ON THE CONVERGENCE OF BROYDEN'S METHOD FOR NONLINEAR SYSTEMS OF EQUATIONS

By

J.E. Dennis Jr.

Technical Report No. 69-48

ON THE CONVERGENCE OF BROYDEN'S METHOD FOR NONLINEAR SYSTEMS OF EQUATIONS

By

J.E. Dennis Jr.

Department of Mathematics
University of Utah
Salt Lake City, Utah 84112

currently on leave at

Department of Computer Science
Cornell University
Ithaca, New York 14850

INTRODUCTION C.G. Broyden [2] suggested an algorithm for iterating to a solution of a system of nonlinear equations which has shown its mettle in dealing with practical problems. The purpose of this paper is to provide a Kantorovich-type analysis for this method. In fact, the analysis is applicable to the entire class of 'single-rank update' methods just to the extent that it seems to heuristically justify the generally superior performance of Broyden's method over the rest of the class.

These methods generate sequences $\{x_n\},\{H_n\}$, one consisting of approximate roots and the other of the corresponding approximate inverse Jacobian matrices. At the nth step one obtains x_{n+1} from x_n and H_n by setting

$$x_{n+1} = x_n - \gamma_n H_n F(x_n) \qquad (1)$$

where $G(x_n) = (f_1(x_n),\ldots,f_n(x_n))^T$ is the residual vector at the point x_n and γ_n is a real number about which more will be said later. H_{n+1} is obtained by using an idea due to Davidon [5]. It is required to satisfy the equation $H_{n+1}y_n \equiv H_{n+1}(F(x_{n+1}) - F(x_n)) = x_{n+1} - x_n$. (2)

This seems a very clever use of the small amount of new information furnished by $F(x_{n+1})$, since it is equivalent to

$$\int_{x_n}^{x_{n+1}} (H_{n+1}J(x) - I)dx = 0,$$

where J denotes the Jacobian matrix. It makes H_{n+1} look like a very reasonable approximation to the inverse Jacobian somewhere between x_n and x_{n+1}, at least in the direction $x_{n+1} - x_n$. For the single-rank methods the choice of H_{n+1} from the class of $N \times N$ matrices satisfying (2) is

$$H_{n+1} = H_n - (H_n y_n + \gamma_n H_n F(x_n)) \frac{d_n^T H}{d_n^T y_n} \qquad (3)$$

If $B_i = H_i^{-1}$ then

$$B_{n+1} = B_n - (y_n + \gamma_n F(x_n) \frac{d_n^T B_n}{\gamma_n d_n^T F(x_n)} \qquad (3')$$

where d_n is, of course, chosen so that $d_n^T y_n \neq 0$.

Broyden [3] has shown that for his method, $d_n = H_n^T H_n F(x_n)$, if $J(x) = L$ a constant matrix, ie, the system is linear, then $\{x_n\}$ and $\{H_n\}$ converges to the x^*, the root and L^{-1} respectively from any x_0 and any H_0 sufficiently close to L^{-1}.

Such a result would be too much to hope for when working with nonlinear systems but in section 2 we will show that the rate of deteriation in the approximation of $J(x_n)$ by B_n depends on the nonlinearity of F in a very simple way.

This will enable us to analize the method as a Newton-like method [7] in section 3. In section 4 we will draw some reasonable though non-rigorous conclusions, based on the results of section 3, about the choice of γ .

THE APPROXIMATION OF $J(x_n)$ BY B_n.

Let x_0 be a point in the domain of F and suppose that the function J which maps x into the Jacobian of F at x satisfies a Lipschitz condition on some set D_0 which contains x_0 in its interior, ie; Let $K \geq 0$ and $D_0 \subset E^N$ be such that D_0 is a convex set and is the closure of its interior. Assume that for every $x,x' \in D_0$, $||J(x)-J(x')|| \leq K||x-x'||$. Here, the norm on the left is the subordinate matrix norm, ie, $||A|| = \sqrt{\rho(A^TA)}$, $\rho(B)$ is the maximum modulus of any eigenvalue of B. The following lemmas will be useful.

LEMMA 1 Let $x,x' \in D_0$, then $||F(x) - F(x') - J(x')(x-x')|| \leq \frac{K}{2}||x-x'||^2$.

Proof: See [4] or [7].

LEMMA 2 Let $x, x' \in E^N$, with $x^Tx'=1$, then $||I-x'x^T||=||x'||\,||x||$.

PROOF See [3] .

LEMMA 3 Let $x,x' \in D_0$ and let B be an $N \times N$ matrix. Let $d \in E^N$ with $d^TF(x) \neq 0$ and let B' be derived from B by

$$B' = B-((F(x')-F(x)-B(x'-x)) \frac{d^TB}{d^TB(x'-x)} .$$ Then define

$$q \equiv \frac{||x'-x||\,||d^TB||}{|d^TB(x'-x)|}, \text{ and } ||B'-J(x')|| \leq q||B-J(x)||+K(1+ q/2)||x'-x||.$$

Proof: $B' - J(x') = B - J(x) + (J(x)(x' - x) - B(x' - x)) \frac{d^T B}{d^T B(x'-x)} + J(x) - J(x') - (J(x)(x' - x) - (F(x') - F(x)) \frac{d^T B}{d^T B(x'-x)}$.

Now, by norming both sides and using Newton's formula, the Lipschitz condition on J, Lemma 1 and the Cauchy and triangle inequalities we obtain

$$||B' - J(x')|| \leq ||B - J(x)|| \cdot ||I - \frac{(x'-x)d^T B}{d^T B(x'-x)}||$$

$$+ K||x' - x|| + \frac{K}{2} \frac{||d^T B||}{||d^T B(x'-x)||} ||x' - x||^2 .$$

(4) now follows from Lemma 2.

We will now apply this lemma to the case when $B = B_n$, $B' = B_{n+1}$, $x = x_n$, $x' = x_{n+1}$, $d = d_n$. In the next theorem, interpet the quotient q_n in the obvious way and define $q_{-1} \equiv 1$.

THEOREM 1 Let $x_0, \ldots, x_{n+1}$, $B_0, \ldots, B_{n+1}$ be generated by any single rank method and let $\{x_i : i=1,\ldots,n+1\} \subset D_0$, then

$$||B_{n+1} - J(x_{n+1})|| \leq (\prod_{j=0}^{n} q_j)||B_0 - J(x_0)|| \qquad (5)$$

$$+ K \sum_{j=0}^{n} (\prod_{i=-1}^{n-j-1} q_i) (1 + q_j/2)||x_{j+1} - x_j|| .$$

Proof: The proof will be by induction on $n + 1$.

Let $n + 1 = 1$ and by making the proper substitutions in (4)

$$||B_1 - J(x_1)|| \leq q_0||B_0 - J(x_0)|| + K(1 + q_0/2)||x_1 - x_0||,$$

which is (5) for $n = 0$. Assume by way of induction that (5) holds for $n + 1 \leq k$. Then again by (4) $||B_{k+1} - J(x_{k+1})|| \leq$

$q_k [||B_k - J(x_k)|| + K(1 + q_k/2)||x_{k+1} - x_k||$. Now by the induction hypothesis, (5) with $n + 1 = k$, $||B_k - J(x_k)|| \leq$

$$\left(\prod_{j=0}^{k-1} q_j\right)||B_0 - J(x_0)|| + K \sum_{j=0}^{k-1} \left(\prod_{i=-1}^{k-1-j-1} q_i\right)(1 + q_j/2)||x_{j+1} - x_j||,$$

and so $||B_{k+1} - J(x_{k+1})|| \leq q_k[\left(\prod_{j=0}^{k-1} q_j\right)||B_0 - J(x_0)|| +$

$$K \sum_{j=0}^{k-1} \left(\prod_{i=-1}^{k-1-j-1} q_i\right)(1 + q_j/2)||x_{j+1} - x_j||)]$$

$+ K(1 + q_k/2)||x_{k+1} - x_k|| =$ (5) with $n + 1 = k + 1$ and the induction is complete.

Now obviously $q_n \geq 1$ and just as obviously any analysis based on (5) needs $q_n = 1$ so that $\prod q_j$ is uniformly bounded. There are various ways of 'fudging' however in order to control the deteriation [1], [6].

The Cauchy inequality tells us that $q_n = 1$ if and only

if there is a constant $c_n \neq 0$ such that $B_n^T d_n = c_n(x_{n+1}-x_n) =$ $c_n' H_n F_n$, ie, $d_n = c_n' H_n^T H_n F_n$, Broyden's choice. If we view a particular method as the selection function for d_n then a glance at (3) and (3') convinces us that two methods are the same if $d_n' = c_n d_n$ for some sequence $\{c_n\}$ of nonzero constants. Hence Broyden's method is the unique single rank method which is naturally of bounded deteriation.

3. A KANTOROVICH-TYPE ANALYSIS FOR BROYDEN'S METHOD

In [7] the author considers the convergence of a class of Newton-like methods of the form $x_{n+1} = x_n - A(x_n)^{-1}F(x_n)$. (6) An immediate corollary of the results there is the following extension of Rheinboldt's Theorem [10].

THEOREM 2 Let F be as above and let A have the property that given any $x \in D_0$, $A(x)$ is an $N \times N$ real matrix. Let δ_0, δ_1 be nonnegative real numbers such that

$$||A(x) - J(x)|| \leq \delta_0 + \delta_1||x - x_0|| \qquad (7)$$

for every $x \in D_0$ and let β and γ be real numbers such that $A(x_0)^{-1}$ exists and $||A(x_0)^{-1}|| \leq \beta$, $||A(x_0)^{-1}F(x_0)|| \leq \gamma$

Then, $1>\beta\delta_o$, $1/2 \geq h' \equiv \frac{\beta K \eta}{(1-\beta\delta_o)^2}$ and $N(x_o, r_o') \subset D_o$ where $r_o' \equiv \frac{1-\sqrt{1-2h'}}{\beta K}(1-\beta\delta_o)$, imply that F has a root $x^* \in \bar{N}(x_o, r_o')$ which is unique in $D_o \cap N(x_o, r_1)$ where $r_1 \equiv \frac{1+\sqrt{1-2h'}}{\beta K}(1-\beta\delta_o)$. Furthermore, $x'_{n+1} = x'_n - A(x_o)^{-1} F(x'_n)$ converges to x^* from any $x'_o \in D_o \cap N(x_o, r_1)$.

If in addition, $1>3\beta\delta_o$, $1/2 \geq h \equiv \frac{1-\sqrt{1-2h}}{(2\delta_1+K)\beta}(1-3\beta\delta_o)$, then the sequence $\{x_n\}$ generated by (6) converges to x^*.

The fact that the theorem ensures convergence of the $\{x'_n\}$ sequence under less strengent conditions than for the $\{x_n\}$ sequence is a characteristic of this type of theorem and results from using the $\{x'_n\}$ sequence to establish the existence and uniqueness of x^*.

If we write (5) for the special case of Broyden's method we obtain

$$||B_{n+1} - J(x_{n+1})|| \leq ||B_o - J(x_o)|| + \frac{3K}{2} \sum_{j=o}^{n} ||x_{j+1} - x_j|| \quad (5)$$

which, except for a wrong-way triangle inequality, looks like a version of (7) with $\delta_o = ||B_o - J(x_o)||$, $\delta_1 = 3K/2$ and the function A only defined at the iteration points. One is immediately led to conjecture

the following theorem for Broyden's method.

THEOREM 3 Let F be as above and let B_0 be a nonsingular $N \times N$ matrix such that $||J(x_0) - B_0|| \leq \delta$, $||H_0|| \leq \beta$ and $||H_0 F(x_0)|| \leq \eta$. Then $1 > \beta\delta$, $1/2 \geq h \equiv \frac{\beta K \eta}{(1-\beta\delta)^2}$ and $N(x_0, r_0') \subset D_0$ where $r_0' = \frac{1-\sqrt{1-2h}}{\beta K}(1-\beta\delta)$ imply that F has a root x^*, $||x_0 - x^*|| \leq r_0'$, and x^* is unique in $D_0 \cap N(x_0, r_1')$ where $r_1' \equiv \frac{1+\sqrt{1-2h}}{\beta K}(1-\beta\delta)$. Furthermore, $x_{n+1}' = x_n' - H_0 F(x_n')$ converges to x^* from any $x_0' \in D_0 \cap N(x_0, r_1')$.

If in addition, $1 > 3\beta\delta$, $1/8 \geq h \equiv \frac{\beta K \eta}{(1-3\beta\delta)^2}$ and $N(x_0, r_0) \subset D_0$ where $r_0 \equiv \frac{1-\sqrt{1-8h}}{4\beta K}(1-3\beta\delta)$, then Broyden's method with $\gamma_n = 1$ for every n, converges to x^*.

This is, in fact, a theorem. One approach which would yield a proof, would be to generalize Theorem 2 to handle the situation when one has a sequence $\{A_n\}$ in place of a function A. defined for every x. We do not choose this path for esthetic reasons.

It seems to the author that writing mathematics is the process of arriving at an acceptable compromise between what one can prove and what one can conveniently state. In most

cases, this is worthwhile since theorems elucidate the structure of the mathematical context within which one is working. In this field it seems that one begins to lose part of the 'meat' as soon as he begins to make his preliminary definitions. The author feels (at least today) that the particular branch of the theory of Newton-like methods developed in [7] and [10] is perhaps at the point when it should be used to directly prove some theorems for specific methods rather than to generate an almost endless variety of possible theorems, like Theorem 2, which would then be matched to specific iterative methods in a manner reminicient of a computerized marriage bureau.

Proof of Theorem 3. The part of Theorem 3 concerning the existence and uniqueness of x^* and the convergence of the x' sequence can be gotten directly from Theorem 2 by, for example, setting $A(x) \equiv B_0$ for every $x \in D_0$. Thus we make the second set of assumptions and proceed to a consideration of the full Broyden sequence. The first step in the proof of Theorem 2 by the techniques of [7] is to show that $A(x)$ is always invertable and to find a scalar function $a(\cdot)$ such that $a(||x-x_0||)^{-1} \geq ||A(x)^{-1}||$ for every $x \in N(x_0, r_0)$. Let us assume that for $n \geq 0$, $\sum_{i=0}^{n} ||x_{i+1}-x_i|| < r_0$. Then from (5') $||H_0 B_{n+1} - I|| \leq \beta ||B_{n+1} - B_0|| \leq \beta [||B_{n+1} - J(x_{n+1})|| +$

$$||J(x_{n+1}) - J(x_0)|| + ||J(x_0) - B_0||]$$

$$\leq \beta(2\delta + \frac{3K}{2} \sum_{j=0}^{n} || x_{j+1} - x_j || + K||x_{n+1} - x_0||$$

$$\leq 2\beta\delta + \frac{5\beta K}{2} \sum_{j=0}^{n} || x_{j+1} - x_j || < 2\beta\delta + \frac{5}{2} \beta K r_0$$

$$= 2\beta\delta + 5/8\ (1-\sqrt{1-8h})(1-3\beta\delta) \leq 3\beta\delta + c(1-3\beta\delta) < 1$$

since $c < 1$. Hence, by the Banach Lemma [7], [8], $(H_0 B_{n+1})^{-1} H_0 = H_{n+1}$ exists and is bounded in norm by

$$\beta(1-2\beta\delta - \frac{5}{2} K \sum_{j=0}^{n} ||x_{j+1} - x_j||)^{-1} .$$

Let $b_0 = \beta^{-1}$ and $h_0 = \beta$, $t_0 = 0$. Define $f(t) = 2Kt^2 - (b_0 - 3\delta)t + b_0$, $h_k = \beta(1-2\beta\delta - \frac{5}{2} K t_k)^{-1}$, $h_k^{-1} = b_k$ and consider the sequence $t_{k+1} = t_k + h_k f(t_k)$. Notice that $r_0 > t_1 = \eta \geq ||H_0 F(x_0)|| = ||x_1 - x_0||$. Suppose now that $0 < t_k < r_0$, then since $f(r_0) = 0$, $r_0 - t_{k+1} = r_0 - t_k + h_k(f(r_0) - f(t_k)) = (r_0 - t_k)(1 + h_k f'(\xi))$, $\xi \in (t_k, r_0)$. Now $f'(\xi) = 3K\xi - (b_0 - 3\delta) < 4Kr_0 - (b_0 - 3\delta)$ $= (1-\sqrt{1-8h})(b_0 - 3\delta) - (b_0 - 3\delta) = -\sqrt{1-8h} \cdot (b_0 - 3\delta) < 0$. Hence $1 + h_k f'(\xi) < 1$ so $r_0 - t_{k+1} < r_0 - t_k$. Furthermore,

$h_k \geq \beta - h_o$, so $h_k f'(\xi) \geq -h_o(b_o - 3\delta) = -1 + 3\beta\delta > -1$ and

$1 + h_k f'(\xi) > 0$. Thus $0 < r_o - t_{k+1} < r_o - t_k$ and

$\{t_k\}$ is a strictly increasing sequence in $[0, r_o]$. Hence

it converges to $t^* \leq r_o$ if $t^* < r_o$, then $t^* = t^* + \beta(1 - 2\beta\delta - 5Kt^*/2)^{-1} f(t^*)$.

and so $f(t^*) = 0$. This ensures that $t^* = r_o$,

since r_o is the smallest root of f in $[0, r_o]$. We have

already that $t_1 \geq ||x_1 - x_o||$. Suppose now that $x_o, x_1, \ldots, x_k$

satisfy the properties that $||x_{i+1} - x_i|| \leq t_{i+1} - t_i$,

$i = 0, 1, \ldots, k-1$. Then $\sum_{j=0}^{k-1} ||x_{i+1} - x_i|| \leq t_k < r_o$ and H_k

exists with $||H_k|| \leq h_k$. Now apply Lemma 1 and (5') :

$$||x_{k+1} - x_k|| \leq h_k [||F(x_k) - F(x_{k-1}) - J(x_{k-1}) \cdot$$

$$(x_k - x_{k-1})|| + ||J(x_{k-1}) - B_{k-1}|| \cdot ||x_k - x_{k-1}||]$$

$$\leq h_k [\tfrac{1}{2} K||x_k - x_{k-1}||^2 + (\delta + \tfrac{3}{2} K \sum_{j=0}^{k-2} ||x_{j-1} - x_j||) \cdot$$

$$||x_k - x_{k-1}||] \leq h_k [\tfrac{1}{2} K(t_k - t_{k-1})^2 + (\delta + \tfrac{3}{2} Kt_{k-1})(t_k - t_{k-1})].$$

$\frac{f(t_{k-1})}{t_k - t_{k-1}} + f'(t_{k-1}) = b_{k-1} + f'(t_{k-1}) = b_o(1 - 2h_o\delta -$

$\frac{5}{2} Kt_{k-1}) + 4Kt_{k-1} - (b_o - 3\delta) = b_o - 2\delta - \frac{5}{2} b_o Kt_{k-1} +$

$4Kt_{k-1} - b_o + 3\delta = \delta + \frac{3}{2} Kt_{k-1}$. We can thus write

$||x_{k+1} - x_k|| \leq h_k[1/2\ K(t_k - t_{k-1})^2 + f'(t_{k-1})(t_k - t_{k-1})$ $+ f(t_{k-1})] = h_k f(t_k) = t_{k+1} - t_k$, since the bracketed expression is the Taylor expansion for the quadratic f. Now we can conclude by induction that $\sum_{j=0}^{k} ||x_{j+1} - x_j|| \leq t_k < r_o$ for every k. Hence $\{x_n\}$ is a Cauchy sequence and must converge to some $x^{**} \in \bar{N}(x_o, r_o)$. It only remains for us to show that $x^{**} = x^*$. From (5') we obtain

$||B_k|| \leq ||J(x_k)|| + \delta + 3/2\ Kt_k \leq ||J(x_o)|| + \delta + 3Kt_k$ $< ||J(x_o)|| + \delta + 3Kr_o \equiv C$. Hence $||F(x^{**})|| = \lim_k$ $||F(x_k)|| = \lim_k ||B_k|| \cdot ||x_{k+1} - x_k|| \leq C \lim_k ||x_{k+1} - x_k|| = 0$.

Now $r_o < r_1'$ so the unicity assertion concerning x^* ensures that $x^* = x^{**}$.

The following corollary is straight forward.

THEOREM 4 Let x^* be a root of the nonlinear system F. Let the first partial derivatives satisfy a Lipschitz condition of order 1 in some open set containing x^* and let $J(x^*)$ be invertable. Under these conditions there is an $\epsilon > 0$ and a $\delta > 0$ such that if x_o is any N-vector and B_o is any matrix satisfying $||x_o - x^*|| < \epsilon$ and $||J(x_o) - B_o|| < \delta$ then Broyden's method $\gamma_n^- - 1$ for all n converges to x^* from x_0.

4. THE CHOICE OF γ_n

In the introduction we made the obvious statement that d_n was chosen so that $d_n^T y_n \neq 0$. Clearly this requirement alone allows us to define H_{n+1} from H_n by (3). The following theorem clarifies this requirement.

THEOREM 5 Let H_n be a non-singular $N \times N$ matrix and let $d_n \in E^N$ such that $d_n^T y_n \neq 0$. Then H_{n+1} is nonsingular if and only if $d_n^T F(x_n) \neq 0$.

PROOF. First let us assume that $d_n^T F(x_n) = 0$. Direct substitution in (3) yields $H_{n+1} F(x_n) = d_n F(x_n)$. Also, $H_{n+1} F(x_{n+1}) = H_n F(x_{n+1}) - (H_n y_n + \gamma_n d_n F(x_n)) \dfrac{d_n^T F(x_{n+1})}{d_n^T y_n} = H_n F(x_{n+1}) - H_n y_n + \gamma_n H_n F(x_n) = (1-\gamma_n) H_n F(x_n)$.

If $\gamma_n = 1$, H_{n+1} is clearly singular. If $\gamma_n \neq 1$, then as above, $H_{n+1} F(x_{n+1}) = (1-\gamma_n) H_{n+1} F(x_n)$ and so $H_{n+1}(F(x_{n+1}) - (1-\gamma_n) F(x_n)) = 0$. Hence, either H_{n+1} is singular or $F(x_{n+1}) = (1-\gamma_n) F(x_n)$. If the latter were true then $d_n^T F(x_{n+1}) = 0$ and $d_n^T y_n = 0$ would result. This would contradict the hypotheses $d_n^T y_n \neq 0$ so H_{n+1} is singular.

Assume now that $H_{n+1}x = 0$ for some $x \neq 0$. Clearly, $d_n^T x \neq 0$ since H_{n+1} agrees with the nonsinguler matrix H_n on the orthogonal complement of d_n. From (3),

$$H_n x = (H_n y_n + \gamma_n H_n F(x_n)) \frac{d_n^T x}{d_n^T y_n} \quad \text{and so} \quad x = (y_n + \gamma_n F(x_n)) \frac{d_n^T x}{d_n^T y_n}$$

$$\text{Hence} \quad d_n^T x = (d_n^T y_n + \gamma_n d_n^T F(x_n)) \frac{d_n^T x}{d_n^T y_n} \quad \text{and so}$$

$\gamma_n d_n^T F(x_n) = 0$. Now if $\gamma_n \neq 0$, $F(x_{n+1}) = F(x_n)$ and $d_n^T y_n = 0$ so we can conclude that d_n is perpendicular to $F(x_n)$.

The previous theorem obviously has a more general analogue in the language of Lemma 3.

The proof of Theorem 5 could have been simplified by using (3') but we wanted the little result that $d_n^T F(x_n) = 0$ implies $H_{n+1}F(x_{n+1}) = (1-\gamma_n)H_n F(x_n)$.

Let us suppose that we are choosing γ_n by some minimisation criterion like $\min_\gamma ||F(x_n - \gamma H_n F(x_n))||$. In a difficult situation we may have to settle for a small γ_n and $F(x_{n+1})$ close to $F(x_n)$. This means that y_n is close to zero and, since the norm of d_n is relatively irrelevant,

$d_n^T y_n$ would be close to zero. The obvious remedy is to set $d_n = y_n$. Now we note that $d_n^T F(x_n) = F(x_{n+1})^T F(x_n)$ $- F(x_n)^T F(x_n)$ will be near zero and so the new direction of search $H_{n+1}F(x_{n+1}) \approx H_n F(x_n)$, the direction we have just searched unsuccessfully. C.G. Broyden confirms that the above heuristics fit his computational experience with $d_n = y_n$. See [2].

In Broyden's successful method, also defined in [2], $d_n = -H_n^T H_n F(x_n)$ and so $d_n^T F(x_n)$ is in no serious danger of going to zero except as a result of convergence, ie, the direction of search will probably not suffer too much from a small γ_n. Of course, $d_n^T y_n$ will be in danger and so the magnatude of the correction will be distorted and it may be desireble to obtain a fresh approximate Jacobian. This seems to explain behavior observed by M.J.D. Powell [9].

A useful strategy for choosing γ would probably be based on the one outlined by A. Goldstein [8] for Newton's method. It would seem reasonable to choose γ by some descent criterion until one feels he has a good approximate root and then switch to $\gamma \equiv 1$.

5. CONCLUDING REMARKS.

In this investigation we have ignored the fact that for many problems the direction d_n will sweep through a basis often enough to make (5) and (5') unduly pessimestic. Powell [9], at additional computational expense, modifies Broyden's correction in such a way that B_n converges to $J(x^*)$ if x_n converges to x^*. There seems little doubt that Powell's modification is justified for many problems.

The results in this paper were obtained while the auther was on leave at the University of Essex Department of Computer Science. The author wishes to thank the department in general and Mr. C.G. Broyden in particular for providing an excellant climate for research.

REFERENCES

[1] J.C.P. Barnes, An algorithm for solving nonlinear equations based on the secant method, Computer Journal, 8 (1964), pp. 66-72.

[2] C.G. Broyden, A class of methods for solving nonlinear simultaneous equations, Math. Comp., 19 (1965) pp.577-593.

[3] _____, The convergence of single-rank quasi-Newton methods, to appear.

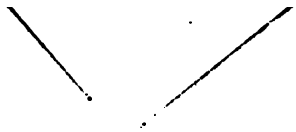
[4] J. Davis, The solution of nonlinear operator equations with critical points, PhD Thesis, Univ. of Wisc. 1966.

[5] W.C. Davidon, Variable metric method for minimization, A.E.C. research and development report ANL-5990 (1959)

[6] _____, Variance algorithm for minimization, Computer Journal, 10 (1968), pp. 406-410.

[7] J.E. Dennis, On the convergence of Newton-like methods, presented at Conf. on Num. Meth. for Nonlinear Algebraic Equations, Univ. of Essex, Jan. 1969.

[8] A.A. Goldstein, Constructive Real Analysis, Harper Row, 1967.

[9] M.J.D. Powell, A hybrid method for nonlinear equations U.K.A.E.R.E. technical paper #364, Jan 1969.

[10] W.C. Rheinboldt, A unified convergence theory for a class of iterative processes, SIAM J.Numer. Anal.,5(1968), pp. 42-63.