Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS. Vol. 15, No 2, August 1966

A Newton-Raphson Method for the Solution of Systems of Equations

ADI BEN-ISRAEL

*Technion-Israel Institute of Technology and Northwestern University**

Submitted by Richard Bellman

INTRODUCTION

The Newton-Raphson method for solving an equation

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

is based upon the convergence, under suitable conditions [1, 2], of the sequence

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

to a solution of (1), where x_0 is an approximate solution. A detailed discussion of the method, together with many applications, can be found in [3].

Extensions to systems of equations

$$\begin{matrix} f_1(x_1, \cdots, x_n) = 0 \\ \\ f_m(x_1, \cdots, x_n) = 0 \end{matrix} \qquad \text{or} \qquad \mathbf{f}(\mathbf{x}) = \mathbf{0} \tag{3}$$

are immediate in case: $m = n$, [1], where the analog of (2) is:[1]

$$\mathbf{x}_{p+1} = \mathbf{x}_p - J^{-1}(\mathbf{x}_p)\,\mathbf{f}(\mathbf{x}_p) \qquad (p = 0, 1, \cdots). \tag{4}$$

Extensions and applications in Banach spaces were given by Hildebrandt and Graves [4], Kantorovič [5-7], Altman [8], Stein [9], Bartle [10], Schröder [11] and others. In these works the Frechet derivative replaces $J(\mathbf{x})$ in (4), yet nonsingularity is assumed throughout the iterations.

* Part of the research underlying this report was undertaken for the Office of Naval Research, Contract Nonr-1228(10), Project NR 047-021, and for the U.S. Army Research Office-Durham, Contract No. DA-31-124-ARO-D-322 at Northwestern University. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government.

[1] See section on notations below.

The modified Newton-Raphson method

$$\mathbf{x}_{p+1} = \mathbf{x}_p - J^{-1}(\mathbf{x}_0)\,\mathbf{f}(\mathbf{x}_p) \qquad (p = 0, 1, \cdots) \tag{5}$$

was extended in [17] to the case of singular $J(\mathbf{x}_0)$, and conditions were given for the sequence

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

to converge to a solution of

$$J^*(\mathbf{x}_0)\,\mathbf{f}(\mathbf{x}) = \mathbf{0}. \tag{7}$$

In this paper the method (4) is likewise extended to the case of singular $J(\mathbf{x}_p)$, and the resulting sequence (12) is shown to converge to a stationary point of $\sum_{i=1}^{m} f_i^2(\mathbf{x})$.

Notations

Let E^k denote the k-dimensional (complex) vector space of vectors $\mathbf{x}$, with the Euclidean norm $\| \mathbf{x} \| = (\mathbf{x}, \mathbf{x})^{1/2}$. Let $E^{m\times n}$ denote the space of $m \times n$ complex matrices with the norm

$$\| A \| = \max \{\sqrt{\lambda} : \lambda \text{ an eigenvalue of } A^*A\},$$

A^* being the conjugate transpose of A.

These norms satisfy [12]:

$$\| A\mathbf{x} \| \leqslant \| A \| \, \| \mathbf{x} \| \qquad \text{for every} \qquad \mathbf{x} \in E^n, \qquad A \in E^{m\times n}.$$

Let $R(A)$, $N(A)$ denote the *range* resp. *null space* of A, and A^+ the generalized inverse of A, [13].

For $\mathbf{u} \in E^k$ and $r > 0$ let

$$S(\mathbf{u}, r) = \{\mathbf{x} \in E^k : \| \mathbf{x} - \mathbf{u} \| < r\}$$

denote the open ball of radius r around $\mathbf{u}$.

The components of a function $f : E^n \to E^m$ are denoted by $f_i(\mathbf{x})$, $(i = 1, \cdots, m)$. The *Jacobian* of f at $\mathbf{x} \in E^n$ is the $m \times n$ matrix

$$J(\mathbf{x}) = \left(\frac{\partial f_i(\mathbf{x})}{\partial x_j}\right), \qquad \begin{pmatrix} i = 1, \cdots, m \\ j = 1, \cdots, n \end{pmatrix}.$$

For an open set $S \subset E^n$, the function $f : E^n \to E^m$ is in the class $C'(S)$ if the mapping $E^n \to E^{m\times n}$ given by $\mathbf{x} \to J(\mathbf{x})$ is continuous for every $\mathbf{x} \in S$ [4].

Results

Theorem 1. *Let* $f : E^n \to E^m$ *be a function,* $\mathbf{x}_0$ *a point in* E^n, *and* $r > 0$ *be such that* $f \in C'(S(\mathbf{x}_0, r))$.

Let M, N *be positive constants such that for all* $\mathbf{u}$, $\mathbf{v}$ *in* $S(\mathbf{x}_0, r)$ *with* $\mathbf{u} - \mathbf{v} \in R(J^*(\mathbf{v}))$:

$$\| J(\mathbf{v})(\mathbf{u} - \mathbf{v}) - \mathbf{f}(\mathbf{u}) + \mathbf{f}(\mathbf{v}) \| \leqslant M \| \mathbf{u} - \mathbf{v} \| \quad (8)$$

$$\| (J^+(\mathbf{v}) - J^+(\mathbf{u}))\, \mathbf{f}(\mathbf{u}) \| \leqslant N \| \mathbf{u} - \mathbf{v} \| \quad (9)$$

and

$$M \| J^+(\mathbf{x}) \| + N = k < 1 \qquad \textit{for all} \qquad \mathbf{x} \in S(\mathbf{x}_0, r) \quad (10)$$

$$\| J^+(\mathbf{x}_0) \| \| \mathbf{f}(\mathbf{x}_0) \| < (1 - k)\, r. \quad (11)$$

Then the sequence

$$\mathbf{x}_{p+1} = \mathbf{x}_p - J^+(\mathbf{x}_p)\, \mathbf{f}(\mathbf{x}_p) \qquad (p = 0, 1, \cdots) \quad (12)$$

converges to a solution of

$$J^*(\mathbf{x})\, \mathbf{f}(\mathbf{x}) = \mathbf{0} \quad (13)$$

which lies in $S(\mathbf{x}_0, r)$.

Proof. Let the mapping $g : E^n \to E^n$ be defined by

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - J^+(\mathbf{x})\, f(\mathbf{x}). \quad (14)$$

Equation (12) now becomes:

$$\mathbf{x}_{p+1} = \mathbf{g}(\mathbf{x}_p) \qquad (p = 0, 1, \cdots) \quad (15)$$

We prove now that

$$\mathbf{x}_p \in S(\mathbf{x}_0, r) \qquad (p = 1, 2, \cdots). \quad (16)$$

For $p = 1$, (16) is guaranteed by (11). Assuming (16) is true for all subscripts $\leqslant p$, we prove it for: $p + 1$. Indeed,

$$\begin{aligned}\mathbf{x}_{p+1} - \mathbf{x}_p &= \mathbf{x}_p - \mathbf{x}_{p-1} - J^+(\mathbf{x}_p)\, \mathbf{f}(\mathbf{x}_p) + J^+(\mathbf{x}_{p-1})\, \mathbf{f}(\mathbf{x}_{p-1}) \\ &= J^+(\mathbf{x}_{p-1})\, J(\mathbf{x}_{p-1})(\mathbf{x}_p - \mathbf{x}_{p-1}) - J^+(\mathbf{x}_p)\, \mathbf{f}(\mathbf{x}_p) + J^+(\mathbf{x}_{p-1})\, \mathbf{f}(\mathbf{x}_{p-1}) \\ &= J^+(\mathbf{x}_{p-1})\, [J(\mathbf{x}_{p-1})(\mathbf{x}_p - \mathbf{x}_{p-1}) - \mathbf{f}(\mathbf{x}_p) + \mathbf{f}(\mathbf{x}_{p-1})] \\ &\qquad + (J^+(\mathbf{x}_{p-1}) - J^+(\mathbf{x}_p))\, \mathbf{f}(\mathbf{x}_p), \end{aligned} \quad (17)$$

where

$$\mathbf{x}_p - \mathbf{x}_{p-1} = J^+(\mathbf{x}_{p-1})\, J(\mathbf{x}_{p-1})(\mathbf{x}_p - \mathbf{x}_{p-1}) \quad (18)$$

follows from

$$\mathbf{x}_p - \mathbf{x}_{p-1} \in R(J^+(\mathbf{x}_{p-1})) = R(J^*(\mathbf{x}_{p-1})) \quad (19)$$

and A^+A being the perpendicular projection on $R(A^*)$ [14]. Setting $\mathbf{u} = \mathbf{x}_p$, $\mathbf{v} = \mathbf{x}_{p-1}$ in (8) and (9), we conclude from (17), (19), and the induction hypothesis that

$$\| \mathbf{x}_{p+1} - \mathbf{x}_p \| \leqslant (M \| J^+(\mathbf{x}_p) \| + N) \| \mathbf{x}_p - \mathbf{x}_{p-1} \| \tag{20}$$

and from (10)

$$\| \mathbf{x}_{p+1} - \mathbf{x}_p \| \leqslant k \| \mathbf{x}_p - \mathbf{x}_{p-1} \|, \tag{21}$$

which implies

$$\| \mathbf{x}_{p+1} - \mathbf{x}_0 \| \leqslant \sum_{j=1}^{p} k^j \| \mathbf{x}_1 - \mathbf{x}_0 \| = \frac{k(1-k^p)}{(1-k)} \| \mathbf{x}_1 - \mathbf{x}_0 \| \tag{22}$$

and, finally, with (11)

$$\| \mathbf{x}_{p+1} - \mathbf{x}_0 \| \leqslant k(1-k^p)\, r < r, \tag{23}$$

which proves (16).

Equation (21) proves indeed that the mapping g is a contraction in the sense:

$$\| \mathbf{g}(\mathbf{x}_p) - \mathbf{g}(\mathbf{x}_{p-1}) \| \leqslant k \| \mathbf{x}_p - \mathbf{x}_{p-1} \| < \| \mathbf{x}_p - \mathbf{x}_{p-1} \| \qquad (p = 1, 2, \cdots). \tag{24}$$

The sequence $\{\mathbf{x}_p\}$, $(p = 0, 1, \cdots)$, converges therefore to a vector $\mathbf{x}_*$ in $S(\mathbf{x}_0, r)$.

$\mathbf{x}_*$ is a solution of (13). Indeed,

$$\begin{aligned} \| \mathbf{x}_* - \mathbf{g}(\mathbf{x}_*) \| &\leqslant \| \mathbf{x}_* - \mathbf{x}_{p+1} \| + \| \mathbf{g}(\mathbf{x}_p) - \mathbf{g}(\mathbf{x}_*) \| \\ &\leqslant \| \mathbf{x}_* - \mathbf{x}_{p+1} \| + k \| \mathbf{x}_p - \mathbf{x}_* \|, \end{aligned} \tag{25}$$

where the right-hand side of (25) tends to zero as $p \to \infty$. But

$$\mathbf{x}_* = \mathbf{g}(\mathbf{x}_*) \tag{26}$$

is equivalent, by (14), to

$$J^+(\mathbf{x}_*)\, \mathbf{f}(\mathbf{x}_*) = \mathbf{0}, \tag{27}$$

which is equivalent to

$$J^*(\mathbf{x}_*)\, \mathbf{f}(\mathbf{x}_*) = \mathbf{0} \tag{28}$$

since $N(A^+) = N(A^*)$ for every $A \in E^{m \times n}$ [14]. Q.E.D.

REMARKS. (a) If $m = n$ and the matrices $J(\mathbf{x}_p)$ are nonsingular, $(p = 0, 1, \cdots)$, then (12) reduces to (4), which converges to a solution of (3). In this case (13) and (3) are indeed equivalent because $N(J^*(\mathbf{x}_*)) = \{\mathbf{0}\}$.

(b) From

$$J^*(\mathbf{x})\, \mathbf{f}(\mathbf{x}) = \frac{1}{2} \operatorname{grad} \left(\sum_{i=1}^{m} f_i^2(\mathbf{x}) \right) \tag{29}$$

it follows that the limit $\mathbf{x}_*$ of the sequence (12) is a stationary point of $\sum_{i=1}^m f_i^2(\mathbf{x})$, which by Theorem 1 exists in $S(\mathbf{x}_0, r)$, where (8)-(11) are satisfied. Even when (3) has a solution in $S(\mathbf{x}_0, r)$, the sequence (12) does not necessarily converge to it, but to a stationary point of $\sum_{i=1}^m f_i^2(\mathbf{x})$ which may be a least squares solution of (3), or a saddle point of the function $\sum_{i=1}^m f_i^2(\mathbf{x})$, etc.

DEFINITION. A point $\mathbf{x}$ is an *isolated point for the function f in the linear manifold L* if there is a neighborhood U of $\mathbf{x}$ such that

$$\mathbf{y} \in U \cap L, \qquad \mathbf{y} \neq \mathbf{x} \Rightarrow \mathbf{f}(y) \neq \mathbf{f}(\mathbf{x}).$$

THEOREM 2. *Let the function $f: E^n \to E^m$ be in the class $C'(S(\mathbf{u}, r))$. Then $\mathbf{u}$ is an isolated point for f in the linear manifold $\{\mathbf{u} + R(J^*(\mathbf{u}))\}$.*

PROOF. Suppose the theorem is false. Then there is a sequence

$$\mathbf{x}_k \in S(\mathbf{u}, r) \cap \{\mathbf{u} + R(J^*(\mathbf{u}))\} \qquad (k = 1, 2, \cdots) \tag{30}$$

such that

$$\mathbf{x}_k \to \mathbf{u} \qquad \text{and} \qquad \mathbf{f}(\mathbf{x}_k) = \mathbf{f}(\mathbf{u}) \qquad (k = 1, 2, \cdots) \tag{31}$$

From $f \in C'(S(\mathbf{u}, r))$ it follows that

$$\| \mathbf{f}(\mathbf{u} + \mathbf{z}_k) - \mathbf{f}(\mathbf{u}) - J(\mathbf{u})\, \mathbf{z}_k \| \leqslant \delta(\| \mathbf{z}_k \|)\, \| \mathbf{z}_k \| \tag{32}$$

where,

$$\mathbf{x}_k = \mathbf{u} + \mathbf{z}_k \qquad (k = 1, 2, \cdots) \tag{33}$$

and

$$\delta(t) \to 0 \qquad \text{as} \qquad t \to 0. \tag{34}$$

Combining (31) and (32) it follows that

$$\left\| J(\mathbf{u}) \frac{\mathbf{z}_k}{\| \mathbf{z}_k \|} \right\| \leqslant \delta(\| \mathbf{z}_k \|). \tag{35}$$

The sequence $\{\mathbf{z}_k\}/\| \mathbf{z}_k \|$, $(k = 1, 2, \cdots)$, consists of unit vectors which by (30) and (33) lie in $R(J^*(\mathbf{u}))$, and by (35) and (34) converge to a vector in $N(J(\mathbf{u}))$, a contradiction. Q.E.D.

REMARKS (a) In case the linear manifold $\{\mathbf{u} + R(J^*(\mathbf{u}))\}$ is the whole space E^n, (i.e., $N(J(\mathbf{u})) = \{\mathbf{0}\}$, i.e., the columns of $J(\mathbf{u})$ are linearly independent), Theorem 2 is due to Rodnyanskii [15] and was extended to Banach spaces by *S*. Kurepa [16], whose idea is used in our proof.

(b) Using Theorem 2, the following can be said about the solution $\mathbf{x}_*$ of (13), obtained by (12):

Corollary. *The limit* $\mathbf{x}_*$ *of the sequence (12) is an isolated zero for the function* $J^*(\mathbf{x}_*)\,\mathbf{f}(\mathbf{x})$ *in the linear manifold* $\{\mathbf{x}_* + R(J^*(\mathbf{x}_*))\}$, *unless* $J(\mathbf{x}_*) = 0$ *the zero matrix.*

Proof. $\mathbf{x}_*$ is in the interior of $S(\mathbf{x}_0, r)$, therefore $f \in C'(S(\mathbf{x}_*, r_*))$ for some $r_* > 0$. Assuming the corollary to be false it follows, as in (35), that

$$\left\| J^*(\mathbf{x}_*)\, J(\mathbf{x}_*) \frac{\mathbf{z}_k}{\| \mathbf{z}_k \|} \right\| \leqslant \delta(\| \mathbf{z}_k \|)\, \| J^*(\mathbf{x}_*) \|, \tag{36}$$

where $\mathbf{u}_k = \mathbf{x}_* + \mathbf{z}_k$, $(k = 1, 2, \cdots)$, is a sequence in $\{\mathbf{x}_* + R(J^*(\mathbf{x}_*))\}$ converging to $\mathbf{x}_*$ and such that

$$\mathbf{0} = J^*(\mathbf{x}_*)\,\mathbf{f}(\mathbf{x}_*) = J^*(\mathbf{x}_*)\,\mathbf{f}(\mathbf{u}_k).$$

Excluding the trivial case: $J(\mathbf{x}_*) = 0$ (see remark below), it follows that $\{\mathbf{z}_k / \| \mathbf{z}_k \|\}$, $(k = 1, 2, \cdots)$, is a sequence of unit vectors in $R(J^*(\mathbf{x}_*))$, which by (36) converges to a vector in $N(J^*(\mathbf{x}_*)\, J(\mathbf{x}_*)) = N(J(\mathbf{x}_*))$, a contradiction. Q.E.D.

Remark: The definition of an isolated point is vacuous if the linear manifold L is zero-dimensional. Thus if $J(\mathbf{x}_*) = 0$, then $R(J^*(\mathbf{x}_*)) = \{\mathbf{0}\}$ and every $\mathbf{x} \in E^n$ is a zero of $J^*(\mathbf{x}_*)\, f(\mathbf{x})$.

Examples

The following examples were solved by the iterative method:

$$\mathbf{x}_{p\alpha+k+1} = \mathbf{x}_{p\alpha+k} - J^+(\mathbf{x}_{p\alpha})\,\mathbf{f}(\mathbf{x}_{p\alpha+k}), \tag{37}$$

where $\alpha \geqslant 0$ is an integer

$$k = \begin{cases} 0, 1, \cdots, \alpha - 1 & \text{if} \quad \alpha > 0 \\ 0, 1, \cdots & \text{if} \quad \alpha = 0 \end{cases}$$

and

$$p = 0, 1, \cdots$$

For $\alpha = 1$: (37) reduces to (12), $\alpha = 0$ yields the modified Newton method of [17], and for $\alpha \geqslant 2$: α is the number of iterations with the modified method [17], (with the Jacobian $J(\mathbf{x}_{p\alpha})$, $p = 0, 1, \cdots$), between successive computations of the Jacobian $J(\mathbf{x}_{p\alpha})$ and its generalized inverse. In all the examples worked out, convergence (up to the desired accuracy) required the smallest number of iterations for $\alpha = 1$; but often for higher values of α less computations (and time) were required on account of computing J and J^+ only once in α iterations. The computations were carried out on Philco 2000. The method of [18] was used in the subroutine of computing the generalized inverse J^+.

EXAMPLE 1. The system of equations is

$$\mathbf{f}(\mathbf{x}) = \begin{cases} f_1(x_1, x_2) = x_1^2 + x_2^2 - 2 = 0 \\ f_2(x_1, x_2) = x_1 - x_2 = 0 \\ f_3(x_1, x_2) = x_1 x_2 - 1 = 0. \end{cases}$$

Equation (13) is

$$J^*(\mathbf{x})\,\mathbf{f}(\mathbf{x}) = \begin{cases} 2x_1^3 + 3x_1x_2^2 - 3x_1 - 2x_2 = 0 \\ 2x_2^3 + 3x_1^2x_2 - 2x_1 - 3x_2 = 0, \end{cases}$$

whose solutions are (0, 0) a saddle point of $\sum f_i^2(\mathbf{x})$ and (1, 1), (− 1, − 1) the solutions of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. In applying (37), derivatives were replaced by differences with $\Delta x = 0.001$.

Some results are

$$\alpha = 3$$

$p\alpha + k$	0	1	2	3
$\mathbf{x}_{p\alpha+k}$	3.000000 2.000000	1.578143 1.355469	1.287151 1.199107	1.155602 1.118148
$\mathbf{f}(\mathbf{x}_{p\alpha+k})$	11.00000 1.00000 5.00000	2.327834 0.222674 1.139125	1.094615 0.088044 0.543432	0.585672 0.037454 0.292134
Σf_i^2	147.0000	6.766002	1.501252	0.429757

$p\alpha + k$	4	5	6	7
$\mathbf{x}_{p\alpha+k}$	1.008390 1.008365	1.000981 1.000980	1.000118 1.000118	1.000000 1.000000
$\mathbf{f}(x_{p\alpha+k})$	0.033649 0.000025 0.016825	0.003924 0.000000 0.001962	0.000472 0.000000 0.000236	0.000000 0.000000 0.000000
Σf_i^2	0.001415	0.000019	0.000000	0.000000

Note the sharp improvement for each change of Jacobian (iterations: 1, 4, and 7).

$\alpha = 5$

$p\alpha + k$	0	5	6	8	9
$\mathbf{x}_{p\alpha+k}$	3.000000 2.000000	1.050657 1.043431	1.001078 1.001078	1.000002 1.000002	1.000000 1.000000
$\mathbf{f}(\mathbf{x}_{p\alpha+k})$	11.00000 1.00000 5.00000	0.192630 0.007226 0.096289	0.004315 0.000000 0.002157	0.000009 0.000000 0.000004	0.000000 0.000000 0.000000
Σf_i^2	147.0000	0.046430	0.000023	0.000000	0.000000

The Jacobian and its generalized inverse were twice calculated (iterations 1, 6), whereas for $\alpha = 3$ they were calculated 3 times.

$\alpha = 10$

$p\alpha + k$	0	10	11	12
$\mathbf{x}_{p\alpha+k}$	3.0 2.0	1.003686 1.003559	1.000008 1.000008	1.000000 1.000000
$\mathbf{f}(\mathbf{x}_{p\alpha+k})$	11.0 1.0 5.0	0.014516 0.000128 0.007258	0.000032 0.000000 0.000016	0.000000 0.000000 0.000000
Σf_i^2	147.0	0.000263	0.000000	0.000000

EXAMPLE 2.

$$\mathbf{f}(\mathbf{x}) = \begin{cases} f_1(x_1, x_2) = x_1^2 + x_2^2 - 2 = 0 \\ f_2(x_1, x_2) = (x_1 - 2)^2 + x_2^2 - 2 = 0 \\ f_3(x_1, x_2) = (x_1 - 1)^2 + x_2^2 - 9 = 0. \end{cases}$$

This is an inconsistent system of equations, whose least squares solutions are (1.000000, 1.914854) and (1.000000, − 1.914854). Applying (37) with $\alpha = 1$ and exact derivatives, resulted in the sequence:

p	0	1	2	3
$\mathbf{x}_p$	10.00000 20.00000	1.000000 12.116667	1.000000 6.209640	1.000000 3.400059
$\mathbf{f}(\mathbf{x}_p)$	498.0000 462.0000 472.0000	145.8136 145.8136 137.8136	37.55963 37.55963 29.55963	10.56041 10.56041 25.60407
$\Sigma f_i^2(x_p)$	684232.0	61515.80	3695.223	229.60009

p	4	5	6	7
$\mathbf{x}_p$	1.000000 2.239236	1.000000 1.938349	1.000000 1.914996	1.000000 1.914854
$\mathbf{f}(\mathbf{x}_p)$	4.014178 4.014178 −3.985822	2.757199 2.757199 −5.242801	2.667212 2.667212 −5.332788	2.666667 2.666667 −5.333333
$\Sigma f_i^2(\mathbf{x}_p)$	48.114030	42.691255	42.666667	42.666667

EXAMPLE 3.

$$\mathbf{f}(\mathbf{x}) = \begin{cases} f_1(x_1, x_2) = x_1 + x_2 - 10 = 0 \\ f_2(x_1, x_2) = x_1 x_2 - 16 = 0. \end{cases}$$

The solutions of this system are (2,8) and (8,2). However, applying (12) with an initial $\mathbf{x}_0$ on the line: $x_1 = x_2$, results in the whole sequence being on the same line. Indeed,

$$J(x_1, x_2) = \begin{pmatrix} 1 & 1 \\ x_2 & x_1 \end{pmatrix} \quad \text{so that for} \quad \mathbf{x}_0 = \begin{pmatrix} a \\ a \end{pmatrix}$$

$$J(\mathbf{x}_0) = \begin{pmatrix} 1 & 1 \\ a & a \end{pmatrix}, \qquad J^+(\mathbf{x}_0) = \frac{1}{2(1+a^2)} \begin{pmatrix} 1 & a \\ 1 & a \end{pmatrix},$$

and consequently $\mathbf{x}_1$ is also on the line: $x_1 = x_2$. Thus confined to: $x_1 = x_2$, the sequence (12) will, depending on the choice of $\mathbf{x}_0$, converge to either (4.057646, 4.057646) or (− 3.313982, − 3.313982). These are the 2 least squares solutions on the line: $x_1 = x_2$.

ACKNOWLEDGMENT

The author gratefully acknowledges the computing facilities made available to him by the SWOPE Foundation.

REFERENCES

1. A. S. HOUSEHOLDER. "Principles of Numerical Analysis," McGraw-Hill, 1953.
2. A. M. OSTROWSKI. "Solution of Equations and Systems of Equations," Academic Press, New York, 1960.
3. R. BELLMAN AND R. KALABA. "Quasilinearization and Nonlinear Boundary-Value Problems." American Elsevier, New York ,1965.
4. T. H. HILDEBRANDT AND L. M. GRAVES. Implicit functions and their differentials in general analysis. *Trans. Amer. Math. Soc.* **29** (1927), 127-153.
5. L. V. KANTOROVIČ. On Newton's method for functional equations. *Dokl. Akad. Nauk SSSR* (N.S.) **59** (1948), 1237-1240; (*Math. Rev.* 9-537), 1948.
6. L. V. KANTOROVIČ. Functional analysis and applied mathematics. *Uspehi Mat. Nauk* (N.S.) **3**, No. 6 (28), (1948), 89-185 (*Math. Rev.* 10-380), 1949.
7. L. V. KANTOROVIČ. On Newton's method. *Trudy Mat. Inst. Steklov.* **28** (1949), 104-144 (*Math. Rev.* 12-419), 1951.
8. M. ALTMAN. A generalization of Newton's method. *Bull. Acad. Polon. Sci.* (1955), 189-193.
9. M. L. STEIN. Sufficient conditions for the convergence of Newton's method in complex Banach spaces. *Proc. Amer. Math. Soc.* **3** (1952), 858-863.
10. R. G. BARTLE. Newton's method in Banach spaces. *Proc. Amer. Math. Soc.* **6** (1955), 827-831.
11. J. SCHRÖDER. Über das Newtonsche Verfahren. *Arch. Rat. Mech. Anal.* **1** (1957), 154-180.
12. A. S. HOUSEHOLDER. "Theory of Matrices in Numerical Analysis." Blaisdell, 1964.
13. R. PENROSE. A generalized inverse for matrices. *Proc. Cambridge Phil. Soc.* **51** (1955), 406-413.
14. A. BEN-ISRAEL AND A. CHARNES. Contributions to the theory of generalized inverses. *J. Soc. Indust. Appl. Math.* **11** (1963), 667-699.
15. A. M. RODNYANSKII. On continuous and differentiable mappings of open sets of Euclidean space. *Mat. Sb.* (N.S.) **42** (84), (1957), 179-196.
16. S. KUREPA. Remark on the (F)-differentiable functions in Banach spaces" *Glasnik Mat.-Fiz. Astronom.* Ser. II **14** (1959), 213-217; (*Math. Rev.* 24-A1030), 1962.
17. A. BEN-ISRAEL. A modified Newton-Raphson method for the solution of systems of equations. *Israel J. Math.* **3** (1965), 94-98.
18. A. BEN-ISRAEL AND S. J. WERSAN. An elimination method for computing the generalized inverse of an arbitrary complex matrix. *J. Assoc. Comp. Mach.* **10** (1963), 532-537.

PRINTED IN BRUGES, BELGIUM, BY THE ST CATHERINE PRESS, LTD.

A Newton-Raphson Method for the Solution of Systems of Equations

ADI BEN-ISRAEL

Technion-Israel Institute of Technology and Northwestern University

Submitted by Richard Bellman

Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS. Vol. 15, No 2, August 1966