A MODIFIED NEWTON-RAPHSON METHOD FOR THE SOLUTION OF SYSTEMS OF EQUATIONS

BY

ADI BEN-ISRAEL*

ABSTRACT

An implicit function theorem and a resulting modified Newton-Raphson method for roots of functions between finite dimensional spaces, without assuming non-singularity of the Jacobian at the initial approximation.

Introduction. The *Newton-Raphson method* for solving an equation

$$f(y) = 0 \tag{1}$$

is based upon the convergence, under suitable conditions, of the sequence

$$y_{p+1} = y_p - \frac{f(y_p)}{f'(y_p)} \qquad p = 0, 1, \cdots \tag{2}$$

to the solution of (1), where y_0 is an initial approximation to that solution. The *modified Newton-Raphson method* uses, instead of (2), the sequence

$$y_{p+1} = y_p - \frac{f(y_p)}{f'(y_0)} \qquad p = 0, 1, \cdots \tag{3}$$

These methods are described in detail in [7], [6] and [4].

Extensions to systems of equations

$$\begin{aligned} f_1(y_1, y_2, \cdots, y_n) &= 0 \\ \vdots \quad & \quad \vdots \\ f_m(y_1, y_2, \cdots, y_n) &= 0 \end{aligned} \tag{4}$$

are immediate in case: $m = n$, e.g. [3] and [4]. The analogs of (2), (3) are respectively:

$$\boldsymbol{y}_{p+1} = \boldsymbol{y}_p - (J(\boldsymbol{y}_p))^{-1}\boldsymbol{f}(\boldsymbol{y}_p) \qquad p = 0, 1, \cdots \tag{5}$$

$$\boldsymbol{y}_{p+1} = \boldsymbol{y}_p - (J(\boldsymbol{y}_0))^{-1}\boldsymbol{f}(\boldsymbol{y}_p) \qquad p = 0, 1, \cdots \tag{6}$$

where $\boldsymbol{y}$ is the vector with components y_j, $j = 1, \cdots, n$

Received February 28, 1965.

* Research supported by the Swope Foundation.

$f(y)$ is the vector with components $f_i(y)$, $i = 1, \cdots, n$

$J(y)$ is the Jacobian matrix, whose (i,j)th element is $\dfrac{\partial f_i(y)}{\partial y_j}$

and y_0 is an initial approximation to a solution of (4).

The composite Newton-Raphson gradient method of Hart and Motzkin [2], is applicable also if $m \neq n$, provided rank $J(y) = n$ at the solution.

In this note the Algorithm (6) is extended to general systems of equations, by using the generalized inverse [8] of the Jacobian matrix. Conditions of convergence, as well as bounds on the convergence rate, are stated in Theorem 2. These are based on theorem 1, which is an implicit function theorem following from a classical result of Hildebrandt and Graves [5, Theorem 3], and is of independent interest.

NOTATIONS.: Let E^k be the k-dimensional vector space with the Euclidean norm $\| x \| = (x,x)^{1/2}$. Let $E^{m\times n}$ be the space of $m \times n$ complex matrices, with the norm $\| A \| = \max\{\sqrt{\lambda}: \lambda$ an eigenvalue of $A^*A\}$, A^* being the conjugate transpose of A. These norms satisfy $\| Ax \| \leqq \| A \| \, \| x \|$ for every $A \in E^{m\times n}$, $x \in E^n$. By $R(A)$,! $N(A)$ we denote the *range space* respectively *null space* of A, and by A^+ the *generalized inverse* of A, [8]. For $x_0 \in E^k$ and a real positive r, $S(x_0, r) = \{x \in E^k; \| x - x_0 \| < r\}$, the open ball of radius r around x_0. The components of a function $F: E^n \to E^m$ are denoted by $f_i(y)$, $i = 1, \cdots, m$. The *Jacobian* of F at $y \in E^n$ is the $m \times n$ matrix

$$J(y) = \left(\frac{\partial f_i(y)}{\partial y_j} \right) \qquad \begin{array}{l} i = 1, \cdots, m \\ j = 1, \cdots, n \end{array}$$

Let Y be an open set in E^n. Following Hildebrandt and Graves [5] we say that a function $F: E^n \to E^m$ is in the class $C'(Y)$ if the mapping: $E^n \to E^{m\times n}$ given by: $y \to J(y)$ is continuous for every $y \in Y$. The modulus of continuity of $J(y)$ at y_0, $\delta(y_0, \varepsilon)$ is defined by $\| y - y_0 \| \leqq \delta(y_0, \varepsilon) \Rightarrow \| J(y) - J(y_0) \| \leqq \varepsilon$.

THEOREM 1. *Let X_0 be an open set in E^p, y_0 a vector in E^n, F a function, $F: X_0 \times S(y_0, r) \to E^m$, T a linear transformation, $T: E^n \to E^m$, M a real positive number, such that:*

(7) $\quad M \| T^+ \| < 1$

(8) $\quad \| T(y_1 - y_2) - F(x, y_1) + F(x, y_2) \| \leqq M \| y_1 - y_2 \|$ *for every* $x \in X_0$ *and* $y_1, y_2 \in S(y_0, r)$ *which satisfy* $y_1 - y_2 \in R(T^*)$

(9) $\quad \| T\|^+ \| F(x, y_0) \| < (1 - M \| T^+ \|) r$ *for every* $x \in X_0$.

Then there is a unique function $y\colon X_0 \to S(y_0, r) \cap \{y_0 + R(T^*)\}$, *which for every* $x \in X_0$ *is the solution of*

$$T^*F(x, y(x)) = 0 \tag{10}$$

Proof. Define a function $G\colon X_0 \times S(y_0, r) \to E^n$ by

$$G(x,y) = y - T^+F(x,y) \tag{11}$$

For every y_1, y_2 such that $y_1 - y_2 \in R(T^*)$ we recall that [8]: $T^+T(y_1 - y_2) = y_1 - y_2$ and therefore

$$G(x,y_1) - G(x,y_2) = T^+\{T(y_1 - y_2) - F(x,y_1) + F(x,y_2)\} \tag{12}$$

From (12) and (8) it follows that:

$$\| G(x,y_1) - G(x,y_2) \| \leqq M \| T^+ \| \| y_1 - y_2 \| \text{ for every } y_1, y_2 \text{ in } S(y_0,r) \text{ which satisfy } y_1 - y_2 \in R(T^*) \tag{13}$$

Consider now the sequence

$$\begin{aligned} y_1(x) &= G(x, y_0) \\ y_{p+1}(x) &= G(x, y_p(x)) \qquad p = 1, 2, \cdots. \end{aligned} \tag{14}$$

Since $R(T^+) = R(T^*)$ it follows that

$$y_{p+1}(x) - y_p(x) \in R(T^*) \text{ for every } x \in X_0, \qquad p = 0, 1, \cdots. \tag{15}$$

Also by (9), for every $x \in X_0$:

$$\| y_1(x) - y_0 \| = (1-k)c \qquad \text{where } k = M \| T^+ \| < 1 \text{ by (7) and } c < r \tag{16}$$

and by induction:

$$\| y_{p+1}(x) - y_p(x) \| \leqq k^p(1-k)c \text{ for every } x \in X_0, \qquad p = 1, 2, \cdots. \tag{17}$$

Thus for every $x \in X_0$ the sequence $\{y_p(x)\}$ converges to a unique vector $y(x)$, which by (15) and (17) lies in $S(y_0, r) \cap \{y_0 + R(T^*\}$. We prove now that for every $x \in X_0$, $y(x)$ is a solution of

$$y(x) = G(x, y(x)) \tag{18}$$

Indeed

$$\begin{aligned} \| y(x) - G(x,y(x)) \| &\leqq \| y(x) - y_{p+1}(x) \| + \| G(x, y_p(x)) - G(x, y(x)) \| \\ &\leqq \| y(x) - y_{p+1}(x) \| + k \| y_p(x) - y(x) \| \\ &\quad \text{which} \to 0 \text{ as } p \to \infty. \end{aligned} \tag{19}$$

The proof is completed by noting that (18) is equivalent, by (11), to

$$T^{+}F(\mathbf{x},\mathbf{y}(\mathbf{x})) = \mathbf{0} \tag{20}$$

which, since $N(T^{+}) = N(T^{*})$, is equivalent to (10). Q.E.D.

REMARKS:

(i) If rank $T^{*} = m$ then Theorem 1 reduces to a well known theorem of Hildebrandt and Graves [5, Theorem 3] restricted to the manifold $\{\mathbf{y}_0 + R(T^{*})\}$.

(ii) Taking Theorem 1 as the basis of a modified Newton-Raphson method for solving the system (4), where $\mathbf{y}$ is the independent variable, we proceed by regarding X_0 and E^p of Theorem 1 both identical with the zero-dimensional vector space.

THEOREM 2. *Let F be a function, $F:E^n \to E^m$, $\mathbf{y}_0$ a vector in E^n, M a real positive number, such that:*

$$F \in C'(S(\mathbf{y}_0, \delta(\mathbf{y}_0, M))) \tag{21}$$

$$M \| J_0^{+} \| < 1 \quad \textit{where } J_0 = J(\mathbf{y}_0) \tag{22}$$

$$\| J_0^{+} \| \| F(\mathbf{y}_0) \| < (1 - M \| J_0^{+} \|)\delta(\mathbf{y}_0, M) \tag{23}$$

Then the sequence

$$\mathbf{y}_{p+1} = \mathbf{y}_p - J_0^{+} F(\mathbf{y}_p) \qquad p = 0, 1, \cdots \tag{24}$$

converges to the unique solution of

$$J_0^{*} F(\mathbf{y}) = \mathbf{0} \tag{25}$$

which lies in $S(\mathbf{y}_0, \delta(\mathbf{y}_0, M)) \cap \{\mathbf{y}_0 + R(J_0^{*})\}$.

Moreover:

$$\| \mathbf{y}_{p+1} - \mathbf{y}_p \| \leqq k_0^p (1 - k_0) \delta(\mathbf{y}_0, M), \qquad \text{where } k_0 = M \| J_0^{+} \|. \tag{26}$$

Proof. Specializing a result of Bartle [1, Lemma 1] we verify that (21) implies that:

$$\| F(\mathbf{y}_1) - F(\mathbf{y}_2) - J(\mathbf{y}_0)(\mathbf{y}_1 - \mathbf{y}_2) \| \leqq M \| \mathbf{y}_1 - \mathbf{y}_2 \| \quad \text{for every } \mathbf{y}_1, \mathbf{y}_2 \in S(\mathbf{y}_0, \delta(\mathbf{y}_0, M)) \tag{27}$$

Taking $T = J(\mathbf{y}_0)$ in Theorem 1, we have conditions (7), (8) and (9) satisfied respectively by (22), (27) and (23). The proof is completed by noting the correspondence between (24), (25), (26) and respectively (14), (10), (17) in Theorem 1. Q.E.D.

REMARKS. (i) If y_0 is chosen so that rank $J(y_0) = m$, then (25) is equivalent to (4) and (24) is a generalization of the Algorithm (6) for the solution of (4).

(ii) The computational efficiency of the Algorithm (24) proposed above, depends upon that of computing the generalized inverse J_0^+, and upon the initial approximation y_0. Methods for computing generalized inverses were recently given by several authors, and the rapid progress in this area may result in favor of the proposed Algorithm.

REFERENCES

1. R. G. Bartle, *Newton's method in Banach spaces*, Proc. Amer. Math. Soc. **6** (1955), 827–831.
2. W. L. Hart and T. S. Motzkin, *A composite Newton-Raphson gradient method for the solution of systems of equations*, Pacific J. Math. **6** (1956), 691–707.
3. P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations*, J. Wiley, 1962.
4. P. Henrici, *Elements of numerical analysis*, J. Wiley, 1964.
5. T. H. Hildebrandt and L. M. Graves, *Implicit functions and their differentials in general analysis*, Trans. Amer. Math. Soc. **29** (1927), 127–153.
6. A. S. Householder, *Principles of Numerical Analysis*, McGraw-Hill, 1953.
7. A. M. Ostrowski, *Solution of Equations and Systems of Equations*, Academic Press, (1960)
8. R. Penrose, *A generalized inverse for matrices* Proc. Cambridge Philos. Soc. **51** 1955, 406–413.

TECHNION—ISRAEL INSTITUTE OF TECHNOLOGY,
HAIFA

Reprinted from
ISRAEL JOURNAL OF MATHEMATICS
Volume 3, Number 2, June 1965