Contemporary Mathematics
Volume 180, 1994

A Bisection Method to Find All Solutions of a System of Nonlinear Equations

PETR MEJZLÍK

June 20, 1994

ABSTRACT. This paper describes an algorithm for the solution of a system of nonlinear equations $F(x) = \theta$, where $F = (f_1, \dots, f_n) : D \subset \mathbf{R}^m \to \mathbf{R}^n$ and D is a compact domain, given that any of the functions f_i is monotonic when restricted to any single variable at an arbitrary point. The algorithm finds an approximation of the solutions as a union of m–dimensional intervals. The computation is based on reduction of the box containing all the solutions, its bisection, and elimination of subintervals which do not contain a solution. The algorithm does not require computation of partial derivatives or their approximations. Its use is illustrated on a model case.

1. Introduction

Let $F = (f_1, \dots, f_n) : [a, b] \subset \mathbf{R}^m \to \mathbf{R}^n$ be a continuous function. We will consider a generalized bisection method to approximate, with certainty, all solutions of a nonlinear system

$$F(x) = \theta, \tag{1.1}$$

where $\theta = (0, \dots, 0)$, given that

$$\Delta_{\delta e^j} f_i(x) \cdot \Delta_{\delta e^j} f_i(y) \geq 0, \quad \text{for all} \quad \delta > 0,\ 1 \leq i \leq n,\ 1 \leq j \leq m,\ x, y \in [a, b], \tag{1.2}$$

where $\Delta_h f(x) = f(x + h) - f(x)$ and e^j is the j-th unit vector.

1991 *Mathematics Subject Classification.* Primary 65H10, 65H20; Secondary 65Y05.
This paper is in final form and no version of it will be submitted for publication elsewhere.

Generalized bisection methods to solve (1.1) can be described using a root inclusion function T_F (defined formally in [4]) with values *true*, *false*, and *unknown*, which has the following properties [5]:
if $T_F(S)$=*true* then there is a unique solution of (1.1) within S.
if $T_F(S)$=*false* then there is no solution of (1.1) within S.
The template of a generalized bisection method can then be written as follows:

Input: Bounded domain $S \subset \mathbf{R}^m$.

(i) **if** $T_F(S) = false$ **then** the result is $\emptyset$.
(ii) (*optional*)
Reduce the domain S to $S' \subset S$ such that
$\{x; F(x) = \theta, x \in S\} = \{x; F(x) = \theta, x \in S'\}$.
(iii) **if** $\text{diam}(S')$ or $\sup\{\|F(x)\|; x \in S'\}$ are sufficiently small,
then S' is the result,
else
(a) Split S' into two subdomains S_1, S_2.
(b) Compute solutions on S_1 and S_2 separately, using this algorithm.

Current bisection methods to solve (1.1) [1–6] use the Jacobian matrix of F, or its approximations, to compute $T_F(S)$. The method presented here does not require computation of any derivatives, taking advantage of the restriction (1.2) on the class of functions considered. Our root inclusion test has only two values (*false* and *unknown*). The refinement of distinct solutions of (1.1) is done by customizations of a precision parameter of the algorithm, instead of testing whether $T_F(S) = true$.

2. The algorithm

The following is a recursive algorithm to solve problem (1.1), given (1.2). The algorithm contains a non–sequential loop ("for all"), instances of the body of which can be computed in any order. The algorithm can be efficiently run in parallel, as it decomposes the computation into independent subproblems.

DEFINITION 2.1. *We define norm*

$$\|x\|_M = \max_i \mid x_i \mid \quad \textit{for all } x \in \mathbf{R}^k, k = 1, 2, \ldots$$

Algorithm D($S, F, \varepsilon, \delta$):

Input: m–dimensional interval $S = [a, b] \subset \mathbf{R}^m$, function $F = (f_1, \ldots, f_n) : S \to \mathbf{R}^n$, and $\varepsilon, \delta > 0$.

(i) Reduce the m–dimensional interval $[a, b] = [(a_1, \ldots, a_m), (b_1, \ldots, b_m)]$, preserving all the roots of F on $[a, b]$:

repeat
$a^0 := a,\ b^0 := b,$
  for all $(i,j) \in \{1,\dots,n\} \times \{1,\dots,m\}$ **do**
    if f_i does not have a root in $[a,b]$
    then return $\emptyset$,
    else $a_j := \min\ \{x_j; x = (x_1,\dots,x_m) \in [a,b], f_i(x) = 0\}$,
      $b_j := \max\ \{x_j; x = (x_1,\dots,x_m) \in [a,b], f_i(x) = 0\}$,
until $\|a - a^0\|_M < \delta\ \wedge\ \|b - b^0\|_M < \delta$.

(ii) Test whether $\|F(x)\|_M < \varepsilon$ on $[a,b]$:

if for each $1 \le i \le n$: $|\min_{[a,b]} f_i(x)|, |\max_{[a,b]} f_i(x)| < \varepsilon$

then **return** $[a,b]$.

(iii) Divide $[a,b]$ into two parts:

Choose $k \in \{1,\dots,m\}$ such that $|a_k - b_k| = \max_{j=1,\dots,m} |a_j - b_j|$,

$S_0 := [a, (b_1,\dots,b_{k-1}, \frac{a_k+b_k}{2}, b_{k+1},\dots,b_m)]$,

$S_1 := [(a_1,\dots,a_{k-1}, \frac{a_k+b_k}{2}, a_{k+1},\dots,a_m), b]$,

return $\mathbf{D}(S_0,F,\varepsilon,\delta) \cup \mathbf{D}(S_1,F,\varepsilon,\delta)$.

To complete the specification, we describe methods to compute expressions in steps (i) and (ii) of algorithm **D**. The following theorem gives a recipe for computation of the minima and maxima in step (i). The problem is reduced to the calculation of the minimal root for a single–variable monotonic continuous function on an interval of finite size, which can be done by bisection. The computation of a minimum only is considered in the theorem, as the corresponding maximum can be computed as

$$\max\{x_j; x \in [a,b], f_i(x) = 0\} = -\min\{x_j; x \in [-a,-b], f_i(-x) = 0\}.$$

THEOREM 2.1. *Let $f : [a,b] \subset \mathbf{R}^m \to \mathbf{R}$ be a continuous function, which fulfills* (1.2), *$j \in \{1,\dots,m\}$, and let $\Delta_{\delta e^j} f(x) \ge 0$ for all $\delta > 0$. Let us define*

$$\begin{aligned} c_k &= b_k \quad \textit{if } \Delta_{\delta e^k} f(x) \ge 0 \textit{ on } [a,b], \textit{ for all } \delta > 0, \\ &= a_k \quad \textit{otherwise}, \\ &\quad (k = 1,\dots,j-1,j+1,\dots,m), \\ \phi_j(t) &= f(c_1,\dots,c_{j-1},t,c_{j+1},\dots,c_m). \end{aligned}$$

Then

(i) ϕ_j *is continuous and non–decreasing on* $[a_j, b_j]$.

(ii) *If* f *has a root within* $[a, b]$, *then*

$$\begin{aligned}\min\{x_j; x \in [a,b], f(x) = 0\} &= \min\{t; \phi_j(t) = 0, t \in [a_j, b_j]\}, \\ &\qquad \textit{if } \phi_j \textit{ has a root in } [a_j, b_j], \\ &= a_j \quad \textit{otherwise.}\end{aligned}$$

PROOF.

(i) The continuity of ϕ_j follows directly from the continuity of f. To prove that ϕ_j is non–decreasing, it is sufficient to note that by (1.2) the term $\Delta_{\delta e^j} f(x)$ does not change sign for any $\delta > 0$, $x \in [a, b]$, and $\mathrm{sgn}(\Delta_{\delta e^j} \phi_j(x)) = \mathrm{sgn}(\Delta_{\delta e^j} f(x))$.

(ii) From the definition of function ϕ_j we see that

$$\phi_j(x_j) \geq f(x) \quad \text{for all } x \in [a, b]. \tag{2.1}$$

Let $x_j^* = \min\{x_j; x \in [a, b], f(x) = 0\}$. Then from the definition of ϕ_j it follows that $x_j^* \leq \min\{t; \phi_j(t) = 0, t \in [a_j, b_j]\}$ (otherwise there would be some $t \in [a_j, x_j^*)$ such that $\phi_j(t) = f(c_1, \dots, c_{j-1}, t, c_{j+1}, c_m) = 0$). If $\phi_j(x_j^*) = 0$, then clearly

$$\min\{x_j; x \in [a, b], f(x) = 0\} = \min\{t; \phi_j(t) = 0, t \in [a_j, b_j]\}.$$

If $\phi_j(x_j^*) \neq 0$, then $\phi_j(t) > 0$ for all $x_j^* \leq t \leq b_j$, which follows from part (i) of this theorem and from (2.1). There is also $\phi_j(t) > 0$ for $a_j \leq t \leq x_j^*$, because $\phi(t) \neq 0$ on $[a_j, x_j^*]$ by the definition of ϕ_j. Thus, $\{t; \phi_j(t) = 0, t \in [a_j, b_j]\} = \emptyset$.

□

In the case when $\Delta_{\delta e^j} f(x) \leq 0$, we can compute the minimum in step (i) of the algorithm as $\min\{x_j; x \in [a, b], f_i(x) = 0\} = \min\{x_j; x \in [a, b], -f_i(x) = 0\}$. The term on the right side can be evaluated by Theorem 2.1 because $\Delta_{\delta e^j}(-f(x)) = -\Delta_{\delta e^j} f(x) \geq 0$.

From Theorem 2.1 it is clear that in step (ii)

$$\begin{aligned}|\max_{[a,b]} f_i(x)| &= \phi_1(b_1) \quad \text{if } \Delta_{\delta e^j} f(x) \geq 0 \text{ on } [a, b], \\ &= \phi_1(a_1) \quad \text{otherwise.}\end{aligned}$$

$|\min_{[a,b]} f_i(x)|$ can be computed similarly.

3. Convergence

THEOREM 3.1. *Let* $F : [a, b] \subset \mathbf{R}^m \to \mathbf{R}^n$ *be a continuous function, which fulfills* (1.2), *and let* $\varepsilon, \delta > 0$. *Let* S^* *be the output of algorithm* $\mathbf{D}$ *for* $F, [a, b], \varepsilon, \delta$, *i.e.,* $S^* = \mathbf{D}(F, [a, b], \varepsilon, \delta)$. *Then*

(i) S^* *contains all the roots of* F *on* $[a, b]$.

(ii) $\|F(x)\|_M < \varepsilon$ *for all* $x \in S^*$.

PROOF. Both parts of the theorem follow directly from the iterational part of algorihm **D** and from the definition of $\|.\|_M$. □

The following theorem shows that it is always possible to distinguish between two distinct solutions of (1.1) using a sufficiently small parameter ε in algorithm **D**.

THEOREM 3.2. *Let* $F : [a,b] \subset \mathbf{R}^m \to \mathbf{R}^n$ *be a continuous function, which fulfills* (1.2), $\delta > 0$, *and let* $\theta \notin F([a,b])$. *Then there is* $\varepsilon > 0$ *such that* $\mathbf{D}([a,b], F, \varepsilon, \delta) = \emptyset$.

PROOF. The interval $[a,b]$ is a compact set, which implies that $F([a,b])$ is also compact, because function F is continuous. The norm $\|.\|_M$ has a minimum on $F([a,b])$, as $\|.\|_M$ is continuous on $\mathbf{R}^n$. Let us have

$$\varepsilon = \min\{\|y\|_M; y \in F([a,b])\} = \min\{\|F(x)\|_M; x \in [a,b]\}.$$

Then $\varepsilon > 0$, as $F(x) = \theta$ has no solution within $[a,b]$. From Theorem 3.1(ii) we have

$$\|F(x')\|_M < \varepsilon = \min\{\|F(x)\|_M; x \in [a,b]\} \quad \text{for all } x' \in \mathbf{D}(F,[a,b],\varepsilon,\delta).$$

Thus, $\mathbf{D}(F,[a,b],\varepsilon,\delta) = \emptyset$. □

4. Example

We illustrate the method with a system of two nonlinear inequalities:

$$\begin{aligned} y &\leq 1/x \\ x^2 + y^2 &\leq 4 \end{aligned} \tag{4.1}$$

for $x, y \in [0.01, 2]$.

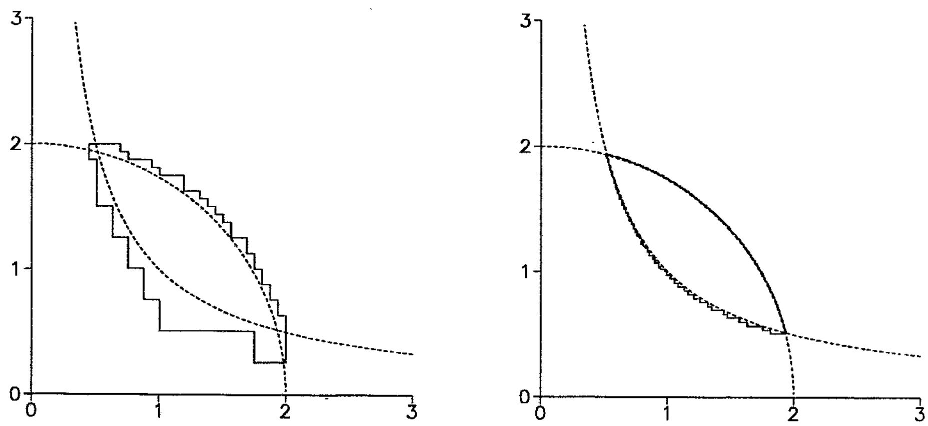
It can be seen that the set of solutions is the "lens" in Figure 1. The system (4.1) is equivalent to the following set of equations:

$$\begin{aligned} f_1(x,y) &= 0 \\ f_2(x,y) &= 0 \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} f_1(x,y) &= 0 \quad \text{for } y \geq \frac{1}{x}, \\ &= \frac{1}{x} - y \quad \text{otherwise}, \\ f_2(x,y) &= 0 \quad \text{for } x^2 + y^2 \leq 4, \\ &= x^2 + y^2 - 4 \quad \text{otherwise}. \end{aligned}$$

The results of computations of solutions of (4.2) by algorithm **D** are shown in Fig. 1.



(a) (b)

FIGURE 1. Output of algorithm **D** for for the example system.
(a) $\varepsilon = 0.5$, $\delta = \infty$, (b) $\varepsilon = 0.05$, $\delta = \infty$.

REFERENCES

1. G. Alefeld and L. Platzöder, *A quadratically convergent Krawczyk-like algorithm*, SIAM J. Numer. Anal. **20** (1983), 210–219.
2. E. R. Hansen, *A globally convergent interval method for computing and bounding real roots*, BIT **18** (1978), 415–424.
3. R. B. Kearfott, *Preconditioners for the interval Gauss–Seidel method*, SIAM J. Numer. Anal. **27** (1990), 804–822.
4. ———, *Abstract generalized bisection and a cost bound*, Math. Comput. **49** (1987), 187–202.
5. ———, *Some tests on generalized bisection*, ACM Trans. Math. Softw. **13** (1987), 197–220.
6. R. E. Moore and S. T. Jones, *Safe starting regions for iterative methods*, SIAM J. Numer. Anal. **14** (1977), 1051–1065.

INSTITUTE OF COMPUTER SCIENCE, MASARYK UNIVERSITY, BUREŠOVA 20, 60200 BRNO, CZECH REPUBLIC
E-mail address: mejzlik@muni.cz