Hint: You may want to use induction on the number of operations required to compute the function.

5. Assuming each operation is either a unary or binary operation (where standard function evaluations are counted as an operation), count the number of additions, subtractions, multiplications, and divisions to perform the reverse mode on a general system set up as System (1.32). Let n be the number of variables and N the number of operations to evaluate $\phi(x_1, x_2, \ldots, x_n)$. Write the operation count in terms of N and n. Does this differ from the common perception of the amount of work required to compute gradients?

6. If $\phi(x) = x^5 + x^4 + x^3 + x$, then write a code list for ϕ in the form of Table 1.2, using the operation codes of Table 1.1. Are such code lists unique? If not, then also write down an alternative code list.

   Note: Different code lists are obtained by doing the operations in a different order. However, functions such as $\tilde{\phi}(x) = x(1+x^2(1+x(1+x)))$, although equivalent in real arithmetic, are considered to be different functions. You may interpret the problem as asking if there are two different code lists that both obey the restrictions on the order of operations defined in the Fortran 90 standard.

7. For $f(x) = x^4 + x^3 + x$ with code list as in Table 1.2 and with entry value `X` $= 2$, manually trace execution of the subroutine `POINT_EVALUATION` in Figure 1.5. (That is, make a list of each statement passed in the program, with values of the variables after they are assigned.)

   The value `FVAL` *upon return should be $f(2)$.*

1.5 INTERVAL NEWTON METHODS AND INTERVAL FIXED POINT THEORY

The books [77], [175] and [191] are good references for interval Newton methods and interval fixed point theory. The reference [77] emphasizes concepts and practical aspects of implementation. An especially good reference for theory and the computational existence and uniqueness tests in §1.5.2 is [175], supplemented with the more recent results on uniqueness with slope matrices that appear in [208]. The reference [7] gives a good survey of interval Newton methods based on interval Gaussian elimination, and also surveys the Krawczyk method. One relatively early use of classical fixed point theory in the analysis of interval methods is in [205].

In this section, univariate interval Newton methods will first be discussed. After that, multivariate interval Newton methods, based upon methods of bounding the solution sets to interval linear systems in §1.2.2 will be introduced. Next, computational existence and uniqueness theory, based on classical fixed-point theorems, will be presented.

1.5.1 Interval Newton Methods

Except for the choice of interval extensions or slopes, there is essentially a single univariate interval Newton method. In contrast, multivariate interval Newton methods vary according to solution method for the linear system, preconditioner, etc. We will discuss the simpler case first.

Univariate Interval Newton Methods

Suppose $f : \boldsymbol{x} = [\underline{x}, \overline{x}] \rightarrow \mathbb{R}$ has a continuous first derivative on $\boldsymbol{x}$, suppose that there exists $x^* \in \boldsymbol{x}$ such that $f(x^*) = 0$, and suppose that $\check{x} \in \boldsymbol{x}$. Then, since the mean value theorem implies

$$\begin{aligned} 0 &= f(x^*) = f(\check{x}) + f'(\xi)(x^* - \check{x}), \\ \text{we have } x^* &= \check{x} - f(\check{x})/f'(\xi) \end{aligned}$$

for some $\xi \in \boldsymbol{x}$. If $\boldsymbol{f}'(\boldsymbol{x})$ is any interval extension of the derivative of f over $\boldsymbol{x}$, then

$$x^* \in \check{x} - f(\check{x})/\boldsymbol{f}'(\boldsymbol{x}) \quad \text{for any } \check{x} \in \boldsymbol{x}.$$

Similarly, if $\mathbf{S}(f, \boldsymbol{x}, \check{\boldsymbol{x}})$ is a slope set for f centered at $\check{x}$ (as in Definition 1.12 on page 27), then, also

$$x^* \in \check{x} - f(\check{x})/\mathbf{S}(f, \boldsymbol{x}, \check{\boldsymbol{x}}) \quad \text{for any } \check{x} \in \boldsymbol{x}.$$

This leads to

Definition 1.15 (Univariate interval Newton methods) *The operators*

$$\boldsymbol{N}(f; \boldsymbol{x}, \check{x}) = \check{x} - f(\check{x})/\boldsymbol{f}'(\boldsymbol{x}) \tag{1.37}$$

or

$$\boldsymbol{N}(f; \boldsymbol{x}, \check{x}) = \check{x} - f(\check{x})/\mathbf{S}(f, \boldsymbol{x}, \check{x}) \tag{1.38}$$

are[30] *termed the* univariate interval Newton methods.

[30] Although methods differ according to the interval extension of the derivative and whether a derivative or slope is used, all such variants will be viewed as a single method here.

Recapitulating, if $x^* \in \boldsymbol{x}$ has $f(x^*) = 0$, then $x^* \in \boldsymbol{N}(f; \boldsymbol{x}, \check{x})$. This fact underlies the use of interval Newton iteration to sharpen bounds on solutions and global optimizers. Another important fact making interval Newton methods efficient for global optimization algorithms is local quadratic convergence. An example of the type of theorem that can be proven[31] is

Theorem 1.14 *Suppose*

(a) $\check{x}^{(k)} = (\underline{x}^{(k)} + \overline{x}^{(k)})/2$ *for all* $\boldsymbol{x}^{(k)}$,

(b) f *has a continuous derivative on an initial interval* $\boldsymbol{x}^{(0)}$,

(c) $|f(x)| \le M_f$ *for* $x \in \boldsymbol{x}^{(0)}$,

(d) $\boldsymbol{N}(f; \boldsymbol{x}^{(0)}, \check{x}^{(0)})$ *is defined with Equation (1.37),*

(e) $\boldsymbol{f}'$ *is an inclusion monotonic interval extension of at least first order in the sense of Definition 1.4 on page 14,*

(f) $\min\{|y|, y \in \boldsymbol{f}'(\boldsymbol{x}^{(0)})\} = m_{f'} > 0$, *and*

(g) $KM_f/m_{f'}^2 = \beta < 1$.

Then

1. $\mathrm{w}\big(\boldsymbol{N}(f; \boldsymbol{x}^{(0)}, \check{x}^{(0)})\big) \le \beta \mathrm{w}\big(\boldsymbol{x}^{(0)}\big)$.
2. *Furthermore, if*
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} \cap \boldsymbol{N}(f; \boldsymbol{x}^{(k)}, \check{x}^{(k)}), \tag{1.39}$$
then $\mathrm{w}\big(\boldsymbol{x}^{(k+1)}\big) = \mathcal{O}\left(\mathrm{w}\big(\boldsymbol{x}^{(k)}\big)\right)^2$.

Proof: Property 1 will be proved first. Only the case $f(\check{x}^{(k)}) > 0$ and $\boldsymbol{f}'(\boldsymbol{x}^{(0)}) = [a, b] > 0$ will be considered, since the other cases are similar. In that case,

$$\begin{aligned}
\mathrm{w}\Big(\boldsymbol{N}(f; \boldsymbol{x}^{(0)}, \check{x}^{(0)})\Big) &= \mathrm{w}\left(\frac{f(\check{x}^{(0)})}{[a, b]}\right) \\
&= \mathrm{w}\left(\left[\frac{f(\check{x}^{(0)})}{b}, \frac{f(\check{x}^{(0)})}{a}\right]\right) = f(\check{x}^{(0)})\frac{b-a}{ab}
\end{aligned}$$

[31] This theorem is just meant to be an example that is easy to understand, not the sharpest or most general result in the literature.

$$\begin{aligned}&\le M_f \mathrm{w}\Big(\boldsymbol{f}'(\boldsymbol{x}^{(0)})\Big)/m_{f'}^2\\&\le KM_f \mathrm{w}\Big(\boldsymbol{x}^{(0)}\Big)/m_{f'}^2 = \beta \mathrm{w}\Big(\boldsymbol{x}^{(0)}\Big).\end{aligned}$$

To prove property 2, we proceed as with property 1, assuming the case $f(\check{x}^{(k)}) > 0$ and $\boldsymbol{f}'(\boldsymbol{x}^{(k)}) = [a_k, b_k] > 0$: $\mathrm{w}\big(\boldsymbol{x}^{(k+1)}\big) \le \mathrm{w}\big(\boldsymbol{N}(f; \boldsymbol{x}^{(k)}, \check{x}^{(k)})\big) = f(\check{x}^{(k)})\frac{b_k - a_k}{a_k b_k}$. However, $\check{x}^{(k)} = x^* + (\check{x}^{(k)} - x^*)$, so $f(\check{x}^{(k)}) = f'(\eta)(\check{x}^{(k)} - x^*) \le b_k \mathrm{w}\big(\boldsymbol{x}^{(k)}\big)/2$ since $x^* \in \boldsymbol{x}^{(k)}$ and $\check{x}^{(k)}$ is the midpoint of $\boldsymbol{x}^{(k)}$. Therefore, $\{f(\check{x}^{(k)})\}\left\{\frac{b_k - a_k}{a_k b_k}\right\} \le \{b_k \mathrm{w}\big(\boldsymbol{x}^{(k)}\big)/2\}\{KM_f \mathrm{w}\big(\boldsymbol{x}^{(k)}\big)/(m_{f'} b_k)\} = \frac{KM_f}{2m_{f'}}\mathrm{w}\big(\boldsymbol{x}^{(k)}\big)^2$, thus proving property 2 for this case. The other cases are similar. □

Equation (1.39) may be used directly to reduce the size of the initial interval. In fact, this equation is applicable even when $0 \in \boldsymbol{f}'(\boldsymbol{x}^{(0)})$, because the set inclusion properties of the Kahan arithmetic of §1.1.4 are the same as those of ordinary interval arithmetic:

Example 1.4 (Extended interval arithmetic in an interval Newton method) *Let $f(x) = x^2 - 4$, and let $\boldsymbol{x}^{(0)} = [-2, 2]$, and $\check{x}^{(0)} = 0$, so the range of f' is $\boldsymbol{f}'(\boldsymbol{x}^{(0)}) = [-4, 4]$. The first iteration of the interval Newton method thus becomes*

$$\begin{aligned}\boldsymbol{N}(f; \boldsymbol{x}^{(0)}, \check{x}^{(0)}) &= 0 - \frac{-4}{[-4,4]} = 0 - ([-\infty, -1] \cup [1, \infty])\\&= ([-\infty, -1] \cup [1, \infty]).\end{aligned}$$

Thus,

$$\boldsymbol{N}(f; \boldsymbol{x}^{(0)}, \check{x}^{(0)}) \cap \boldsymbol{x}^{(0)} = [-2, -1] \cup [1, 2] = \boldsymbol{x}^{(1,1)} \cup \boldsymbol{x}^{(1,2)}.$$

One of $\boldsymbol{x}^{(1,1)} = [-2, -1]$ and $\boldsymbol{x}^{(1,2)} = [1, 2]$, say $\boldsymbol{x}^{(1,1)}$, is now put on a stack for later processing, and the interval Newton method is iterated with $\boldsymbol{x}^{(1,2)}$. The latter, implemented in the software system of §2.2.1 below, gives the iterates[32] in Table 1.3. There, the last column, labelled "rat.", represents the ratios $\mathrm{w}\big(\boldsymbol{x}^{(k)}\big)/\mathrm{w}\big(\boldsymbol{x}^{(k-1)}\big)^2$.

Provided the case $f(\check{x}) = 0$ with $0 \in \boldsymbol{f}'(\boldsymbol{x})$ does not occur, the univariate interval Newton method combined with Kahan–Novoa–Ratz arithmetic and deferred processing of additional boxes often converges to small boxes containing all of the roots of f. Such a univariate interval Newton method algorithm can be summarized in

[32] The last digits in $x^{(k)}$ and $\check{x}^{(k)}$ have been rounded out for rigorous inclusions.

k	$\boldsymbol{x}^{(k)}$	$\check{x}^{(k)}$	$\mathrm{w}\left(\boldsymbol{x}^{(k)}\right)$	rat.
1	[0.9999999999999997, 2.0000000000000005]	1.5000000000000000	1.00	—
2	[1.9374999999999990, 2.3750000000000019]	2.1562500000000004	4.38E-01	4E-1
3	[1.9886592741935471, 2.0195312500000010]	2.0040952620967740	3.09E-02	2E-1
4	[1.9999724292486506, 2.0000354537727550]	2.0000039415107027	6.30E-05	7E-2
5	[1.9999999999417791, 2.0000000000659869]	2.0000000000038831	1.24E-10	3E-2
6	[1.9999999999999988, 2.0000000000000010]	2.0000000000000000	2.00E-15	1E+5

Table 1.3 Illustration of quadratic convergence of the univariate interval Newton method for $f(x) = x^2 - 4$

Algorithm 5 (Univariate interval Newton search, similar to [77, §7.4])

INPUT: f, $\boldsymbol{x}$, and a tolerance ϵ.

OUTPUT: A list $\mathcal{C}$ of narrow intervals of length at most ϵ that possibly contain roots.

1. *Initialize:* $\mathcal{U} \leftarrow \{\boldsymbol{x}\}$.
2. DO *while* $\mathcal{U} \neq \emptyset$.
   - (a) *Remove an interval* $\boldsymbol{x}$ *from* $\mathcal{U}$.
   - (b) DO *while* $\mathrm{w}(\boldsymbol{x}) > \epsilon$.
     - i. *Compute* $\boldsymbol{N}(f;\boldsymbol{x},\check{x}) \cap \boldsymbol{x}$ *for some* $\check{x} \in \boldsymbol{x}$.
     - ii. IF $\boldsymbol{N}(f;\boldsymbol{x},\check{x}) \cap \boldsymbol{x}$ *consists of a single interval*

       THEN

       $\boldsymbol{x} \leftarrow \boldsymbol{N}(f;\boldsymbol{x},\check{x}) \cap \boldsymbol{x}$.

       ELSE

       Put one of the intervals into $\mathcal{U}$.

       Set the other interval to $\boldsymbol{x}$.

       END IF

     END DO
   - (c) *Store* $\boldsymbol{x}$ *in* $\mathcal{C}$.

   END DO

End Algorithm 5

A relative width $\mathrm{w}(\boldsymbol{x})/\max\{|\underline{x}|,|\overline{x}|\}$ may be used in Step 2b. For rigor, all of the computations to obtain $\boldsymbol{N}(f;\boldsymbol{x},\check{x})$, including evaluation of $f(\check{x})$, should be done with outwardly rounded interval arithmetic.

For a summary of the theory underlying Algorithm 5, as well as for extensive additional commentary, see [77, Chapter 7]. Algorithm 5 is simplified, and does not include steps to verify existence or uniqueness within the boxes in $\mathcal{U}$. Such techniques are explained in the general, multidimensional context in §1.5.2.

Algorithm 5 should function satisfactorily for a variety of f, although more sophisticated techniques are necessary for problems such as the Gritton problem studied in [114], in which there is a combination of serious overestimation in the derivative values and cancellation error.

Multivariate Interval Newton Methods

Although analogous to the univariate interval Newton methods above, multivariate interval Newton methods have more variety. Different univariate interval Newton methods of formulas (1.37) and (1.38) vary in the choice of base point $\check{x}$ and interval extension of the derivative or slope. In multivariate interval Newton methods, not only can the base point $\check{X}$ and interval extension $\boldsymbol{F}'(\boldsymbol{X})$ or $\mathbf{S}(\boldsymbol{F},\boldsymbol{X},\check{\boldsymbol{X}})$ be varied, but there is a wider variety of ways of bounding the solution sets of the resulting linear systems than in the univariate case; cf. §1.2.2. Furthermore, there are additional issues in dealing with interval slopes [120, §3.3] and [210]. Finally, there are questions concerning preconditioning the related interval linear systems. Differences in how interval Newton methods are used in computational existence and uniqueness tests are discussed later in this chapter in §1.5.2. Extensive new results on theory and practice of preconditioning systems are presented later in Chapter 3. Further discussion (besides that in §1.2.2) of the differences between the Krawczyk method, interval Gaussian elimination, and the interval Gauss–Seidel method also occurs below and in Chapter 3.

Multivariate interval Newton methods are closely tied to the methods for linear systems discussed in §1.2. In particular, multivariate interval Newton methods involve iterative execution of algorithms to bound the solution sets of linear systems, except that the linear systems are of the special form

$$\boldsymbol{A}(X-\check{X})=-F(\check{X}),$$

where $\check{X}$ is an approximation to a root of F (and hence $\|F(\check{X})\|$ is small), and $\boldsymbol{A}$ is either $\boldsymbol{F}'(\boldsymbol{X})$ or $\mathbf{S}(\boldsymbol{F},\boldsymbol{X},\check{\boldsymbol{X}})$. Thus, interval Newton methods involve