Interval Methods That Are Guaranteed to Underestimate
(and the resulting new justification of Kaucher arithmetic)

V. Kreinovich, V. M. Nesterov, and N. A. Zheludeva

Introduction

One of the main objectives of interval computations is, given the function \( f(x_1, ..., x_n) \), and \( n \) intervals \( \bar{x}_1 = [\bar{x}_1^-, \bar{x}_1^+] \), ..., \( \bar{x}_n = [\bar{x}_n^-, \bar{x}_n^+] \), to compute the interval \( \bar{y} = f(\bar{x}_1, ..., \bar{x}_n) \).

Traditional methods of interval computation compute an enclosure \( Y \supseteq \bar{y} \) for the desired interval \( \bar{y} \).

Problem. The main problem with these methods is that this enclosure is sometimes close to \( \bar{y} \); sometimes it is a gross overestimation. It would be nice to get an idea how close \( Y \) is to \( \bar{y} \).

The need for an underestimating method. To solve this problem, it is desirable to develop a method that would produce an inner estimate for \( \bar{y} \), i.e., an interval \( Y \) such that \( y \subseteq \bar{y} \).

1 Our Idea

Why do interval computations overestimate? Let us consider a textbook example when naive interval computations overestimate: computing the range \( f(\bar{x}) \) of the function \( f(x) = x - x^2 \), for \( \bar{x} = [0, 1] \). This computation consists of two steps: first, we compute the interval of possible values of \( x^2 = x \ast x \) by multiplying the interval \( [0, 1] \) with itself: \( [0, 1] \ast [0, 1] = [0, 1] \). Then, we find an enclosure for the desired range as a difference between the two intervals \( Y = [0, 1] - [0, 1] = [-1, 1] \). This is a gross overestimation of the actual range \( \bar{y} = [0, 0.25] \).

In this case, the interval \( X_1 = \bar{x}_1 \) of possible values of \( x_1 = x \) is \([0, 1]\); the interval \( X_2 = \bar{x}_2 = [0, 1] \) of possible values of \( x_2 = x^2 \) has been computed precisely. However, when we estimate the range of the difference \( g(x_1, x_2) = x_1 - x_2 \), we overestimate, because we are using the formula of interval computations that is based on the assumption
that the set $X$ of possible pairs $(x_1, x_2)$ coincides with the entire box $X_1 \times X_2$. In our example, $x_1$ and $x_2$ are related and therefore, not all the values from this box are possible: $X \subset X_1 \times X_2$. In general, if we have the intervals $\tilde{x}_i$ for $x_i$ computed correctly, we know that this set $X$ satisfies the following two properties:

1) First, $X$ is a connected set (since it is an image of an interval under a continuous function).

2) Second, that (since we have intervals for $x_1$ and $x_2$ correctly), the projection $\pi_i(X)$ of the set $X$ on $i$–th axis is exactly $\tilde{x}_i$.

So, the actual range is equal to $g(X)$ for one of the sets that satisfies the conditions 1) and 2). If we only know the intervals $\tilde{x}_i$ and we do not know which set $X$ we are dealing with, we can be sure that the desired range contains the intersection of the sets $g(X)$ for all possible sets $X$.

We can follow this idea step-by-step, and on each step of the calculations, we will get two intervals: the traditional interval $Z$ that contains the actual range $\tilde{z}$ of the corresponding intermediate quantity $z$, and the new interval $z$ that is contained in $\tilde{z}$.

**Historical comment.** Such pairs were proposed in [2, 3, 4] under the name of a twin (for recent applications of twins, see, e.g., [8, 11]). A similar notion of “uncertainty of systematic uncertainty” has also been proposed in [6] in slightly different terms, but, as shown in Artbauer [1], it is essentially a twin. A similar idea of describing uncertainty by two intervals was proposed in [9, 10].

Initially, we have $X_1 = x_1 = \tilde{x}_1$. When we go to the next computation step, we want to compute a similar pair. Let us describe this idea formally.

## 2 Definitions, Proposed Method, and the Main Result

In the following text, boldface letters will denote intervals, and $x^-$ and $x^+$ will indicate the lower and upper bounds of an interval $x$.

**Definition 1.** By a twin, we mean a pair $\mathcal{X} = (x, X)$ of intervals for which $x \subseteq X$. (The interval $x$ may be empty.)

**Comments.**

1. In particular, every interval $x$ can be viewed as a twin $(x, x)$.

2. The expression $[a, b]$ denotes the set $\{x \mid a \leq x \leq b\}$. So, if $a > b$, the expression $[a, b]$ will denote an empty set.
3. In the original papers [2, 3, 4], twins were denoted by square brackets: \([x, X]\).
We decided to use parentheses instead; the reason for these new notations is that we are using both twins and intervals, and we want to avoid (as much as possible) any confusion between twins and intervals.

**Definition 2.** Let \(n\) twins \(X_i = (x_i, X_i)\), \(1 \leq i \leq n\), be given.

- We say that a set \(X \subseteq X_1 \times \ldots \times X_n\) is possible if \(X\) is connected and for each \(i\), its projection \(\pi_i(X) = \{x_i | (x_1, \ldots, x_i, \ldots, x_n) \in X\}\) on \(i\)-th axis satisfies the property \(x_i \subseteq \pi_i(X) \subseteq X_i\).

- Let \(g(x_1, \ldots, x_n)\) be a continuous function. We define \(g(X_1, \ldots, X_n)\) as a twin \(G = (g, G)\), where \(G = g(X_1, \ldots, X_n)\), and \(g\) is an intersection of the sets \(g(X)\) for all possible sets \(X\).

**Main Theorem.** Let \(f(x_1, \ldots, x_n)\) be the result of a sequence \(g^{(1)}, g^{(2)}, \ldots, g^{(N)}\) of elementary operations \(+, -, *, /, \) and let intervals \(\hat{x}_1, \ldots, \hat{x}_n\) be given. Then, if we start with \(n\) twins \(X_i = (\hat{x}_i, \hat{x}_i)\), and follow the same sequence of operations \(g^{(1)}, g^{(2)}, \ldots, g^{(N)}\) on twins, then, at the end, we get a twin \(Y = (y, \overline{y})\) for which \(y \subseteq f(\hat{x}_1, \ldots, \hat{x}_n) \subseteq \overline{Y}\).

**Proof:** similar to standard interval computations, by induction over the total number of computation steps.

## 3 Computations

To apply our main idea, we must be able to compute twins corresponding to basic arithmetic operations. In contrast to naive interval computations, finding \(g_r\) requires minimization over many sets \(X_i\) and is, therefore, not very straightforward. We have found the explicit expressions for arithmetic operations \(g(x_1, x_2)\). Let us first consider the case when the twins are just intervals:

**Proposition 1.**

- When \(g\) is increasing in both variables (e.g., if \(g = +\), or if \(g = *\) and both arguments are positive), then:
  \[ g([x_1^-, x_1^+], [x_2^-, x_2^+]) = [\min(g(x_1^-, x_2^-), g(x_1^+, x_2^-)), \max(g(x_1^-, x_2^+), g(x_1^+, x_2^+))]. \]

- When \(g\) is increasing in \(x_1\) and decreasing in \(x_2\) (e.g., for \(g = -\)):
  \[ g([x_1^-, x_1^+], [x_2^-, x_2^+]) = [\min(g(x_1^-, x_2^-), g(x_1^+, x_2^+)), \max(g(x_1^-, x_2^+), g(x_1^+, x_2^-))]. \]

- When \(g = *\), \(0 \notin [x_1^-, x_1^+]\), and \(0 \in [x_2^-, x_2^+]\), then \(g([x_1, x_1], [x_2, x_2]) = [x_1^* x_2^-, x_1^- x_2^+]. \)

- When \(g = *\), \(0 \in x_1\), and \(0 \in x_2\), then \(g([x_1, x_1], [x_2, x_2]) = \{0\}. \)
Comment. The resulting operations are not new; the exact same operations appear in Kaucher arithmetic proposed (for somewhat different reasons) in [5]: Namely, they coincide with operations between so-called proper and improper intervals in this arithmetic (see also [2, 3, 4]). To avoid misunderstanding, we must point out that Kaucher arithmetic is more general than these formulas; it describes three possible cases:

- operations between proper intervals; these operations are identical with traditional interval arithmetic;
- operations between improper intervals; these operations are equivalent to operations of traditional interval arithmetic;
- operations between proper and improper intervals; these operations are radically different from traditional interval arithmetic.

Similar formulas were later proposed and analyzed by S. Markov (see, e.g., [7]) as “arithmetic of directed intervals”.

In this paper, we propose a new justification for the new (radically different) formulas of Kaucher arithmetic. From the viewpoint of the above-formulated problem, both Kaucher and Markov proposed heuristic methods that “underestimate” the desired interval but do not prove that the resulting estimates are the best that we can get. Proposition 1 gives a mathematical proof of these estimates being the best; this proof is not so easy as the proofs of many algebraic results about interval arithmetic because we have to consider all possible connected sets.

Proof. Let us first consider the case when $g$ is increasing in both variables. W.l.o.g., we can assume that $g(x_1^-, x_2^+) \leq g(x_1^+, x_2^-)$. We will first prove that if $X$ is possible, then $g(X) \supseteq [g(x_1^-, x_2^+), g(x_1^+, x_2^-)]$. Indeed, since $X$ is possible, and the twin is an interval, $\pi_1(X) = [x_1^-, x_1^+]$; hence, there exists a point $x \in X$ for which $\pi_1(x) = x_1^-$. In other words, $(x_1^-, x_2) \in X$ for some $x_2 \in [x_2^-, x_2^+]$. Hence, $g(x_1^-, x_2) \in g(X)$. But $g$ is increasing, so, $g(X) \ni g(x_1^-, x_2) \leq g(x_1^-, x_2^+)$. Similarly, for some $x_2^+, g(x_1^+, x_2^-) \leq g(x_1^+, x_2^+) \in g(X)$. Since $X$ is connected and $g$ is continuous, the set $g(X)$ is also connected. Therefore, $g(X)$ contains the entire interval $[g(x_1^-, x_2^+), g(x_1^+, x_2^-)]$.

To show that the intersection of all $g(X)$ is exactly $[g(x_1^-, x_2^+), g(x_1^+, x_2^-)]$, we produce a possible set $X$ for which $g(X)$ is exactly this interval: $X = g^{-1}([g(x_1^-, x_2^+), g(x_1^+, x_2^-)])$. That this set is connected follows from the fact that $g(x_1, x_2)$ is continuous and monotonic in both variables. (For $g = +$ and $g = *$, this conclusion can be also obtained in a very straightforward manner.)

The proof for the case when $g$ is increasing in $x_1$ and decreasing in $x_2$ is similar.

Let us now consider the case when $g = *$, $0 \not\in x_1$, and $0 \in x_2$. In this case, for $X = (x_1^+ \times x_2^+) \cup (x_1^- \times \{0\})$, $g(X)$ is exactly the desired interval. Vice versa, if $X$ is possible, then, die to $s_2(X) = x_2$, we have $(x_1, x_2) \in X$ for some $x_1 \in [x_1^-, x_1^+]$. Hence, $g(X) \ni x_1 \ast x_2^+ \geq x_1^- \ast x_2^+$. Similarly, for some $x_1^+$, we have $g(X) \ni x_1^+ \ast x_2^- \leq x_1^+ \ast x_2^-$. Since the set $g(X)$ is connected, it thus contains the desired interval $[x_1^- \ast x_2^+, x_1^+ \ast x_2^-]$. 
Finally, let us show that when \( g = *, 0 \in x_1 \), and \( 0 \in x_2 \), then \( G = \{0\} \). Indeed, if we take as \( X \) all points from the box \( x_1 \times x_2 \) for which \( x_1 * x_2 \geq 0 \), we get a possible set with \( g(X) \subseteq [0, \infty) \). The set \( X' \) of all points from this box for which \( x_1 * x_2 \leq 0 \) is also possible, and \( g(X) \subseteq (-\infty, 0] \). The intersection of these two sets is \( \{0\} \), so, \( G \) (the intersection of all such sets \( g(X) \)) is contained in \( \{0\} \). To complete the proof, it is sufficient to show that \( 0 \in g(X) \) for all possible sets \( X \). Indeed, since \( \pi_1(x) = x_1 \geq 0 \), we have \( (0, x_2) \in X \) for some \( x_2 \), hence, \( 0 = 0 * x_2 \in g(X) \). Q.E.D.

**Example.** For \( x - x^2 \), we have \( X_1 = X_2 = [0, 1] \) and hence, \( G = [\min(0 - 0, 1 - 1), \max(0 - 0, 1 - 1)] = \{0\} \). So, \( Y = \{0, -1, 1\} \).

**General case: idea behind the computations.** The general case easily follows from the case when twins are intervals, if we take into consideration that \( X \) is possible for \( n \) twins \( (x_i, X_i) \) iff \( X \) is possible for some intervals \( \hat{x}_i \) for which \( x_i \subseteq \hat{x}_i \subseteq X_i \). Therefore, to find the intersection \( g \) of images \( g(X) \) for all possible \( X \), it is sufficient to find such intersections (i.e., to find \( g((\hat{x}_1, \hat{x}_1), \ldots) \)) for all \( \hat{x}_i \), and then, to take the intersection of the resulting intersections. For the cases when we know the explicit expressions for \( g_i(\chi_i, \ldots) \) for interval \( \chi_i \), we can thus get explicit expressions for the general case:

**PROPOSITION 2.**

- When \( g \) is increasing in both variables (e.g., if \( g = + \), or if \( g = * \) and both arguments are positive), then
  \[
g_i(\chi_1, \chi_2) = [\min(g(x_1^-, x_2^+), g(x_1^+, x_2^-)), \max(g(x_1^-, x_2^+), g(x_1^+, x_2^-))].
\]

- When \( g \) is increasing in \( x_1 \) and decreasing in \( x_2 \) (e.g., for \( g = - \)):
  \[
g_i(\chi_1, \chi_2) = [\min(g(x_1^-, x_2^-), g(x_1^+, x_2^+)), \max(g(x_1^-, x_2^-), g(x_1^+, x_2^+))].
\]

**Comments.**

1. A (reasonably clumsy) explicit expression can also be written for \( * \) for the case when intervals are not necessarily all positive or all negative (so that \( * \) is not monotonic).

2. The reader should be cautioned that the resulting operations are, in general, **different** from the operations of twin arithmetic proposed (on purely algebraic grounds) in [2, 3, 4]. This difference is in line with the fact (mentioned after Proposition 1) that when twins are intervals, our formulas coincide with only a particular case of Kaucer arithmetic.

**Recommendations.** Our numerical experiments has shown that this method gives the best (= closest to \( \hat{y} \)) underestimates \( y \) when one of the input intervals \( \hat{x}_i \) is much wider than the others (i.e., e.g., if one of the measurements that lead to \( \hat{x}_i \) was much less accurate than the others).
Acknowledgments. This work was partially supported by NSF Grants No. CDA-9015006 and EEC-9322370, by NASA Grant No. NAG 9-757, and by the German Science Foundation. The authors are greatly thankful to Svetoslav Markov and Sergey Shary for valuable discussions, and to the anonymous referee for important suggestions.

References


Addresses:

V. KREINOVIČ, Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, USA, E-mail: vladik@cs.utep.edu.

V. M. NESTEROV, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 14-th line, 39, V.O., St.Petersburg, 199178, Russia, E-mail: nest@nit.spb.su.

N. A. ZHELUDAEVA, Chapygina 5–50, St. Petersburg 22, 197022, Russia.