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

J.E. Dennis, Jr.

Technical Report

No. 69-46

October 1969

Department of Computer Science
Cornell University
Ithaca, New York 14850

ON THE LOCAL 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

The purpose of this paper is to present an alternate to the proof given in [3] of the local convergence of Broyden's method. The result was stated there as a simple corollary of a Kantorovich-type theorem for the method. Here we allow ourselves to assume the existence of a root for the system of equations in question and as a result we are able to slightly relax the requirements on the partial derivatives of the system and greatly simplify the proof.

We will confine the description of the method to bare essentials and refer the reader to [1] or [3] for more details.

Let $f_1,\ldots,f_N$ be real valued functions of N real variables and consider the problem of constructing a sequence $\{x_n\}$ of real N-tuples such that $\{x_n\}$ converges to $x^* \in E^N$ and $f_i(x^*) = 0$, for every $i = 1,\ldots,N$ i.e.,

$$F(x^*) = \begin{pmatrix} f_1(x^*) \\ \vdots \\ f_N(x^*) \end{pmatrix} = 0 .$$

* Currently on leave at Cornell University, Department of Computer Science, Ithaca, New York 14850.

Broyden's method for generating $\{x_n\}$ is as follows.

i) Choose the vector x_o and the nonsingular matrix B_o to approximate x^* and the Jacobian of the system at x_o (or x^*). Let H_o denote B_o^{-1} .

ii) Set $x_{n+1} = x_n - H_n F(x_n)$

and $H_{n+1} = H_n - H_n F(x_{n+1}) \dfrac{(H_n F(x_n))^T H_n}{(F(x_n)H_n)^T H_n (F(x_{n+1}) - F(x_n))}$.

The corresponding sequence $B_n = H_n^{-1}$ is not computed but it is given by

$$B_{n+1} = B_n - F(x_{n+1}) \frac{(H_n F(x_n))^T}{(H_n F(x_n))^T H_n F(x_n)} .$$

In [3] we assumed that the Jacobian matrix of F satisfied a Lipschitz condition of order one, i.e, $||J(x) - J(y)|| \leq K||x - y||$. Here, we will need to assume only that for x sufficiently close to x^* , $J(x)$ exists as a continuous function of x and

$$||J(x) - J(x^*)|| \leq K||x - x^*|| \tag{1}$$

This would be implied, for example, by an assumption that the first partials exist and are continuous at x^* and that the differential quotients at x^* of the first partials are uniformly bounded.

<u>LEMMA 1</u>. Let Ω_o be an open, convex set containing x^* such that (1) holds for every $x \in \Omega_o$. Let B be an $N \times N$ real matrix, $H = B^{-1}$ and let $x' = x - HF(x)$. If $x' \in \Omega_o$ and B' is defined by

$$B' = B - F(x') \frac{(HF(x))^T}{||HF(x)||^2},$$

then
$$||B' - J(x^*)|| \leq ||B - J(x^*)|| + K(||x' - x^*||^2 + ||x - x^*||^2) \frac{||x'-x||}{||x'-x||^2}.$$

<u>Proof</u>.
$$B' - J(x^*) = B - J(x^*) - (F(x') - F(x) - B(x' - x)) \frac{(x-x')^T}{||x-x'||^2}.$$

Now
$$\begin{aligned} F(x') - F(x) - B(x' - x) = {} & F(x') - F(x^*) - J(x^*)(x' - x^*) \\ & -(F(x) - F(x^*) - J(x^*)(x - x^*)) \\ & + (J(x^*) - B)(x' - x), \end{aligned}$$

Hence,

$$\begin{aligned} ||B' - J(x^*)|| \leq {} & ||B - J(x^*)|| \cdot \left|\left| I - \frac{(x'-x)(x'-x)^T}{||x'-x||^2} \right|\right| \\ & + ||F(x') - F(x^*) - J(x^*)(x-x^*)|| \frac{||x'-x||}{||x'-x||^2} \\ & + ||F(x) - F(x^*) - J(x^*)(x-x^*)|| \frac{||x'-x||}{||x'-x||^2} \\ \leq {} & ||B-J(x^*)|| + K(||x'-x^*||^2 + ||x-x^*||^2) \frac{||x'-x||}{||x'-x||} \end{aligned}$$

Here we have used the fact (See [2]) that

$||I - \frac{(x'-x)(x'-x)^T}{||x'-x||^2}|| = 1$ as well as (See [4], pg. 658)

$||F(x) - F(x^*) - J(x^*)(x - x^*)|| \leq K||x - x^*||^2$ by (1) .

We now present the main results of the paper.

THEOREM 1. Let x^* be a root of F and let Ω_o be an open convex neighborhood of x^* in which F satisfies (1). Under these conditions, if $J(x^*)$ is invertible, then there exist real positive numbers ϵ and δ such that if B_o if a real $N \times N$ matrix, $||B_o - J(x^*)|| \leq \delta$, and $x_o \in \Omega_o$, $||x_o - x^*|| < \epsilon$, Broyden's method converges to x^* from this starting point.

Proof. Let β bound $||J(x^*)^{-1}||$. Choose $\delta \leq \frac{1}{5\beta}$ and $\epsilon \leq \frac{\delta}{5K}$ such that $N(x_o, \epsilon) \subset \Omega_o$. Now select x_o , B_o as above. It is clear from the Banach Lemma [4] that H_o exists and is bounded in norm by $\beta(1 - \beta\delta)^{-1}$. The choice of ϵ ensure that $F(x_o)$ exists and so x_1 exists.

$$\begin{aligned}
e_1 \equiv ||x_1 - x^*|| &= ||x_o - x^* - H_o(F(x_o) + F(x^*))|| \\
&\leq ||-H_o|| \cdot ||F(x_o) - F(x^*) - B_o(x_o - x^*)|| \\
&\leq ||H_o|| \, [||F(x_o) - F(x^*) - J(x^*)(x_o - x^*)|| \\
&\qquad + ||J(x^*) - B_o|| - ||x_o - x^*||] \\
&< \frac{\beta}{1-\beta\delta} (Ke_o^2 + \delta e_o) \leq \frac{\beta}{1-\beta\delta} (Ke_o + \delta)e_o \\
&\leq \frac{\beta}{1-\beta\delta} (\frac{\delta}{5} + \delta)e_o \leq \frac{6}{5} \frac{\beta\delta}{1-\beta\delta} e_o < \frac{1}{2} e_o ,
\end{aligned}$$

Hence $F(x_1)$ and B_1 exist. Now

$$||x_1 - x_o|| \geq ||x_o - x^*|| - ||x_1 - x^*|| \geq e_o - \frac{1}{2} e_o = \frac{1}{2} e_o \text{ .}$$

Applying Lemma 1 and the above we obtain

$$||B_1 - J(x^*)|| \leq \delta + K(\frac{1}{4} e_o^2 + e_o^2) \frac{||x_1-x_o||}{||x_1-x_o||^2}$$

$$\leq \delta + \frac{5}{2} Ke_o \cdot \frac{1}{2} e_o \frac{||x_1-x_o||}{||x_1-x_o||^2} \leq \delta + \frac{5}{2} Ke_o$$

$$\leq \frac{3}{2} \delta < 2\delta \text{ .}$$

Hence $||J(x^*)^{-1}B_1 - I|| < 2\beta\delta$ so H_1 exists and is bounded in norm by $\beta(1 - 2\beta\delta)^{-1}$ so x_2 is defined. Assume by way of induction that $x_1,\ldots,x_n$, $H_1,\ldots,H_{n-1}$ all exist and $e_k \leq \frac{1}{2} e_{k-1}$, $||B_k - J(x^*)|| \leq (2 - (\frac{1}{2})^k)\delta$, $k \leq n$. Then $||J(x^*)^{-1}B_n - I|| \leq (2 - (\frac{1}{2})^k)\beta\delta < \frac{2}{5}$ and so H_n exists by the Banach Lemma and is bounded in norm by $\beta(1 - 2\beta\delta)^{-1}$, This ensures that x_{n+1} exists.

Now

$$e_{n+1} \leq ||H_n|| \cdot [||F(x_n) - F(x^*) - J(x^*)(x_n-x^*)||+||J(x^*)-B_n||e_n]$$

$$\leq \beta(1 - 2\beta\delta)^{-1} [Ke_n^2 + (2 - (\frac{1}{2})^n)\delta e_n]$$

$$\leq \beta(1 - 2\beta\delta)^{-1} [(\frac{1}{2})^n \frac{\delta}{5} + (2 - (\frac{1}{2})^n)\delta] e_n$$

$$< \beta(1 - 2\beta\delta)^{-1} 2\delta e_n < \frac{1}{2} e_n \text{ .}$$

We complete the induction by applying Lemma 1 to write

$$||B_{n+1} - J(x^*)|| \le ||B_n - J(x^*)|| + K(e_{n+1}^2 + e_n^2)||x_{n+1} - x_n||^{-1}$$

$$\le (2 - (\tfrac{1}{2})^n)\delta + \tfrac{5}{2} Ke_n$$

$$\le (2 - (\tfrac{1}{2})^n + \tfrac{5}{2} \cdot (\tfrac{1}{2})^n \cdot \tfrac{1}{5})\delta$$

$$= (2 - (\tfrac{1}{2})^n + (\tfrac{1}{2})^{n+1})\delta$$

$$= (2 - (\tfrac{1}{2})^{n+1})\delta .$$

We have used here that, as above, $\frac{1}{2} e_n \le ||x_{n+1} - x_n||$.

Hence the sequence $\{x_n\}$, $\{H_n\}$ exists and $e_n \le (\frac{1}{2})^n e_o$ so Broyden's method converges.

In the previous lemma and theorem, we note that if $K = 0$, i.e., the system is linear, then $||B_n - J(x^*)|| \le ||B_o - J(x^*)||$ and $\varepsilon =$ "∞" . This is simply a reflection of the fact, indicated by the expression for e_1 in terms of e_o , that for a linear system, $\delta \le \frac{1}{2\beta}$ is sufficient to ensure convergence from any x_o , See [2].

It is simple to change the hypotheses of Theorem 1 so that B_o is required to be close to $J(x_o)$ instead of $J(x^*)$.

THEOREM 2. Under the hypotheses of Theorem 1, there exist real positive numbers δ' , ε' such that if $x_o \in \Omega_o$, $||x_o - x^*|| \le \varepsilon'$, and B_o is a real $N \times N$ matrix, $||B_o - J(x_o)|| \le \delta'$, then Broyden's method converges to x^* from this starting point.

<u>Proof</u>. Let ε and δ be as in the previous theorem. Select $\varepsilon' \leq \varepsilon$ and δ' such that $\delta' + K\varepsilon' \leq \delta$. Now let $||x_o - x^*|| < \varepsilon'$, $||B_o - J(x_o)|| < \delta'$. Then,

$$||B_o - J(x^*)|| \leq ||B_o - J(x_o)|| + ||J(x_o) - J(x^*)||$$

$\leq \delta' + K\varepsilon' \leq \delta$ and $||x_o - x^*|| < \varepsilon$ so the result follows from Theorem 1.

Much of this work was accomplished while the author was on leave at the University of Essex, England, during 1968-69.

REFERENCES

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

[2] __________, The convergence of single-rank quasi-Newton Methods, to appear.

[3] J. E. Dennis, On the convergence of Broyden's method for nonlinear systems of equations, Technical Report No. 69-48, Department of Computer Science, Cornell University, October 1969.

[4] L. V. Kantorovich, G. P. Akilov, Functional Analysis in Normed Spaces, Pergamon Press, New York, 1964.