Validated Solutions of Nonlinear Equations

Rohn, Jiří
1995

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Validated Solutions of Nonlinear Equations

Jiří Rohn

Technical report No. 641

June 16, 1995

Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic
phone: (+422) 66414244 fax: (+422) 8585789
e-mail: uivt@uivt.cas.cz

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Validated Solutions of Nonlinear Equations[1]

Jiří Rohn[2]

Technical report No. 641
June 16, 1995

Abstract

An existence and uniqueness check for systems of nonlinear equations is given.

Keywords
Nonlinear equations, solution, existence

[1]This work was supported by the Czech Republic Grant Agency under grant GAČR 201/95/1484

[2]Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague (rohn@uivt.cas.cz), and Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic (rohn@kam.ms.mff.cuni.cz)

1 Introduction

We consider here a system of n nonlinear equations in n unknowns

$$f_i(x_1, \ldots, x_n) = 0 \qquad (i = 1, \ldots, n)$$

which we write simply as

$$f(x) = 0. \tag{1.1}$$

The main goal is to establish a verifiable sufficient condition for existence and uniqueness of solution of (1.1) in a given region $X \subset \mathrm{I\!R}^n$ (later chosen as an interval vector, $X = [x - d, x + d]$). Let $R(x) : X \to \mathrm{I\!R}^{n \times n}$ be any mapping such that $R(x)$ is a nonsingular matrix for each $x \in X$. Then (1.1) is equivalent to

$$x = x - R(x)f(x)$$

in X, and for

$$F(x) := x - R(x)f(x)$$

we obtain an equivalent fixed-point equation

$$x = F(x) \tag{1.2}$$

which we shall consider in the sequel. In the main result of this report (Theorem 1) we give an existence and uniqueness test for (1.2) together with a method for constructing a nested sequence of interval vectors containing the solution x^* of (1.2) and tending to x^* (hence, we have a verified enclosure of x^* at each iteration). The result is preceded by an auxiliary lemma which may be of independent interest. A simple nonexistence test is given in Theorem 2; a result of such type is necessary when using a branch–and–bound method (see e.g. Ratschek and Rokne [1]) for finding all solutions of (1.2) in some region X for which the conditions under which Theorem 1 works are not yet satisfied. In the last section we briefly consider the connection of these results to the material of our earlier report [2] on validated solutions of linear equations.

2 Lemma

Lemma 1 *Let $\{x_j\}_{j=0}^{\infty}$ and $\{d_j\}_{j=0}^{\infty}$ be vector (or scalar) sequences satisfying*

$$|x_j - x_{j+1}| \leq d_j - d_{j+1} \tag{2.1}$$

for each j, and let $d_j \to d^*$. Then we have:

1) $x_j \to x^*$,

2) $x^* \in [x_j - d_j + d^*, x_j + d_j - d^*]$ for each j,

3) the sequence of intervals $\{[x_j - d_j + d^*, x_j + d_j - d^*]\}_{j=0}^{\infty}$ is nested.

Proof. 0) For each $j \geq 0$ and $m \geq 1$, from (2.1) we have $|x_j - x_{j+m}| \leq \sum_{k=j}^{j+m-1} |x_k - x_{k+1}| \leq \sum_{k=j}^{j+m-1}(d_k - d_{k+1}) = d_j - d_{j+m}$, hence

$$|x_j - x_{j+m}| \leq d_j - d_{j+m}. \tag{2.2}$$

1) Let $\varepsilon > 0$. Since $\{d_j\}$ is convergent, there exists a j such that $|d_j - d_{j+m}| = d_j - d_{j+m} < \varepsilon$ for each $m \geq 1$. Then from (2.2) we have

$$|x_j - x_{j+m}| < \varepsilon$$

for each $m \geq 1$, hence $\{x_j\}$ is a cauchian sequence, thus $x_j \to x^*$.

2) For each $j \geq 0$, taking $m \to \infty$ in (2.2), we obtain

$$|x_j - x^*| \leq d_j - d^*,$$

which implies

$$x^* \in [x_j - d_j + d^*, x_j + d_j - d^*].$$

3) It follows from (2.1) that $x_j - d_j \leq x_{j+1} - d_{j+1}$ and $x_{j+1} + d_{j+1} \leq x_j + d_j$ for each j, hence the sequence of intervals $[x_j - d_j + d^*, x_j + d_j - d^*]$ is nested. ∎

3 Main result

Theorem 1 *Let F map an interval vector $[x-d, x+d] \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and let there exist a nonnegative matrix H with the following properties:*

(i) $|F(x') - F(x'')| \leq H|x' - x''|$ *for each* $x', x'' \in [x-d, x+d]$,

(ii) $|x - F(x)| < (I - H)d$.

Then the equation

$$\tilde{x} = F(\tilde{x}) \tag{3.1}$$

has a unique solution x^* in $[x-d, x+d]$, and the sequence of interval vectors $\{[x_j - d_j, x_j + d_j]\}_{j=0}^{\infty}$ defined by

$$\begin{aligned} x_0 &= x, \\ d_0 &= d \end{aligned}$$

and

$$\begin{aligned} x_{j+1} &= F(x_j), \\ d_{j+1} &= H d_j \end{aligned} \tag{3.2}$$

$(j = 0, 1, \ldots)$ *is nested, satisfies*

$$x^* \in [x_j - d_j, x_j + d_j]$$

for each j, and $x_j \to x^*$, $d_j \to 0$.

Proof. Since H is nonnegative, from (ii) we have $Hd < d$ and $d > 0$, hence $\varrho(H) < 1$ and $H^j \to 0$, $(I-H)^{-1} \geq 0$. We shall prove by induction that the sequences $\{x_j\}$ and $\{d_j\}$ satisfy

$$|x_j - x_{j+1}| \leq d_j - d_{j+1} \tag{3.3}$$

for each j. For $j = 0$ we have $|x_0 - x_1| = |x - F(x)| < (I-H)d = d_0 - d_1$ due to (ii). Let (3.3) hold for some $j \geq 0$, then

$$|x_{j+1} - x_{j+2}| = |F(x_j) - F(x_{j+1}| \leq H|x_j - x_{j+1}| \leq H(d_j - d_{j+1}) = d_{j+1} - d_{j+2},$$

which concludes the inductive proof of (3.3). Since $d_j = H^j d_0 \to 0$, Lemma 1 implies that $x_j \to x^*$, $x^* \in [x_j - d_j, x_j + d_j]$ for each j (in particular, $x^* \in [x-d, x+d]$), and the sequence $\{[x_j - d_j, x_j + d_j]\}_{j=0}^{\infty}$ is nested. Since F is continuous in $[x-d, x+d]$ due to (i), taking $j \to \infty$ in (3.2) we obtain that x^* solves (3.1). Let $\tilde{x}$ be any solution to (3.1). Then

$$|\tilde{x} - x^*| = |F(\tilde{x}) - F(x^*)| \leq H|\tilde{x} - x^*|,$$

hence

$$(I-H)|\tilde{x} - x^*| \leq 0$$

and premultiplying this inequality by the nonnegative matrix $(I-H)^{-1}$ yields $|\tilde{x} - x^*| \leq 0$, hence $\tilde{x} = x^*$. Thus x^* is the unique solution of (3.1) in $[x-d, x+d]$. ■

The assumption (i) is satisfied if F is differentiable and

$$\left|\frac{\partial F_i}{\partial x_j}(x')\right| \leq H_{ij}$$

holds for each $x' \in [x-d, x+d]$. Thus the values H_{ij} can be computed using interval extensions of the partial derivatives, see [1].

4 Nonexistence test

Theorem 2 *Let F map an interval vector $[x-d, x+d] \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and let it satisfy the assumption (i) of Theorem 1 for some nonnegative matrix H. If there exists an $x' \in [x-d, x+d]$ satisfying*

$$|x'_i - F_i(x')| > 2((I+H)d)_i \tag{4.1}$$

for some i, then the equation

$$\tilde{x} = F(\tilde{x})$$

does not have a solution in $[x-d, x+d]$.

Proof. Assume to the contrary that $x^* = F(x^*)$ holds for some $x^* \in [x-d, x+d]$. Then we have

$$|x' - F(x')| = |x' - x^* + F(x^*) - F(x')| \leq (I+H)|x' - x^*| \leq 2(I+H)d,$$

which contradicts (4.1). ■

5 The linear case

For a system of linear equations

$$Ax = b$$

(A square), using a nonsingular matrix R, we can write equivalently

$$x = (I - RA)x + Rb$$

and for $F(x) := (I - RA)x + Rb$ we have

$$|F(x') - F(x'')| \leq |I - RA| \cdot |x' - x''|,$$

hence the assumption (i) of Theorem 1 is satisfied for

$$H := |I - RA|.$$

Thus Theorem 1 generalizes both the "enclosure theorem" and the "refinement theorem" given for the linear case in our earlier paper [2]. Like in [2], Brouwer's fixed-point theorem has not been used in the proof of the main result.

Bibliography

[1] H. Ratschek and J. Rokne, New Computer Methods for Global Optimization, Ellis Horwood Ltd., Chichester (England) 1988

[2] J. Rohn, Validated Solutions of Linear Equations, Technical Report No. 620, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague 1995, 11 p.