ICCM-2004 WORKSHOPS P. 307–313

Krawczyk operator revised

Sergey P. Shary*

Abstract. For solution existence tests based on Brouwer fixed point theorem, we propose modifications that make use, first, of the idea of bicentered interval extension of functions and, second, of the restriction of the test domain to the boundary of the box only. Being applied both separately and in combination with each other, they allow one to substantially increase the efficacy of computational procedures for verified solution of equations systems by interval techniques.

Introduction

In nonlinear analysis, the following fact is well-known

Brouwer fixed point theorem [1]. *Let $D \subseteq \mathbb{R}^n$ be a convex compact set. If $\Phi : D \to \mathbb{R}^n$ is a continuous function that maps D into itself, i. e.*

$$\Phi(D) \subseteq D, \tag{1}$$

then Φ has a fixed point $x^\star$ on D, such that $x^\star = \Phi(x^\star)$.

Every system of n equations and n unknown variables

$$F(x) = 0, \tag{2}$$

$F(x) = \big(F_1(x), F_2(x), \ldots, F_n(x)\big)^\top$, $x = (x_1, x_2, \ldots, x_n)^\top$, can be reduced to equivalent recurrent form

$$x = \Phi(x) \tag{3}$$

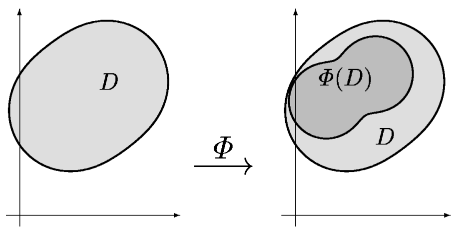

Figure 1. Illustration of Brouwer fixed point theorem

with $\Phi(x) = x - \Lambda F(x)$, where Λ is a nonsingular $n \times n$-matrix. Because of this, Brouwer fixed point theorem and its generalizations have been often used for proving existence of solutions to equations and systems of equations. These applications were, nevertheless, mainly of theoretical character since finding the image of a set under the action of a mapping is not an easy task. It could be resolved analytically in very few special cases. The situation radically changed after appearance of interval analysis, a mathematical discipline that makes it possible to operate, on computers, sets of small and moderate constructive complexity, through establishing arithmetical and analytical operations, relations, etc., between these sets as individual entities.

*Institute of computational technologies SB RAS.

1. Interval methods for the solution of equations

In our text, the interval notation adheres to the recently adopted project of the international standard [5]. Specifically, we designate intervals and interval objects (vectors, matrices, functions) by boldface letters. $\mathbb{IR}$ stands for classical interval arithmetic [7, 8] or its support, that is, the set of closed intervals of the real axis $\mathbb{R}$. $\mathbb{IR}^n$ means the set of n-dimensional interval vectors, whose geometric images are axes aligned boxes in $\mathbb{R}^n$.

In interval analysis, estimating the image of a set under the action of a mapping takes a specific form, being associated with a problem of computing the so-called *interval extension* of a function.

Definition 1 [7, 8]. Interval function $\boldsymbol{f} : \mathbb{IR}^n \to \mathbb{IR}^m$ is called interval extension of a real function $f : \mathbb{R}^n \to \mathbb{R}^m$, if

1) $\boldsymbol{f}(x) = f(x)$ for any $x \in \mathbb{R}^m$ from the domain of f,
2) $\boldsymbol{f}(\boldsymbol{x})$ is inclusion monotonic, i.e., $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}^n$, $\boldsymbol{x} \subseteq \boldsymbol{y} \Rightarrow \boldsymbol{f}(\boldsymbol{x}) \subseteq \boldsymbol{f}(\boldsymbol{y})$.

Therefore, if $\boldsymbol{f}(\boldsymbol{x})$ is an interval extension of the function $f(x)$, then always

$$\{f(x) \mid x \in \boldsymbol{x}\} \subseteq \boldsymbol{f}(\boldsymbol{x}),$$

and we get an outer (by superset) estimate of the range of f over the box $\boldsymbol{x} \in \mathbb{IR}^n$. Constructing interval extensions of functions is one of the most important problems that interval analysis deals with, and its various aspects have been being under investigation since 1960 up to now. It makes sense to present the first result on the subject, which is often called "the main theorem of interval arithmetic":

Theorem 1 [7, 8]. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a rational function of the arguments $(x_1, x_2, \dots, x_n) = x$. If, for a box $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$, defined is the result $\boldsymbol{f}(\boldsymbol{x})$ of substituting the intervals $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ instead of the arguments of $f(x)$ and further performing the interval arithmetic operations, then*

$$\{f(x) \mid x \in \boldsymbol{x}\} \subseteq \boldsymbol{f}(\boldsymbol{x}),$$

that is, $\boldsymbol{f}(\boldsymbol{x})$ contains the range of values of the function $f(x)$ on $\boldsymbol{x}$.

The interval extension of the rational function $f(x)$, whose construction is described in the main theorem of interval arithmetic, is referred to as *natural interval extension* $\boldsymbol{f}_{nat}(\boldsymbol{x})$, and its values can be computed by elementary means. At the same time, using the natural interval extensions often leads to very crude results when estimating the ranges of functions. In this connection, more advanced forms of interval extensions have been developed, and one of the most popular among them is the so-called *centered form*. For a function $f : \mathbb{R}^n \to \mathbb{R}$, the centered form of the interval extension looks as follows:

$$\boldsymbol{f}_c(\boldsymbol{x}, \tilde{x}) = f(\tilde{x}) + \sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{x}, \tilde{x})(\boldsymbol{x}_i - \tilde{x}_i),$$

where $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ is a "center" point, $\boldsymbol{g}_i(\boldsymbol{x}, \tilde{x})$ are intervals that depend on $\tilde{x}$ and $\boldsymbol{x}$ in general. In particular, $\boldsymbol{g}_i(\boldsymbol{x}, \tilde{x})$ may be interval enclosures of the ranges of partial derivatives $\partial f(x)/\partial x_i$ over $\boldsymbol{x}$. The interested reader can draw the further information from the books [4, 7, 8], which expound the construction of interval extensions of functions in detail.

If the tools for computing interval extensions of functions are available, we can bypass the difficulties arising in the practical verification of the inclusion (1) in Brouwer fixed point theorem, providing that,

- first, we restrict ourselves to considering the domains D in the form of interval boxes, that is, requiring $D \in \mathbb{IR}^n$, and,
- second, we change the exact range of values of the function Φ over D (it may have a complicated shape) to its outer estimate through interval extension.

Following this way, one can derive the solution existence tests by Krawczyk, Moore, Qi, etc., which are very popular in the modern interval analysis. In the sequel, the first one of these, proposed by West German mathematician Rudolf Krawczyk in 1969 [6], is of special interest for us. It is usually introduced as follows:

Definition 2. Let some rules be defined that assign, to any box $\boldsymbol{x} \in \mathbb{IR}^n$, a point $\tilde{x} \in \boldsymbol{x}$ and a real $n \times n$-matrix Λ, while interval $n \times n$-matrix $\boldsymbol{G}$ is an enclosure for the derivative $F'(x)$ of the function $F : \mathbb{R}^n \to \mathbb{R}^n$ over the box $\boldsymbol{x}$. The mapping

$$\mathcal{K} : \mathbb{IR}^n \times \mathbb{R}^n \to \mathbb{IR}^n,$$

defined by the rule

$$\mathcal{K}(\boldsymbol{x}, \tilde{x}) := \tilde{x} - \Lambda F(\tilde{x}) + (I - \Lambda \boldsymbol{G})(\boldsymbol{x} - \tilde{x}),$$

is called interval *Krawczyk operator* for the function F.

Krawczyk operator is nothing but the centered form of the interval extension of the mapping $\Phi(x) = x - \Lambda F(x)$, which emerges in the right-hand side of the equations system (2) after it is reduced to the reccurrent form (3).

The following important statements concerning Krawczyk operator are valid [6–8]:

(i) every solution $x^\star$ of the system $F(x) = 0$ within the box $\boldsymbol{x}$ is also contained in $\mathcal{K}(\tilde{x}, \boldsymbol{x})$, so that $x^\star \in \boldsymbol{x} \cap \mathcal{K}(\boldsymbol{x}, \tilde{x})$;

(ii) if $\boldsymbol{x} \cap \mathcal{K}(\boldsymbol{x}, \tilde{x}) = \emptyset$, then the box $\boldsymbol{x}$ contains no solutions of the equations system $F(x) = 0$;

(iii) if $\mathcal{K}(\boldsymbol{x}, \tilde{x}) \subseteq \boldsymbol{x}$, then the box $\boldsymbol{x}$ contains, with certainty, at least one solution of the system $F(x) = 0$;

(iv) if $\mathrm{int}\ \boldsymbol{x} = \{x \in \mathbb{R}^n \mid \underline{\boldsymbol{x}}_i < x_i < \overline{\boldsymbol{x}}_i \text{ for every } i\}$ is the interior of the box $\boldsymbol{x}$, $\tilde{x} \in \mathrm{int}\ \boldsymbol{x}$ and $\mathcal{K}(\boldsymbol{x}, \tilde{x}) \subseteq \mathrm{int}\ \boldsymbol{x}$, then the matrix $\boldsymbol{G}$ is strongly nonsingular and $\mathcal{K}(\boldsymbol{x}, \tilde{x})$ contains exactly one solution of the system $F(x) = 0$.

Using the result of the item (i), we can reduce the box which is suspected to have a solution. The item (ii) provides us with an instrument for sifting unpromising boxes that does not have solutions. Conversly, the items (iii) and (iv), which follow from Brouwer fixed point theorem and further fine results, enables one to prove the existence of the solutions and even their uniqueness.

2. Bicentered Krawczyk operator

As far as Krawczyk operator is a centered form interval extension of the mapping $\Phi(x)$ from (3), it is amenable to all the modifications that can be applied to the centered forms in general. A promising way to improve the quality of the enclosures computed by the centered form of interval extensions is varying the center $\tilde{x}$, and a "final" result in the direction is

Baumann theorem [5, 8]. *Let a "cut-off function"* cut : $\mathbb{R} \times \mathbb{IR} \to \mathbb{R}$ *be defined as*

$$\operatorname{cut}(x, \boldsymbol{x}) = \begin{cases} \overline{\boldsymbol{x}}, & \textit{if} \quad x \geq \overline{\boldsymbol{x}}, \\ \underline{\boldsymbol{x}}, & \textit{if} \quad x \leq \overline{\boldsymbol{x}}, \\ \tilde{x}, & \textit{otherwise.} \end{cases}$$

Let also, for a smooth function $f : \mathbb{R}^n \to \mathbb{R}$,

$$p_i = \operatorname{cut}\left(\frac{\operatorname{mid}(f'(\boldsymbol{x}))_i}{\operatorname{rad}(f'(\boldsymbol{x}))_i}, [-1, 1]\right), \qquad i = 1, 2, \ldots, n,$$

and n-vectors $\check{x} = (\check{x}_i)$, $\hat{x} = (\hat{x}_i)$ *be such that*

$$\check{x}_i = \operatorname{mid} \boldsymbol{x}_i - p_i \operatorname{rad} \boldsymbol{x}_i, \quad \hat{x}_i = \operatorname{mid} \boldsymbol{x}_i + p_i \operatorname{rad} \boldsymbol{x}_i,$$

where $f'(\boldsymbol{x})$ *is an interval enclosure of gradients for* $x \in \boldsymbol{x}$, mid *and* rad *are mid-point and radius of an interval. Then*

1) *the lower endpoint of the centered form,* $\underline{\boldsymbol{f}_c(\boldsymbol{x}, \tilde{x})}$, *attains its maximum for* $\tilde{x} = \check{x}$,
2) *the upper endpoint of the centered form,* $\overline{\boldsymbol{f}_c(\boldsymbol{x}, \tilde{x})}$, *attains its minimum for* $\tilde{x} = \hat{x}$,
3) *the radius of the centered form,* rad $\boldsymbol{f}_c(\boldsymbol{x}, \tilde{x})$, *attains its minimum for* $\tilde{x} = \operatorname{mid} \boldsymbol{x}$.

Baumann theorem naturally begets the so-called *bicentered form* of interval extensions of functions in which we take the intersection of two ordirnary centered interval extensions computed with respect to two optimal centers $\check{x}$ and $\hat{x}$:

$$\boldsymbol{f}_{bic}(\boldsymbol{x}) = \boldsymbol{f}_c(\boldsymbol{x}, \check{x}) \cap \boldsymbol{f}_c(\boldsymbol{x}, \hat{x}).$$

And the same idea can be applied to Krawczyk operator as well! In doing so, each separate component with the number i should be supplied with its own optimal centers $\check{x}^i$ and $\hat{x}^i$, $i = 1, 2, \ldots, n$.

An outline of the algorithm for computing bicentered form of the interval Krawczyk operator looks as follows:

1. Compute "shifts" of the centers

$$p_{ij} = \mathrm{cut}\left(\frac{\mathrm{mid}(\Phi'_{ij}(\boldsymbol{x}))_i}{\mathrm{rad}(\Phi'_{ij}(\boldsymbol{x}))_i}, [-1,1]\right), \qquad i,j = 1,2,\ldots,n,$$

   where $\Phi'_{ij}(\boldsymbol{x})$ is an interval enclosure, over $\boldsymbol{x}$, of the derivative $\partial\Phi_i/\partial x_j$ of the function $\Phi(x) = x - \Lambda F(x)$.

2. Compute the "shifted centers"

$$\check{x}^i_j = \mathrm{mid}\,\boldsymbol{x}_j - p_{ij}\,\mathrm{rad}\,\boldsymbol{x}_j, \quad \hat{x}^i_j = \mathrm{mid}\,\boldsymbol{x}_j + p_{ij}\,\mathrm{rad}\,\boldsymbol{x}_j, \qquad i,j = 1,2,\ldots,n.$$

3. Compute the centered forms proper:

$$\mathcal{K}_i(\boldsymbol{x}, \check{x}^i) = (\check{x}^i_i - \Lambda F(\check{x}^i)) + (I - \Lambda\boldsymbol{G})(\boldsymbol{x} - \check{x}^i),$$
$$\mathcal{K}_i(\boldsymbol{x}, \hat{x}^i) = (\hat{x}^i_i - \Lambda F(\hat{x}^i)) + (I - \Lambda\boldsymbol{G})(\boldsymbol{x} - \hat{x}^i), \qquad i = 1,2,\ldots,n.$$

4. Compute the bicentered Krawczyk operator:

$$\mathcal{K}_{bic}(\boldsymbol{x}) = \mathcal{K}(\boldsymbol{x}, \check{x}) \cap \mathcal{K}(\boldsymbol{x}, \hat{x})$$

The statements (i)–(iv) that substantiate application of Krawczyk operator for testing solutions of equations will remain valid for bicentered Krawczyk operator too. However, we should pay a special attention to the choice of the "preconditioning" matrix Λ: it must not be equal to the inverse of the midpoint of the derivative enclosure matrix $F'(\boldsymbol{x})$, insofar as all the shifts p_{ij} are zeros in this case, and the advantages of the bicentered form are unrealized.

3. Boundary Krawczyk operator

Brouwer fixed point theorem can be substantially strengthened:

Strengthened Brouwer fixed point theorem [3, 9]. *Let $D \subseteq \mathbb{R}^n$ be a convex compact set. If $\Phi : \mathbb{R}^n \supseteq D \to \mathbb{R}^n$ is a continuous function that maps the boundary ∂D of the set D into D itself, i.e.*

$$\Phi(\partial D) \subseteq D, \tag{4}$$

then Φ has a fixed point $x^\star$ on D, such that $x^\star = \Phi(x^\star)$.

The above formulation is really stronger than the original Brouwer fixed point theorem since it embraces a wider class of mapplings to be verified with respect for the existence of the fixed point. It is applicable for functions that may transform

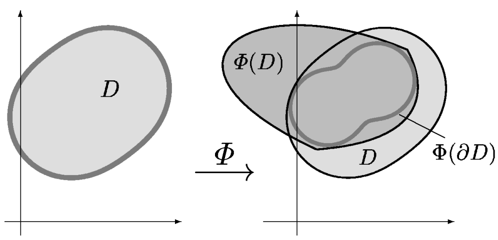

Figure 2. Illustration for strengthened Brouwer fixed point theorem

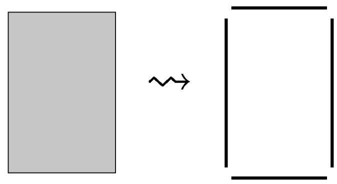

Figure 3. Only boundary to be checked instead of the whole box ...

D into a set stretching outside D itself, and the function's behavior only on the boundary proves to be essential (Figure 2).

Additionally, when testing the solutions by interval techniques, checking the condition (4) is much more beneficial than (1) from the accuracy reasons. The point is that, for any form of interval extension, the accuracy of the interval evaluation crucially depends on the width of the domain box. Usually, the excess width of the interval enclosure is proportional to a certain power of the width of the box, namely, the first power for the natural interval extension, the second power for centered form, and so on. Meanwhile, the boundary of a box is always "thinner" than the box itself (Figure 3).

As applied to Krawczyk operator in the property (iii), we suffice to check out its action not on the entire box $\boldsymbol{x}$, but only on its boundary made up from $2n$ pieces

$$\begin{array}{ll}
(\underline{\boldsymbol{x}}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n), & (\overline{\boldsymbol{x}}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n), \\
(\boldsymbol{x}_1, \underline{\boldsymbol{x}}_2, \ldots, \boldsymbol{x}_n), & (\boldsymbol{x}_1, \overline{\boldsymbol{x}}_2, \ldots, \boldsymbol{x}_n), \\
\ddots & \ddots \\
(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \underline{\boldsymbol{x}}_n), & (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \overline{\boldsymbol{x}}_n).
\end{array}$$

Each one of the above is an $(n-1)$-dimensional interval box having both the dimension and width smaller than those of $\boldsymbol{x}$. Therefore, we will get a more sensitive variant of the existence test (iii) at the price of some additional labor (approximately $2n$ times larger).

The idea of bicentered Krawczyk operator may also be implemented to result in further efficiency improvement.

References

[1] Nikaido H. Convex Structures and Economic Theory. – New York: Academic Press, 1968.

[2] Baumann E. Optimal centered forms // BIT. – 1988. – Vol. 28. – P. 80–87.

[3] Kantorovich L.V., Akilov G.P. Functional Analysis: 3rd edition. – Moscow: Nauka, 1984 (in Russian).

[4] Kearfott R.B. Rigorous Global Search: Continuous Problems. – Dordrecht: Kluwer, 1996.

[5] Kearfott R.B., Nakao M.T., Neumaier A., Rump S.M., Shary S.P., van Hentenryck P. Standardized notation in interval analysis // Reliable Computing. – submitted (see http://www.mat.univie.ac.at/ neum/software/int).

[6] Krawczyk R. Newton-Algorithmen zur Besstimmung von Nullstellen mit Fehlerschranken // Computing. – 1969. – Vol. 4. – P. 187–201.

[7] Moore R.E. Methods and Applications of Interval Analysis. – Philadelphia: SIAM, 1979.

[8] Neumaier A. Interval Methods for Systems of Equations. – Cambridge: Cambridge University Press, 1990.

[9] Opoitsev V.I. Nonlinear Systemostatics. – Moscow: Nauka, 1986 (in Russian).