ON AN INTERVAL ARITHMETIC AND ITS APPLICATIONS

Svetoslav Markov

University of Sofia / Bulgaria

Abstract. It is our point of view that familiar interval arithmetic defined by $A*B=\{a * b:\ a \in A,\ b \in B\}, * \in \{+,-,\times,:\}$ is inefficient in certain respects. For instance, it is not in a position to produce exact representations of sets of the form $\{f(x,y,\dots,z): x \in X, y \in Y, \dots, z \in Z\}$ even for simple functions f of one variable. We make use of another interval arithmetic which is very convenient for computer computations and for construction of interval algorithms. As an example we consider a method for the construction of interval expressions for sets of the form $\{f(x): x \in [x_1, x_2]\}$, where f is an elementary function.

1. Introduction

Interval arithmetic is becoming rapidly popular as a perspective weapon against round-off errors. By interval arithmetic we mean an algebraic system consisting of the set IR of all closed intervals on the real line R together with some binary operations on IR. The idea of this mathematical tool is as follows. Instead of the usual numerical algorithms for solving certain mathematical problem we use interval algorithms based upon interval arithmetic. Such interval algorithms are designed to produce as final results intervals (instead of numbers) which include the exact solution of the original problem. This idea leads to the following two requirements with respect to interval algorithms: 1) the interval algorithm should be reliable in the sense that such an inclusion of the exact solution must indeed take place in an arbitrary situation; 2) the interval algorithm should be good in the sensethat the resulting interval bounds must be as sharp as possible.

There are still some difficulties related to the realization of these two requirements which prevent the wide utilization of interval arithmetic and interval algorithms. The first requirement creates difficulties with the implementation of interval arithmetic on computers since the widly-available hardware and software systems do not provide the necessary downwardly (or upwardly) directed roundings. The second requirement leads, in our opinion, to the need of a sufficiently rich interval arithmetic structure, and this is the topic we want to discuss in some detail.

Familiar interval arithmetic [3] is quite inefficient in certain situations. Thus this arithmetic is not able to produce exact interval representations of sets of the form $\{f(x,y,\dots): x \in [x_1, x_2] = X, y \in Y, \dots$ even for simple monotonic functions of one variable (consider the simple example $\{2x-x^2: x \in X\}$, where $X=[x_1,x_2] \subset [0,1]$). In this paper we present an interval arithmetic, which is very simple and more efficient than the familiar one in certain such situations.

2. An interval arithmetic

We shall denote the elements of IR, that is the compact intervals on R, by capital Roman letters A,B,C,... . Real numbers will be denoted by a,b,c,... or by

For the end-points of an interval A we use the following two types of notations:
i) a_1 and a_2 are the end-points of A such that $a_1 \le a_2$;
ii) a_μ and a_ν are the end-points of A such that $|a_\mu| < |a_\nu|$.

Remark. Note that the numbers a_μ and a_ν are not defined in the situation when A is a symmetric interval of the form $A=[-a, a]$, $a>0$.

By $\alpha \vee \beta$ or $[\alpha \vee \beta]$ we shall mean an interval A with end-points α and β. This notation is very useful in situations when α and β are some variable quantities for which it is not known whether $\alpha \le \beta$ or $\alpha > \beta$ takes place, so that we are not in a position to decide which one of the equalities $A=[\alpha,\beta]$ or $A=[\beta,\alpha]$ should be written. This useful operation "$\vee$", called join, can be extended for intervals A,B as follows: $A \vee B = [\min\{a_1,b_1\}, \max\{a_2,b_2\}]$.

CH1630-3/81/0000/0274$00.75 © 1981 IEEE

The width of A is denoted by $w(A) = a_2-a_1$. The number $|A|=\max\{|a_1|, |a_2|\}$ is is called the absolute value of A.

We are now in position to introduce the following interval arithmetic on IR:

$$\begin{aligned} A+B &= [a_1+b_1 \vee a_2+b_2], \\ A-B &= [a_1-b_1 \vee a_2-b_2], \\ A\cdot B &= [a_\mu b_\mu \vee a_\nu b_\nu], \\ A/B &= [a_\mu/b_\mu \vee a_\nu/b_\nu]. \end{aligned} \tag{1}$$

Consider these operations in some detail. The sum of A and B is the interval with end-points a_1-b_1 and a_2-b_2, that is the interval $[a_1-b_1, a_2-b_2]$. In the other three formulas we can not say which end-point is the left one and which is the right one for arbitrary A and B. Thus, the difference A-B is the interval with end-points a_1-b_1 and a_2-b_2, that is, $A-B = [a_1-b_1, a_2-b_2]$ in the situation when $w(A) \geq w(B)$ and $A-B=[a_2-b_2, a_1-b_1]$ in the situation $w(A) < w(B)$.

The product AB and the quotient A/B of two intervals are not defined in the situation when some of the intervals involved are symmetric, since then the numbers $a_\mu, b_\mu, a_\nu, b_\nu$ are not defined. However, we shall extend definition (1) for the case when the interval A is a degenerate interval of zero width, $w(A)=0$, and B is a symmetric interval. Since the set of degenerate intervals $A=[\alpha, \alpha]$ is isomorphic to the reals we shall make no distinction between such intervals and the real numbers, writing thus $A=[\alpha,\alpha]=\alpha$. We now add to (1) the following definition: if B is a symmetric interval, that is, $b_1=-b_2$, then

$$\begin{aligned} \alpha B &= [\alpha b_1 \vee \alpha b_2], \\ \alpha/B &= [\alpha/b_1 \vee \alpha/b_2]. \end{aligned} \tag{1'}$$

Obviously the interval A+(-B) is not equal to A-B, in general, which means that our arithmetic is not linear. Nevertheless the interval arithmetic (IA) defined by (1)-(1') possesses a simple algebraic structure, which has been already studied in some detail.

In addition, we introduce several other convenient notations. The interval $(-1)A=[-1,-1]A=[-a_2,-a_1]$ will be briefly denoted by -A. Further, we set: $A\ominus B=A+(-B)=[a_1-b_2, a_2-b_1]$, $A\oplus B=A-(-B)=[a_1+b_2 \vee a_2+b_1]$, $A\odot B=A(1/B)=[a_\mu/b_\nu \vee a_\nu/b_\mu]$, $A\otimes B=A/(1/B)=[a_\mu b_\nu \vee a_\nu b_\mu]$; $A^0=1$, $A^n=\underbrace{A\cdot A\cdot \ldots \cdot A}_{n \text{ times}}$.

3. A comparisson with familiar interval arithmetic

This section might be of some interest only for readers acquainted with well-known interval interval arithmetic [3]. We shall make now a comparisson between our interval arithmetic (IA) defined by (1)-(1') and the familiar interval arithmetic (FIA) defined by $A*B=\{a*b: a\in A, b\in B\}$, $*\in\{+,-,\times,:\}$, that is:

$$\begin{aligned} A+B &= \{a+b:\ a\in A,\ b\in B\}, \\ A-B &= \{a-b:\ a\in A,\ b\in B\}, \\ AB &= \{ab:\ a\in A,\ b\in B\}, \\ A/B &= \{a/b:\ a\in A,\ b\in B\}. \end{aligned} \tag{FIA}$$

The first operation in (FIA) is the same as the first operation in (IA). The second operation in (FIA) is a composite operation, $A-B=A+(-B)$, and thus is of no interest. The third operation in (FIA):

$$AB=\{ab:\ a\in A,\ b\in B\} \tag{FIA3}$$

coincides with the third operation in (IA) in the situation when neither A nor B contain zero. The operation (FIA3) possesses a serious insufficiency: it is rather complicated when presented in terms of the end-points of the operands (such presentation can be found in [3], p.12). We note that end-point presentations are important for the calculations. The third operation in (IA) has the advantage of being simply representable by means of end-points and, therefore more efficient for computation (especially via computers). The only advantage of (FIA3) is that it gives a simple expression for the set $\{ab:\ a\in A, b\in B\}$ for all possible A and B. In (IA) it is true that $AB=\{ab: a\in A, b\in B\}$ if A and B does not contain zero. The general formula, valid for arbitrary A,B, is a bit more complicated. However, in practice we usually work with small intervals and the situation when an interval contains zero is comparatively rare. This says that the mentioned advantage of (FIA3) is minor. There are many other cases when the third operation in (IA) is more convenient.

With regard to the fourth operation in (FIA) it should be noted that in (FIA) we have $A/B=A(1/B)$, showing that if 1/B is defined then A/B is a composite operation (multiplication of A by 1/B). The operation A/B in (FIA) is almost the same as the composite operation $A\odot B=A(1/B)$ in (IA). Indeed, in (IA) it is true that $\{a/b: a\in A, b\in B\}=A\odot B$ in the situation $A\not\ni 0, B\not\ni 0$.

Finally we note that (FIA) is based upon only two "and a half" operations: A+B, AB and 1/B, whereas (IA) is based on four independent operations. The power of these four operations is demonstrated in the next section. At any case the structure of (IA) is richer then the structure of (FIA) which will become clear from the presented appli- cations. As an application we shall con- der in section 6 some interval algor for computation of sets of the form

$\{f(x):x \in X\}$ where f is an elementary function. We shall first consider some basic theorems which allow us to obtain exact interval representations of sets of the form $\{f(x)*g(x):x \in X\}$, $* \in \{+,-,\times,:\}$, by the assumption that the sets $\{f(x):x \in X\}$ and $\{g(x):x \in X\}$ are known.

4. Construction of interval expressions for ranges of monotone functions

Assume that f is continuous and monotone (c.m.) on $D \in IR$. For any $X=[x_1,x_2] \subset D$ the set $\{f(x):x \in X\}$ is an interval which is easy to compute: $\{f(x):x \in X\}=[f(x_1) \vee f(x_2)]$; this interval will be further denoted by $F(X)$. We thus obtain for every c.m. on D function f an interval function F on interval argument, that is a mapping $F:D \to IR$.

Consider two functions f,g which are c.m. on D. We shall distinguish between the following two situations: i) both functions f,g are monotone increasing on D or both are monotone decreasing on D; we shall then say that f and g are equimonotone (e.m.) on D; ii) one of the functions f,g is monotone increasing on D and the other is monotone decreasing on D; in this case we say that f and g are differently monotone on D (d.m. on D).

Our applications given in section 6 are based on the following propositions[2]:

Proposition 1. The functions f,g and f+g are c.m. on D. Then for every interval $X \subset D$

$$\{f(x)+g(x):x \in X\}=\begin{cases} F(X)+G(X), & \text{if f and g are e.m. on D,} \\ F(X) \oplus G(X), & \text{if f and g are d.m. on D.} \end{cases}$$

Proposition 2. The functions f,g and f-g are c.m. on D. Then for every interval $X \subset D$

$$\{f(x)-g(x):x \in X\}=\begin{cases} F(X)-G(X), & \text{if f and g are e.m. on D,} \\ F(X) \ominus G(X), & \text{if f and g are d.m. on D.} \end{cases}$$

Proposition 3. The functions f and g are such that $|f|,|g|$ and fg are c.m. on D. Then for every $X \subset D$

$$\{f(x)g(x):x \in X\}=\begin{cases} F(X) \cdot G(X), & \text{if f and g are e.m. on D,} \\ F(X) \otimes G(X), & \text{if f and g are d.m. on D.} \end{cases}$$

Proposition 4. The functions f and g are such that $|f|,|g|$ and f/g are c.m. on D. Then for every X D

$$\{f(x)/g(x):x \in X\}=\begin{cases} F(X)/G(X), & \text{if f and g are e.m. on D,} \\ F(X):G(X), & \text{if f and g are d.m. on D.} \end{cases}$$

Propositions 1-4 can be used for definition of our interval operations, analogous to the definition $X*Y=\{x*y:x \in X, y \in Y\}$ in (FIA). Indeed, consider the following definition of addition of intervals: given two intervals X,Y, choose an interval T and two continuous and e.m. on T functions $x=x(t)$, $y=y(t)$, such that $x(T)=X$, $y(T)=Y$, that is, $\{x(t):t \in T\}=X$, $\{y(t):t \in T\}=Y$. Then we have $X+Y=\{x(t)+y(t):t \in T\}$.

Subtraction of X,Y can be defined analogously as follows. Take an interval T and two continuous and e.m. on T functions $x=x(t)$ and $y=y(t)$, such that $x(t)-y(t)$ is monotone on T, and $x(T)=X, y(T)=Y$. Then we have $X-Y=\{x(t)-y(t): t \in T\}$.

Similarly, the other operations in (IA) can be defined.

5. Interval series

The concept of an interval series can be introduced as follows.

Definition. If we are given an interval sequence $\{A_n\}_{n=1}^{\infty}$, then the interval sequence $\{S_n\}_{n=1}^{\infty}$, with $S_1=A_1$, $S_{n+1}=S_n \pm A_{n+1}$ is called an interval series $A_1 \pm A_2 \pm A_3 \pm \dots$.

When S_n converges to the interval S (in the sense of the norm $|\cdot|$) we say that the series $A_1 \pm A_2 \pm \dots$ converges and its sum is S.

Examples:

$$1+A \oplus A^2/2!+A^3/3! \oplus A^4/4!+\dots;$$
$$B-B^2/2+B^3/3-B^4/4+\dots;$$
$$1 \ominus C^2/2! \oplus C^4/4! \ominus C^6/6! \oplus \dots .$$

It will be shown below that the sums of the above series are corresspondingly equal to the intervals

$\{e^x:x \in A\}$, for $A \subset [-1,0]$,
$\{\ln(1+x):x \in B\}$, for $B \subset [0,1]$, and
$\{\cos x:x \in C\}$, for $C \subset [0, \pi/4]$.

Note that the interval arithmetic operation σ_n between the n-th partial sum S_n and the term A_{n+1} can be either + or -, $\sigma_n \in \{+,-\}$ (and hence $\sigma_n \in \{+,-,\oplus,\ominus\}$). Thus an interval series is determined by means of the sequences of terms $\{A_n\}_{n=1}^{\infty}$ and the sequence of signs $\{\sigma_n\}_{n=1}^{\infty}$, $\sigma_n = \pm$.

In what follows we shall obtain interval series expressions for sets of the form $\{\sum_{n=1}^{\infty} a_n x^n: x \in X\}$. We note first that the equality $\{x^{2k=1}: x \in X\} = X^{2k+1}$, $k=0,1,\dots$, holds true for any $X \in IR$, whereas the equality $\{x^{2k}:x \in X\} = X^{2k}$ is true if $X \geq 0$.

6. Interval computation of elementary functions

Consider the interval exponential function $e^X=\{e^x:x \in X\}$. It is well-known that for $X \geq 0$ we have

$$e^X = 1 + X/1! + X^2/2! + X^3/3! + \ldots \quad (2)$$

Remark. The practical significans of expressions like (2) is straightforward. If we truncate them up to the n-th summand and add the corresponding truncation error interval term, then we obtain a formula which can be used for the numerical computation of the corresponding function f (in this case $f(x)=e^x$) with an a posteriori control over the total computational error in $f(x)$, that is: inherent error in the argument + truncation error + roundoff error (providing that the necessary directed roundings are performed). Such formulas provide us with simple algorithms for the computation of the elementary functions for interval software (cf.[5]). This remark applies for any of the interval series expressions below.

Formula (2) is not valid for $X \leq 0$. Moreover, as is shown in [4], it does not produce sharp bounds for the value of exp X and therefore is of little use for the practical computation of exp x for $x < 0$.

We shall next obtain exact interval expressions for exp X for nonpositive intervals $X \leq 0$. First, we shall give a formula for $-1 \leq X \leq 1$. We note the fact that the partial sums $\sum_{n=0}^{\infty} x^n/n!$ are monotone increasing or decreasing on $[-1,0]$ according to whether n is odd or even. Applying consecutively Proposition 1 we obtain:

$$e^X = 1 + X \oplus X^2/2! + X^3/3! \oplus X^4/4! + \ldots, \quad -1 \leq X \leq 0.$$

We shall now obtain a formula for e^X for arbitrary $X \leq 0$. To this end we rearrange the terms in the Taylor series for exp x as follows:

$$e^x = x + 1 + x^3/3! + x^2/2! + x^5/5! + x^4/4! + x^7/7! + \ldots,$$

and notice that all partial sums in the above series are monotone increasing functions on $(-\infty,0]$. Then, by means of Proposition 1 we obtain for $X \leq 0$:

$$e^X = X \oplus 1 + X^3/3! \oplus X^2/2! + X^5/5! \oplus X^4/4! + X^7/7! \oplus \ldots \quad (3)$$

Remark. Formulas (2) and (3) can be used for the computation of exp X for any X. Indeed, if $X=[x_1,x_2] \ni 0$ we set $X = X' \vee X'' = [x_1,0] \vee [0,x_2]$ and then have exp X $=(\exp X') \vee (\exp X'')$ which can be computed by means of (2) and (3). Similarly, any other formula of the form $F(X)=\{G(X), X \leq a; H(x), X \geq a$ can be used for the computation of $F(X)$ for $X \ni a$.

Consider now the logarithmic function $\ln(1+x)$. For $x \in (-1,1]$ we have $\ln(1+x) = x - x^2/2 + x^3/3 - x^4/4 + \ldots$. Using the fact that all partial sums are monotone decreasing on $(-1,1]$ we obtain for the range $\{\ln(1+x) : x \in X\} = \ln(1+X)$:

$$\ln(1+X) = \begin{cases} X - X^2/2 + X^3/3 - X^4/4 + \ldots, & \text{for } 0 \leq X \leq 1, \\ X \ominus X^2/2 + X^3/3 \ominus X^4/4 + \ldots, & \text{for } -1 < X \leq 0. \end{cases}$$

Let us present some interval expressions for the sine and cosine functions. The argument X can be reduced suitably in advance; it is sufficient to obtain interval formulas that are valid for $0 \leq X \leq \pi/4$.

Consider the cosine function $\cos x = \sum_{k=0}^{\infty} (-1)^k x^{2k}/(2k)!$. Noticing that all partial sums are monotone decreasing functions on $[0,\sqrt{6}]$ and using Propositions 1 and 2, we obtain for the range $\cos X = \{\cos x : x \in X\}$:

$$\cos X = 1 \ominus X^2/2! \oplus X^4/4! \ominus X^6/6! \oplus X^8/8! - \ldots \quad (4)$$

for $X \subset [0,\sqrt{6}]$. Since $\sqrt{6} > \pi/4$ formula (4) can be used for the computation of cos x for any X.

The fact that the functions $\sum_{k=0}^{n} (-1)^k x^{2k}/(2k)!$ are monotone increasing on $[-\sqrt{6},0]$ for every n leads us to the conclusion that (4) holds true for $-\sqrt{6} \leq X \leq 0$ as well.

Consider now the computation of $\sin X = \{\sin x : x \in X\}$. Using the fact, that the partial sums in the Taylor expansion of sin x are monotone increasing functions on $[-\sqrt{2},\sqrt{2}]$ for every fixed number n of the summands, we obtain

$$\sin X = X - X^3/3! + X^5/5! - X^7/7! + X^9/9! - \ldots$$

for $-\sqrt{2} \leq X \leq \sqrt{2}$.

Using similar arguments for $(1+X)^a = \{(1+x)^a : x \in X\}$, when $0 < a < 1$ and $-1 \leq X \leq 1$, we obtain

$$(1+X)^a = \begin{cases} 1 + aX - \frac{a(a-1)}{2} X^2 + \ldots, & 0 \leq X \leq 1, \\ 1 + aX \ominus \frac{a(a-1)}{2} X^2 + \ldots, & -1 \leq X \leq 0. \end{cases}$$

For similar considerations for functions of many variables see [1].

REFERENCES

1. Dimitrova, N., S. Markov, Über die intervallarithmetische Berechnung des Wertebereichs einer Function mit Anwendungen, Freiburger Intervall-Berichte, 1981 (to appear).

2. Markov, S, Some applications of extended interval arithmetic to interval iterations, Computing Suppl. 2, 69-84 (1980).

3. Moore, R., Methods and applications of interval analysis, SIAM, 1979.

4. Ris, F.N., Tools for the analysis of interval arithmetic, in: Interval mathematics (ed. K. Nickel), Springer, 1975, 75-98.

5. Yohe, J.M., Portable software fot interval arithmetic, Computing Suppl. 2, 211-229 (1980).

REFEREES

D.P. Aggrawal
D.E. Atkins
D. Banerji
J. Bannister
J. Barlow
K. Bowyer
I.N. Chen
C.K. Chow
P. Chow
J. Dallon
J. Demmel
J. Ellis
M.D. Ercegovac
D. Fisher
R. Geist
R.T. Gregory
V.C. Hamacher
E. Hansen
T.E. Hull
K. Hwang
M.J. Irwin
H. Kobayashi
I. Koren
P. Kornerup
D.J. Kuck
M. Leuze
P. Ligomenides
M. Mace
G. Mago
F. Murray
R. Newcomb
R. Owens
C.H. Papachristou
D. Ramm
T.M. Rao
T.R.N. Rao
J.E. Robertson
J. Rosenberg
J. Rusnak
J. Salz
A. Sameh
G. Sinaki
G.L. Srinivasan
G. Taylor
S. Utku
M. Valtorta
M. Varansi
W. Wallach
R. Wagner
O. Watanuki
R. Willoner

AUTHOR INDEX