}) - \bar{B})e_j\|^2 \\ &= \sum_{i=1}^p \sum_{j \in c_i} \|(F'(\bar{x}) - J_i)e_j\|^2. \end{aligned} \quad (2.10)$$

Using (2.3) and Lipschitz condition (2.8), we obtain

$$\begin{aligned} &\sum_{j \in c_i} \|(F'(\bar{x}) - J_i)e_j\|^2 \\ &= \sum_{j \in c_i} \left\| (F'(\bar{x}) - \int_0^1 F'(\bar{x} - g_i + t(g_i - g_{i-1}))dt)e_j \right\|^2 \\ &\leq \sum_{j \in c_i} \left(\alpha_j \int_0^1 \|g_i - t(g_i - g_{i-1})\| dt \right)^2 \\ &\leq \sum_{j \in c_i} \alpha_j^2 \left(\int_0^1 (1-t) \|g_i\| dt + \int_0^1 \|g_{i-1}\| t dt \right)^2 \\ &\leq \sum_{j \in c_i} \alpha_j^2 \left(\frac{1}{2} \|s\| + \frac{1}{2} \|s\| \right)^2 = \|s\|^2 \sum_{j \in c_i} \alpha_j^2. \end{aligned} \quad (2.11)$$

It follows from (2.10) and (2.11) that

$$\|F'(\bar{x}) - \bar{B}\|_F^2 \leq \|s\|^2 \sum_{i=1}^p \sum_{j \in c_i} \alpha_j^2 = \alpha^2 \|s\|^2. \quad (2.12)$$

Then (2.9) follows from (2.12).

In (2.1), to determine $\bar{B}$ uniquely, we assume that $s_j \neq 0$, $j = 1, \dots, n$. However, sometimes it may happen that $s_i = 0$ for some $1 \leq i \leq n$. If this happens, then the i th column of $\bar{B}$ can not be determined uniquely by (2.1). In this case, let

$$\Omega_1 = \{i \in \{1, 2, \dots, n\} : s_i \neq 0\},$$

and let

$$\Omega_2 = \{1, 2, \dots, n\} \setminus \Omega_1.$$

Now we deal with the general case in such a way that if $j \in \Omega_1$, then the j th column of $\bar{B}$ is determined uniquely by (2.1). If $j \in \Omega_2$, then we let the j th column of $\bar{B}$ be equal to the j th column of B . In practice, if $|s_j|$ is too close to zero the cancellation errors will become significant. Therefore, there should be a lower bound θ for $|s_j|$. Now the SFD algorithm with a global strategy can be stated as follows:

Algorithm 2.2. Given a consistent partition of the columns of the Jacobian, which divides the set $\{1, 2, \dots, n\}$ into p subsets $c_1, c_2, \dots, c_p$, and given $x^{-1}, x^0 \in R^n$ such that $s_i^{-1} \equiv x_i^0 - x_i^{-1} \neq 0$, $i=1, 2, \dots, n$, at each step $k \geq 0$:

(1). Set

$$g_i^{k-1} = \sum_{j=1}^i \sum_{l \in c_j} s_l^{k-1} c_l, \quad i=1, 2, \dots, p-1, \quad g_0^{k-1} = 0.$$

where $s^{k-1} = x^k - x^{k-1}$.

(2). Compute $F(x^k - g_i^{k-1})$, $i=0, 1, \dots, p-1$, and set

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

where $F(x^k - g_p^{k-1}) = F(x^{k-1})$.

(3). If $(l, m) \in M$ and $|s_m^{k-1}| \geq \theta$, then set

$$b_{lm}^k = \frac{c_l^T y_i^{k-1}}{s_m^{k-1}}, \quad (2.13)$$

otherwise set

$$b_{lm}^k = b_{lm}^{k-1} ,$$

where $m \in c_i$, $i=1, 2, \dots, p$.

(4). Solve $B_k s^k = -F(x^k)$.

(5). Choose x^{k+1} by $x^{k+1} = x^k + s^k$, or by a global strategy.

(6). Check for convergence.

The SFD algorithm is also an update algorithm, and the update can be written as

$$\bar{B} = B \sum_{j \in \Omega_2} c_j c_j^T + \sum_{i=1}^p \sum_{j \in c_i \cap \Omega_1} J_i c_j c_j^T , \quad (2.14)$$

The following result shows that the SFD algorithm is a secant algorithm.

Lemma 2.3. $\bar{B}$ satisfies the secant equations

$$\bar{B} d_i = y_i , \quad i = 1, \dots, p , \quad (2.15)$$

and (2.15) implies

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

The proof of Lemma 2.3 is straightforward.

Suppose that we have finished the k th step of the iteration. Then the information we have is x^k , $F(x^k)$, B_k , and x^{k+1} . Let

$$d_i^k = \sum_{j \in c_i} s_j^k c_j , \quad i = 1, \dots, p ,$$

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

and

$$J_i^{k+1} = \int_0^1 F'(x^{k+1} - g_i^k + t d_i^k) dt , \quad i = 1, \dots, p. \quad (2.16)$$

Then by (2.6),

$$J_i^{k+1} c_j = B_{k+1} c_j , \quad (2.17)$$

where $j \in c_i$, $i=1, 2, \dots, p$.

Theorem 2.4. Assume that F' satisfies Lipschitz condition (2.8). Let $\{x^j\}_{j=0}^{k+1}$ and $\{B_j\}_{j=0}^{k+1}$ be generated by Algorithm 2.2. Suppose that $\{x^{j+1} - g_i^j, i=0, 1, 2, \dots, p\}_{j=0}^k \subset D$. If $s_i^k = 0$ appears

consecutively in at most m steps for any specific $1 \leq i \leq n$, then for $k \geq m$,

$$\|F'(x^{k+1}) - B_{k+1}\|_F \leq \alpha \sum_{j=k-m}^k \|x^{j+1} - x^j\|. \quad (2.18)$$

Proof. By the hypothesis of the theorem, given k , for any $1 \leq i \leq n$, there exists at least one integer $0 \leq j \leq m$ such that $s_i^{k-j} \neq 0$. Let $j(k, i)$ be the smallest one of these integers. Then,

$$B_{k+1}e_i = B_{k-j(k,i)+1}e_i.$$

Let $i \in c_l$. Then,

$$B_{k-j(k,i)+1}e_i = J_l^{k-j(k,i)+1}e_i$$

by (2.17). Therefore,

$$\begin{aligned} & \|(F'(x^{k+1}) - B_{k+1})e_i\| \\ &= \|(F'(x^{k+1}) - B_{k-j(k,i)+1})e_i\| \\ &\leq \|(F'(x^{k+1}) - F'(x^{k-j(k,i)+1}))e_i\| + \|(F'(x^{k-j(k,i)+1}) - B_{k-j(k,i)+1})e_i\| \\ &= \|(F'(x^{k+1}) - F'(x^{k-j(k,i)+1}))e_i\| + \|(F'(x^{k-j(k,i)+1}) - J_l^{k-j(k,i)+1})e_i\| \\ &\leq \alpha_i \|x^{k+1} - x^{k-j(k,i)+1}\| + \alpha_i \|x^{k-j(k,i)+1} - x^{k-j(k,i)}\| \\ &\leq \alpha_i \sum_{l=k-j(k,i)}^k \|x^{l+1} - x^l\| \\ &\leq \alpha_i \sum_{l=k-m}^k \|x^{l+1} - x^l\|. \end{aligned}$$

Hence,

$$\begin{aligned} \|F'(x^{k+1}) - B_{k+1}\|_F^2 &= \sum_{i=1}^n \|(F'(x^{k+1}) - B_{k+1})e_i\|^2 \\ &\leq \left(\sum_{j=k-m}^k \|x^{j+1} - x^j\| \right)^2 \sum_{i=1}^n \alpha_i^2. \end{aligned} \quad (2.19)$$

Then, (2.18) follows from (2.19).

3. A Kantorovich-Type Analysis.

The following estimate for the SFD algorithm is sharper than that for Broyden's algorithm given by Dennis [4].

Theorem 3.1. Assume that F' satisfies Lipschitz condition (2.8) and that $\{x^k\}$ and $\{B_k\}$ are generated by Algorithm 2.2 with $\|x^{-1} - x^0\| \leq \delta$. If $\{x^{j+1} - g_j^j, j=0, 1, \dots, p\}_{j=0}^k \subset D$, then

$$\|F'(x^{k+1}) - B_{k+1}\|_F \leq \alpha \sum_{j=0}^k \|x^{j+1} - x^j\| + \alpha \delta. \quad (3.1)$$

Proof. Inequality (3.1) can be obtained immediately by setting $m = k+1$ in (2.18).

Theorem 3.2. Let F' satisfy Lipschitz condition (2.8). Suppose that $x^{-1}, x^0 \in D$, and that B_0 , generated by x^{-1} and x^0 , is a nonsingular $n \times n$ matrix such that

$$\begin{aligned} \|x^{-1} - x^0\| \leq \delta, \quad \|B_0^{-1}\|_F \leq \beta, \quad \|B_0^{-1}F(x^0)\| \leq \eta, \\ h = \frac{\alpha^2 \beta \eta}{(1 - 3\alpha\beta\delta)^2} \leq \frac{1}{6}, \end{aligned} \quad (3.2)$$

and

$$\alpha\beta\delta < \frac{1}{3}.$$

If $\bar{S}(x^0, 2t^*) \subset D$, where

$$t^* = \frac{1 - 3\alpha\beta\delta}{3\alpha\beta} (1 - \sqrt{1 - 6h}), \quad (3.3)$$

then $\{x^k\}$, generated by Algorithm 2.2 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

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

Proof. Consider the scalar iteration

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

where

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

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

$$t_{k+1} - t_k = 3\beta \left[\frac{\alpha}{2} (t_k - t_{k-1}) + \alpha t_{k-1} + \alpha\delta \right] (t_k - t_{k-1}), \quad k = 1, 2, \dots.$$

From this equation, it follows that $\{t_k\}$ is a monotonically increasing sequence and

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

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

Now, by induction we will prove the following estimate:

$$||x^{k+1} - x^k|| \leq t_{k+1} - t_k , \quad k = 0, 1, 2, \dots . \quad (3.4)$$

For $k = 0$, we have

$$||x^1 - x^0|| \leq \eta = t_1 - t_0 .$$

Suppose that (3.4) holds for $k = 0, 1, 2, \dots, m-1$. Then

$$||x^m - x^0|| \leq t^* .$$

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

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

Using Theorem 3.1, we have

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

Then by Banach lemma,

$$||B_m^{-1}|| \leq \frac{\beta}{1-2/3} = 3\beta .$$

Hence,

$$\begin{aligned} &||x^{m+1} - x^m|| \\ &\leq ||B_m^{-1}||_F ||F(x^m) - F(x^{m-1}) - B_{m-1}(x^m - x^{m-1})|| \\ &\leq 3\beta \left[\frac{\alpha}{2} ||x^m - x^{m-1}|| + \alpha \sum_{i=0}^{m-2} ||x^{i+1} - x^i|| + \alpha\delta \right] ||x_m - x_{m-1}|| \\ &\leq 3\beta \left[\frac{\alpha}{2} ||(t_m - t_{m-1}) + \alpha t_{m-1} + \alpha\delta|| (t_m - t_{m-1}) = t_{m+1} - t_m . \end{aligned}$$

This completes the induction. By (3.4), 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 Ortega and Rheinboldt's Theorem 12.6.4 [8, p.425] by setting $A(x) \equiv B_0$.

4. Local Convergence Properties.

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

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

Theorem 4.1. Assume that $F:D \subset R^n \rightarrow R^n$ satisfies (4.1) and that F' satisfies Lipschitz condition (2.8). Let $\{x^k\}$ be generated by Algorithm 2.2. without any global strategy. Then there exist $\epsilon, \delta > 0$ such that if $x^{-1}, x^0 \in D$ satisfy

$$\|x^0 - x^*\| < \epsilon, \quad \|x^{-1} - x^0\| \leq \delta,$$

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

Proof. Notice that we can choose ϵ small enough so that $\|B_0^{-1}F(x^0)\|$ is also small such that $h < \frac{1}{6}$ and that $\bar{S}(x^0, 2t^*) \subset D$, where h and t^* are defined by (3.2) and (3.3) respectively.

When δ is small enough, we also have

$$\alpha\beta\delta < \frac{1}{3},$$

where β is defined in Theorem 3.2. Therefore, by Theorem 3.2,

$$\{x^{k+1} + g_i^k, i = 1, 2, \dots, p\} \subset D, \quad k=0, 1, \dots$$

Thus, from (2.14) and the proof of Theorem 2.1, we have

$$\begin{aligned} & \|F'(x^*) - \bar{B}\|_F^2 \\ &= \sum_{i \in \Omega_1} \|(F'(x^*) - \bar{B})e_i\|^2 + \sum_{i \in \Omega_2} \|(F'(x^*) - B)e_i\|^2 \\ &= \sum_{i=1}^p \sum_{j \in \epsilon_i \cap \Omega_1} \|(F'(x^*) - J_i)e_j\|^2 + \sum_{i \in \Omega_2} \|(F'(x^*) - B)e_i\|^2 \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^p \sum_{j \in \epsilon_i \cap \Omega_1} \left\| \int_0^1 [F'(x^*) - F'(\bar{x} - g_i + t(g_i - g_{i-1}))] e_i dt \right\|^2 \\
&+ \sum_{i \in \Omega_2} \|(F'(x^*) - B)e_i\|^2 \\
&\leq \alpha^2(\|x^* - \bar{x}\| + \|\bar{x} - x\|)^2 + \|F'(x^*) - B\|_F^2 \\
&\leq \alpha^2(2\|x^* - \bar{x}\| + \|x^* - x\|)^2 + \|F'(x^*) - B\|_F^2.
\end{aligned}$$

Therefore,

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

where

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

Notice that by Theorem 2.1 and Lipschitz condition (2.8),

$$\begin{aligned}
&\|F'(x^*) - B_0\|_F \\
&\leq \|F'(x^*) - F'(x^0)\|_F + \|F'(x^0) - B_0\|_F \\
&\leq \alpha\|x^* - x^0\| + \alpha\|x^0 - x^{-1}\| \\
&\leq \alpha(\epsilon + \delta)
\end{aligned}$$

Thus, by Dennis and More's [5] Theorem 5.1, we know that $\{x^k\}$ converges at least q -linearly to x^* .

According to Dennis and More's [5] Theorem 3.1, to get q -superlinear convergence, we need only to prove that

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - F'(x^*))(x^{k+1} - x^k)\|}{\|x^{k+1} - x^k\|} = 0. \quad (4.2)$$

If for all $1 \leq i \leq n$, $s_i^k = 0$ appears consecutively in at most m steps, then by Theorem 2.4 it is easy to show that

$$\lim_{k \rightarrow \infty} \|B_k - F'(x^*)\|_F = 0. \quad (4.3)$$

Thus, (4.2) follows immediately from (4.3).

Otherwise, let

$$A_2 = \{i \in \{1, 2, \dots, n\} : \text{For any } k > 0, \text{ there exists at least one integer } m > k \text{ such that } s_i^m \neq 0\},$$

and let $A_1 = \{1, \dots, n\} \setminus A_2$. Then

$$B_k - F'(x^*) = \sum_{i \in A_1} (B_k - F'(x^*))e_i e_i^T + \sum_{i \in A_2} (B_k - F'(x^*))e_i e_i^T.$$

From the definition of A_1 , there exists a large integer k_0 such that $s_i^k = 0$ for all $i \in A_1$ and $k > k_0$.

Therefore,

$$\sum_{i \in A_1} (B_k - F'(x^*))e_i e_i^T (x^{k+1} - x^k) = 0, \quad (4.4)$$

for $k > k_0$. Now we show that

$$\lim_{k \rightarrow \infty} \left\| \sum_{i \in A_2} (B_k - F'(x^*))e_i e_i^T \right\|_F = 0. \quad (4.5)$$

In the first part of the proof, we proved that $\lim_{k \rightarrow \infty} \|x^k - x^*\| = 0$. This implies that given $\epsilon > 0$,

there exists an integer K such that

$$\|x^k - x^*\| < \frac{\epsilon}{3\alpha}, \quad \forall k > K.$$

By the definition of A_2 , there exists an integer K_1 , which depends on K , such that for every $i \in A_2$, there exists at least one integer $0 < j < K_1$ such that $s_i^{K+j} \neq 0$. Let $\bar{K} = K + K_1$. For $k > \bar{K}$ and $i \in A_2$, define

$$j(k, i) = \min\{j : s_i^{k-j} \neq 0\}.$$

Then $k - j(k, i) > K$. Let $i \in c_l$, $1 \leq l \leq p$, we have that

$$B_k e_i = B_{k-j(k,i)+1} e_i = J_i^{k-j(k,i)+1} e_i.$$

Thus, by Lipschitz condition (2.8),

$$\begin{aligned}
& ||(B_k - F'(x^*))e_i||^2 \\
&= ||(J_t^{k-j(k,i)+1} - F'(x^*))e_i||^2 \\
&= ||\int_0^1 (F'(x^{k-j(k,i)+1}) - g_i^{k-j(k,i)} + t(g_i^{k-j(k,i)} - g_{i-1}^{k-j(k,i)})) - F'(x^*))e_i dt||^2 \\
&\leq (\int_0^1 \alpha_i ||x^{k-j(k,i)+1} - g_i^{k-j(k,i)} + t(g_i^{k-j(k,i)} - g_{i-1}^{k-j(k,i)}) - x^*|| dt)^2 \\
&\leq \alpha_i^2 (||x^{k-j(k,i)+1} - x^*|| + \frac{1}{2} ||g_i^{k-j(k,i)}|| + \frac{1}{2} ||g_{i-1}^{k-j(k,i)}||)^2 \\
&\leq \alpha_i^2 (||x^{k-j(k,i)+1} - x^*|| + ||x^{k-j(k,i)+1} - x^{k-j(k,i)}||)^2 \\
&\leq \alpha_i^2 (2||x^{k-j(k,i)+1} - x^*|| + ||x^{k-j(k,i)} - x^*||)^2 \\
&< \alpha_i^2 (2\frac{\epsilon}{3\alpha} + \frac{\epsilon}{3\alpha})^2 = \alpha_i^2 \frac{\epsilon^2}{\alpha^2} .
\end{aligned}$$

Then

$$\begin{aligned}
& ||\sum_{i \in A_2} (B_k - F'(x^*))e_i e_i^T||_F^2 . \\
&= \sum_{i \in A_2} ||(B_k - F'(x^*))e_i||^2 \\
&< \frac{\epsilon^2}{\alpha^2} \sum_{i \in A_2} \alpha_i^2 \leq \epsilon^2 .
\end{aligned}$$

Therefore,

$$||\sum_{i \in A_2} (B_k - F'(x^*))e_i e_i^T||_F < \epsilon .$$

This completes the proof of (4.5).

By (4.4) and (4.5),

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \frac{|| (B_k - F'(x^*))(x^{k+1} - x^k) ||}{||x^{k+1} - x^k||} \\
&= \lim_{k \rightarrow \infty} \frac{||\sum_{i \in A_2} (B_k - F'(x^*))e_i e_i^T (x^{k+1} - x^k)||}{||x^{k+1} - x^k||} \\
&\leq \lim_{k \rightarrow \infty} ||\sum_{i \in A_2} (B_k - F'(x^*))e_i e_i^T||_F = 0 .
\end{aligned}$$

Theorem 4.2. Assume that F , x^{-1} , x^0 and $\{x^k\}$ satisfy the hypotheses of Theorem 4.1. If, for any $1 \leq i \leq n$, $s_i^k = 0$ appears consecutively in at most m steps, then the r -convergence order is not less than τ , where τ is the unique positive root of

$$t^{m+2} - t^{m+1} - 1 = 0. \quad (4.6)$$

In particular, if $s_i^k \neq 0$, $i = 1, \dots, n$, $k = 1, 2, 3, \dots$, then $\tau = \frac{1 + \sqrt{5}}{2} \approx 1.618$.

Proof. Notice that (4.3) implies that there exist k_0 and $\beta > 0$ such that $\|B_k^{-1}\| \leq \beta$ for all $k \geq k_0$. Thus, by Theorem 2.4,

$$\begin{aligned} \|x^{k+1} - x^*\| &= \|x^k - x^* - B_k^{-1}F(x^k)\| \\ &\leq \|B_k^{-1}\|_F \{ \|F(x^k) - F(x^*) - F'(x^*)(x^k - x^*)\| \\ &\quad + (\|F'(x^*) - F'(x^k)\|_F + \|F'(x^k) - B_k\|_F) \|x^k - x^*\| \} \\ &\leq \beta \left\{ \frac{3}{2} \alpha \|x^k - x^*\| + \alpha \sum_{j=k-m-1}^{k-1} \|x^{j+1} - x^j\| \right\} \|x^k - x^*\| \\ &\leq \frac{5}{2} \alpha \beta \left(\sum_{j=k-m-1}^k \|x^j - x^*\| \right) \|x^k - x^*\|. \end{aligned} \quad (4.7)$$

Thus, the desired result follows from Ortega and Rheinboldt's Theorem 9.2.9 [8, p.291].

5. Numerical Results.

We computed some examples with tridiagonal Jacobians by the CPR algorithm and Algorithm 2.2. In this section, we compare the numerical results from the two algorithms. The global strategy we used in computing the examples is the line search with backtracking strategy (see Dennis and Schnabel [6]. If $p^k = -B_k^{-1}F(x^k)$ is not a descent direction, then we try $-p^k$. If it is not a descent direction either, then the algorithm fails. The stopping test we used is

$$\max_{1 \leq i \leq n} \frac{|x_i^{k+1} - x_i^k|}{\max\{|x_i^{k+1}|, \text{typ}x_i\}} \leq \epsilon, \quad (5.1)$$

and we choose $\text{typ}x_i = 10^{-8}$ and $\epsilon = 10^{-5}$. For the lower bound of $|s_j|$, we choose

$$\theta = \sqrt{\text{macheps}} \|s\|.$$

We used double precision, and the machine precision is $2.22d-16$.

Example 5.1 is new, and it can be seen to be an extension of the Rosenbrock [9] function (also see Moré, Garbow and Hillstom [7]) to nonlinear system of equations with a tridiagonal

structure. Example 5.2 was given by Broyden [1] (also see Moré, Garbow and Hillstom [7]).

The results are shown in the tables below, where IT is the number of iterations, NF is the number of function ($F(x)$) evaluations, and LN is the number of line searches in which the step length $\lambda < 1$. ND is the number of nondecrease directions. ZR is the number of the iterations that there exists an integer j such that $|s_j| < \theta$. x_0 is the initial guess.

Example 5.1.

$$f_1(x) = 8(x_1 - x_2^2),$$

$$f_j(x) = 16x_j(x_j^2 - x_{j-1}) - 2(1 - x_j) + 8(x_j - x_{j+1}^2), \quad j = 2, \dots, n-1,$$

$$f_n(x) = 16x_n(x_n^2 - x_{n-1}) - 2(1 - x_n),$$

$$n = 9$$

$$x_1 = (-1, -1, \dots, -1)^T, \quad x_2 = (-0.5, -0.5, \dots, -0.5)^T, \quad x_3 = (2, 2, \dots, 2)^T.$$

Algorithms	$x_0 = x_1$				
	IT	NF	LN	ND	ZR
CPR	22	88	15	0	0
Alg. 2.2	24	73	13	4	0

Table 5.1 (a).

Algorithms	$x_0 = x_2$				
	IT	NF	LN	ND	ZR
CPR	22	88	15	0	0
Alg. 2.2	22	67	13	1	0

Table 5.1 (b).

Algorithms	$x_0 = x_3$				
	IT	NF	LN	ND	ZR
CPR	8	32	0	0	0
Alg. 2.2	10	31	0	0	0

Table 5.1 (c).

Example 5.2 (Broyden tridiagonal function).

$$f_i(x) = (3 - 2x_i)x_i - x_{i-1} - 2x_{i+1} + 1,$$

$$x_0 = x_{n+1} = 0,$$

$$n = 9,$$

$$x1 = (-1, -1, \dots, -1)^T, \quad x2 = (-0.3, 0.3, \dots, -0.3, 0.3)^T,$$

$$x3 = (-10, -10, \dots, -10)^T.$$

Algorithms	$x0=x1$				
	IT	NF	LN	ND	ZR
CPR	5	20	0	0	0
Alg. 2.2	6	19	0	0	0

Table 5.2 (a).

Algorithms	$x0=x2$				
	IT	NF	LN	ND	ZR
CPR	5	20	0	0	0
Alg. 2.2	6	19	0	0	0

Table 5.2 (b).

Algorithms	$x0=x3$				
	IT	NF	LN	ND	ZR
CPR	8	32	0	0	0
Alg. 2.2	10	31	0	0	0

Table 5.2 (c).

6. Concluding Remarks.

We have presented an algorithm for solving sparse nonlinear systems of equations. This algorithm is based on consistent partitions of the columns of the Jacobians, and it is a combination of the CPR-CM algorithm and a secant algorithm. This algorithm incorporates the advantages of the CPR-CM algorithm and secant algorithms in such a way that it reduces by one the number of function evaluations required by the CPR-CM algorithm at each iteration, and it has

good local convergence properties. We have shown that the SFD algorithm is locally q -superlinearly convergent, and that under reasonable assumptions, the r -convergence order of the SFD algorithm is not less than $\frac{1+\sqrt{5}}{2}$, which is the r -convergence order of the one dimensional secant algorithm. Our numerical results indicate that when p , the number of the groups in a partition of the columns of the Jacobian, is not large, the SFD algorithm is probably more efficient than the CPR-CM algorithm.

The idea exploited here can also be used with Powell and Toint's [10] work, which will lead to a method for unconstrained optimization problems. This will be our future work.

Acknowledgement. This paper is part of my Ph. D. thesis. I would like to express my deepest thanks to my advisor John Dennis. He read this paper several times and made many helpful suggestions and corrections which improved this paper a lot. I am also very grateful to him for his support, his guidance and his encouragement.

References

- [1]. Broyden, C.G., A class of methods for solving nonlinear simultaneous equations, *Math. Comp.*, 19, pp. 577-593.
- [2]. Coleman, T.F., and J.J. Moré, Estimation of sparse Jacobian and graph coloring problems, *SIAM J. Numer. Anal.*, 20, pp. 187-209.
- [3]. Curtis, A.R., M.J.D. Powell and J.K. Reid, On the estimation of sparse Jacobian matrices, *IMA J. Appl. Math.*, Vol. 13, pp. 117-119.
- [4]. Dennis, J.E., Jr. and J.J. Moré, Quasi-Newton methods, motivation and theory, *SIAM Review*, Vol. 19, No. 1 (1977).
- [5]. Dennis, J.E., Jr., On the convergence of Broyden's method for nonlinear systems of equations, *Math. Comp.*, Vol. 25, No. 115 (1971), pp. 559-567.
- [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]. Moré, J.J., B.S. Garbow and K.E. Hillstom, Testing unconstrained optimization software, *ACM Transactions on Mathematical Software*, Vol. 7, No. 1 (1981), pp. 17-41.
- [8]. Ortega, J.M., and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York (1970).
- [9]. Rosenbrock, H.H., An automatic method for finding the greatest or least value of a function, *Comput. J.* 5 (1962), pp.147-151.
- [10]. Powell, M.J.D., and P.H.L. Toint, On the estimation of sparse Hessian matrices, *SIAM J. Numer. Anal.*, Vol. 16 (1979), pp. 1060-1074.