Numer. Math. 34, 247 – 270 (1980)

Numerische Mathematik

Global Optimization Using Interval Analysis– The Multi-Dimensional Case

Eldon Hansen

Lockheed Missiles and Space Company
Sunnyvale, CA 94086, USA

Summary. We show how interval analysis can be used to compute the global minimum of a twice continuously differentiable function of n variables over an n-dimensional parallelopiped with sides parallel to the coordinate axes. Our method provides infallible bounds on both the globally minimum value of the function and the point(s) at which the minimum occurs.

Subject Classification: AMS(MOS): 65K05, 90C30.

1. Introduction

Consider the function $f(x)$ in C^2 of n variables $x_1, \ldots, x_n$. We shall describe a method for computing the minimum value f^* of $f(x)$ over a box $X^{(0)}$. A box is defined to be a closed rectangular parallelopiped with sides parallel to the coordinate axes. We assume the number of points in $X^{(0)}$ at which $f(x)$ is globally minimum is finite. Our method provides infallible bounds on f^* and on the point(s) x^* for which $f(x^*)=f^*$. That is, our algorithm produces bounds on x^* and f^* which are always correct despite the presence of rounding errors. How sharp these bounds can be depends on the function f and the precision of the computer used.

For a highly oscillatory function f, our algorithm could be prohibitively slow. Presumably this will always be the case for any future global optimization algorithm. However, our algorithm is sufficiently fast for 'reasonable' functions.

We assume that interval extensions (see [8]) of f and its derivatives are known. This is the case if every function in terms of which f and its derivatives are defined have known rational approximations with either uniform or rational error bounds for the arguments of interest.

Since the initial box can be chosen as large as we please, our algorithm actually solves the unconstrained minimization problem provided it is known that the solution occurs in some finite region (which we enclose in the initial box).

0029-599X/80/0034/0247/$04.80

There is a common misconception among researchers in optimization that it is impossible to obtain infallible bounds on x^* and f^* computationally. The argument is that we can only sample $f(x)$ and a few derivatives of $f(x)$ at a finite number of points. It is possible to interpolate a function having the necessary values and derivatives values at these points and still have its global minimum at any other arbitrary point. The fallacy of this argument is that interval analysis can provide bounds on a function over an entire box; that is over a continuum of points. It is only necessary to make the box sufficiently small in order to make the bounds arbitrarily sharp. This is what our algorithm does. It narrows the region of interest until the bound is as sharp as desired (subject to roundoff restrictions).

In a previous paper [5], we gave a method of this type for the one-dimensional case. The method never failed to converge provided $f'(x)$ and $f''(x)$ had only a finite number of isolated zeros. Our method for the n-dimensional problem appears to always converge also; but we have not yet attempted to prove it. When it does converge, there is never a question that x^* and f^* satisfy the computed bounds.

Recently, R.E. Moore [9] published a method for computing the range of a rational function of n variables over a bounded region. (See also [14].) Although he does not note the fact, his method will serve to bound the global minimum value f^* of a rational function. However, our algorithm is more efficient. Moreover, it is designed to bound x^* as well as f^*.

We suggest the reader read the previous paper [5] before the current one. The one-dimensional case therein serves as an easier introduction. However, the current paper is essentially self contained. It would be better if the reader had some familiarity with the rudiments of interval analysis such as can be found in the first three chapters of [8]. However, we shall review some of its relevant properties.

Our method will find the global minimum (or minima). Because of computer limitations of accuracy, it may also find near-global minima such that rounding errors prevent determination of which is the true minimum. However, if the termination criteria are sufficiently stringent, our algorithm will always eliminate a local minimum whose value is substantially larger than f^*.

Our algorithm is composed of four separate parts. One part uses an interval version of Newton's method to find stationary points. A second part eliminates points of $X^{(0)}$ where f is greater than the smallest currently known value $\bar{f}$.

A third part of our algorithm tests whether f is monotonic in a sub-box X of $X^{(0)}$. If so, we delete part or all of X depending on whether X contains boundary points of $X^{(0)}$.

A fourth part checks for convexity of f in a sub-box X of $X^{(0)}$. If f is not convex anywhere in X, there cannot be a stationary minimum of f in X.

The first part of the algorithm, if used alone, would find all stationary points in $X^{(0)}$. The second part serves to eliminate stationary points where $f > f^*$. Usually they are eliminated before they are found with any great accuracy. Hence computational effort is not wasted using the first part to accurately find an unwanted stationary point. The second part also serves to eliminate boundary points of $X^{(0)}$ and to find a global minimum if it occurs on the boundary.

The second part of the algorithm used alone would find the global minimum (or minima) but its asymptotic convergence is relatively slow compared to that of the Newton method. Hence the latter is used also. The third and fourth parts of the algorithm merely improve convergence.

2. Interval Analysis

The tool which allows us to be certain we have bounded the global minimum is interval analysis. We bound rounding errors by using interval arithmetic. More importantly, however, we use interval analysis to bound the range of a function over a box.

Let $g(x)$ be a rational function of n variables $x_1, \ldots, x_n$. On a computer, we can evaluate $g(x)$ for a given x by performing a sequence of arithmetic operations involving only addition, subtraction, multiplication, and division.

Let X_i $(i=1, \ldots, n)$ be closed intervals. If we use X_i in place of x_i and perform the same sequence of operations using interval arithmetic (see [8]) rather than ordinary real arithmetic, we obtain a closed interval $g(X)$ containing the range

$$\{g(x): x_i \in X_i (i=1, \ldots, n)\}$$

of $g(x)$ over the box X. This result will not be sharp, in general, but if outward rounding (see [8]) is used, then $g(X)$ will always contain the range. The lack of sharpness results from other causes besides roundoff. With exact interval arithmetic, the lack of sharpness diappears as the widths of the intervals decrease to zero.

If $g(x)$ is not rational, we assume an algorithm is known for computing an interval $g(X)$ containing the range of $g(x)$ for $x \in X$. Methods for deriving such algorithms are discussed in [8]).

3. Taylor's Theorem

We shall use interval analysis in conjunction with Taylor's theorem in two ways. First, we expand f as

$$f(y)=f(x)+(y-x)^T g(x)+\tfrac{1}{2}(y-x)^T H(x, y, \xi)(y-x) \tag{3.1}$$

where $g(x)$ is the gradient of $f(x)$ and has components $g_i(x)=\partial f(x)/\partial x_i$. The quantity $H(x, y, \xi)$ is the Hessian matrix to be defined presently. For reasons related to the use of interval analysis, we shall express it as a lower triangular matrix instead of a symmetric matrix so that there are fewer terms in the quadratic form involving $H(x, y, \xi)$.

We define the element in position (i, j) of $H(x, y, \xi)$ as

$$h_{ij}=\begin{cases} \partial^2 f/\partial x_i^2 & \text{for } j=i (i=1, \ldots, n), \\ 2\partial^2 f/\partial x_i \partial x_j & \text{for } j<i (i=1, \ldots, n; j=1, \ldots, i-1), \\ 0 \quad \text{otherwise.} \end{cases} \tag{3.2}$$

The arguments of h_{ij} depend on i and j. If we expand f sequentially in one of its variables at a time, we can obtain the following results illustrating the case $n=3$

$$H(x, y, \xi)=\begin{bmatrix} h_{11}(\xi_{11}, x_2, x_3) & 0 & 0 \\ h_{21}(\xi_{21}, x_2, x_3) & h_{22}(y_1, \xi_{22}, x_3) & 0 \\ h_{31}(\xi_{31}, x_2, x_3) & h_{32}(y_1, \xi_{32}, x_3) & h_{33}(y_1, y_2, \xi_{33}) \end{bmatrix}.$$

Assume $x_i \in X_i$ and $y_i \in X_i$ for $i=1, \ldots, n$. Then $\xi_{ij} \in X_j$ for each $j=1, \ldots, i$. For general n, the arguments of H_{ij} are $(y_1, \ldots, y_{j-1}, \xi_{ij}, x_{j+1}, \ldots, x_n)$. Other arrangements of arguments could be obtained by reordering the indices.

Let x be a fixed point in X. Then for any point $y \in X$,

$$H(x, y, \xi) \in H(x, X, X);$$

that is, for $i \geqq j$,

$$h_{ij}(y_1, \ldots, y_{j-1}, \xi_{ij}, x_{j+1}, \ldots, x_n) \in h_{ij}(X_1, \ldots, X_j, x_{j+1}, \ldots, x_n).$$

In the sequel, we shall shorten notation and use $H(\xi)$ to denote $H(x, y, \xi)$ and $H(X)$ to denote $H(x, X, X)$.

The purpose of this particular Taylor expansion is to obtain real (non-interval) quantities for as many arguments of the elements of $H(X)$ as possible. The standard Taylor expansion would have intervals for all arguments of all elements of $H(X)$. This type of expansion was introduced in [3]. A more general approach of this kind is discussed in [4].

The other Taylor expansion we shall want is of the gradient g. Each element $g_i (i=1, \ldots, n)$ of g can be expanded as

$$\begin{aligned} g_i(y) = g_i(x) &+ (y_1 - x_1) J_{i1}(\eta_1, x_2, \ldots, x_n) + (y_2 - x_2) J_{i2}(y_1, \eta_2, x_3, \ldots, x_n) \\ &+ (y_3 - x_3) J_{i3}(y_1, y_2, \eta_3, x_4, \ldots, x_n) + \ldots + (y_n - x_n) J_{in}(y_1, \ldots, y_{n-1}, \eta_n), \end{aligned}$$

where (3.3)

$$J_{ij} = \partial^2 f / \partial x_i \, \partial x_j \quad (i, j = 1, \ldots, n).$$

This Jacobian matrix J and the Hessian H introduced above are, of course, essentially the same. However, they will be evaluated with different arguments depending on whether we are expanding f or g. Also, H is lower triangular while J is a full matrix.

Let $J(x, y, \eta)$ denote the Jacobian matrix with elements $J_{ij}(y_1, \ldots, y_{j-1}, \eta_j, x_{j+1}, \ldots, x_n)$. Then

$$g(y) = g(x) + J(x, y, \eta)(y - x). \tag{3.4}$$

If $x \in X$ and $y \in X$, then $\eta_i \in X_i$ for all $i=1, \ldots, n$. Hence

$$g(y) \in g(x) + J(x, X, X)(y - x). \tag{3.5}$$

We shall again shorten notation and denote $J(x, y, \eta)$ by $J(\eta)$ and $J(x, X, X)$ by $J(X)$.

Note that the elements of $H(X)$ on and below the diagonal have the same arguments as the corresponding elements of $J(X)$. Thus we need only calculate $J(X)$; then $H(X)$ follows easily.

4. The Approximate Value of the Global Minimum

As we proceed with our algorithm, we shall evaluate $f(x)$ at various points in $X^{(0)}$. Let $\bar{f}$ denote the currently smallest value of f found so far. The very first step is to evaluate f at the center of $X^{(0)}$. This value serves as the first one for $\bar{f}$.

One part of our algorithm deletes sub-boxes of $X^{(0)}$ wherein $f > \bar{f}$ since this implies $\inf f > f^*$. (See Sect. 7.)

In practice we cannot generally evaluate $f(x)$ exactly because of rounding errors. Hence we do the evaluation using interval arithmetic. Suppose we obtain the interval $[f^L, f^R]$. Then we know that $f(x) \leqq f^R$ and hence that $\bar{f} \leqq f^R$. Hence when we evaluate $f(x)$, we update $\bar{f}$ by replacing it by f^R only if f^R is less than the previous value of $\bar{f}$. In this way, we assure that $\bar{f}$ is always an upper bound for f^*.

5. A Test for Convexity

As our algorithm proceeds, we dynamically subdivide $X^{(0)}$ into sub-boxes. Let X denote such a sub-box. We evaluate $h_{ii}(X_1, \ldots, X_n)$ for $i = 1, \ldots, n$, where h_{ii} is the diagonal element of the Hessian. Note that every argument of h_{ii} is an interval and hence the resulting interval contains the value of $h_{ii}(x)$ for every $x \in X$. That is, if $[u_i, v_i]$ denotes the computed interval $h_{ii}(X_1, \ldots, X_n)$, then

$$h_{ii}(x) \in [u_i, v_i]$$

for all $x \in X$.

Suppose we find $v_i < 0$ for some value of i. Then $h_{ii}(x) < 0$ for every $x \in X$. Hence there is no point in X at which the real (non-interval) Hessian is positive definite. Hence f is not convex and cannot have a minimum which is a stationary point in X. Hence we can delete all of X except for any boundary points of $X^{(0)}$ which might lie in X.

When we evaluate $h_{ii}(X_1, \ldots, X_n)$, we may find that the left endpoint $u_i \geqq 0$ for all $i = 1, \ldots, n$. When this occurs, we know from inclusion monotonicity (see [8]) that we will find each $u_i \geqq 0$ for any sub-box of X. Hence we could save some computational effort by noting when a box is a sub-box of one for which $u_i \geqq 0$ for all $i = 1, \ldots, n$. We would skip this test for such a box.

Note that an element h_{ii} with arguments $(X_1, \ldots, X_n)$ is not obtained when we compute $H(X)$ since the diagonal elements of $H(X)$ have arguments different from $(X_1, \ldots, X_n)$ except for the element in position (n, n). Hence our test for convexity requires recalculation of the diagonal of the Hessian.

6. The Interval Newton Method

For each sub-box X of $X^{(0)}$ that our algorithm generates, we can apply an interval Newton method to the gradient g. Such methods seek the zeros of g and

hence the stationary points of f. Such a method produces from X a new box or boxes $N(X)$. Any points in X not in $N(X)$ cannot contain a zero of g and can be discarded unless they are boundary points of $X^{(0)}$.

These methods, in effect, solve (3.5) for points y where $g(y)=0$. The first such method was derived by Moore [8]. Variants of Moore's method can be found in [3, 8, 12, 13]. The most efficient variant is described below. Krawczyk's method [8] is a suitable alternative to the method in [6]. Discussions of Krawczyk's method can be found in [10] and [11].

We now give a brief synopsis of our method. We wish to solve the set of equations

$$g(x)+J(\xi)(y-x)=0 \tag{6.1}$$

for the set of points y obtained by letting ξ range over X. We shall find a subset Y of X containing this set.

Let J_c be the matrix whose element in position (i,j) is the midpoint of the corresponding interval element $J_{ij}(X)$ of the Jacobian $J(X)$. Let B be an approximate inverse of J_c. As pointed out in [3], a useful first step in solving for Y is to multiply (6.1) by B giving

$$Bg(x)+BJ(\xi)(y-x)=0. \tag{6.2}$$

Note that the product $BJ(\xi)$ approximates the identity matrix. However it may be a very poor approximation when X is a large box.

We 'solve' (6.2) by a process similar to a single sweep of the Gauss-Seidel method. Write

$$BJ(X)=L+D+U$$

where L, D, and U are the lower triangular, diagonal, and upper triangular part of $BJ(X)$, respectively. The interval matrix

$$D^{-1}=\operatorname{diag}[1/D_{11}, 1/D_{22}, \ldots, 1/D_{nn}] \tag{6.3}$$

contains the inverse of every matrix in D. The box Y 'solving' (6.2) is obtained as

$$Y=x-D^{-1}[Bg(x)+L(Y-x)+U(X-x)]. \tag{6.4}$$

When obtaining the component Y_i of Y, the components $Y_1, \ldots, Y_{i-1}$ appearing in the right member of this equation have already been obtained.

This formulation presupposes that the intervals D_{ii} $(i=1, \ldots, n)$ do not contain zero. When X is a small box, $BJ(X)$ is closely approximated by the identity matrix and hence D is also. However, for X large, it is possible to have $0 \in D_{ii}$ for one or more values of i. This case is easily handled. We simply use an extended interval arithmetic which allows division by an interval containing zero. A detailed discussion of this new method will be published elsewhere.

Note that we cannot allow the Newton procedure to delete boundary points of $X^{(0)}$ since the global minimum need not be a stationary point if it occurs on the boundary. We discuss this point further in Sect. 10.

If we were to use this Newton method only, we would in general find stationary points of f which were not minima. Moreover, we would find local minima which were not global minima. To avoid this, we use an additional procedure to delete points where f exceeds the smallest known value $\bar{f}$. This procedure is described in the next section.

In some applications, it may be desirable to find all the stationary points of f in a given box. This can be done using the Newton method alone or in conjunction with the monotonicity check of Sect. 9. If, in addition, the convexity check of Sect. 5 were used, all stationary points except maximum would be found.

7. Bounding f

We now consider how to delete points $y \in X$ where we know $f(y) > \bar{f}$ and hence where $f(y)$ is not a global minimum. We retain the complementary set which is a sub-box (or sub-boxes) $Y \subset X$ wherein $f(y)$ may be $\leqq \bar{f}$.

As pointed out in [5], if we only wish to bound f^* and not x^*, we can delete points where

$$f(y) > \bar{f} - \varepsilon_1 \tag{7.1}$$

for some $\varepsilon_1 > 0$. We can allow ε_1 to be nonzero only if we do not need to know the point(s) x^* at which f is globally minimum.

We want to retain points where (7.1) is not satisfied. From (3.1), this is the case for points y if

$$f(x) + (y-x)^T g(x) + \tfrac{1}{2}(y-x)^T H(\xi)(y-x) \leqq \bar{f} - \varepsilon_1$$

because the left member equals $f(y)$. Denote

$$E = \bar{f} - f(x) - \varepsilon_1.$$

Then

$$\tilde{y}^T g(x) + \tfrac{1}{2}\tilde{y}^T H(\xi)\,\tilde{y} \leqq E \tag{7.2}$$

where $\tilde{y} = y - x$. We shall use this relation to reduce X in one dimension at a time to yield the sub-box(es) Y resulting from deleting points where $f(y) > \bar{f} - \varepsilon_1$.

We shall illustrate the process for the case $n = 2$. The higher dimensional case follows in the same way. For $n = 2$, (7.2) becomes

$$\tilde{y}_1 g_1(x) + \tilde{y}_2 g_2(x) + \tfrac{1}{2}[\tilde{y}_1^2 h_{11}(\xi) + \tilde{y}_1 \tilde{y}_2 h_{21}(\xi) + \tilde{y}_2^2 h_{22}(\xi)] \leqq E. \tag{7.3}$$

We first wish to reduce X in the x_1-direction. Thus we solve this relation for acceptable values of y_1. After collecting terms in y_1, we replace y_2 by X_2. In the higher dimensional case we would also replace y_i by X_i for all $i = 3, \ldots, n$. We also replace ξ by X (since $\xi \in X$). We obtain

$$\tilde{y}_1[g_1(x) + \tfrac{1}{2}\tilde{X}_2 h_{21}(X)] + \tfrac{1}{2}\tilde{y}_1^2 h_{11}(X) + \tilde{X}_2 g_2(x) + \tfrac{1}{2}\tilde{X}_2^2 h_{22}(X) - E \leqq 0 \tag{7.4}$$

where $\tilde{X}_2 = X_2 - x_2$.

We solve this quadratic for the interval or intervals of points y_1 as described below. Call the resulting set Z_1. Since we are only interested in points with $y_1 \in X_1$, we compute the desired set Y_1 as $Y_1 = X_1 \cap Z_1$.

For the sake of argument, suppose Y_1 is a single interval. We can then try to reduce X_2 the same way we (hopefully) reduced X_1 to get Y_1. We again rewrite (7.3). This time we replace y_1 by Y_1 and (as before) ξ by X. We could obtain better results by replacing ξ_1 by Y_1 rather than X_1 but this would require re-evaluation of the elements of H. We obtain

$$\tilde{y}_2[g_2(x)+\tfrac{1}{2}\tilde{Y}_1 h_{21}(X)]+\tfrac{1}{2}\tilde{y}_2^2 h_{22}(X)+\tilde{Y}_1 g_1(x)+\tfrac{1}{2}\tilde{Y}_1^2 h_{11}(X)-E\leqq 0 \quad (7.5)$$

where $\tilde{Y}_1 = Y_1 - x_1$.

If the solution set Y_2 is strictly contained in X_2, we could replace X_2 by Y_2 in (7.4) and solve for a new Y_1. We have not tried to do this in practice. Instead, we start over with the box Y in place of X as soon as we have tried to reduce each X_i to Y_i $(i=1, \ldots, n)$. Note this means we re-evaluate $H(X)$.

We now consider how to solve the quadratic equation (7.4) or (7.5). These have the general form

$$A+Bt+Ct^2\leqq 0 \quad (7.6)$$

where A, B, and C are intervals and we seek values of t satisfying this inequality.

Denote $C=[c_1, c_2]$ and let c be an arbitrary point in C. Similarly, let $a\in A$ and $b\in B$ be arbitrary. Suppose t is such that (7.6) is violated; that is $Q(t)>0$, where

$$Q(t)=a+bt+ct^2.$$

If this is true for $c=c_1$, then it is true for all $c\in C$. Hence if we wish to find the complementary values of t where (7.6) might hold we need only consider

$$A+Bt+c_1 t^2\leqq 0. \quad (7.7)$$

If $c_1=0$, this relation is linear and the solution set T is as follows: Denote $A=[a_1, a_2]$ and $B=[b_1, b_2]$. Then the set of solution points t is

$$T=\begin{cases} [-a_1/b_2, \infty] & \text{if } a_1\leqq 0, b_2<0, \\ [-a_1/b_1, \infty] & \text{if } a_1>0, b_1<0, b_2\leqq 0, \\ [-\infty, \infty] & \text{if } a_1\leqq 0, b_1\leqq 0\leqq b_2, \\ [-\infty, -a_1/b_2]\cup[-a_1/b_1, \infty] & \text{if } a_1>0, b_1<0<b_2, \\ [-\infty, -a_1/b_1] & \text{if } a_1\leqq 0, b_1>0, \\ [-\infty, -a_1/b_2] & \text{if } a_1>0, b_1\geqq 0, b_2>0, \\ \text{empty set} & \text{if } a_1>0, b_1=b_2=0. \end{cases}$$

Recall that we will intersect T with X_i for some value of i. Thus although T may be unbounded, the intersection is bounded.

If $c_1\neq 0$, the quadratic (7.6) may have no solution or it may have a solution set T composed of either one or two intervals. In the latter case, the intervals may be semi-infinite. However, after intersecting T with X_i, the result is finite.

Denote

$$Q_1(t)=a+bt+c_1t^2$$

where $a\in A$, $b\in B$, and c_1 is the left endpoint of C. We shall delete points t where $Q_1(t)>0$ for all $a\in A$ and $b\in B$. Thus we retain a set T of points where $Q_1(t)\leqq 0$, as desired. But we also retain (in T) points where, for fixed t, $Q_1(t)>0$ for some $a\in A$ and $b\in B$ and $Q_1(t)\leqq 0$ for other $a\in A$ and $b\in B$. This same criterion was used to obtain T when $c_1=0$. This assures that we shall always retain points in X_i where $f(x)$ is a minimum.

Denote

$$q_1(t)=\begin{cases}a_1+b_2t+c_1t^2 & \text{if } t\leqq 0,\\ a_1+b_1t+c_1t^2 & \text{if } t\geqq 0\end{cases}$$

and

$$q_2(t)=\begin{cases}a_2+b_1t+c_1t^2 & \text{if } t\leqq 0,\\ a_2+b_2t+c_1t^2 & \text{if } t\geqq 0.\end{cases}$$

Then we can write the interval quadratic as

$$\begin{aligned}Q_1(t)&=[a_1,a_2]+[b_1,b_2]\,t+c_1t^2\\&=[q_1(t),q_2(t)].\end{aligned}$$

Thus for any finite t, $q_1(t)$ is a lower bound for $Q_1(t)$ and $q_2(t)$ is an upper bound for $Q_1(t)$ for any $a\in A$ and any $b\in B$.

For a given value of t, if $q_1(t)>0$, then $Q_1(t)>0$ for all $a\in A$ and $b\in B$. Hence we need only to solve the real quadratic equation $q_1(t)=0$ in order to determine intervals wherein, without question, $Q_1(t)>0$. This is a straightforward problem.

The function $q_1(t)$ is continuous but its derivative is discontinuous at the origin when $b_1\neq b_2$ which will generally be the case in practice. Hence we must consider the cases $t\leqq 0$ and $t\geqq 0$ separately.

If $c_1>0$, the curve $q_1(t)$ is convex for $t\leqq 0$ and convex for $t\geqq 0$. Consider the case $t\leqq 0$. If $q_1(t)$ has real roots, then $Q_1(t)>0$ outside these roots, provided $t\leqq 0$. Hence, we retain the interval between these roots. We need only examine the discriminant of $q_1(t)$ to determine whether the roots are real or not. Hence it is a simple procedure to determine which part (if any) of the half line $t\leqq 0$ can be deleted. The same procedure can be used for $t\geqq 0$.

For $c_1<0$, $q_1(t)$ is concave for $t\leqq 0$ and for $t\geqq 0$. In this case we can delete the interval (if any) between the roots of $q_1(t)$ in each half line. The set T is the complement of this interval. It is composed of two semi-infinite intervals.

In determining T for either the case $c_1<0$ or in the case $c_1>0$, it is necessary to know whether the discriminant of $q_1(t)$ is non-negative or not. Denote

$$\Delta_1=b_1^2-4a_1c_1,\qquad \Delta_2=b_2^2-4a_1c_1.$$

These are the discriminants of $q_1(t)$ when $t\geqq 0$ and $t\leqq 0$, respectively.

When we compute Δ_1 or Δ_2, we shall make rounding errors. Thus we should compute them using interval arithmetic to bound these errors. When computing $\Delta_i=(i=1,2)$, suppose we obtain the interval

$$\Delta_i^I=[\Delta_i^I,\Delta_i^R]\qquad (i=1,2).$$

We use the appropriate endpoint of Δ_1^I or Δ_2^I to determine T which assures that we never delete a point t where $Q_1(t)$ could be non-positive. Thus we use the endpoint of Δ_1^I or Δ_2^I which yields the larger set T.

When we compute the roots of $q_1(t)$, we shall make rounding errors. Hence we compute them using interval arithmetic and again use the endpoints which yield the larger set T to assure we do not delete a point in X_i where f is a minimum.

For $i=1$ and 2, denote

$$R_i^{\pm} = (-b_i \pm \Delta_i^{1/2})/(2c_1)$$

and

$$S_i^{\pm} = 2a(-b_i \pm \Delta_i^{1/2}).$$

Note that $R_i^+ = S_i^-$ and $R_i^- = S_i^+$. As is well known, the rounding error is less if we compute a root in the form R_i^+ rather than in the form S_i^- when $b_i < 0$. The converse is true when $b_i > 0$. Similarly, the rounding error is less when using R_i^- rather than S_i^+ when $b_i > 0$. Hence we compute the roots of $q_1(t)$ as R_i^+ and S_i^+ when $b_i < 0$ and as R_i^- and S_i^- when $b_i > 0$.

Note that computing $R_i^{\pm}$ or $S_i^{\pm}$ involves taking the square root of the interval Δ_i^I. In exact arithmetic this would be the real quantity Δ_i. We would never be computing roots of $q_1(t)$ when Δ_i was negative. Hence if we find that the computed result Δ_i^I contains zero, we can replace it by its non-negative part. Thus we will never try to take the square root of an interval containing negative numbers.

Given any interval I, let I^L and I^R denote its left and right endpoint, respectively. We use this notation below. Using the above prescriptions on how to compute the set T, we obtain the following results:

For $b_1 \geqq 0$ and $c_1 > 0$,

$$T = \begin{cases} \emptyset \text{ (the empty set)} & \text{if } \Delta_2^R < 0, \\ [(R_2^-)^L, (S_2^-)^R] & \text{if } a_1 > 0 \text{ and } \Delta_2^R \geqq 0, \\ [(R_2^-)^L, (S_1^-)^R] & \text{if } a_1 \leqq 0. \end{cases} \tag{7.8}$$

For $b_2 \leqq 0$ and $c_1 > 0$,

$$T = \begin{cases} \emptyset & \text{if } \Delta_1^R < 0, \\ [(S_1^+)^L, (R_1^+)^R] & \text{if } a_1 > 0 \text{ and } \Delta_1^R \geqq 0, \\ [(S_2^+)^L, (R_1^+)^R] & \text{if } a_1 \leqq 0. \end{cases} \tag{7.9}$$

For $b_1 \leqq 0 \leqq b_2$ and $c_1 > 0$,

$$T = \begin{cases} \emptyset & \text{if } \max(\Delta_1^R, \Delta_2^R) < 0, \\ [(R_2^-)^L, (S_2^-)^R] & \text{if } |b_1| < b_2 \text{ and } \min(\Delta_1^R, \Delta_2^R) \leqq 0 \\ & \quad \leqq \max(\Delta_1^R, \Delta_2^R), \\ [(S_1^+)^L, (R_1^+)^R] & \text{if } |b_1| > b_2 \text{ and } \min(\Delta_1^R, \Delta_2^R) \leqq 0 \\ & \quad \leqq \max(\Delta_1^R, \Delta_2^R), \\ [(R_2^-)^L, (S_2^-)^R] \cup [(S_1^+)^L, (R_1^+)^R] & \text{if } a_1 > 0 \text{ and } \min(\Delta_1^R, \Delta_2^R) > 0 \\ [(R_2^-)^L, (R_1^+)^R] & \text{if } a_1 \leqq 0. \end{cases} \tag{7.10}$$

For $b_1 \geqq 0$ and $c_1 < 0$,

$$T = \begin{cases} [-\infty, (S_2^-)^R] \cup [(R_1^-)^L, \infty] & \text{if } a_1 > 0, \\ [-\infty, (S_1^-)^R] \cup [R_1^-)^L, \infty] & \text{if } a_1 \leqq 0 \leqq \Delta_1^L, \\ [-\infty, \infty] & \text{if } \Delta_1^L < 0. \end{cases} \tag{7.11}$$

For $b_2 \leqq 0$ and $c_1 < 0$,

$$T = \begin{cases} [-\infty, (R_2^+)^R] \cup [(S_1^+)^L, \infty] & \text{if } a_1 > 0, \\ [-\infty, (R_2^+)^R] \cup [(S_2^+)^L, \infty] & \text{if } a_1 \leqq 0 \leqq \Delta_2^L, \\ [-\infty, \infty] & \text{if } \Delta_2^L < 0. \end{cases} \tag{7.12}$$

For $b_1 \leqq 0 \leqq b_2$ and $c_1 < 0$,

$$T = \begin{cases} [-\infty, (S_2^-)^R] \cup [(S_1^+)^L, \infty] & \text{if } a_1 \geqq 0, \\ [-\infty, \infty] & \text{if } a_1 < 0. \end{cases} \tag{7.13}$$

Note that if $c_1 > 0$, then Δ_2 can be negative only if $a_1 > 0$. Hence the condition $\Delta_2 < 0$ implies $a_1 > 0$. This, and similar cases, has been used to shorten the conditional statements in the above expressions for T.

We have seen that the solution of the quadratic inequalities such as (7.4) or (7.5) can be an interval Z_i or two semi-infinite intervals, say $Z_i^{(1)}$ and $Z_i^{(2)}$. The desired solution set Y_i is obtained by intersection with X_i. In the former case, $Y_i = X_i \cap Z_i$ can be empty or a single interval. In the latter case,

$$Y_i = X_i \cap (Z_i^{(1)} \cup Z_i^{(2)})$$

can be empty, a single interval, or two intervals. We now consider the logistics of handling these cases.

The quadratic inequality to be solved for Z_i will have quadratic term $\frac{1}{2}\tilde{y}_i^2 h_{ii}(X)$ so the interval C in (7.6) is $\frac{1}{2}h_{ii}(X)$ and the left endpoint is $c_1^{(i)} = [\frac{1}{2}h_{ii}(X)]^L$. If $c_1^{(i)} \geqq 0$ the solution set is a single interval. But if $c_1^{(i)} < 0$, it is two semi-infinite intervals and it may be that Y_i will be two intervals. This would complicate the process of finding $Y_{i+1}, \ldots, Y_n$. Thus we proceed as follows.

Let I_1 denote the set of indices i for which $c_1^{(i)} \geqq 0$ and I_2 denote the set of indices i for which $c_1^{(i)} < 0$. We first find Y_i for each $i \in I_1$. We then begin to find Y_i for $i \in I_2$. Let $j \in I_2$ be such that Y_j is composed of two intervals, say $Y_j^{(1)}$ and $Y_j^{(2)}$. Then X_j is the smallest interval containing both $Y_j^{(1)}$ and $Y_j^{(2)}$. When finding Y_i for the remaining values of i, we use X_j in place of Y_j.

After finding all Y_i for $i = 1, \ldots, n$, we wish to use the fact that we can delete the interval, say Y_j^c, between $Y_j^{(1)}$ and $Y_j^{(2)}$. We would like to do this for all the values of j for which Y_j was two intervals. However, it could be that this occurred for all the indices $j = 1, \ldots, n$. After deleting the interior interval Y_j^c in each dimension, the resulting set would be composed of 2^n boxes. For large n, this is too many boxes to handle separately. Hence we delete only a few (one, two, or three) of the largest of the intervals Y_j^c. We then process each of the new boxes separately.

We would like to prevent the generation of long, narrow boxes. Thus a good choice of which Y_j^c to delete is the one(s) corresponding to the component for which the smallest interval containing both $Y_j^{(1)}$ and $Y_j^{(2)}$ is largest. However, we have chosen to delete the largest interval Y_j^c.

Let us call the process we have described in this section the *quadratic method.* We can combine the quadratic method with the Newton method. It is desirable to do this as we now explain.

If the left endpoint of $H_{ii}(X)$ is negative, then the quadratic method can give rise to two new intervals $Y_i^{(1)}$ and $Y_i^{(2)}$ in place of X_i. When trying to improve X_{i+1} (say), it is impractical to use $Y_i^{(1)}$ and $Y_i^{(2)}$ separately and we use X_i, instead. Thus the improvement of X_i is of no help when trying to improve X_{i+1}, etc. Similarly, when applying the Newton step, if $J_{ii}(X)$ contains zero as an interior point, we can obtain two subintervals in place of X_i. Again, we cannot conveniently use this fact in the remaining part of the Newton step.

We would like to do those steps first which are of help in subsequent steps. Hence the following sequence is suggested. First try to improve X_i by the quadratic method for each value of $i=1, \ldots, n$ for which the left endpoint of $H_{ii}(X)$ is positive. Then apply the interval Newton method to the (old or new) components for which $0 \notin BJ_{ii}(X)$ $(i=1, \ldots, n)$. Next use the quadratic method for those components for which the left endpoint of $H_{ii}(X)$ is non-positive. Finally, complete the Newton step for those components with $0 \in BJ_{ii}(X)$.

At each stage of either method, when trying to improve the i-th component of the box, we use the currently best interval for the other components. This may be the smallest interval containing two disjoint intervals in some cases. In fact it would be possible for the quadratic method and the Newton method to each delete disjoint sub-intervals for a given component. This would give rise to three sub-intervals to be retained. However, it seems better to simplify this case and only delete the larger of the two sub-intervals.

When both methods are completed, we may have several components divided into two sub-intervals. If so, we find the one for which the largest interior sub-interval has been deleted. We replace all the others by the smallest sub-interval containing the two disjoint parts. We then divide the remaining part of the current box into two sub-boxes by deleting the sub-interval for the component in question. We could do this for more than one component, but each deletion would double the number of boxes. It seems better to keep the number of boxes small.

8. Choice of ε_1

Suppose we want to bound the value f^* of the global minimum to within a tolerance ε_1 but we do not care where f takes on this value. Then, as pointed out in Sect. 7, we can delete points y where

$$f(y) > \bar{f} - \varepsilon_1. \tag{8.1}$$

Once we have found a point $\bar{x}$ where $f(\bar{x}) = \bar{f}$ is such that $\bar{f} - f^* \leqq \varepsilon_1$, our algorithm will eventually delete all of $X^{(0)}$ if we use (8.1). However, $\bar{x}$ may be far

from the point x^* where f is globally minimal. When all of $X^{(0)}$ is deleted, we will know that

$$\bar{f} - \varepsilon_1 \leqq f^* \leqq \bar{f}.$$

Choosing $\varepsilon_1 > 0$ will speed up our algorithm. However, if we wish to obtain good bounds on x^*, we must choose $\varepsilon_1 = 0$. We then terminate our algorithm when the remaining set of points is sufficiently small. See Sect. 13 for a termination procedure.

9. Monotonicity

Another step in our algorithm makes use of the monotonicity of f. Suppose, for example, the i-th component $g_i(x)$ of the gradient is non-negative for all $x \in X$. Then the smallest value of $f(x)$ for $x \in X$ must occur for x_i equal to the left endpoint of X_i.

To make use of a fact such as this, we evaluate $g_i(X_1, \ldots, X_n)$. The resulting interval, which we denote by $[\sigma_i, \omega_i]$, contains $g_i(x)$ for all $x \in X$. Denote $X_i = [x_i^L, x_i^R]$. If $\sigma_i \geqq 0$, $f(x)$ is smallest (in X) for $x_i = x_i^L$. Hence we can delete all of X except the points with $x_i = x_i^L$. If $\sigma_i > 0$, $f(x)$ cannot have a stationary point in X. Hence we can delete *all* of X unless the boundary at $x_i = x_i^L$ contains boundary points of the initial box $X^{(0)}$. Similar results occur if $\omega_i \leqq 0$ or if $\omega_i < 0$.

We evaluate $g_i(X_1, \ldots, X_n)$ for $i = 1, \ldots, n$ and reduce the dimensionality of X for any value of i for which $\sigma_i \geqq 0$ or $\omega_i \leqq 0$. Of course, we delete all of X, if possible.

It is possible that we can reduce X in every dimension in this way. If so, only a single point, say $\tilde{x}$, remains. In this case, we evaluate $f(\tilde{x})$. If $f(\tilde{x}) > \bar{f}$, we can eliminate $\tilde{x}$ and hence all of X is deleted by the process. If $f(\tilde{x}) \leqq \bar{f}$, we reset $\bar{f}$ equal to $f(\tilde{x})$. In this latter case, X is again deleted; but we store $\tilde{x}$ for future reference.

10. Boundary Points

The process just described in Sect. 9 can sometimes eliminate points of the boundary of $X^{(0)}$ which lie in X. Suppose that for some X, we find $g_j(X_1, \ldots, X_n) > 0$ for some $j = 1, \ldots, n$. Then we can delete all of X except for any boundary points of $X^{(0)}$ occurring at $x_j = x_j^L$. Any other boundary points of $X^{(0)}$ which are in X are thus deleted.

The quadratic method of Sect. 7 deletes any point x where $f(x) > \bar{f}$ whether x lies on the boundary of $X^{(0)}$ or not. However, the Newton method of Sect. 6 and the procedure in Sect. 5 (which considers convexity) cannot delete any boundary points of $X^{(0)}$.

Suppose we apply a step of the Newton method to a box X and obtain a new box X' contained in X. A simple way to proceed is to retain the smallest box containing both X' and all boundary points of $X^{(0)}$ which are in X. This will

generally save points of X outside X' thus reducing the efficiency of the procedure. In fact, it may be that the smallest box containing the boundary points of $X^{(0)}$ which are in X is X itself. If this were the case, we would bypass the Newton step for the box X. This approach would rely upon the methods of Sects. 7 and 9 to delete boundary points of $X^{(0)}$.

This same idea can be used for the method of Sect. 5. If f is not convex in X, we can simply replace X by the smallest (perhaps degenerate) box, say $\bar{X}$, containing the boundary points of $X^{(0)}$ which lie in X. In this case, either $\bar{X}=X$ or else $\bar{X}$ is a degenerate box of dimension less than that of X.

Suppose we are given a box X. For this approach, if $\bar{X}=X$, we do not apply either the Newton method or the convexity test. We could use the Newton method in this case also or we might bypass its use whenever X contains boundary points of $X^{(0)}$.

A more straightforward procedure is to simply express the boundary of $X^{(0)}$ as $2n$ separate (degenerate) boxes of dimension $n-1$. The interior of $X^{(0)}$ can then be treated as a box wherein the global minimum must be a stationary point. However, the $(n-1)$-dimensional faces of $X^{(0)}$ have $(n-2)$-dimensional boundaries which must, in turn, be separated from the interiors, and so on. Finally the vertices of $X^{(0)}$ would have to be separated. These vertices alone are 2^n points. Even for moderate values of n, this separation process produces too many (degenerate) boxes. Thus it is better, in general, not to try to separate the boundaries from $X^{(0)}$.

These two approaches represent extreme cases. Intermediate methods might be used wherein the boundaries of $X^{(0)}$ in a given box X are separated off under special circumstances.

It should, perhaps, be pointed out that the Newton method can delete boundary points of $X^{(0)}$ under certain circumstances. Suppose our algorithm has produced a degenerate box X which is all or part of a face of $X^{(0)}$. In this degenerate $(n-1)$-dimensional box, the Newton method can delete points which are not in the $(n-2)$-dimensional boundary of the face of $X^{(0)}$. Such deleted points are, of course, on the boundary of $X^{(0)}$.

In some examples, we shall know a priori that the global minimum is a stationary point. In this case we are free to delete boundary points by any of our procedures.

11. The List of Boxes

When we begin our algorithm, we shall have a single box $X^{(0)}$. We apply the four procedures described in Sects. 5, 6, 7, and 9 to this box. It is possible that none of these procedures can delete any of $X^{(0)}$. If so, we divide $X^{(0)}$ in half in a direction of its maximum width. We put one of these new boxes in a list L to be processed and work on the other. These and subsequent boxes may also have to be subdivided, thus adding to the list L of boxes yet to be processed. If boundary points are handled appropriately, the four procedures described in Sects. 5, 6, 7 and 9 can each produce more than one new sub-box; and all but one are added to the list. Thus the number of boxes in the list tends to grow initially.

Eventually, however, the boxes become small and often a box is entirely eliminated. Thus the number of boxes in the list eventually decreases to one, or just a few, or to none at all when ε_1 is chosen to be nonzero.

12. Subdividing a Box

In the initial stages of our algorithm, we shall be applying it to large boxes. For example, we begin by applying it to the entire initial box $X^{(0)}$. Thus it could be that, for a given box X, none of the procedures described in Sects. 5, 6, 7, and 9 can delete any of X. When this occurs, we wish to subdivide X.

We could subdivide each component X_i of X into two parts. But this would give rise to 2^n sub-boxes. To prevent generation of too many sub-boxes, we shall divide only one component in half. It is best to subdivide the largest component X_i to prevent generation of a long, narrow box.

Suppose we divide $X_i=[x_i^L, x_i^R]$ in half giving two new boxes X' and X'' whose i-th components are $X_i'=[x_i^L, \bar{x}_i]$ and $X_i''=[\bar{x}_i, x_i^R]$, respectively, where $\bar{x}_i=(x_i^L+x_i^R)/2$. The boxes X' and X'' have a boundary in common at $x_i=\bar{x}_i$. If f had a global minimum on this common boundary, we would subsequently find it twice. This is unlikely to be the case. To avoid having the same points in two boxes, we could define one of them in terms of a half open interval. Thus we could define $X_i'=[x_i^L, \bar{x}_i)$.

It is simpler to always use closed intervals. The extra work of keeping track of whether an interval contains a given endpoint is probably not worth the effort. In practice, we have elected to avoid this problem. Thus we have always used closed intervals only. In general, this does not cause the algorithm to find a given global minimum more than once.

13. Termination

If we have chosen $\varepsilon_1>0$, we can continue our algorithm until $X^{(0)}$ is entirely eliminated. As pointed out in Sect. 8, we then have f^* bounded to within a error ε_1. In this case, we do not obtain a bound on x^*.

If $\varepsilon_1=0$, we cannot eliminate all of $X^{(0)}$ since we always retain a box or boxes containing the point(s) x^* where $f(x^*)=f^*$. As pointed out above, we might also retain a box or boxes wherein f has a value very near f^* but no value equal to f^*.

Suppose that at some stage in our algorithm, the list L contains s boxes. Denote these boxes by $X^{(1)}, \ldots, X^{(s)}$. Let $X_j^{(i)}$ denote the interval defining the j-th component (dimension) of $X^{(i)}$. Let $w(X_j^{(i)})$ denote the width of the interval $X_j^{(i)}$ and let

$$w_i = \max_{1 \leqq j \leqq n} [w(X_j^{(i)})].$$

That is, w_i is the maximum dimension of $X^{(i)}$.

We could continue processing boxes in the list L by our algorithm until

$$\sum_{i=1}^{s} w_i \leqq \varepsilon_2 \tag{13.1}$$

for some $\varepsilon_2 > 0$. This is provided ε_2 is chosen large enough that the prescribed precision is attainable using (say) single precision arithmetic. However, it is more convenient computationally to require only that

$$w_i \leqq \varepsilon_2 \tag{13.2}$$

for each $i = 1, \ldots, s$. If $\varepsilon_1 > 0$, we set $\varepsilon_2 = 0$ for convenience.

Thus whenever a new box $X^{(i)}$ is obtained by our algorithm we can check whether (13.2) is satisfied. If so, we no longer apply our algorithm to $X^{(i)}$ (except as discussed below). If $X^{(i)}$ contains a point x^* where f is a global minimum, then the location of x^* is bounded. In fact, if $x^{(i)}$ is the center of $X^{(i)}$ and (13.2) holds for $X^{(i)}$, then

$$|x_j^{(i)} - x_j^*| \leqq \varepsilon_2/2 \qquad (j = 1, \ldots, n).$$

Let s denote the number of boxes remaining and denote the boxes by $X^{(i)}$ $(i = 1, \ldots, s)$. As a final step, we want to assure that f^* is bounded sufficiently sharply. We do this as follows.

For each $i = 1, \ldots, s$ we evaluate $f(X^{(i)})$; that is we evaluate f with interval arguments $X_j^{(i)}$ $(j = 1, \ldots, n)$. The result, say $[F_i^L, F_i^R]$ contains the range of f for all $x \in X^{(i)}$, but will not be sharp, in general. However, if ε_2 was chosen to be small, the interval result should be 'close to sharp' since ε_2 is an upper bound on the largest dimension of any box $X^{(i)}$; and the smaller $X^{(i)}$ is, the sharper $[F_i^L, F_i^R]$ is. (See [8].) Therefore, it is generally not necessary to use special procedures to sharpen the computed interval.

Since $[F_i^L, F_i^R]$ contains the range of $f(x)$ for all $x \in X^{(i)}$, we have

$$F_i^L \leqq f(x) \leqq F_i^R$$

for any $x \in X^{(i)}$. Denote

$$\underline{f} = \min_{1 \leqq i \leqq s} F_i^L$$

Then

$$\underline{f} \leqq f(x)$$

for any x in any of the boxes $X^{(i)}$ $(i = 1, \ldots, s)$. Therefore, since any global minimum must occur at a point x^* lying in one of the boxes $X^{(i)}$, we have $\underline{f} \leqq f^*$. But also $f^* \leqq \bar{f}$ (as discussed above) and hence

$$\underline{f} \leqq f^* \leqq \bar{f}. \tag{13.3}$$

We thus have bounds on f^*. However, they may not be sharp enough since our specified requirement is to bound f^* to within, say, ε_0; and it may be that $\bar{f} - \underline{f} > \varepsilon_0$. If this is the case, we shall improve our bounds. To do this, we find a

value j for which $\underline{f}=f_j^L$. We apply our main algorithm to $X^{(j)}$. This will increase f_j^L, in general. It might also decrease $\bar{f}$. For exact interval arithmetic, this must decrease $\bar{f}-f_j^L$ since $X^{(j)}$ will be reduced in size (even if it is merely subdivided). Repeating this step for each j such that $\underline{f}=f_j^L$, we must decrease $\bar{f}-\underline{f}$ (at least, if exact interval arithmetic is used) and hence eventually have

$$\bar{f}-\underline{f}\leqq\varepsilon_0$$

so that f^* is bounded to sufficient accuracy since (13.3) holds.

Because of rounding errors, we cannot reduce $\bar{f}-\underline{f}$ arbitrarily, in practice. Hence we assume ε_0 is chosen commensurate with achievable accuracy using (say) single precision arithmetic.

We also require that

$$F_i^R-F_i^L\leqq\varepsilon_3$$

for each box $X^{(i)}$ $(i=1,\ldots,s)$. For convenience, we can choose $\varepsilon_3=\varepsilon_0$ to reduce the number of quantities to be specified. Note that $F_i^L\leqq\bar{f}$ since otherwise $f(x)>\bar{f}$ for all $x\in X^{(i)}$ in which case $X^{(i)}$ can be deleted. Hence

$$F_i^R\leqq\bar{f}+\varepsilon_3\leqq\underline{f}+\varepsilon_0+\varepsilon_3\leqq f^*+\varepsilon_0+\varepsilon_3$$

for every $i=1,\ldots,s$. That is, every remaining box contains a point x at which $f(x)$ differs from f^* by no more than $\varepsilon_0+\varepsilon_3$.

14. The Steps of the Algorithm

We now describe the steps involved in our algorithm. Initially, the list L of boxes to be processed consists of a single box $X^{(0)}$. In general, divide the list L into the list L_1 of intervals $X^{(i)}$ satisfying the condition $w_i\leqq\varepsilon_2$ (see (13.2)) and a list L_2 which do not satisfy this condition.

We assume we have evaluated f at the center of $X^{(0)}$ and thus obtained an initial value for $\bar{f}$. The subsequent steps are to be done in the following order except as indicated by branching:

(1) Of the boxes in L_2, choose one which has been in L_2 longest. Call it X. If L_2 is empty, go to step (11) if $\varepsilon_1>0$. If L_2 is empty and $\varepsilon_1=0$ choose a box which has been in L_1 longest and go to step 2. If both L_1 and L_2 are empty, print the bounds $\bar{f}-\varepsilon_1$ and $\bar{f}$ on f^* and stop.

(2) Check for monotonicity. Evaluate $g(X)$ as described in Sect. 9. For $i=1,\ldots,n$, if $g_i(X)>0$ (<0) and the boundary of X at $x_i=x_i^L$ $(=x_i^R)$ does not contain a boundary point of $X^{(0)}$, delete X and go to step (1). Otherwise, if $g_i(X)\geqq 0$ $(\leqq 0)$, replace $X_i=[x_i^L, x_i^R]$ by $[x_i^L, x_i^L]$ $([x_i^R, x_i^R])$. Rename the result X again.

(3) Test for non-convexity as in Sect. 5. Let X' denote the smallest box in X containing all the boundary points of $X^{(0)}$ which lie in X. If $X'=X$ go to step 4. Otherwise, evaluate $h_{ii}(X_1,\ldots,X_n)$ for $i=1,\ldots,n$. If the resulting interval is strictly negative for any value of i, replace X by X'. If X' is empty, go to step 1. If X' is not empty, put it in the list L and go to step 1.

(4) Begin use of the quadratic method of Sect. 7. For those values of $i = 1, \ldots, n$ for which the left endpoint of $H_{ii}(X)$ is non-negative, solve the quadratic for the interval Y_i to replace X_i. Rename the result X_i.

(5) Begin use of the Newton method. For those values of $i = 1, \ldots, n$ for which $0 \notin [BJ(X)]_{ii}$, solve for the new interval to replace X_i. Rename the result X_i. For a given value of i, omit this step if a reduction of X_i will delete boundary points of $X^{(0)}$ from the box X.

(6) Complete the quadratic method. For those values of i not used in step 4, solve the quadratic for Y_i. If Y_i is a single interval, replace X_i by Y_i, renaming it X_i. Otherwise save Y_i for use in step 8.

(7) Complete the Nowton method. For those values of i not used in step 5, solve for the new set (say) Y_i'. For a given value of i, omit this step if a reduction of X_i will delete boundary points of $X^{(0)}$ from the box X. If Y_i' is a single interval, replace X_i by Y_i', renaming it X_i.

(8) Combine the results from the quadratic and Newton methods for those components X_i for which both methods divided X_i into two sub-intervals. That is, find the intersection Y_i'' of Y_i from step 6 and Y_i' from step 7. If Y_i'' is composed of three intervals, replace it by either Y_i or Y_i', whichever has the smallest intersection with X_i. Of all the Y_i'', save the one (or two or three) which deletes the largest subinterval of X_i. That is, save that Y_i'' whose complement in X_i is largest. Let j be its index. For all relevant values of $i \neq j$, replace Y_i'' by X_i, that is, ignore the fact that part of X_i could be deleted.

(9) If Y_j'' exists; that is, if at least one interval X_j was divided into two sub-intervals, say $Y_j^{(1)}$ and $Y_j^{(2)}$, subdivide the box X into two sub-boxes. These sub-boxes will have the same components X_i as X except one will have j-th component $Y_j^{(1)}$ and the other will have j-th component $Y_j^{(2)}$. If no such Y_j'' exists, we may wish to subdivide the current box. Let X denote the box chosen in step 1 and let X'' denote the current box resulting from applying steps 2 through 8 to X. If the improvement of X'' over X is so small that (say)

$$w(X'') > 0.75\, w(X),$$

then divide X'' in half in its greatest dimension.

(10) Evaluate f at the center of the box or boxes resulting after step 9. Update $\bar{f}$ as described in Sect. 4. Put the box(es) in the list L and go to step 1.

(11) Evaluate $f(X^{(i)})$ for each remaining box $X^{(i)}$ in L. Denote the result by $[F_i^L, F_i^R]$. If $F_i^R - F_i^L > \varepsilon_0$ for any value of i, use $X^{(i)}$ for X and go to step 4. If $F_i^R - F_i^L \leqq \varepsilon_0$ for all $i = 1, \ldots, s$, find

$$\underline{f} = \min_{1 \leqq i \leqq s} F_i^L.$$

Then print the bounds $\underline{f}$ and $\bar{f}$ on f^* and stop.

For $\varepsilon_1 > 0$, these steps bound f^* to within an error ε_1. If $\varepsilon_1 = 0$, they found f^* to within ε_0; they bound x^* to within ε_2; and they assure that for any point x in any final box, $f(x)$ exceeds f^* by no more than $\varepsilon_0 + \varepsilon_3$.

In this step, we sometimes branch to step 4. Note that we could go to step 2, but it is unlikely that either step 2 or step 3 will be helpful. This is because we expect each box remaining at this stage to contain a minimum.

15. A Numerical Example

We now illustrate the steps of our algorithm. We shall consider the so-called three hump camel function.

$$f(x)=2x_1^2-1.05x_1^4+\tfrac{1}{6}x_1^6-x_1x_2+x_2^2 \tag{15.1}$$

which has three minima and two saddle points. The gradient $g(x)$ has components

$$\begin{aligned} g_1(x)&=4x_1-4.2x_1^3+x_1^5-x_2,\\ g_2(x)&=2x_2-x_1. \end{aligned} \tag{15.2}$$

The interval Jacobian $J(X)$ (see Sect. 3) has elements

$$\begin{aligned} J_{11}(X)&=4-12.6X_1^2+5X_1^4,\\ J_{12}(X)&=J_{21}(X)=-1,\\ J_{22}(X)&=2. \end{aligned} \tag{15.3}$$

As described in [4], a better formulation for $J(X)$ could be derived which would give smaller intervals, in general. However, we shall use the simpler form given here.

Suppose that the box we choose in step 1 has first component $X_1=[1, 1.1]$. In step 2 we find that, whatever X_2 is,

$$h_{11}(X_1, X_2)=[-5.196, -2.55].$$

Since this is strictly negative, we know that f does not have a minimum in the interval X_1 for any value of x_2. Hence if X does not contain a boundary point of $X^{(0)}$, we can delete all of X.

Now suppose the box chosen in step 1 has components $X_1=[2, 3]$ and $X_2=[0, 1]$. We find

$$h_{11}(X_1, X_2)=[-29.4, 358.6], \qquad h_{22}(X_1, X_2)=2.$$

Since neither interval is negative, we cannot say that f is not convex in X. Hence we go to step 3.

In step 3, we evaluate $g(X)$ obtaining

$$g_1(X)=[-74.4, 221.4], \qquad g_2(X)=[-3, 0].$$

We see that $g_2(x)$ is non-positive for all $x\in X$ and hence f is smallest in X for $x_2=1$. Thus we can replace X by the degenerate box X' with components $X_1'=[2, 3]$ and $X_2'=[1, 1]$.

If the box X had components $X_1=[0, 1]$ and $X_2=[2, 3]$, we would have obtained

$$g_1(X)=[-7.2, 3], \qquad g_2(X)=[3, 6].$$

In this case $g_2(x)$ is strictly positive and we can eliminate all of X unless the boundary of X at $x_2=2$ contains a boundary point of $X^{(0)}$. Suppose $X_1^{(0)}=[0, 1]$

and $X_2^{(0)}=[1,3]$. Then $X^{(0)}$ has boundary points at $x_2=2$ for $x_1=0$ and 1. We could thus delete all of X except the points (0, 2) and (1,2). This is simple to do in this two-dimensional problem. In higher dimensions, it might be simpler to retain the entire boundary at $x_2=2$.

Now suppose that X is given by $X_1=X_2=[0,1]$. Then $g_1(X)=[-5.2,5]$ and $g_2(X)=[-1,2]$ so that we do not have monotonicity. Therefore, we do step 4 which involves the quadratic method of Section 7.

For this box, we obtain

$$H_{11}(X)=J_{11}(X)=[-8.6,9], \quad H_{21}(X)=2J_{21}(X)=-2, \quad H_{22}(X)=J_{22}(X)=2.$$

The center of the box is at $x=(0.5,0.5)$. We wish to evaluate $f(x)$ and $g(x)$. We cannot obtain $f(x)$ exactly using finite precision decimal arithmetic. Let us use five significant decimal digits and evaluate $f(x)$ using interval arithmetic to bound rounding errors. Thus we replace the coefficient 1/6 by [0.16666, 0.16667] and obtain

$$f(x)=[0.43697, 0.43699].$$

We also obtain

$$g(x)=\begin{bmatrix}[1.0062, 1.0063] \\ 0.5\end{bmatrix}.$$

Suppose we have previously obtained $\bar{f}=0.2$ and that we choose $\varepsilon_1=0$. To do step 4, we wish to solve (7.2) for points $y\in X$ where we know that $f(y)>\bar{f}$ does not hold and hence $f(y)\leqq\bar{f}$ might hold. If we first tried to solve for Y_1, we would rewrite (7.2) in the form (7.4). However, the left endpoint of $H_{11}(X)$ is less than zero. Hence, solving (7.4) would give rise to two semi-infinite intervals. Therefore, we defer this operation until step 6 and first solve for Y_2 which will be a single interval since the "left endpoint" of $H_{22}(X)$ is positive.

We solve for Y_2 using (7.5). As we have not yet solved for Y_1, we use X_1 in its place. Substituting into (7.5), we obtain

$$[-1.2412, 1.9652]+[0,1]\,\tilde{y}_2+\tilde{y}_2^2\leqq 0.$$

From Eq. (7.8), the solution set is

$$\tilde{Z}_2=[-1.7212, 1.1141].$$

Hence

$$Z_2=x_2+\tilde{Z}_2=[-1.2212, 1.6141]$$

and

$$Y_2=Z_2\cap X_2=X_2.$$

Thus we have not deleted any of X_2.

Next we do step 5 which applies that part of the Newton method which generates a single interval. The interval Jacobian is

$$J(X)=\begin{bmatrix}[-8.6,9] & -1 \\ -1 & 2\end{bmatrix}.$$

The center of this interval matrix is

$$J_c = \begin{bmatrix} 0.2 & -1 \\ -1 & 2 \end{bmatrix}$$

whose inverse is (approximately)

$$B = \begin{bmatrix} -3.3333 & -1.6667 \\ -1.6667 & -0.33333 \end{bmatrix}.$$

For simplicity of exposition, we shall compute $BJ(X)$ explicitly. We obtain from (6.2) (with ξ replaced by X),

$$\begin{bmatrix} [-4.1877, -4.1872] \\ [-1.844, \quad -1.8436] \end{bmatrix} + \begin{bmatrix} [-28.334, 30.334] & -0.0001 \\ [-14.668, 14.668] & [1, 1.0001] \end{bmatrix} (y-x) = 0. \quad (15.4)$$

We try to improve X_2 first rather than X_1 because $[BJ(X)]_{11}$ contains zero while $[BJ(X)]_{22}$ does not. Thus the first equation of (15.4) gives rise to two new intervals while the second equation does not.

The second equation is

$$[-1.844, -1.8436] + [-14.668, 14.668](X_1 - x_1) + [1, 1.0001](y_2 - x_2) = 0$$

where we have replaced y_1 by X_1. Solving for y_2, we obtain the interval

$$Z_2 = [-4.9904, 8.678].$$

The intersection $Y_2 = Z_2 \cap X_2$ equals X_2 so again no improvement has been made.

Step 6 prescribes that we use the quadratic method to try to improve those components of X not solved for in step 4. We solve (7.5) for points y_1 where $f(y) \leqq \bar{f}$. We would use Y_2 in place of X_2; but they are equal. Substituting into (7.5), we obtain

$$[0.08697, 0.83699] + [0.5062, 1.5063]\,\tilde{y}_1 + [-4.3, 4.5]\,\tilde{y}_1^2 \leqq 0.$$

Using equation (7.11), we find that this quadratic has the solution set

$$\tilde{Z}_1 = [-\infty, -0.10093] \cup [0.42555, \infty].$$

Thus

$$Z_1 = \tilde{Z}_1 + x_1 = [-\infty, 0.39907] \cup [0.92555, \infty]$$

and

$$Y_1 = Z_1 \cup X_1 = [0, 0.39907] \cup [0.92555, 1]$$

note we have eliminated a subinterval of length 0.52648 from X_1.

In step 7, we use the Newton method to try to improve X_1. We solve the first equation of (15.4),

$$[-4.1877, -4.1872] + [-28.334, 30,334](y_1 - x_1) - 0.0001(X_2 - x_2) = 0,$$

where we have now replaced Y_2 by X_2. Solving for y_1, we obtain the two semi-infinite intervals

$$Z_1^{(3)} = [-\infty, 0.35223], \quad Z_1^{(4)} = [0.63803, \infty].$$

Their intersections with X_1 are (say) $Y_1^{(3)}$ and $Y_1^{(4)}$ where

$$Y_1^{(3)} = [0, 0.35223], \quad Y_1^{(4)} = [0.63803, 1].$$

We wish to combine the results obtained using the quadratic method and the Newton method. Thus we retain the intersection of

$$Y_1^{(1)} \cup Y_1^{(2)} \quad \text{and} \quad Y_1^{(3)} \cup Y_1^{(4)}$$

which is

$$[0, 0.35223] \cup [0.92555, 1].$$

We have deleted a substantial portion of the original box X. The remaining points compose the two boxes

$$\begin{bmatrix} [0, 0.35223] \\ [0, 1] \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} [0.92555, 1] \\ [0, 1] \end{bmatrix}.$$

In subsequent steps, our method would be applied separately to each of these boxes.

16. Computational Results

We now describe some computational results obtained using the algorithm described above. The computations were done on the Amdahl 470V/6-II computer. In each case, we assumed it was known that the global minimum occurred in the interior of the initial box. This speeds up the algorithm since boundary points need not get special treatment.

This is consistent with the fact that we are really considering the unconstrained case. We intend to treat the constrained case in a later paper.

We give results for only one example which typifies the problems in two and three dimensions which we have used. The example is the three hump camel function given by (15.1).

This function has its global minimum at the origin. It has two local minima at approximately $[\pm 1.75, \pm 0.87]$ and two saddle points at approximately $[\pm 1.07, \pm 0.535]$. Our initial box was defined by $X_1 = X_2 = [-2, 4]$ which contains all these points. We chose $\varepsilon_1 = 0$ and $\varepsilon_0 = \varepsilon_2 = \varepsilon_3 = 10^{-4}$.

We find that after eight steps of our algorithm, we have six sub-boxes in our list. In the next step a sub-box is entirely eliminated and after the fifteenth step, only one sub-box remains but its width exceeds ε_2. After an additional step, we obtain final box

$$X = \begin{bmatrix} [-5.78 \times 10^{-7}, 7.11 \times 10^{-6}] \\ [-2.91 \times 10^{-7}, 3.56 \times 10^{-6}] \end{bmatrix}.$$

Here and in the following, we record results to only three decimals. This box satisfies the error criterion requiring its width to be less than ε_2. Evaluating f at the center of this box, we obtain $\bar{f} = 1.12 \times 10^{-10}$.

As prescribed in step 11, we evaluate $f(X)$ and obtain $[-1.24 \times 10^{-11}, 7.57 \times 10^{-10}]$. Thus $\underline{f} = -1.24 \times 10^{-11}$ and

$$f^* \in [\underline{f}, \bar{f}] = [-1.24 \times 10^{-11}, 1.12 \times 10^{-10}].$$

Since $\bar{f} - \underline{f} < \varepsilon_0$, we have f^* bounded to the prescribed tolerance. If we approximate f^* by

$$(\underline{f} + \bar{f})/2 = 4.98 \times 10^{-11},$$

then we know that the error is at most 6.22×10^{-11} in magnitude.

If we approximate x^* by the center $(3.27 \times 10^{-6}, 1.63 \times 10^{-6})$, then we know that the error in x_1^* is less than 3.85×10^{-6} and the error in x_2^* is less than 1.93×10^{-6}.

We have obtained x^* to far more accuracy than required because of the rapid rate of convergence of the interval Newton method used. The bound on f^* is much better than required simply because a given error bound on x^* automatically yields a much better bound on f^*.

We also used this example with an initial box of width 2×10^6. This case required 46 steps to run to completion. This illustrates that if we use a very large box to assure containment of x^*, the computing time need not increase drastically.

17. Conclusion

We have presented an algorithm for solving the unconstrained minimization problem assuming we have an initial box which is known to contain the minimum.

It would certainly be possible to construct a highly oscillatory function for which our algorithm would be prohibitively slow. However, it has converged adequately rapidly for the test problems on which we have tried it. (See Sect. 16.)

We have assumed $f(x) \in C^2$. The global minimization problem can also be solved for $f(x) \in C^1$. In this case, the Newton method cannot be used. The quadratic method can be replaced by a corresponding linear method in which we find points y at which $f(y) \leqq \bar{f}$. This is done by noting that if $x \in X$ and $y \in X$, then

$$f(y) = f(x) + (y - x)^T g(\xi)$$

for some $\xi \in X$. Thus we can solve for the approximate points y from

$$f(x) + (y - x)^T g(X) \leqq \bar{f}.$$

For the problems of low dimension on which we have used this method, it was less efficient then the quadratic method described in Sect. 7. We do not know the relative efficiencies for large n.

It is possible to solve the global optimization case when $f(x)$ is only

continuous but not differentiable. However, our algorithm is very slow. It entails a different approach that we hope to describe in another paper.

The nonlinear constrained optimization problem can also be solved by interval methods. An extension of our algorithm is required. Our experience in this case is for hand calculations only. A difficulty exists (currently) when it is difficult to find a point in the neighborhood of x^* which is without question, feasible.

One of the virtues of interval arithmetic is that it is usually possible to formulate an iterative algorithm in such a way that it stops automatically when the best possible result has been obtained for the finite precision arithmetic used. We plan to do this for our algorithm and thus preclude the need for specifying ε_0, ε_1, ε_2, and ε_3.

Acknowledgments. The author is greatly indebted to Thomas Kratzke and Saumyendra Sengupta for their assistance in programming and debugging the computer program with which the data in this paper was obtained.

This research was supported by the U.S. Air Force Office of Scientific Research under grant F49620-76-C-0003. The United State Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

References

1. Dixon, L.C.W., Szegö, G.P.: Towards global optimization. Amsterdam: North Holland, 1975
2. Dixon. L.C.W., Szegö, G.P.: Towards global optimization 2. Amsterdam: North Holland, 1977
3. Hansen, Eldon: On solving systems of equations using interval arithmetic. Math. Comp. **22**, 374–384 (1968)
4. Hansen, Eldon: Interval forms of Newton's method. Computing **20**, 153–163 (1978)
5. Hansen, Eldon: Global optimization using interval analysis – the one-dimension case. Jour. Optimiz. Theo. Applic. **29**, 331–344 (1979)
6. Hansen, Eldon, and Roberta Smith: Interval arithmetic in matrix computations, II. SIAM J. Num. Anal. **4**, 1–9 (1967)
7. Krawczyk, R.: Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken. Computing **4**, 187–201 (1969)
8. Moore, R.E.: Interval analysis. New York: Prentice-Hall (1966)
9. Moore, R.E.: On computing the range of a rational function of n variables over a bounded region. Computing **16**, 1–15 (1976)
10. Moore, R.E.: A test for existence of solutions to nonlinear systems. SIAM J. Num. Anal. **14**, 611–615 (1977)
11. Moore, R.E.: A computational test for convergence of iterative methods for nonlinear systems. SIAM J. Num. Anal. **15**, 1194–1196 (1978)
12. Moore, R.E.: Methods and applications of interval analysis. Philadelphia: SIAM (1979)
13. Nickel, Karl: On the Newton method in interval analysis. Mathematics Research Center Report 1136, University of Wisconsin (1971)
14. Skelboe, S.: Computation of rational interval functions, BIT **14**, 87–95 (1974)

Received June 12, 1978