1.1 INTERVAL ARITHMETIC

Prior to widespread use of interval arithmetic, bounds on the range of a function were sometimes obtained with Lipschitz constants or moduli of continuity. Such computations can be considered special cases of computation of *interval enclosures*. Judicious use of *interval arithmetic* allows such range bounds to be computed efficiently, without extensive ad hoc analysis. Additionally, interval arithmetic, with *directed roundings*, can provide mathematically rigorous results from floating point operations on computers.

1.1.1 Real Interval Arithmetic

Introduced in its modern form by R. E. Moore [163], real interval arithmetic is based on arithmetic within the set of closed intervals of real numbers. If $\boldsymbol{x} = [\underline{x}, \overline{x}]$ and $\boldsymbol{y} = [\underline{y}, \overline{y}]$, then the four elementary operations for such *idealized interval arithmetic* obey

$$\boldsymbol{x} \operatorname{op} \boldsymbol{y} = \{x \operatorname{op} y \mid x \in \boldsymbol{x} \text{ and } y \in \boldsymbol{y}\} \quad \text{for op } \in \{+, -, \times, \div\} \tag{1.4}$$

Thus, the image of each of the four basic interval operations is the *exact range* of the corresponding real operation. Although Equation (1.4) characterizes these operations mathematically, interval arithmetic's usefulness is due to the *operational definitions*. For example,

$$\begin{aligned}
\boldsymbol{x} + \boldsymbol{y} &= [\underline{x} + \underline{y}, \overline{x} + \overline{y}], && (1.5)\\
\boldsymbol{x} - \boldsymbol{y} &= [\underline{x} - \overline{y}, \overline{x} - \underline{y}], && (1.6)\\
\boldsymbol{x} \times \boldsymbol{y} &= [\min\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}] && (1.7)\\
\frac{1}{\boldsymbol{x}} &= [1/\overline{x}, 1/\underline{x}] \qquad \text{if } \underline{x} > 0 \text{ or } \overline{x} < 0 && (1.8)\\
\boldsymbol{x} \div \boldsymbol{y} &= \boldsymbol{x} \times 1/\boldsymbol{y} && (1.9)
\end{aligned}$$

Here, reciprocation, and hence ordinary interval division $\boldsymbol{x}/\boldsymbol{y}$ is undefined when $0 \in \boldsymbol{y}$. However, under certain circumstances (as in §1.1.4 and §1.5.1 below), it makes sense to define such quotients in an *extended* or *Kahan–Novoa–Ratz arithmetic*. Also, when multiplication is actually implemented on a computer, a somewhat more complicated alternative to Equation (1.7), such as [165, (2.20), p. 12], is often used, since it is faster on average.

Although the ranges of interval arithmetic operations are exactly the ranges of the corresponding real operations, this is not so if the operations are composed. For example, if

$$f(x) = x^2 - x,$$

then evaluating f over $\boldsymbol{x} = [0, 1]$ with interval arithmetic gives

$$[0, 1]^2 + [0, 1] = [0, 1] - [0, 1] = [-1, 1],$$

but the range of f over $[0, 1]$ is $[-1/4, 0]$; thus, the interval computation *overestimates* the range. This is due to the fact that the variable x is implicitly assumed to vary independently in the term x^2 and the term $-x$, so that e.g. numbers such as $1^2 - 0$, in addition to $0^2 - 0$ and $1^2 - 1$, are included in the interval result. This phenomenon is termed *interval dependency.* Due to interval dependency, algebraic expressions that are equivalent in real arithmetic give different results when evaluated in interval arithmetic. For example, if f above is written as $x(x - 1)$, then

$$[0, 1]\big([0, 1] - 1\big) = [0, 1][-1, 0] = [-1, 0] \neq [-1, 1].$$

In this case, the expression $x(1 - x)$, obtained using Horner's method, gives a sharper bound on the range than the power representation, although that is not always the case[1]. Theorem 1.4 in §1.1.7 below states an additional property of interval dependency. Also, Theorem 1.3 below asserts that *evaluating a function with interval arithmetic always gives bounds on the range of the function.* Here, we can state

Theorem 1.1 *Interval arithmetic is* subdistributive *in the sense that, if $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are intervals, then*

$$\boldsymbol{x}(\boldsymbol{y} + \boldsymbol{z}) \subseteq \boldsymbol{x}\boldsymbol{y} + \boldsymbol{x}\boldsymbol{z}.$$

Thus, although addition or multiplication of intervals is commutative and associative, the distributive laws do not hold. Furthermore, although there is an additive identity $[0, 0]$ and a multiplicative identity $[1, 1]$, additive and multiplicative inverses do not exist. For instance, the example

$$[1, 2] - [1, 2] = [-1, 3]$$

[1] One study related to this is [192].

illustrates that cancellation does not fully occur in subtraction as defined by Equation (1.4). However, a fifth operation, *cancellation subtraction*, is appropriate in various contexts. Cancellation subtraction is defined by the equation

$$[\underline{x}, \overline{x}] \ominus [\underline{y}, \overline{y}] = [\underline{x} - \underline{y}, \overline{x} - \overline{y}]. \tag{1.10}$$

For example, if interval sums $s_j = \sum_{i=1, i\neq j}^{n} \boldsymbol{x}_i$ are required for each j, the overall sum $s = \sum_{i=1}^{n} \boldsymbol{x}_i$ can first be computed, then the individual s_j can be obtained with cancellation subtraction.

Additional elementary properties of interval arithmetic are listed in [8], [77], [83], [165], and [175].

1.1.2 Notation

Throughout, boldface will denote intervals, lower case will denote scalar quantities, and upper case will denote vectors and matrices. Brackets "$[\cdot]$" will delimit intervals while parentheses "$(\cdot)$" will delimit vectors and matrices. Underscores will denote lower bounds of intervals and overscores will denote upper bounds of intervals. Corresponding lower case letters will denote components of vectors. For example, we may have:

$$\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)^T,$$

where $\boldsymbol{x}_i = [\underline{x}_i, \overline{x}_i]$. If $\boldsymbol{X}$ is an interval vector or matrix, then $\boldsymbol{X} = [\underline{X}, \overline{X}]$, where $\underline{X}$ is the vector or matrix whose components are lower bounds of corresponding components of $\boldsymbol{X}$, and $\overline{X}$ is the vector or matrix whose components are upper bounds of corresponding components of $\boldsymbol{X}$. For example, if $\boldsymbol{X} = ([1, 2], [3, 4])^T$, then $\underline{X} = (1, 3)^T$ and $\overline{X} = (2, 4)^T$.

The symbol $\check{x}$ will denote a representative point, usually in $\boldsymbol{x}$ and often its center. Similarly, $\check{X}$ will denote a representative point for the box $\boldsymbol{X}$. The actual center, or midpoint, of an interval $\boldsymbol{x}$ will be denoted by $\mathrm{m}(\boldsymbol{x})$, and the vector or matrix whose entries are midpoints of the entries of the vector or matrix $\boldsymbol{X}$ will be denoted by $\mathrm{m}(\boldsymbol{X})$. The *magnitude* of an interval is defined as $|\boldsymbol{x}| = \max\{|\underline{x}|, |\overline{x}|\}$. The magnitude of an interval vector or matrix will be interpreted componentwise:

$$|\boldsymbol{X}| = (|\boldsymbol{x}_1|, |\boldsymbol{x}_2|, \ldots, |\boldsymbol{x}_n|)^T,$$

while the norm of an interval vector is defined as $\|\boldsymbol{X}\| = \big\| |\boldsymbol{X}| \big\|_\infty$. The *mignitude* of an interval $\boldsymbol{x}$ will be defined by $<\boldsymbol{x}> = \min_{x \in \boldsymbol{x}} |x|$.

The *width* of an interval $\boldsymbol{x}$ is denoted by $\mathrm{w}(\boldsymbol{x}) = \overline{x} - \underline{x}$. The width of an interval vector $\boldsymbol{X}$, denoted by $\mathrm{w}(\boldsymbol{X})$, is defined componentwise. We use $\mathrm{w}(\boldsymbol{X})$ in the context of $\|\mathrm{w}(\boldsymbol{X})\| = \|\mathrm{w}(\boldsymbol{X})\|_\infty$.

The set of intervals will be denoted $\mathbb{IR}$, the set of n-dimensional interval vectors, also called *boxes*, will be denoted by $\mathbb{IR}^n$, and the set of m by n matrices whose entries are intervals will be denoted by $\mathbb{IR}^{m \times n}$. The set corresponding to $\mathbb{IR}^2$, when identified with rectangles in the complex plane, will be denoted by $\mathbb{IC}$.

Arithmetic operations involving both real numbers and intervals will occur. In these, such as $[1,2] + 1 = [2,3]$, the real number 1 is interpreted as a *thin interval* $[1,1]$. (Thin intervals are simply intervals of width zero.)

The topological interior of a set $\mathbf{D}$ will be denoted by $\mathrm{int}(\mathbf{D})$. Unless otherwise stated, such sets (and in particular, intervals) will be considered to be topologically closed. The boundary of a set $\mathbf{D}$ will be denoted by $\boldsymbol{\partial}\mathbf{D}$.

If f is a function defined over an interval $\boldsymbol{x}$, then $\boldsymbol{f}^u(\boldsymbol{x})$ denotes[2] the range of f over $\boldsymbol{x}$. The ranges of functions $\phi : \mathbb{R}^n \to \mathbb{R}$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ are similarly denoted as $\boldsymbol{\phi}^u$ and $\boldsymbol{F}^u$, respectively. The gradient of ϕ is denoted by $\nabla\phi$, while the generic notation $F'(X)$ represents the m by n Jacobi matrix of a function $F : \mathbb{R}^n \to \mathbb{R}^m$.

Comparison of intervals will be as sets. For example $\boldsymbol{x} < \boldsymbol{y}$ will mean that every element of $\boldsymbol{x}$ is less than every element of $\boldsymbol{y}$. Comparison of interval vectors $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{IR}^n$ will be componentwise, in the sense of a relation as in [133, pp. 129–130]. For example, $\boldsymbol{X} \le \boldsymbol{Y}$ will mean that $\boldsymbol{x}_i \le \boldsymbol{y}_i$ for i between 1 and n, while $\boldsymbol{X} < \boldsymbol{Y}$ will mean $\boldsymbol{x}_i \le \boldsymbol{y}_i$ for i between 1 and n, and $\boldsymbol{x}_i \ne \boldsymbol{y}_i$ for at least one i. The convex hull of two intervals $\boldsymbol{x}$ and $\boldsymbol{y}$ is an interval that will be denoted by $\boldsymbol{x} \,\underline{\cup}\, \boldsymbol{y}$.

Often, sequences of intervals $\boldsymbol{x}_i$, as well as interval vectors and interval matrices, will be considered. This leads to

Definition 1.1 *Let $\{\boldsymbol{x}_i\}_{i=1}^\infty = \{[\underline{x}_i, \overline{x}_i]\}_{i=1}^\infty$ be a sequence of intervals. The sequence is said to* converge *to an interval $\boldsymbol{x} = [\underline{x}, \overline{x}]$ provided $\underline{x}_i \to \underline{x}$ and $\overline{x}_i \to \boldsymbol{x}_i$ in the standard topology on the real line. Similarly, a sequence of intervals is said to converge to a real number r provided $\underline{x}_i \to r$ and $\overline{x}_i \to r$. A sequence of vectors or matrices is said to converge to a vector or matrix if each*

[2] The notation is suggestive of "united" interval extension, a term first used by Moore.

of its components or coefficients converges to the corresponding component or coefficient of the limit vector or matrix.

For example, we may say that a set of interval matrices $\boldsymbol{A}_k$ converges to the identity matrix I.

1.1.3 Rounded Interval Arithmetic

Computations in §1.1.1 illustrated how real interval arithmetic could compute rigorous bounds on the ranges of functions, assuming infinitely precise computations. However, if such an arithmetic is implemented on the computer using, for example, equations (1.5) through (1.9), then the default rounding may result in non-rigor. For example, the range of the expression $[0.123, 0.456] + [0.0116, 0.0214]$ is $[0.1346, 0.4774]$. But this would be rounded to $[0.135, 0.477]$ with three digit decimal arithmetic and rounding to nearest, and $[0.1346, 0.4774] \not\subseteq [0.135, 0.477]$. Nonetheless, with *directed rounding*, such bounds can be computed rigorously. In particular, if instead of rounding to nearest, the lower bound of the interval is rounded down to the largest machine number less than the exact result and the upper bound is rounded up to the smallest machine number greater than the actual result, then the computed interval necessarily contains the exact range. In our example, the result would be $[0.134, 0.475]$, and $[0.1346, 0.4664] \subset [0.134, 0.475]$. This process is called *outward rounding*, and the resulting widening of the intervals is called *roundout error*.

Thus, with directed rounding, a *machine interval arithmetic* can be defined, such that the result of the four elementary operations contains the result that would be obtained with real interval arithmetic. Mathematical development of this concept is found in [147] and other works of Kulisch *et al.* For notational simplicity, a distinction between the space $\mathbb{IR}$ and the corresponding set of machine intervals will not be made explicit, but should be inferred from the context.

The IEEE binary floating point standard [227] prescribes three rounding modes: nearest, round down, and round up. Thus, rounding modes suitable for interval arithmetic are available on a wide variety of machines. However, standardized programming language support for switching of rounding modes is not available, although the Fortran 90 intrinsic NEAREST [4, p. 636] comes close. Transportable implementations of interval arithmetic thus often resort to mixed assembly language programming. Notwithstanding, if the tightest possible bounds on ranges are not required, then rounding down and rounding up can be simulated. The main requirement is that it be known by how many units in the last place[3] (ULP's) the results of the elementary operations $+$, $-$, $\times$ and $\div$ can be in error. (Note that the IEEE standard [227, §4] specifies that all stored digits be correct, which translates to 1/2 ULP accuracy.) The ACM TOMS algorithm[4] INTLIB implements such a *simulated directed rounding*, and provides rigorous interval arithmetic in portable FORTRAN-77 on a wide variety of machines, including some common computers without IEEE arithmetic. See §2.1.2 below.

Computing dot products with the formula $\boldsymbol{X} \circ \boldsymbol{Y} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{y}_i$ and interval addition sometimes results in significant overestimation of the range of $\boldsymbol{X} \circ \boldsymbol{Y}$, especially when the widths of the components of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are small. Because such dot products are important in obtaining *a posteriori* rigorous error bounds to solutions of linear systems of equations, Kulisch *et al.* [147] have proposed considering $\circ$ as a fifth elementary operation, having the same accuracy as $+$, $-$, $\times$ and $\div$. This *maximally accurate dot product* is implemented in the excellent "SC" and "XSC" languages [243] such as FORTRAN-SC [21, 237] (known commercially as ACRITH-XSC), Pascal-SC [188], C-XSC [134], and Pascal-XSC [71]. It enables computation of tight and rigorous bounds to the solutions of very ill-conditioned linear systems. However, the accurate dot product gives less benefit when the widths of the component intervals of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are large in relation to roundoff error, as is the case in most (but not all) of the interval computation in optimization and solution of nonlinear systems.

1.1.4 Kahan–Novoa–Ratz Arithmetic

Although the interval quotient $\boldsymbol{x}/\boldsymbol{y}$ is undefined in ordinary interval arithmetic when $0 \in \boldsymbol{y}$, an extension to interval arithmetic for this case is useful e.g. in eliminating regions in branch and bound algorithms in which critical points cannot exist. The arithmetic on infinite intervals presented here differs from previous arithmetics on infinite intervals, since it corrects inconsistencies [198]. We call this arithmetic *Kahan–Novoa–Ratz arithmetic*, since Kahan first proposed an arithmetic on infinite intervals in the summer school lecture notes [102], since Kahan used extended interval arithmetic effectively in his continued fraction research, and since Novoa [180] and Ratz [198] separately proposed the corrections that make it consistent for use in nonlinear equations and optimization. In Kahan's arithmetic, the set of real intervals $[a, b] \in \mathbb{IR}$ is augmented by the

[3] For example, in a three-digit decimal system, with numbers of the form 0.123×10^4, one unit in the last place is $.001 \times 10^0 = 10^{-3}$.

[4] cf. §2.1

set of complements $]a, b[= [-\infty, a] \cup [b, \infty]$. In Kahan–Novoa–Ratz arithmetic, division of two ordinary intervals $\boldsymbol{x}$ and $\boldsymbol{y}$ with $0 \in \boldsymbol{y}$ is still defined by the principle in 1.4, operationally as follows:

$$\frac{[\underline{x}, \overline{x}]}{[\underline{y}, \overline{y}]} = \begin{cases} \boldsymbol{x}[1/\overline{y}, 1/\underline{y}] & \text{if } 0 \notin \boldsymbol{y} \\ [-\infty, \infty] & \text{if } 0 \in \boldsymbol{x} \text{ and } 0 \in \boldsymbol{y}, \\ [\overline{x}/\underline{y}, \infty] & \overline{x} < 0 \text{ and } \underline{y} < \overline{y} = 0, \\ [-\infty, \overline{x}/\overline{y}] \cup [\overline{x}/\underline{y}, \infty] & \text{if } \overline{x} < 0 \text{ and } \underline{y} < 0 < \overline{y} \\ [-\infty, \overline{x}/\overline{y}] & \text{if } \overline{x} < 0 \text{ and } 0 = \underline{y} < \overline{y}, \\ [-\infty, \underline{x}/\underline{y}] & \text{if } 0 < \underline{x} \text{ and } \underline{y} < \overline{y} = 0. \\ [-\infty, \underline{x}/\underline{y}] \cup [\underline{x}/\overline{y}, \infty] & \text{if } 0 < \underline{x} \text{ and } \underline{y} < 0 < \overline{y} \\ [\underline{x}/\overline{y}, \infty] & \text{if } \underline{x} < 0 \text{ and } 0 = \underline{y} < \overline{y} \\ \emptyset & \text{if } 0 \notin \boldsymbol{x} \text{ and } 0 = \boldsymbol{y}. \end{cases} \tag{1.11}$$

For example, according to Formula (1.11), $[2, 3]/[-3, 4] = [-\infty, -2/3] \cup [1/2, \infty]$, where $[-\infty, -2/3] \cup [1/2, \infty]$ represents the actual range of x/y, $x \in [2, 3]$, $y \in [-3, 4]$.

Based on the lecture notes [102], arithmetic on intervals and their complements based on formulas similar to (1.11) is published, fully developed, in [149]. This arithmetic is useful in continued fraction computations, since the reciprocal of an interval complement[5] is an ordinary interval. However, for most[6] purposes here it suffices merely to use Formula (1.11), then to intersect the result with an ordinary interval, to obtain zero, one or two ordinary intervals; an example occurs in Example 1.4 on page 53 below.

Kahan–Novoa–Ratz arithmetic, defined on intervals and their complements, should not be confused with other extensions of real interval arithmetic, designed for different purposes. For example, *Kaucher arithmetic* is used in [217] to obtain both *inner estimates* and *outer estimates*[7] for the solution sets of linear interval systems. Yet a third structure, Markov arithmetic, purported to be useful in reducing the overestimation of ranges with ordinary interval arithmetic, is described in [157, 158]. Each of these arithmetics is at times called *extended interval arithmetic*.

[5] also termed an *extended interval*

[6] except perhaps for evaluation of some special objective functions

[7] An inner estimate of the range of a function is a set contained within the range, while an outer estimate contains the range.

1.1.5 Complex Interval Arithmetic

It is possible to identify rectangles of complex numbers

$$\boldsymbol{z} = \{x + iy \mid x \in \boldsymbol{x} \text{ and } y \in \boldsymbol{y}\}$$

with boxes in $\mathbb{IR}^2$, and to define interval operations componentwise. Thus, a complex interval $\boldsymbol{z}$ is identified with the interval vector $[\boldsymbol{x}, \boldsymbol{y}]^T \sim \boldsymbol{x} + i\boldsymbol{y}$. So interpreted, complex interval operations become

$$\begin{aligned} \boldsymbol{z}_1 + \boldsymbol{z}_2 &= \boldsymbol{x}_1 + \boldsymbol{x}_2 + i(\boldsymbol{y}_1 + \boldsymbol{y}_2) &(1.12)\\ &= [\underline{x}_1 + \underline{x}_2, \overline{x}_1 + \overline{x}_2] + i[\underline{y}_1 + \underline{y}_2, \overline{y}_1 + \overline{y}_2], &(1.13)\\ \boldsymbol{z}_1 \times \boldsymbol{z}_2 &= \boldsymbol{x}_1\boldsymbol{x}_2 - \boldsymbol{y}_1\boldsymbol{y}_2 + i(\boldsymbol{x}_1\boldsymbol{y}_2 + \boldsymbol{x}_2\boldsymbol{y}_1), &(1.14) \end{aligned}$$

where $\boldsymbol{z}_1 = \boldsymbol{x}_1 + i\boldsymbol{y}_1$ and $\boldsymbol{z}_2 = \boldsymbol{x}_2 + i\boldsymbol{y}_2$, where multiplication of the component intervals is according to Formula (1.7), and where $-$ and $\div$ are defined similarly. The problem is that, although the set $\boldsymbol{x}_1\boldsymbol{x}_2 - \boldsymbol{y}_1\boldsymbol{y}_2 + i(\boldsymbol{x}_1\boldsymbol{y}_2 + \boldsymbol{x}_2\boldsymbol{y}_1)$ corresponds to a rectangle in $\mathbb{R}^2$, the actual range

$$\{(x_1 + iy_1)(x_2 + iy_2) \mid x_1 \in \boldsymbol{x}_1,\ y_1 \in \boldsymbol{y}_1,\ x_2 \in \boldsymbol{x}_2,\ y_2 \in \boldsymbol{y}_2\}$$

is less regularly shaped. Thus, the product $\boldsymbol{z}_1\boldsymbol{z}_2$ does *not* represent the exact range, and there is overestimation even in a single elementary operation. This can be viewed as complex interval dependency arising from the relationship given by the Cauchy–Riemann equations.

An alternate arithmetic can be defined on the set of *disks*, rather than rectangles, in the complex plane. First introduced by Henrici [59], this *circular arithmetic* results in less overestimation in multiplication[8], but may be more difficult to implement. It has been proposed [246] that a combination of both rectangular and circular arithmetic can be used, to reduce overall overestimation. Both rectangular and circular arithmetic are explained in [8, Ch. 5].

Despite the intrinsic overestimation, complex interval arithmetic can be useful in practice. Although, with interval arithmetic, problems should be solved in the real domain if possible (rather than recast in the complex plane), rectangular arithmetic, recommended for its simplicity, can be used effectively. If rectangular arithmetic is not implemented as a data type, either intrinsically or through operator overloading[9], functions $w = f(z)$, $f : \mathbb{IC} \rightarrow \mathbb{IC}$

[8] in the sense of ratio of area of the interval product to the area of the range of the product operation

[9] cf. §1.4.4

may be explicitly rewritten in terms of their components as $F : \mathbb{IR}^2 \rightarrow \mathbb{IR}^2$. For example, to use a multidimensional interval branch and bound method to find all roots of $f(z) = z^2 - 2$ in the complex plane, f could be written as $F(x, y) = [f_1(x, y), f_2(x, y)]^T = [0, 0]^T$, where $f_1(x, y) = x^2 - y^2 - 2$ and $f_2(x, y) = xy$.

1.1.6 Extensions of Standard Functions

An example in §1.1.1 illustrated how interval arithmetic could be used to obtain bounds on the range of a function that could be evaluated as a sequence of the four elementary operations. This section introduces, in elementary terms, some of the fundamental techniques for obtaining the ranges of functions, for libraries to compute interval extensions.

Throughout, the term *standard function* will appear. It will mean a set of functions, such as those specified in the `FORTRAN-77` standard, that are bundled with a particular compiler or software package. Libraries for evaluation of such functions must be supplied by the package provider. Such standard functions generally include exponentials, logarithms, and a selection of trigonometric functions and their inverses.

Definition 1.2 *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function computable as an expression, algorithm or computer program involving the four elementary arithmetic operations, then a* natural interval extension *of f, whose value over an interval $\boldsymbol{x}$ is denoted by $\boldsymbol{f}(\boldsymbol{x})$, is obtained by replacing each occurrence of x by the interval $\boldsymbol{x}$ and by executing all operations according to formulas (1.5) through (1.9).*

It is not hard to show

Theorem 1.2 *If $\boldsymbol{f}$ is any natural interval extension of f, and $\boldsymbol{x} \in \mathbb{IR}$ is contained within the domain of $\boldsymbol{f}$, then $\boldsymbol{f}(\boldsymbol{x})$ contains the range $\boldsymbol{f}^u(\boldsymbol{x})$ of f over $\boldsymbol{x}$.*

This property is crucial to interval computations, so it should be a part of the *definition* of an interval extension.

Definition 1.3 *A function* $\boldsymbol{f} : \mathbb{IR} \to \mathbb{IR}$ *is said to be an* interval extension *of* $f : \mathbb{R} \to \mathbb{R}$ *provided*

$$\{f(x) | x \in \boldsymbol{x}\} \subseteq \boldsymbol{f}(\boldsymbol{x})$$

for all intervals $\boldsymbol{x} \subset \mathbb{IR}$ *within the domain of* $\boldsymbol{f}$. *(Interval extensions are defined similarly for* $F : \mathbb{R}^n \to \mathbb{R}^m$.*)*

In fact, natural interval extensions can be obtained for virtually *any* function that can be represented with a computer program. To this end, sharp bounds on the ranges of the standard functions are desired. Such bounds can be computed from any expansion (rational, Taylor series, etc.) that has an explicit formula for the error term. For example, suppose $\boldsymbol{x} \in \mathbb{IR}$ and $\check{x} \in \boldsymbol{x}$. Then, for any $x \in \boldsymbol{x}$,

$$\cos(x) = \cos(\check{x}) - (x - \check{x})\sin(\check{x}) - \frac{(x-\check{x})^2}{2}\cos(c)$$

for some c between $\check{x}$ and x. Translating to intervals, we obtain the inclusion

$$\cos(x) \in \cos(\check{x}) - (\boldsymbol{x} - \check{x})\sin(\check{x}) - \frac{(\boldsymbol{x}-\check{x})^2}{2}\left[\min_{c\in\boldsymbol{x}}\{\cos(c)\}, \max_{c\in\boldsymbol{x}}\{\cos(c)\}\right].$$

Specifically, if $\boldsymbol{x} = [-0.1, 0.1]$, we may use $\check{x} = 0$ to obtain the interval extension value

$$\begin{aligned}
\mathbf{cos}^u([-0.1, 0.1]) &\subseteq \cos(0) - [-0.1, 0.1]\sin(0) - \frac{([-0.1, 0.1]^2)}{2}[0.9, 1] \\
&= 1 - [0, 0.005][0.9, 1] = 1 - [0, 0.005] \\
&= [0.995, 1] = \mathbf{cos}([-0.1, 0.1]),
\end{aligned}$$

whereas the range of cos over $[-0.1, 0.1]$, given to ten digits and rounded out, is $[0.995004164, 1]$.

Remark 1.1 *Above,* $[-0.1, 0.1]^2$ *is given as* $[0, 0.01]$, *not as* $[-0.1, 0.1][-0.1, 0.1] = [-0.01, 0.01]$. *Generally, integer powers should be viewed as standard functions with exact ranges (to within roundout), not as repeated multiplication.*

Ranges of the standard functions can thus be approximated with any desired accuracy, within the limits of the floating point system. Monotonicity can also be used, as appropriate, to piece together the range from efficiently obtained narrow enclosures for the endpoints of $\boldsymbol{x}$. Argument reductions and other considerations, such as those explained in [32], are also important in construction of libraries of interval values of functions. Taylor series are often used, due to their simplicity and ease of including the truncation error term. In fact, Braune [23] and Krämer [138] describe how both roundoff and truncation error can be taken into account in Taylor series values, given machine parameters, so that only floating point arithmetic is necessary when accumulating the series terms, to obtain rigorous bounds on ranges. Along other lines, [15] contains recipes for computing ranges of the various standard functions, based on rational approximations.

There are additional issues associated with interval values of certain standard functions. For example, the range of the inverse trigonometric functions should be viewed as an infinite sequence of intervals in some applications, and the range of some trigonometric functions contains branch points. However, most of the standard functions available with programming language compilers for scientific computation can be considered as functions $\omega : \mathbb{IR} \rightarrow \mathbb{IR}$ such that $\boldsymbol{\omega}(\boldsymbol{x})$ is, to within roundout, the exact range of ω over $\boldsymbol{x}$.

Detailed discussion of techniques for computing the ranges of functions appears in [39]. Particulars on construction of the `FORTRAN-77` library described in §2.1 appears in [86, 124, 123].

1.1.7 Properties of Interval Extensions

Some elementary properties of interval extensions were already introduced as Theorems 1.1 and 1.2. Amplification and additional properties useful in formulating and coding objective functions and systems of equations appear here.

Theorem 1.3 *Suppose that an interval extension $\boldsymbol{F}(\boldsymbol{X})$ of a function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is obtained by a sequence of computation of rigorous bounds on ranges $\boldsymbol{\omega}_j(\boldsymbol{x})$ of standard functions $\{\omega_j\}_{j=1}^{p}$ as in §1.1.6 interspersed with outwardly rounded interval arithmetic involving the four elementary operations, as in formulas (1.5) through (1.9). Then this natural interval extension $\boldsymbol{F}(\boldsymbol{X})$ contains the range of F over the box $\boldsymbol{X}$.*

An early criticism of interval arithmetic was that the bounds on the ranges so obtained were too wide or pessimistic to be of value. This *is* frequently the case when interval arithmetic is applied *naively*, but a good understanding of its properties leads to appropriate and effective algorithms. The next several facts deal with *how sharply* interval extensions enclose the range of a function.

Theorem 1.4 *Suppose $\phi : \mathbb{R}^n \to \mathbb{R}$, is formally written as a sequence of computations of standard functions $\{\omega_j\}_{j=1}^q$ interspersed with the four basic arithmetic operations. Suppose that in this expression, each of the n variables $\{x_i\}_{i=1}^n$ occurs formally only once. Then, if $\boldsymbol{\phi}(\boldsymbol{X})$ is evaluated with exact interval arithmetic and computation of the exact ranges of each ω_j, the resulting interval enclosure will be the exact range $\boldsymbol{\phi}(\boldsymbol{X})$.*

Thus, the range e.g. of $f(x) = x^2 - 2$ over $[-2, 2]$ is $[-2, 2]^2 - 2 = [0, 4] - [2, 2] = [-2, 2]$, and the range of $\phi(x_1, x_2) = x_1 x_2$ over the box $\left([-1, 1], [-1, 1]\right)^T$ is $[-1, 1][-1, 1] = [-1, 1]$.

Theorem 1.4 and subdistributivity (Theorem 1.1) suggest that, to maximize the sharpness of an interval extension, the defining expressions or algorithms should be rewritten to minimize the number of occurrences of each variable or subexpression. Indeed, that is a reasonable heuristic in many[10] instances.

Computing the exact range of an arbitrary function ϕ over a box $\boldsymbol{X}$ is similar to optimizing ϕ over $\boldsymbol{X}$. However, in both cases asymptotic properties of interval bounds on the range of ϕ as $\|\mathrm{w}(\boldsymbol{X})\| \to 0$ are helpful.

Definition 1.4 *Let $\boldsymbol{F}(\boldsymbol{X})$ denote an interval extension of $F : \mathbb{R}^n \to \mathbb{R}^m$ evaluated over a box $\boldsymbol{X}$, and let $\boldsymbol{F}^u(\boldsymbol{X})$ denote the exact range of f over $\boldsymbol{X}$. If there is a constant K, independent of the box $\boldsymbol{X}$ such that*

$$\mathrm{w}\big(\boldsymbol{F}(\boldsymbol{X})\big) - \mathrm{w}\big(\boldsymbol{f}^u(\boldsymbol{X})\big) \le K \mathrm{w}(\boldsymbol{X})^\alpha \tag{1.15}$$

for all boxes $\boldsymbol{X}$ with $\mathrm{w}(\boldsymbol{X})$ sufficiently small and fixed $\alpha > 0$, then we say that $\boldsymbol{F}$ is an order α inclusion function for F. *When α is 1 or 2, we call the inclusion* first order *or* second order, *respectively.*

For clarity in what follows, we now restate Definition 1.2 more generally.

Definition 1.5 *If $F : \mathbb{R}^n \to \mathbb{R}^m$ is a function computable as an expression, algorithm or computer program involving the four elementary arithmetic operations interspersed with evaluations of standard functions $\{\omega_j\}_{j=1}^q$, then a*

[10] but not all

natural interval extension *of F, whose value over an interval vector $\boldsymbol{X}$ is denoted by $\boldsymbol{F}(\boldsymbol{X})$, is obtained by replacing each occurrence of each component x_i of X by the corresponding interval component $\boldsymbol{x}_i$ of $\boldsymbol{X}$, by executing all operations according to formulas (1.5) through (1.9), and by computing the exact ranges of the standard functions.*

Remark 1.2 *We will refer to natural interval extensions even when machine arithmetic is used. However, the standard function ranges are then merely enclosures of the actual ranges that are sharp to within roundout error and the bounds on the truncation error. The order of such an interval extension is then only approximate, and not valid when the widths of the independent variables are on the order of the distance between machine numbers.*

Theorem 1.5 *([8], [165], [190, §1.5]) Natural interval extensions are first order.*

Second order extensions are somewhat more desirable, and indeed, almost indispensable in some contexts. This has been pointed out in [114] for solving nonlinear systems and in [52], [125] and [51] for optimization. Second-order extensions may be obtained, as in the examples in §1.1.6, by series expansions and bounding the range of derivatives.

Definition 1.6 *Suppose $\phi : \mathbf{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ has continuous derivatives, $\boldsymbol{X} \subseteq \mathbf{D}$ and $\check{X} \in \boldsymbol{X}$. Then the* mean value extension *for ϕ over $\boldsymbol{X}$ and centered at $\check{X}$ is defined by*

$$\boldsymbol{\phi}_2(\boldsymbol{X}, \check{X}) := \phi(\check{X}) + \boldsymbol{\nabla\phi}(\boldsymbol{X})(\boldsymbol{X} - \check{X}), \tag{1.16}$$

where $\boldsymbol{\nabla\phi}(\boldsymbol{X})$ is a componentwise interval enclosure for the range of $\nabla\phi$ over $\boldsymbol{X}$.

It follows from the mean value theorem and properties of interval arithmetic that $\boldsymbol{\phi}_2(\boldsymbol{X}, \check{X})$ is an enclosure for the range of $\boldsymbol{\phi}$ over $\boldsymbol{X}$. Furthermore:

Theorem 1.6 *(originally in [142]; see also [190, Ch. 2, esp. Theorem 2.3]) Suppose that the components of $\boldsymbol{\nabla\phi}$ are interval extensions of $\nabla\phi$ of order at least one. Then $\boldsymbol{\phi}_2$ is an order-2 interval extension of ϕ.*

For example, suppose $\phi(x) = x^2 - 2$, $\boldsymbol{x} = [0,1]$ and $\check{x} = 1/2$, so that $\nabla\boldsymbol{\phi}(\boldsymbol{x}) = 2\boldsymbol{x}$ is the exact range of the derivative of ϕ over $\boldsymbol{x}$. Then the natural interval extension obtained from the particular expression $x^2 - 2$ is, as in §1.1.1, $\boldsymbol{\phi}_{\rm n}([0,1]) = [-1,1]$, while the mean value extension is

$$\begin{aligned}\boldsymbol{\phi}_2([0,1],1/2) &= -\frac{1}{4} + [0,2]\left([0,1] - \frac{1}{2}\right) \\ &= [-5/4, 3/4]\end{aligned}$$

Thus $w\big(\boldsymbol{\phi}_{\rm n}([0,1])\big) = w\big(\boldsymbol{\phi}_2([0,1],1/2)\big) = 2$, whereas the width of the range is $w\big(\boldsymbol{\phi}^u([0,1])\big) = 1/4$. This illustrates that a second order extension may not be superior to a first order extension (and may actually be substantially wider than a natural extension) when the widths of the arguments are large. It also illustrates that the extensions are *different*: $\boldsymbol{\phi}_{\rm n}([0,1]) \cap \boldsymbol{\phi}_2([0,1],1/2) = [-1,1] \cap [-5/4, 3/4] = [-1, 3/4]$ is also an enclosure for the range of ϕ over $[0,1]$ that is sharper than either $\boldsymbol{\phi}([0,1])$ or $\boldsymbol{\phi}_2([0,1],1/2)$ separately.

If on the other hand $\boldsymbol{x} = [0.49, 0.51]$ and $\check{x} = 0.5$, then the natural interval extension is $\boldsymbol{\phi}_{\rm n}([0.49,0.51]) = [-0.2699, -0.2299]$, $\boldsymbol{\phi}_2([0.49,0.51],0.5) = [-0.2602, -0.2398]$, $\boldsymbol{\phi}^u([0.49,0.51]) = [-0.25, -0.2499]$, and a hint of the convergence behavior of $\boldsymbol{\phi}_{\rm n}$ and $\boldsymbol{\phi}_2$ is seen.

Unfortunately, although it would be desirable in global optimization algorithms, it appears difficult to obtain interval extensions of order higher than 2 for arbitrary functions; cf. [41] and the discussion in [51, Ch. 5]: finding a third order extension may require solving a quadratic programming problem which may not always have all positive eigenvalues.

The mean value form is a special case of a *centered form* for an interval inclusion for a function. Alternate centered forms and a discussion of the theory appear in [190]. Such forms may be appropriate to obtain tighter bounds in specific cases. For our purposes[11] the natural extension $\boldsymbol{\phi}$ and mean value extension $\boldsymbol{\phi}_2$ will be adequate. However, we will use interval vectors other than $\nabla\boldsymbol{\phi}(\boldsymbol{X})$ in Equation (1.16) to obtain tighter but still-rigorous bounds; cf. §1.3 below. In algorithms, a particular natural interval extension will be denoted by $\boldsymbol{\phi}$ (or $\boldsymbol{f}$ or $\boldsymbol{F}$), while a generalized mean value extension will be denoted by $\boldsymbol{\phi}_2$, to imply that any first order and any second order extension will do.

[11] since we will supply generic interpreters to compute ϕ and ϕ_2, since order higher than 2 cannot be achieved merely with more terms, and since other mechanisms in global optimization algorithms can supply tighter bounds

A final property, often assumed in the literature to facilitate convergence proofs, is

Definition 1.7 *An interval extension* $\boldsymbol{\phi}$ *is* inclusion monotonic *or* inclusion isotonic *provided* $\boldsymbol{Y} \subseteq \boldsymbol{X}$ *implies* $\boldsymbol{\phi}(\boldsymbol{X}) \subseteq \boldsymbol{\phi}(\boldsymbol{Y})$.

Theorem 1.7 *Natural interval extensions as in Definition 1.5 are inclusion monotonic, provided exact interval arithmetic is used and exact ranges are computed for the standard functions.*

Inclusion monotonicity has been suggested as a stopping criterion for interval iterative algorithms: violation of it indicates that roundout error is predominating.

Although inclusion monotonicity is convenient, in our opinion the asymptotic properties (i.e. the order) are more crucial. Some of our constructions and algorithmic processes, although effective, lead to interval extensions that are not inclusion monotonic.

1.1.8 Exercises

1. On page 4, it was seen that the natural interval extension corresponding to $f_1(x) = x(x-1)$ was sharper than the natural interval extension corresponding to $f_2(x) = x^2 - x$ over the interval $[0,1]$, even if the exact range of x^2 is used, i.e. $\boldsymbol{f}_1([0,1]) \subset \boldsymbol{f}_2([0,1])$. Is this also true for the interval $[-1,1]$, i.e. is $\boldsymbol{f}_1([-1,1]) \subset \boldsymbol{f}_2([-1,1])$? What can you say about Horner's method and factoring interval extensions? Can you make recommendations concerning which natural interval extensions are appropriate or how they should be used? You may wish to consult [192].
2. *(roundout error)* In rounded interval arithmetic, an increase in interval widths due to directed roundings occurs under the same circumstances as roundoff error in traditional floating point computations. Consider an outwardly rounded interval arithmetic based on directed roundings with a four decimal digit floating point system. (For example, the outwardly rounded result of $[.1234 \times 10^0, .1234 \times 10^0] + [.1111 \times 10^{-1}, .1111 \times 10^{-1}]$ would be $[.1345 \times 10^0, .1346 \times 10^0] \ni .13451$.) Now examine the difference quotient $Q(h) = \big(f(x_0 + h) - f(x_0)\big)/h$ with $f(x) = x^2$ and $x_0 = 1$.

Representing x_0 as a four-digit decimal interval $[.1000 \times 10^0, .1000 \times 10^0]$ and using outwardly rounded arithmetic, form a table whose columns are h, $Q(h)$ and $\mathrm{w}(Q(h))$ and whose rows correspond to $h = .1$, $h = .05$, $h = .01$, and $h = .005$. Perform a directed rounding into four digit intervals after each of the three operations ($x \rightarrow x^2$, $y - x$, and y/x).

3. For the example on page 16, form a table with columns $\boldsymbol{x}$, $\boldsymbol{\phi}$, $\boldsymbol{\phi}_2$, $\boldsymbol{\phi}^u$, $\mathrm{w}(\boldsymbol{\phi})$, $\mathrm{w}(\boldsymbol{\phi}_2)$ and $\mathrm{w}(\boldsymbol{\phi}^u)$ for $\boldsymbol{x}$ centered about $\check{x} = 1/2$ and with widths 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6}. Referring to Definition 1.4, form and label additional columns of the table containing ratios of differences of widths to illustrate from the computations that $\boldsymbol{\phi}$ is order 1 and $\boldsymbol{\phi}_2$ is order 2.

4. Prove Theorem 1.7.

   Hint: This may be done by induction on the number of elementary operations (including evaluation of standard functions).

1.2 INTERVAL LINEAR SYSTEMS

Within the scope of this book, interval linear systems of equations are important in interval Newton methods, the primary device for verifying existence and uniqueness of roots and critical points. In particular, interval Newton methods can be viewed as computational analogues of the Brouwer fixed point theorem, and the solution set of an interval linear system must be bounded to carry out an interval Newton method. Also, interval linear systems of equations are useful in computing rigorously verified bounds on the actual solutions of real systems of equations; see [70, Chapter 10] and the references therein.

Generally, interval linear systems are of the form

$$\boldsymbol{A}X = \boldsymbol{B}, \tag{1.17}$$

where $\boldsymbol{A} \in \mathbb{IR}^{n \times n}$ and $\boldsymbol{B} \in \mathbb{IR}^n$, although the form most commonly encountered in interval Newton methods is

$$\boldsymbol{A}(X - \check{X}) = \boldsymbol{B}_0, \tag{1.18}$$

where $\check{X} \in \mathbb{R}^n$ and, usually, $\|\mathrm{w}(\boldsymbol{B}_0)\|$ is small; the special form of Equation (1.18) can be exploited, as explained in Chapter 3 below.

The concept of regularity is important in proving uniqueness of solutions.