J. Numer. Math., Vol. 0, No. 0, pp. 1–6 (2006)

Prepared using jnm.sty *[Version: 02.02.2002 v1.2]*

Existence of a Unique Zero of Nonlinear Systems

P. G. BAO,* F. ANTON,† and J. G. ROKNE*‡

2006

Abstract — Existence statements for zeros of nonlinear equations are established via the interval Newton method generalizing previous theorems requiring regularity of the inclusion for the Jacobian.

Keywords: Newton's method, non-linear equations, existence of zeros

1. INTRODUCTION

In the following, intervals are denoted in uppercase letters, vector intervals are denoted in bold uppercase letters, while matrix intervals are denoted in underlined uppercase characters. However, the uppercase letter J has been used for the Jacobian. Sets of intervals are denoted in mathematical calligraphic characters. Let $\mathbb{R}$ be the set of reals, $\mathscr{I}$ the set of compact intervals on $\mathbb{R}$. $\mathbb{R}^n$ and $\mathscr{I}^n$ are the sets of n-dimensional vectors with elements from $\mathbb{R}$ and $\mathscr{I}$ respectively. Real and interval matrices are defined analogously. If $X \in \mathscr{I}$ then $X = [lbX, ubX]$, $w(X) = |ubX - lbX|$, and so on. A more extensive discussion of interval analysis notation is for example found in Ratschek-Rokne[6].

The interval Newton method for finding a zero of a function $f : \mathbb{R}^n \to \mathbb{R}^n$ was introduced by Moore[3]. A general form of the method is defined as follows:

Let f be as above and assume that an interval vector $\mathbf{X} \in \mathscr{I}^n$ and a point vector $x \in \mathbf{X}$ are given. Then the method consists of determining an inclusion $\mathbf{N}_{\underline{A}}(x, \mathbf{X})$ for the solution set

$$\mathbf{S}_{\underline{A}}(x, \mathbf{X}) = \{y : f(x) + A(y - x) = 0 \text{ for some } A \in \underline{A}\} \tag{1.1}$$

where $\underline{A}$ is an interval inclusion for the Jacobian of $f(x)$ over $\mathbf{X}$. If a zero ξ of $f(x) = 0$ exists in $\mathbf{X}$, it was shown in Moore[3] that $\xi \in \mathbf{N}_{\underline{A}}(x, \mathbf{X})$ with $w(\mathbf{N}_{\underline{A}}(x, \mathbf{X}) \cap \mathbf{X}) \leqslant$

---

*Information Technology, University of Southern Florida, Lakeland, Florida

†Department of Informatics and Mathematical Modelling, Technical University of Denmark, Building 321, DK-2800 Kgs. Lyngby and
Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton New Brunswick, Canada

‡Department of Computer Science,The University of Calgary, Calgary, Alberta, Canada, T2N-1N4
This work was supported by NSERC grant No. 69-1266.

$w(\mathbf{X})$, which means that (1.1) may be used as an iterative method for determining ξ.

A variety of interval Newton methods are now generated depending on the interval inclusion A of the Jacobian and on how the inclusion to (1.1) is computed. For an introduction to some of the methods see Hansen-Walster[2] and for a more extensive survey see Alefeld-Herzberger[1].

In Hansen-Walster[2],

Proposition 1.1. *If $\mathbf{N}_{\underline{A}}(x,\mathbf{X}) \subseteq \mathbf{X}$ for some $x \in \mathbf{X}$ then there exist a solution of $f(x) = 0$ in $\mathbf{X}$.*

is discussed. This proposition is true for many interval Newton methods and systems of equations $f(x)$. The first proof was given in Nickel[5] for a special interval Newton method. The most general proof is given in Neumaier[4] where the proposition was verified for a general class of interval Newton methods. The proof depends, however, on establishing à priori the invertibility of all matrices $A \in \underline{A}$.

The proposition above was noted to be a consequence of

CONJECTURE 1 *Let $\mathbf{S}_{\underline{A}}(x,\mathbf{X})$ denote the solution set of (1.1). If $\mathbf{S}_{\underline{A}}(x,\mathbf{X}) \subset \mathbf{X}$ then there exist a unique solution of $f(x) = 0$ in $\mathbf{S}_{\underline{A}}(x,\mathbf{X})$.*

In this note a stronger form of the above conjecture is proven as Theorem 3.2. The consequence of this theorem is then, that both the conjecture, and the proposition above are true, if we assume that the function f is not constant on $\mathbf{X}$ with respect to any of the variables.

2. THE INTERVAL HULL

It is now assumed that the function $f(x)$ is continuously differentiable in $\mathbf{X}$. In this section we discuss a property of the interval hull of the Jacobian $J(x)$ of $f(x)$.

Let $g(x)$ be a scalar valued function. Then the range of g over an interval $\mathbf{X}$ is defined as $\overline{g}(\mathbf{X}) = \{g(x) : x \in \mathbf{X}\} \in \mathscr{I}$. Similarly if f is a vector valued function then the range is defined as $\overline{f}(\mathbf{X}) = \{f(x) : x \in \mathbf{X}\}$. The interval hull $\Box f(\mathbf{X})$ is defined to be the smallest interval vector containing the range.

It is easy to see that the interval hull is inclusion monotonic (see Ratschek-Rokne[6] for definition of inclusion monotonicity).

Let $f(x)$ be a vector valued function. Let $J(x)$ be the Jacobian matrix of $f(x)$. For all $x \in \mathbf{X}$ we have

$$f(y) = f(x) + J(\xi)(y - x)$$

for some $\xi \in y \vee x$ (where $x \vee y$ denotes the line segment joining x and y). Replacing ξ by X it follows that the inclusions

$$f(y) \in f(x) + \overline{J}(\mathbf{X})(y-x) \subseteq f(x) + \Box J(\mathbf{X})(y-x) \subseteq f(x) + \underline{A}(y-x)$$

are valid where $\underline{A}$ is an interval matrix containing $\Box J(\mathbf{X})$.

We now prove a theorem showing the importance of the interval hull of the Jacobian in the interval Newton algorithm.

Theorem 2.1. *Let $f(x)$ be a vector valued function. Let $\underline{A}$ be any interval matrix such that for all $x, y \in \mathbf{X}$*

$$f(y) \in f(x) + \underline{A}(y-x),$$

Then

$$\Box J(\mathbf{X}) \subseteq \underline{A}. \tag{2.1}$$

Proof. Let A_{ij} be an element of $\underline{A}$. We only need to prove that

$$\Box J_{ij}(\mathbf{X}) \subseteq A_{ij}$$

for all i, j, $1 \leqslant i, j \leqslant n$. In fact we have

$$f(y) = f(x) + J(\xi)(y-x),$$

where $\xi \in x + [0,1](y-x)$. Let $x^* \in \mathbf{X}$ be a point such that $J_{ij}(x^*) = ub[\Box J_{ij}(\mathbf{X})]$. Let $x = x^*$ and $|y - x^*| \leqslant \delta$ such that $|J_{ij}(y) - ub[\Box J_{ij}(\mathbf{X})]| \leqslant \varepsilon$. It follows that

$$f(y) = f(x^*) + J(\xi)(y-x^*) \in f(x^*) + \underline{A}(y-x^*)$$

from the assumption.

Since $|\xi - x^*| \leqslant \delta$, $|J_{ij}(\xi) - ub[\Box J_{ij}(\mathbf{X})]| \leqslant \varepsilon$. From the assumption, $J_{ij}(\xi) \in A_{ij}$. Thus, $ub[\Box J_{ij}(\mathbf{X})] - \varepsilon \leqslant ubA_{ij}$. Then, by making ε go towards 0 and passing to the limit, we can conclude that

$$ub[\Box J_{ij}(\mathbf{X})] \leqslant ubA_{ij}.$$

Similarly we can prove

$$lb[\Box J_{ij}(\mathbf{X})] \geqslant lbA_{ij}.$$

Therefore formula (2.1) is true and so is the theorem.

3. THE EXISTENCE AND UNIQUENESS OF A ZERO

Let $b = (b_1, ..., b_n) := f(x)$, and $A =: (a_{ij})_{i=1,...,n; j=1,...,n}$. In this section we will prove the following theorem.

Theorem 3.1. *If for any $i \in [1,n]$ such that $b_i \neq 0$, $a_{i1} = a_{i2} = \cdots = a_{in} = 0$ does not hold and if $\mathbf{N}_{\underline{A}}(x, \mathbf{X}) \subseteq \mathbf{X}$ for some $x \in \mathbf{X}$, then there exists at most one zero of $f(x) = 0$ in $\mathbf{N}_{\underline{A}}(x, \mathbf{X})$.*

We will first prove the following two theorems (3.2 and 3.3) and then show that the theorem (3.1) above is a direct corollary of these two theorems. This theorem together with Theorem 2.1 establishes the conjecture in the case where for any $i \in [1,n]$ such that $b_i \neq 0$, $a_{i1} = a_{i2} = \cdots = a_{in} = 0$ does not hold, i.e. in the case f is not constant with respect to the variables on which $f(x)$ vanishes.

The theorem below is the main theorem of this paper.

Theorem 3.2. *Let $\mathbf{S}_{\underline{A}}(x,\mathbf{X})$ be the solution set*

$$\{y : f(x) + J(y-x) = 0 \text{ for some } J \in \overline{J}(\mathbf{X})\}.$$

then the following is true.

(i) If x^* is a zero of $f(x) = 0$ then for all $x \in \mathbf{X}$, $x^* \in \mathbf{S}_{\underline{A}}(x,\mathbf{X})$.

(ii) If all $A \in \overline{J}(\mathbf{X})$ are non-singular, there is at most one zero of $f(x)$ in $\mathbf{X}$.

(iii) If $f(x) = 0$ for all $x \in \mathbf{X}$ such that $J(x)$ is singular, then there is at most one zero of $f(x)$ in $\mathbf{X}$ or $f(x) \equiv 0$ in a closed convex domain $\mathbf{D}$ and $f(x)$ has no zero in $\mathbf{X} \setminus \mathbf{D}$.

(iv) If for any $i \in [1,n]$ such that $b_i \neq 0$, $a_{i1} = a_{i2} = \cdots = a_{in} = 0$ does not hold and if

$$\mathbf{S}_{\underline{A}}(x,\mathbf{X}) \subseteq \mathbf{X}$$

for some $x \in int(\mathbf{X})$ then there exists an unique $y^* \in \mathbf{S}_{\underline{A}}(x,\mathbf{X}) \subseteq \mathbf{X}$ with $f(y^*) = 0$.

Proof. Let x,y be any two points in $\mathbf{X}$. From the mean value theorem it follows that

$$f(y) - f(x) = J(x+t(y-x))(y-x), \tag{3.1}$$

for some $t \in (0,1)$. Since $x+t(y-x) \in \mathbf{X}$, $J(x+t(y-x) \in \overline{J}(\mathbf{X})$.

In particular, let y be a zero of $f(x)$ then

$$f(x) + J(x+t(y-x))(y-x) = 0$$

which means $y \in \mathbf{S}_{\underline{A}}(x,\mathbf{X})$. Hence (i) holds.

If all $A \in \overline{J}(\mathbf{X})$ are non-singular then for the $A \in \underline{A}$ that corresponds to the solution y it follows that

$$y = x - A^{-1} f(x) \tag{3.2}$$

and therefore if x is another zero of $f(x)$ then $y = x$. Hence (ii) holds as well.

Assume now that $f(x) = 0$ when $x \in \mathbf{X}$ and $J(x)$ is singular. If x,y are two distinct zeros of $f(x)$ then equation (3.1) implies that $J(x+t(y-x))(y-x) = 0$ for $t \in (0,1)$. Since $y-x \neq 0$, the matrix $J(x+t(y-x))$ must be singular, which implies that $f(x+t(y-x)) = 0$. That is, for any pair of zeros x,y, there is $t \in (0,1)$ such that $x+t(y-x)$ is a zero of $f(x)$. Now assume that for some $t_0 \in (0,1)$, $f(x+t_0(y-x)) \neq 0$. Let (α,β) be the maximum interval containing t_0 such that $f(x+t(y-x)) \neq 0$ for $t \in (\alpha,\beta)$. Then $x+\alpha(y-x)$ and $x+\beta(y-x)$ are two zeros of $f(x)$. From the discussion above there is a zero $x+t^*(y-x)$ for some $t^* \in (\alpha,\beta)$. Thus $f(x+t(y-x)) = 0$ for all $t \in (0,1)$. Since for $t = 0$, $f(x+t(y-x)) = f(x) = 0$ and for $t = 1$, $f(x+t(y-x)) = f(y) = 0$, $f(x+t(y-x)) = 0$ for all $t \in [0,1]$. That is, for any two zeros of $f(x)$, $f(z) = 0$ on the line segment $z \in x + [0,1](y-x)$. Thus $f(x) \equiv 0$ in a closed convex domain $\mathbf{D}$ and $f(x) \neq 0$ in $\mathbf{X} \setminus \mathbf{D}$. Hence (iii) holds.

We will show by contradiction that if for any $i \in [1,n]$ such that $b_i \neq 0$, $a_{i1} = a_{i2} = \cdots = a_{in} = 0$ does not hold and if $\mathbf{N}_{\underline{A}}(x,\mathbf{X}) \subseteq \mathbf{X}$ for some $x \in \mathbf{X}$, then no matrix in $\underline{A}$ is singular.

Let $A \in \underline{A}$ be singular. Then, the system of linear equations $f(x) + A(y-x) = 0$ in y is underdetermined (it has fewer independent equations than variables). If $f(x) = 0$, the system reduces to $A(y-x) = 0$ and there is an infinity of solutions. If $f(x) \neq 0$, since for any $i \in [1,n]$ such that $b_i \neq 0$, $a_{i1} = a_{i2} = \cdots = a_{in} = 0$ does not hold, the system has an infinite number of solutions. Thus, in all the cases, the system has an infinite number of solutions. Since the system of equations is underdetermined and there an infinite number of solutions, the co-dimension of the solution set is at least one and the solution set is not bounded. Thus, $\mathbf{N}_{\underline{A}}(x,\mathbf{X})$ is not compact. Therefore, $\mathbf{N}_{\underline{A}}(x,\mathbf{X})$ (non compact) can not be a subset of an interval (i.e. a compact set).

Now suppose that

$$f(y) = \tau f(x). \tag{3.3}$$

Let $x \in int(\mathbf{X})$ and $\mathbf{S}_{\underline{A}}(x,\mathbf{X}) \subseteq \mathbf{X}$. Since A can not be singular, $A^{-1}f(x) = x - (x - A^{-1}f(x)) \in x - \mathbf{X}$, and (3.2) derives

$$y \in x - (1-\tau)(x - \mathbf{X}) = \tau x - (1-\tau)\mathbf{X}. \tag{3.4}$$

(3.4) implies that $y \in int(\mathbf{X})$ if (3.3) holds for $\tau \in (0,1]$.

By the inverse function theorem applied to f, in an open set constructed in the following way: $\tau \in (\sigma,1)$ where $\sigma < 1$, we can state that for all $\tau \in (\sigma,1)$, (3.3) has a solution $y = \Phi(\tau) \in \mathbf{X}$. Let σ^* be the smallest nonnegative number such that (3.3) has a solution for $\tau \in (\sigma^*,1)$. Since $\mathbf{X}$ is compact, there is an accumulation point $y^* \in \mathbf{X}$ such that $\Phi(\tau) \to y^*$ for $\tau \to \sigma^*$. Since $f(x)$ is continuous, we have $f(y^*) = \sigma^* f(x)$. If $\sigma^* > 0$ then $y^* \in int(\mathbf{X})$. By applying the inverse function theorem with a similar construction, we can derive a $\sigma_1 < \sigma^*$ such that (3.3) has a solution for all $\tau \in (\sigma_1,1)$, which contradicts to the choice of σ^*. Hence $\sigma^* = 0$ and y^* is a zero of $f(x)$ in $\mathbf{X}$. By (ii), this zero is unique.

Theorem 3.3. *Let $\underline{A}$ be an interval matrix such that for all $x,y \in \mathbf{X}$*

$$f(y) \in f(x) + \underline{A}(y-x).$$

Let $\mathbf{S}_{\underline{A}}(x)$ denote the solution set

$$\{y : f(x) + A(y-x) = 0 \text{ for some } A \in \underline{A}\}.$$

Then $\mathbf{S}_{\underline{A}}(x,\mathbf{X}) \subseteq \mathbf{S}_{\underline{A}}(x)$.

Proof. From Theorem 2.1, $\overline{J}(\mathbf{X}) \subseteq \square J(\mathbf{X}) \subseteq \mathbf{X}$, thus

$$\mathbf{S}_{\underline{A}}(x,\mathbf{X}) = \{y : f(x) + \overline{J}(\mathbf{X})(y-x) = 0\} \subseteq \{y : f(x) + A(y-x) = 0 \text{ for } A \in \underline{A}\} = \mathbf{S}_{\underline{A}}(x).$$

4. CONCLUSION

In Hansen-Walster[2] a conjecture P regarding the zeroes of nonlinear equations computed via the interval Newton method was suggested. In this note the conjecture has been proven as a consequence of Theorem 3.2.

REFERENCES

1. Alefeld, G. and Herzberger, J.: *Introduction to Interval Computations.* Translated by J. Rokne. Academic Press, New York (1983).
2. Hansen, E. R. and Walster, G. W.: *Nonlinear Equations and Optimization.* Computers and Mathematics with Applications 25, pp. 125-145 (1993).
3. Moore, R.E.: *Interval analysis.* Prentice-Hall Inc., New Jersey (1966).
4. Neumaier, A.: *Interval Iteration for Zeros of Systems of Equations.* BIT 25, 256-273 (1985).
5. Nickel, K.: *On the Newton method in interval analysis.* Mathematics Research Center Report 1136. University of Wisconsin (1971).
6. Ratschek, H. and Rokne, J.: *Geometric Computations with Interval and New Robust Methods.* Horwood Publishing, Chichester (2003).