A New Super-Convergent Inclusion Function Form and its Use in Global Optimization

P. S. V. Nataraj[1] and K. Kotecha[2]

[1]Systems and Control Engineering Group
EE Department, IIT Bombay 400076, India
Email: nataraj@ee.iitb.ernet.in
[2]Department of IT
GH Patel College of Engineering and Technology
Vallabh Vidyanagar, 388120, Gujarat, India
Email: kk@gcet.ac.in

Abstract

Recently, Lin and Rokne [10] introduced the so-called Taylor-Bernstein form as an inclusion function form for multidimensional functions. This form was theoretically shown to have the super-convergence property. Here, we present an improvement of Lin and Rokne's Taylor-Bernstein form to make it more effective in practice. We test and compare the super-convergence behavior of the proposed form with that of Lin and Rokne's Taylor-Bernstein form and also with that of the Taylor model of Berz *et al.* [3]. We obtain super-convergence of orders up to 9 with the proposed form. Moreover, with the proposed form we quite easily obtain such high orders of super-convergence for up to $5 - \dim$ problems.

We also investigate the use of higher order inclusion functions in the Moore-Skelboe (MS) algorithm of interval analysis (IA) for unconstrained global optimization. We use the improved TB form as an inclusion function in a prototype MS algorithm and also modify the cut-off test and termination condition in the algorithm. We test and compare on several examples the performances of the proposed algorithm, the MS algorithm, and the MS algorithm with the Taylor model of Berz *et al.* [3] as inclusion function. The results of these (preliminary) tests indicate that the proposed algorithm with the improved TB form as inclusion function is quite effective for low to medium dimension problems studied.

1 Introduction

An important problem in interval analysis is the construction of inclusion functions having the property of so-called *super-convergence* (i.e., having a convergence order that is greater than quadratic) for multidimensional functions.

Such inclusion functions have applications in the solutions of equations, optimization, quadrature, and others. The first paper in the literature concerning super-convergence is that of Herzberger [6], who shows that super-convergence can be obtained for a certain class of intervals. However, his requirement on the function is unrealistically strong. Cornelius and Lohner [4] propose the interpolation and remainder forms for multidimensional functions that enable any convergence order to be obtained in theory. However, in practice, convergence order of at most 4 or 5 is recommended even for unidimensional functions, see [4] and [15, pg. 9]. The same holds for the improved version of these forms for unidimensional functions, as proposed by Neumaier in [14, sec. 2.4]. Alefeld and Lohner [1] propose centered forms with super-convergence for *unidimensional* functions. However, because of the strong condition on the functional representation, these higher order centered forms have limited practical value [1, pg. 8]. Lin and Rokne [10] propose super-convergent forms that combine Taylor and Bernstein (or B-spline) forms for multidimensional functions. However, for small domains these forms become computationally very demanding, even for unidimensional functions, see [10, pg. 108]. Berz *et al.* [3, 12] propose the so-called Taylor models for multidimensional functions. Although the accuracy of the so-called remainder interval part of the Taylor model increases in a super-convergent fashion, the Taylor model itself is known to exhibit only quadratic convergence see Kearfott and Arazyan [9].

We propose in this work a new inclusion function form having the super-convergence property for multidimensional functions. The proposed inclusion function form uses Bernstein polynomials for bounding the range of the polynomial obtained from the Taylor form of the function f. The Bernstein algorithm is combined with the Taylor form to obtain the resulting so-called Taylor-Bernstein form as an inclusion function form of f. The proposed Taylor-Bernstein form has some important differences (in the practical way it is constructed) from the Taylor-Bernstein form of Lin and Rokne [10].

We numerically investigate the super-convergence property of the above inclusion function forms on some benchmark examples. The selected examples are of low to medium dimensions. For all our computations, we use a PC/Pentium III 800 MHz 256 MB RAM machine with a FORTRAN 90 compiler, and version 8.1 of the COSY-INFINITY package of Berz *et al.* [2, 7]. We also investigate the performance of the Taylor model as an inclusion function form in these examples. With the proposed form, we quite easily obtain super-convergence (of orders up to 9) for low to medium dimensional problems. To our knowledge, it is perhaps for the first time that super-convergence of such high orders has actually been demonstrated on multidimensional problems. Moreover, the new super-convergent form can be constructed on a computer with the fully automated algorithm presented, without any need for hand computations.

We next use the new super-convergent form to solve the following optimization problem. Let $\Re$ be the set of reals, $\mathbf{X} \subseteq \Re^l$ be a right parallelepiped parallel to the axes (also called as a box), and $f : \mathbf{X} \to \Re$ be a $m + 1$ times differentiable function for some positive integer m. Let $\bar{f}(\mathbf{X})$ denote the set of all values of f on $\mathbf{X}$. We seek global optimization algorithms that are able to efficiently determine arbitrarily good lower bounds for the minimum of $\bar{f}(\mathbf{X})$.

Many algorithms based on interval analysis (IA) are available to solve this global optimization problem, see for example, [5], [8], [16] and the references cited therein. A basic branch and bound algorithm of IA is the so-called Moore-Skelboe (MS) algorithm [16]. Although the MS algorithm is reliable, it is somewhat slow to converge in 'difficult' problems, when inclusion functions of first and sometimes even second orders are used. Faster convergence could possibly be obtained with higher order inclusion functions, and it is of interest in this work to investigate their effectiveness in some such 'difficult' problems.

Our proposed algorithm for global optimization uses the new super-convergent form having high order convergence, and we therefore expect to obtain faster convergence with this form. The new form also allows us to make the cut-off test and termination condition more effective, and we incorporate these modifications in the proposed algorithm. Since this global optimization algorithm involves using the new **T**aylor - **B**ernstein form in **M**oore-**S**kelboe algorithm, we call it as Algorithm TBMS.

We can also have the Taylor model of Berz *et al.* as an inclusion function form in the MS algorithm as done, for instance, in the preliminary work in [9]. We call such an algorithm as Algorithm TMS.

We test and compare the performance of the proposed algorithm with that of Algorithms TMS and MS on six 'difficult' examples. These preliminary tests indicate that Algorithms TMS and TBMS are quite effective compared to Algorithm MS, for lower accuracy problems. For higher accuracy problems, Algorithm TBMS is the most effective one. The best overall choice, in terms of the number of iterations, space-complexity, and speed seems to be Algorithm TBMS with a medium Taylor order $m = 4$.

2 Numerical Results for super convergence

We numerically investigate the super-convergence property of the above inclusion function forms on some benchmark examples.

In each example, we compute the intervals

$F_{TM}(\mathbf{X})$ – using Taylor model of Berz *et al.* [11], computed with the COSY-INFINITY package.

$F_{LR}(\mathbf{X})$ – using Taylor-Bernstein form of Lin and Rokne, computed with Algorithm LR.

$F_{TB}(\mathbf{X})$ – using the proposed Taylor-Bernstein form, computed with Algorithm TB.

$F_{inner}(\mathbf{X})$ – using *inner* estimates of the range computed with the well-known Moore-Skelboe optimization algorithm of interval analysis (see, for instance, [16]).

Let $\mathbf{X} = [a, b], \mathbf{Y} = [c, d]$ be any two intervals. Then, following [4], as a measure of the overestimation we use the Hausdorff metric

$$\mathcal{H}(\mathbf{X}, \mathbf{Y}) = |[a, b], [c, d]| = \max\{|a - c|, |b - d|\}$$

Consider a sequence of nested intervals $\left\{\mathbf{X}^{(i)}\right\}$. We wish to find and compare for each form, the reduction in overestimation with decrease in the domain interval widths. Consider first the form F_{TM}. Let

$$\mathcal{H}_{TM}\left(\mathbf{X}^{(i-1)}\right) := \mathcal{H}\left(\bar{f}(\mathbf{X}^{(i-1)}), F_{TM}\left(\mathbf{X}^{(i-1)}\right)\right) \tag{1}$$

As a measure of the reduction in overestimation obtained with form F_{TM} over successive nested intervals $\mathbf{X}^{(i-1)}$ and $\mathbf{X}^{(i)}$, we use the ratio

$$\mathcal{R}_{TM}\left(\mathbf{X}^{(i-1)}, \mathbf{X}^{(i)}\right) := \frac{\mathcal{H}_{TM}\left(\mathbf{X}^{(i-1)}\right)}{\mathcal{H}_{TM}\left(\mathbf{X}^{(i)}\right)} = \frac{\mathcal{H}\left(\bar{f}(\mathbf{X}^{(i-1)}), F_{TM}\left(\mathbf{X}^{(i-1)}\right)\right)}{\mathcal{H}\left(\bar{f}(\mathbf{X}^{(i)}), F_{TM}\left(\mathbf{X}^{(i)}\right)\right)}$$

Define

$$\mathcal{R}^*\left(\mathbf{X}^{(i-1)}, \mathbf{X}^{(i)}\right) := \left(\frac{w\left(\mathbf{X}^{(i-1)}\right)}{w\left(\mathbf{X}^{(i)}\right)}\right)^{m+1}$$

If F_{TM} is an inclusion function form having convergence order $m+1$, then

$$\mathcal{R}_{TM}\left(\mathbf{X}^{(i-1)}, \mathbf{X}^{(i)}\right) \to \mathcal{R}^*\left(\mathbf{X}^{(i-1)}, \mathbf{X}^{(i)}\right) \tag{2}$$

(where the tending is from above) for "small" enough $w\left(\mathbf{X}^{(i-1)}\right), w\left(\mathbf{X}^{(i)}\right)$.

In practice, the exact range $\bar{f}$ is generally difficult to compute, so the overestimation can be generally found relative only to some *inner* estimate F_{inner} of the range. However, we can easily show that if the $(m+1)$- th convergence order property holds relative to F_{inner}, then it implies that the same holds relative to the exact range $\bar{f}$. That is, it is sufficient if we can show the $(m+1)$- th convergence order property with F_{inner} used in place of $\bar{f}$ in above definitions. To avoid introducing more notation, in the sequel we use the quantities given in (1) through (2), with F_{inner} replacing $\bar{f}$ throughout.

Similarly, we can define $\mathcal{H}_{LR}$, $\mathcal{H}_{TB}, \mathcal{R}_{LR}, \mathcal{R}_{TB}$ for the forms F_{LR} and F_{TB}. For brevity of notation, we drop the arguments $\mathbf{X}^{(i-1)}, \mathbf{X}^{(i)}$ of all $\mathcal{H}$ and $\mathcal{R}$.

Example 1. Trigonometric function [13, problem 26]. The $4-\dim$ function is

$$f(x) = \sum_{i=1}^{4} f_i(x)^2, f_i(x) = 4 - \sum_{j=1}^{4} \cos x_j + i(1-\cos x_i) - \sin x_i$$

The domain is $\mathbf{X}^{(i)} = \left[1.75 + 2^{-i}\left[-1, 1\right]\right]^4$.

For Taylor order $m = 2$:

i	0	1	2	3
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-0}$	$2*2^{-1}$	$2*2^{-2}$	$2*2^{-3}$
$\mathcal{H}_{TM}$	$4E+2$	$9E+1$	$2E+1$	$5E+0$
$\mathcal{H}_{LR}$	$3E+2$	$3E+1$	$3E+0$	$*$
$\mathcal{H}_{TB}$	$3E+2$	$3E+1$	$3E+0$	$3E-1$
$\mathcal{R}^*$	$-$	8	8	8
$\mathcal{R}_{TM}$	$-$	4.9	4.5	4.2
$\mathcal{R}_{LR}$	$-$	10.5	9.5	$-$
$\mathcal{R}_{TB}$	$-$	10.5	9.5	8.8

i	4	5	6	7
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-4}$	$2*2^{-5}$	$2*2^{-6}$	$2*2^{-7}$
$\mathcal{H}_{TM}$	$1E+0$	$3E-1$	$7E-2$	$2E-2$
$\mathcal{H}_{LR}$	$*$	$*$	$*$	$*$
$\mathcal{H}_{TB}$	$3E-2$	$4E-3$	$5E-4$	$7E-5$
$\mathcal{R}^*$	8	8	8	8
$\mathcal{R}_{TM}$	4.1	4.1	4.0	4.0
$\mathcal{R}_{LR}$	–	–	–	–
$\mathcal{R}_{TB}$	8.4	8.2	8.1	8.1

For Taylor order $m=4$:

i	0	1	2	3
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-0}$	$2*2^{-1}$	$2*2^{-2}$	$2*2^{-3}$
$\mathcal{H}_{TM}$	$4E+2$	$9E+1$	$2E+1$	$5E+0$
$\mathcal{H}_{LR}$	$1E+1$	$2E-1$	$*$	$*$
$\mathcal{H}_{TB}$	$1E+1$	$2E-1$	$5E-3$	$1E-4$
$\mathcal{R}^*$	–	32	32	32
$\mathcal{R}_{TM}$	–	4.9	4.4	4.2
$\mathcal{R}_{LR}$	–	56.3	–	–
$\mathcal{R}_{TB}$	–	56.3	50.4	44.5

i	4	5	6	7
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-4}$	$2*2^{-5}$	$2*2^{-6}$	$2*2^{-7}$
$\mathcal{H}_{TM}$	$1E+0$	$3E-1$	$7E-2$	$2E-2$
$\mathcal{H}_{LR}$	$*$	$*$	$*$	$*$
$\mathcal{H}_{TB}$	$3E-6$	$8E-8$	$2E-9$	$7E-11$
$\mathcal{R}^*$	32	32	32	32
$\mathcal{R}_{TM}$	4.1	4.1	4.0	4.0
$\mathcal{R}_{LR}$	–	–	–	–
$\mathcal{R}_{TB}$	39.8	36.5	34.3	30.6

For Taylor order $m=6$:

i	0	1	2	3
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-0}$	$2*2^{-1}$	$2*2^{-2}$	$2*2^{-3}$
$\mathcal{H}_{TM}$	$4E+2$	$9E+1$	$2E+1$	$5E+0$
$\mathcal{H}_{LR}$	$2E+0$	$9E-3$	$*$	$*$
$\mathcal{H}_{TB}$	$2E+0$	$9E-3$	$6E-5$	$4E-7$
$\mathcal{R}^*$	–	128	128	128
$\mathcal{R}_{TM}$	–	4.9	4.4	4.2
$\mathcal{R}_{LR}$	–	189.0	–	–
$\mathcal{R}_{TB}$	–	189.0	167.6	151.3

i	4	5	6	7
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-4}$	$2*2^{-5}$	$2*2^{-6}$	$2*2^{-7}$
$\mathcal{H}_{TM}$	$1E+0$	$3E-1$	$7E-2$	$2E-2$
$\mathcal{H}_{LR}$	$*$	$*$	$*$	$*$
$\mathcal{H}_{TB}$	$3E-9$	$3E-11$	$7E-12$	$7E-12$
$\mathcal{R}^*$	128	128	128	128
$\mathcal{R}_{TM}$	4.1	4.1	4.0	4.0
$\mathcal{R}_{LR}$	$-$	$-$	$-$	$-$
$\mathcal{R}_{TB}$	140.5	99.2	3.6	1.0

For Taylor order $m = 8$:

i	0	1	2	3
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-0}$	$2*2^{-1}$	$2*2^{-2}$	$2*2^{-3}$
$\mathcal{H}_{TM}$	$4E+2$	$9E+1$	$2E+1$	$5E+0$
$\mathcal{H}_{LR}$	$5E-1$	$6E-5$	$*$	$*$
$\mathcal{H}_{TB}$	$5E-1$	$6E-5$	$8E-8$	$1E-10$
$\mathcal{R}^*$	$-$	512	512	512
$\mathcal{R}_{TM}$	$-$	4.9	4.4	4.2
$\mathcal{R}_{LR}$	$-$	828.6	$-$	$-$
$\mathcal{R}_{TB}$	$-$	828.6	734.6	623.2

i	4	5	6	7
$w\left(\mathbf{X}^{(i)}\right)$	$2*2^{-4}$	$2*2^{-5}$	$2*2^{-6}$	$2*2^{-7}$
$\mathcal{H}_{TM}$	$1E+0$	$3E-1$	$7E-2$	$2E-2$
$\mathcal{H}_{LR}$	$*$	$*$	$*$	$*$
$\mathcal{H}_{TB}$	$8E-12$	$7E-12$	$7E-12$	$7E-12$
$\mathcal{R}^*$	512	512	512	512
$\mathcal{R}_{TM}$	4.1	4.1	4.0	4.0
$\mathcal{R}_{LR}$	$-$	$-$	$-$	$-$
$\mathcal{R}_{TB}$	17.2	1.1	1.0	0.9

3 Numerical Tests for Global Optimization

We test and compare the performances of Algorithms TBMS, TMS, and MS on various examples. Here we present one 3-dim example.

Example 2. Bard function [13, problem 8]. The three dimensional function is

$$f(x) = \sum_{i=1}^{15} f_i(x)^2, \quad f_i(x) = y_i - \left(x_1 + \frac{u_i}{v_i x_2 + w_i x_3}\right),$$

$$u_i = i, v_i = 16 - i, w_i = \min(u_i, v_i)$$

where,

i	1	2	3	4	5	6	7	8
y_i	0.14	0.18	0.22	0.25	0.29	0.32	0.35	0.39
i	9	10	11	12	13	14	15	
y_i	0.37	0.58	0.73	0.96	1.34	2.10	4.39	

We take the initial domain as $([-0.25, 0.25], [0.01, 2.5], [0.01, 2.5])$.The performances of the various Algorithms are as under.

		TBMS			
Order, m	Accuracy	Iterations	Time, s	Max. LL	Final LL
2	10^{-3}	406	16.64	74	45
	10^{-5}	520	32.13	74	7
4	10^{-3}	191	35.00	38	7
	10^{-5}	202	60.65	38	1
6	10^{-3}	162	67.80	38	2
	10^{-5}	165	90.22	38	1
8	10^{-3}	157	79.90	38	2
	10^{-5}	159	92.03	38	1

		TMS			
Order, m	Accuracy	Iterations	Time, s	Max. LL	Final LL
2	10^{-3}	3145	76.13	822	772
	10^{-5}	$*$	> 3600	$*$	$*$
4	10^{-3}	3124	86.13	818	772
	10^{-05}	$*$	> 3600	$*$	$*$
6	10^{-3}	3123	122.81	818	772
	10^{-05}	$*$	> 3600		
8	10^{-3}	3122	181.05	818	772
	10^{-5}	$*$	> 3600	$*$	$*$

MS				
Accuracy	Iterations	Time, s	Max. LL	Final LL
10^{-03}	6122	466.56	1643	1622
10^{-05}	$*$	> 3600	$*$	$*$

The global minima found using each of the algorithms is $8.21487....E - 03$.

4 Summary

We proposed a new super-convergent inclusion form for multidimensional functions form and quite easily obtained super-convergence (of order up to 9) for low to medium dimensional problems. We also tested and found the new form to be quite effective in all six global optimization problems that were selected for the tested, in terms of number of iterations, space-complexity and speed.

References

[1] G. Alefeld and R. Lohner, "On higher order centered forms", *Computing*, 1985, Vol. 35, pp. 177–184.

[2] M. Berz and J. Hoefkens, *COSY INFINITY Version 8.1 Programming Manual*, Technical Report MSUCL-1196, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 2001.

[3] M. Berz and G. Hoffstatter, "Computation and application of Taylor polynomials with interval remainder bounds", *Reliable Computing*, 1998, Vol. 4, pp. 83–97.

[4] H. Cornelius and R. Lohner, "Computing the range of values of real functions with accuracy higher than second order", *Computing*, 1984, Vol. 33, pp. 331–347.

[5] E. Hansen, *Global optimization using interval analysis*, Marcel Dekker, 1992.

[6] J. Herzberger, "Zur approximation des wertebereiches reeller funktionen durch intervallausdrucke", *Computing, Supplement*, 1977, Vol. 1, pp. 57–64.

[7] J. Hoefkens, *Rigorous Numerical Analysis with High-Order Taylor Models*, Ph.D. thesis, Michigan State University, East Lansing, Michigan, USA, 2001; also avaliable as Technical Report MSUCL-1217.

[8] R. B. Kearfott, *Rigorous global search: continuous problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[9] R. B. Kearfott and A. Arazyan, "Taylor series models in deterministic global optimization", In: *Proc. 3rd Int. Conf. and Workshop on Automatic Differentiation*, June 2000.

[10] Q. Lin and J. G. Rokne, "Interval approximation of higher order to the ranges of functions", *Computers Math. Applic.*, 1996, Vol. 31, No. 7, pp. 101–109.

[11] K. Makino and M. Berz, "Remainder differential algebras and their applications", In: M. Berz, C. Bischof, G. Corliss, and A. Griewank (eds.), *Computational differentiation: techniques, applications, and tools*, SIAM, 1996, pp. 63–75.

[12] K. Makino and M. Berz, "Efficient control of the dependency problem based on Taylor model methods", *Reliable Computing*, 1999, Vol. 5, pp. 3–12.

[13] J. J. More, B. S. Garbow, and K. E. Hillstrom, "Testing unconstrained optimization software", *ACM Trans. Mathematical Software*, 1981, Vol. 7, No. 1, pp. 17–41.

[14] A. Neumaier, *Interval methods for systems of equations*, Cambridge University Press, Cambridge, England, 1990.

[15] H. Ratschek and J. Rokne, *Computer methods for the range of functions.* Ellis Horwood Ltd, Chichester, 1984.

[16] H. Ratschek and J. Rokne, *New computer methods for global optimization*, Wiley, New York, 1988.