Computing 38, 275–280 (1987)

Computing

© by Springer-Verlag 1987

An Interval Version of Shubert's Iterative Method for the Localization of the Global Maximum

Z. Shen and **Y. Zhu**, Nanjing

Received July 3, 1986

Abstract — Zusammenfassung

An Interval Version of Shubert's Iterative Method for the Localization of the Global Maximum. Using the "bisection rule" of Moore, a simple algorithm is given which is an interval version of Shubert's iterative method for seeking the global maximum of a function of a single variable defined on a closed interval $[a, b]$. The algorithm which is always convergent can be easily extended to the higher dimensional case. It seems much simpler than and produces results comparable to that proposed by Shubert and Basso.

AMS (MOS) Subject Classifications: 65 K 05, 90 C 30.

Key words: Interval analysis, global maximum, iterative method.

Eine Intervallversion der iterativen Methode von Shubert zur Lokalisierung des globalen Maximums. Unter Verwendung der „Bisektionsregel" von Moore wird ein Algorithmus angegeben, der eine Intervallversion der iterativen Methode von Shubert zur Bestimmung des globalen Maximums einer Funktion einer Veränderlichen auf dem abgeschlossenen Intervall $[a, b]$ darstellt. Der Algorithmus konvergiert immer; er kann leicht auf den höherdimensionalen Fall ausgedehnt werden. Er erscheint viel einfacher als der Algorithmus von Shubert und Basso, ergibt aber vergleichbare Ergebnisse.

1. Introduction

Shubert [7] proposed a deterministic sequential method seeking the global maximum of a function of a single variable defined on a closed interval $[a, b]$. Suppose that $f(x)$, $x \in [a, b]$ is a Lipschitz continuous function with the Lipschitz constant C. Let

$$\phi = \max_{x \in [a, b]} f(x)$$

be the global maximum of the function f and let

$$\Phi = \{x \mid x \in [a, b], f(x) = \phi\}$$

be the set of all x for which the maximum is attained.

Define the sampling sequence $x_0, x_1, x_2, \ldots$ of points from $[a, b]$ recursively as follows:

$$x_0 \in [a, b],$$

$$x_{n+1} \text{ such that } F_n(x_{n+1}) = M_n, \quad n = 1, 2, \ldots$$

where

$$M_n = \max_{x \in [a,b]} F_n(x)$$

and

$$F_n(x) = \min_{k=0,1,\ldots,n} \{f(x_k) + C|x - x_k|\}.$$

Theorem 1: *As $n \to \infty$, then*

$$f(x_n) \to \phi, \quad M_n \downarrow \phi$$

and

$$\inf_{x \in \Phi} |x - x_n| \to 0.$$

Based on Shubert's method, Basso [2] proposed a sequence of localizations converging to the set Φ and presented an example which showed that the sequence of localizations produced by Shubert's method is not always converging to the set Φ in the class of Lipschitz continuous functions. Basso also suggested some modifications to Shubert's method called Jacobi and Gauss-Seidel type methods.

In this paper, we suggest a simple algorithm using the "bisection rule" of Mooré [1], [5] and the idea of constructing a sequence of upper bounds of the function. The algorithm which is always convergent seems much simpler than that proposed by Shubert and Basso, and is also designed to find the localization of the global maximum.

2. The Basic Idea

Consider the problem of finding the localization of the global maximum. We also suppose that $f(x)$, $x \in [a,b]$ is a Lipschitz continuous function with constant C. For each interval $X = [\underline{X}, \bar{X}] \subseteq X^{(0)} = [a,b]$,

$$F(X) = f(m(X)) + C(X - m(X))$$

is an interval extension of $f(x)$, where $m(X) =$ midpoint of X. Thus, if $x \in X \subseteq X^{(0)}$, then

$$f(x) \in F(X)$$

and

$$f(x) \leqq b_0 = \overline{F(X)}.$$

Suppose we subdivide the interval X into N parts uniformly so that X can be written as the union of N intervals

$$X^{(1)}, X^{(2)}, \ldots, X^{(N)}, \quad X = \bigcup_{k=1}^{N} X^{(k)}.$$

Let

$$F(X^{(i)}) = f(m(X^{(i)}) + C(X^{(i)} - m(X^{(i)}))).$$

and

$$M_N = \max_{i=1,2,\ldots,N} \overline{F(X^{(i)})}, \tag{1}$$

then it is clear that

$$\max_{x \in [a, b]} f(x) \leqq M_N$$

and we have

Theorem 2: *For arbitrary* $\varepsilon > 0$, *there exists an* N *such that*

$$M_n - \max_{x \in [a, b]} f(x) < \varepsilon.$$

Proof: It follows from (1) that, for some k_0, we have

$$M_N = \overline{F(X^{(k_0)})}$$

and for each $x \in X^{(k_0)}$, we have

$$f(x) \in F(X^{(k_0)}),$$

hence

$$\overline{F(X^{(k_0)})} - f(x) \leqq C w(X^{(k_0)}) = C \frac{w(X)}{N}.$$

Take N large enough and note that

$$\overline{F(X^{(k_0)})} \geqq \max_{x \in [a, b]} f(x) \geqq f(x),$$

the theorem is proved.

From Theorem 2 we can compute arbitrarily sharp upper bounds for the function $f(x)$. Let $N \to \infty$, the convergence

$$M_N \to \max_{x \in [a, b]} f(x)$$

will be obtained, and at the same time, we will have

$$\inf_{k=1, 2, \ldots, N} |m(X^{(k)}) - \hat{x}| \to 0 \text{ for each } \hat{x} \in \Phi.$$

In the following, using the bisection rule of Moore, we will propose some efficient means for doing this.

To find the global maximum in the initial interval $X^{(0)}$, we generate a list of subintervals whose union must contain the global maximum. The elements in the list are of the form $(X, F(X))$, each element is the result of a bisection of some previous interval in the list, the elements are arranged in decreasing order of their upper bounds, i.e. $(X, \overline{F(X)})$ comes before $(Y, \overline{F(Y)})$ in the list if $\overline{F(X)} \geqq \overline{F(Y)}$. Thus, the first element in the list always corresponds to the largest current upper bound interval.

Suppose we subdivide an interval X into two parts by bisecting it so that $x = X^{(1)} \cup X^{(2)}$. If we find that $\overline{f(X^{(1)})} < v$ where v is the value of $f(x)$ for some x in $X^{(2)}$, say the midpoint of $X^{(2)}$, then we can exclude the interval $X^{(1)}$ from further consideration in seeking an upper bound for $\max f(x)$. We call this the midpoint test.

Thus, we do not list an interval at all if it can be determined by the midpoint test that the interval cannot contain the global maximum of f on $X^{(0)}$. We also delete all elements from the list which would follow such an element if these were inserted in proper order. This will certainly make a vast reduction in the amount of the elements in the list. Moreover, using the midpoint test, each interval X which does not contain the global maximum point will eventually be deleted from the list after a finite number of bisections. In fact, if X does not contain the global maximum point, there must exist a point $x_0 \in X^{(0)}$, $x_0 \notin X$, such that

$$f(x_0) > \max_{x \in X} f(x).$$

Suppose after a finite number of bisections, $X = \bigcup_{k=1}^{N} X^{(k)}$ with the width $w(X^{(k)})$, $k=1,2,\ldots,N$ small enough such that

$$f(x_0) > f\left(m(X^{(k)})\right) + C\left(X^{(k)} - m(X^{(k)})\right), \; k=1,2,\ldots,N,$$

then all the intervals $X^{(k)}$, $k=1,2,\ldots,N$ as well as X will be deleted from the list by the midpoint test.

With these ideas, we give the following algorithm for finding the global maximum.

Step 1. Set $X = X^{(0)}$, $F(X) = f\left(m(X)\right) + C\left(X - m(X)\right)$ and $b_0 = \overline{F(X)}$;

Step 2. To begin with, the list is empty;

Step 3. Bisect X: $X = X^{(1)} \cup X^{(2)}$;

Step 4. Compute $F(X^{(i)}) = f\left(m(X^{(i)})\right) + C\left(X^{(i)} - m(X^{(i)})\right)$, $i=1,2$ and set $b = \max_{i=1,2} \overline{F(X^{(i)})}$;

Step 5. If $|b - b_0| < \varepsilon$ for given $\varepsilon > 0$, then take b as the global maximum, $m(X^{(1)})$ the global maximum point if $b = \overline{F(X^{(1)})}$ or $m(X^{(2)})$ the global maximum point if $b = \overline{F(X^{(2)})}$; print message; otherwise proceed with Step 7;

Step 6. If the list is empty, then stop; otherwise proceed with Step 8;

Step 7. Enter the items $\left(X^{(1)}, \overline{F(X^{(1)})}\right)$, $\left(X^{(2)}, \overline{F(X^{(2)})}\right)$ in proper order in the list;

Step 8. Set X = the argument (first member of the pair $\left(X, \overline{F(X)}\right)$ of the first item in the list (with largest $\overline{F(X)}$)· and remove the item $\left(X, \overline{F(X)}\right)$ from the list;

Step 9. Set $b_0 = \overline{F(X)}$;

Step 10. Apply the midpoint test on X to delete all the unnecessary elements from the list (if any);

Step 11. Return to Step 3.

Following the above discussion, we have

Theorem 3: *Using the rounded arithmetic, the algorithm given by Step* 1 – 11 *always converges in a finite number of computation steps to the global maximum.*

3. Numerical Results

In this section, we will report on two examples run by Shubert, and compare the numerical results of Shubert's method and the interval method. It can be observed in these examples that the interval method produced results comparable to Shubert's method.

Now, the following abbreviation will be used:

IT: Number of iterations
MM: Global maximum of the majorant
IM: Global maximum obtained by the interval method
k: Number of the intervals in the list
LOC: First interval in the list-localization

Example 1: $f(x) = (-3x + 1.4)\sin(18x)$, $x \in [0, 1.2]$. This function attains its global maximum $\alpha = 1.489072539$ at the unique point $x = 0.9660858$.

Table 1

IT	Shubert's method		Interval method		
	MM	k	IM	LOC	k
3	9.7361872	2	6.790100	[.0000, .3000]	4
9	2.7081936	8	3.447776	[.6000, .7500]	9
21	1.7691646	4	1.909890	[.0750, .1125]	14
31	1.5582689	8	1.642454	[.9750, .9844]	15
51	1.5030947	24	1.525208	[.9797, .9844]	23
84	1.4932404	49	1.499191	[.9627, .9633]	39
87	1.4929375	48	1.498102	[.9574, .9586]	42
132	1.4905091	83	1.492092	[.9667, .9668]	70
164	1.4900031	109	1.490728	[.9636, .9637]	92
208	1.4895926	133	1.490192	[.9674, .9675]	124
300	1.4893297	185	1.489532	[.9649, .9650]	185

Example 2: $f(x) = \sum_{p=1}^{5} p \sin((p+1)x + p)$, $x \in [-10, 10]$. This is a strongly oscillatory function whose global maximum $\alpha = 12.031249442$ is attained at the points -6.7746, -0.4914, 5.7981.

Table 2

IT	Shubert's method		Interval method		
	MM	k	IM	LOC	k
3	348.9460713	2	52.54837	[.0000, 5.0000]	4
9	87.1645598	8	25.55259	[2.5000, 5.0000]	10
33	25.7210251	32	12.52493	[4.3750, 4.6875]	9
85	13.5460293	28	12.06613	[−0.4688, −0.4492]	39
117	12.4417040	48	12.04523	[−6.7676, −6.7651]	56
175	12.1457273	84	12.03702	[−0.4932, −0.4926]	91
283	12.0621904	140	12.03285	[−6.7834, −6.7822]	170
348	12.0502472	177	12.03248	[−0.4929, −0.4927]	212
499	12.0390173	284	12.03181	[−6.7722, −6.7720]	311
588	12.0362783	341	12.03162	[−0.4887, −0.4886]	360

4. Extension

The method mentioned above can be easily extended to the higher dimensional case. Instead of bisecting an interval, we will bisect a region X. We choose, as our direction in which to bisect a region X, the first direction in which X has maximum width.

5. Acknowledgment

I wish to express my gratitude to Prof. Dr. K. Nickel for his invaluable help and suggestions. I am also thankful to Drs. J. Garloff and A. Neumaier for their helpful comments. Specially, they reminded me of Hansen's and Ratschek's papers [3], [4], [6]. These papers also provide infallible bounds on both the global maximum and the localizations of the global maximum of a function.

References

[1] Asaithambi, N. S., Shen, Z., Moore, R. E.: On computing the range of values. Computing *28*, 225−237 (1982).
[2] Basso, P.: Iterative methods for the localization of the maximum. SIAM J. Numer. Anal. *19*, 781−792 (1982).
[3] Hansen, E.: Global optimization using interval analysis − the one-dimensional case. J. Optim. Theory Appl. *29*, 331−344 (1979).
[4] Hansen, E.: Global optimization using interval analysis − the multi-dimensional case. Numer. Math. *34*, 247−270 (1980).
[5] Moore, R. E.: Methods and applications of interval analysis. SIAM Philadelphia, 1979.
[6] Ratschek, H.: Inclusion functions and global optimization. Mathematical Programming, to appear (1985).
[7] Shubert, B. O.: A sequential method seeking the global maximum of a function. SIAM J. Numer. Anal. *9*, 379−388 (1972).

Prof. Z. Shen
Y. Zhu
Department of Mathematics
Nanjing University
Nanjing
The People's Republic of China

Verleger: Springer-Verlag KG, Mölkerbastei 5, A-1010 Wien. — Herausgeber: Prof. Dr. Hans J. Stetter, Institut für Angewandte und Numerische Mathematik der Technischen Universität Wien, Wiedner Hauptstraße 6—10, A-1040 Wien. — Redaktion: Wiedner Hauptstraße 6—10, A-1040 Wien. — Hersteller: Satz Austro-Filmsatz Richard Gerin, Zirkusgasse 13, A-1020 Wien, Druck Paul Gerin, Zirkusgasse 13, A-1021 Wien. — Verlagsort: Wien. — Herstellungsort: Wien. — Printed in Austria.