Computing 35, 177 – 184 (1985)

Computing

On Higher Order Centered Forms

G. Alefeld and **R. Lohner**, Karlsruhe

Received January 25, 1985

Dedicated to Professor R. Albrecht on the occasion of his 60th birthday

Abstract — Zusammenfassung

On Higher Order Centered Forms. If the real-valued mapping f has a representation of the form $f(x)=f(c)+(x-c)^n h(x)$, $x \in X$, then we introduce an interval expression which approximates the range of values of f over the compact interval X with order $n+1$. The well known centered form is the special case $n=1$ of this result.

AMS Subject Classification: 65G10.

Key words: Interval arithmetic, range of values.

Über zentrierte Formen höherer Ordnung. Für den Fall, daß f eine Darstellung der Form $f(x)=f(c)+(x-c)^n h(x)$, $x \in X$, besitzt, geben wir eine intervallmäßige Auswertung an, die den Wertebereich über dem kompakten Intervall X mit der Ordnung $n+1$ approximiert. Für $n=1$ erhält man die bekannten Aussagen über die zentrierte Form.

1. Introduction

A fundamental property of interval arithmetic is the fact that it allows to include the range of values of (rational) functions. It is well known that the distance of such a result to the exact range is strongly dependent on the representation of the function. For details see, for example, the discussion in Chapter 3 of [1]. In this paper we discuss a generalization of the following well known facts:

Let the real valued function f be defined on the real compact interval $X \subseteq \mathbb{R}$. Assume that for some $c \in X$ the function f can be represented as

$$f(x)=f(c)+(x-c)\cdot h(x), \quad x \in X=[x_1, x_2], \tag{1}$$

where h is a continuous real function defined on X. Let $d(X)=x_2-x_1$ be the *width* (or diameter) of the interval X and denote by $q(A,B)$ the Hausdorff *distance* of two compact intervals $A=[a_1,a_2]$ and $B=[b_1,b_2]$:

$$q(A,B)=\max(|a_1-b_1|, |a_2-b_2|). \tag{2}$$

Assume furthermore that $h(X)$ denotes a real compact interval with the properties that $h(x) \in h(X)$, $x \in X$, and that

$$d(h(X)) \leq \sigma \cdot d(X) \tag{3}$$

where σ is a real nonnegative constant.

Then it is well known (see [1], [3], [5], [6], [7], for example) that the real compact interval $f(X)$ defined by

$$f(X) := f(c) + (X - c) \cdot h(X) \tag{4}$$

has the so-called *quadratic approximation property* which means the following:

Let

$$W(f, X) = \{f(x) \mid x \in X\} \tag{5}$$

be the range of values of f over X. $W(f, X)$ is a compact real interval. It then holds that

$$q(W(f, X), f(X)) \leq \kappa \cdot (d(X))^2 \tag{6}$$

where κ is a real nonnegative constant. Furthermore $W(f, X) \subseteq f(X)$. The representation (1) of $f(x)$ is usually called *a centered form of f*. See, for example, [1], [3], [5], [6], [7]. In this note we show that (6) is only a special case of a more general result. We first consider a simple example.

2. Example

Let

$$f(x) = x^2 \cdot \frac{x-1}{x+1} = f(c) + (x-c)^2 \cdot h(x),$$

$$X = [-\alpha, \alpha], \ \alpha < 1$$

where

$$c = 0 \ \text{ and } \ h(x) = \frac{x-1}{x+1}.$$

A simple discussion yields for the range of f over $X = [-\alpha, \alpha]$ the interval

$$W(f, X) = \left[-\frac{\alpha^2(1+\alpha)}{1-\alpha}, 0\right].$$

Defining

$$[-\alpha, \alpha]^2 := W(x^2, [-\alpha, \alpha]) = [0, \alpha^2]$$

yields

$$f_S(X) := [-\alpha, \alpha]^2 \cdot \frac{[-\alpha, \alpha] - 1}{[-\alpha, \alpha] + 1} = \left[\frac{-\alpha^2(1+\alpha)}{1-\alpha}, 0\right],$$

where the index S stands for square. On the other hand defining $[-\alpha, \alpha]^2$ as the product of the two factors $[-\alpha, \alpha]$ and $[-\alpha, \alpha]$ we get

$$f_P(X) := [-\alpha, \alpha] \cdot [-\alpha, \alpha] \cdot \frac{[-\alpha, \alpha]-1}{[-\alpha, \alpha]+1} = \left[\frac{-\alpha^2(1+\alpha)}{1-\alpha}, \frac{\alpha^2(1+\alpha)}{1-\alpha}\right],$$

where the index P stands for product.

Obviously it holds that

$$q\big(W(f, X), f_P(X)\big) = \frac{\alpha^2(1+\alpha)}{1-\alpha} = 0\,\big((d(X))^2\big)$$

whereas (in this example even

$$q\big(W(f, X), f_S(X)\big) = 0$$

from which it follows, of course, that)

$$q\big(W(f, X), f_S(X)\big) = 0\,\big((d(X))^3\big).$$

3. Results

The result of the example from the preceding chapter is covered by the following main result of this paper.

Theorem: *Let the real valued function f be defined on a compact interval $X \subseteq \mathbb{R}$. Assume that for some $c \in X$ and for some integer $n \geq 1$ the function f can be represented as*

$$f(x) = f(c) + (x-c)^n \cdot h(x) \tag{7}$$

where h is a continuous function defined on X. Assume that $h(X)$ is a real compact interval for which

$$h(x) \in h(X), \quad x \in X \tag{8}$$

and

$$d\big(h(X)\big) \leq \sigma \cdot d(X) \tag{9}$$

hold. Defining the compact interval $f(X)$ by

$$f(X) := f(c) + W\big((x-c)^n, X\big) \cdot h(X) \tag{10}$$

then it follows that

$$f(x) \in f(X), \quad x \in X \tag{11}$$

and

$$q\big(W(f, X), f(X)\big) \leq \kappa \cdot \big(d(X)\big)^{n+1} \tag{12}$$

where κ is a nonnegative constant. □

The special case $n=1$ of this theorem is the quadratic approximation property discussed in Chapter 1.

Before we go into the details of a proof of this theorem we include some *remarks*.

a) Because of (12) we call (7) *a centered form of order* $n+1$. (Note that so-called standard centered forms of higher order were also introduced in Chapter 2.4 of [7]. These forms, however, do not exhibit higher than second order of convergence.) Methods which include the range of values with arbitrarily high order were already introduced in [2]. See also [4].

b) Because of $c \in X = [x_1, x_2]$ we have $0 \in X - c$ and therefore

$$W((x-c)^n, X) = \begin{Bmatrix} [0, \max\{(x_1-c)^n, (x_2-c)^n\}], & n=2,4,\ldots \\ [(x_1-c)^n, (x_2-c)^n], & n=1,3,5,\ldots \end{Bmatrix}. \tag{13}$$

Hence the interval $W((x-c)^n, X)$ which is needed in (10) can easily be computed.

c) The inequality (9) holds, for example, if $h(x)$ is a rational function and if the interval $h(X)$ is computed by replacing the variable x by the interval X and by performing all operations following the rules of interval arithmetic.

d) The example from Chapter 2 shows that in general (12) does not hold if instead of (10) one uses the interval

$$f_P(X) = f(c) + (X-c)^n \cdot h(X)$$

where $(X-c)^n$ is computed as the product of n intervals $X-c$.

e) If f is a polynomial whose derivative has a zero c of order $n \geq 1$ (with $n < m =$ degree of f) then

$$f(x) = f(c) + (x-c)^n \cdot h(x)$$

where

$$h(x) = \frac{f^{(n)}(c)}{n!} + \frac{f^{(n+1)}(c)}{(n+1)!}(x-c) + \ldots + \frac{f^{(m)}(c)}{m!}(x-c)^{m-n}.$$

Of course such a c is not known in general (if it exists at all) and therefore even for polynomials it is not an easy task to find a representation (7) of $f(x)$ with $n > 1$. (Therefore it must be stressed that the theorem is only of limited practical value.) If on the other hand the derivative of f has no zero in X then computing the exact range is trivial. □

For the proof of the theorem we need the following result whose proof is given in the appendix of this paper.

Lemma: *Let the assumptions of the theorem hold. If*

$$|h(w)| = \min_{x \in X} |h(x)|$$

then

$$f(c) + W((x-c)^n, X) \cdot h(w) \subseteq f(c) + \{(x-c)^n \cdot h(x) \mid x \in X\} = W(f, X). \tag{14}$$

□

Proof of the Theorem:

A) It holds that

$$f(x) = f(c) + (x-c)^n \cdot h(x) \in f(c) + W((x-c)^n, X) \cdot h(X) = f(X)$$

which is (11). It follows that $W(f,X)\subseteq f(X)$ and therefore that

$$q\big(W(f,X),f(X)\big)\le d\big(f(X)\big)-d\big(W(f,X)\big). \tag{15}$$

See [1], Chapter 2, Theorem 11.

B) From the inclusion relation (14) of the Lemma it follows that

$$d\big(W(f,X)\big)\ge |h(w)|\cdot d\big(W((x-c)^n,X)\big).$$

See [1], Chapter 2, (9) and (14).

C) From the definition of $f(X)$ by (10) we get that

$$d\big(f(X)\big)\le d\big(h(X)\big)\cdot |W((x-c)^n,X)|+d\big(W((x-c)^n,X)\big)\cdot |h(X)|$$

where the absolute value $|\cdot|$ of an interval $A=[a_1,a_2]$ is defined to be $|A|:=q(A,[0,0])$.

See [1], Chapter 2, Definition 8, (10) and (12).

D) It holds that $W((x-c)^n,X)\subseteq (X-c)^n$ where on the right hand side the product of n factors, all equal to $X-c$, has to be performed. Therefore, since $d(X)=|X-X|$,

$$|W((x-c)^n,X)|\le |(X-c)^n|\le |X-X|^n=\big(d(X)\big)^n$$

and

$$d\big(W((x-c)^n,X)\big)\le d((X-c)^n)\le \big(d(X)\big)^n.$$

See the remarks at the end of Chapter 2 in [1] for the last $\le$-sign.

E) From the triangle inequality for the Hausdorff distance it follows that

$$\begin{aligned}|h(X)|&=q\big(h(X),[0,0]\big)\\&\le q\big(h(X),h(w)\big)+q\big(h(w),[0,0]\big)\\&=q\big(h(X),h(w)\big)+|h(w)|\end{aligned}$$

and therefore that

$$|h(X)|-|h(w)|\le q\big(h(X),h(w)\big).$$

Since $h(w)\in h(X)$ it follows that

$$q\big(h(X),h(w)\big)\le d\big(h(X)\big)$$

(see [1], Chapter 2, Theorem 11) and therefore that

$$|h(X)|-|h(w)|\le d\big(h(X)\big).$$

F) Using B) – E) and (12) from Theorem 9 in Chapter 2 of [1] we get from (15) that

$$\begin{aligned}q\big(W(f,X),f(X)\big)\le{}& d\big(h(X)\big)\cdot |W((x-c)^n,X)|+\\&+d\big(W((x-c)^n,X)\big)\cdot |h(X)|-\\&-|h(w)|\cdot d\big(W((x-c)^n,X)\big)\le\\&\le 2\cdot d\big(h(X)\big)\cdot \big(d(X)\big)^n.\end{aligned}$$

Using the assumption (9) we get the assertion (12) of the theorem with $\kappa=2\sigma$.□

4. Remark

It is interesting to note that the proof of the theorem could also be performed by using the concept of a remainder form of f on X introduced by Cornelius and Lohner in [2]. Let therefore $f(X)$ again be defined by (10) and assume that $h_0 \in h(X)$ and define the compact real interval $F(X)$ by

$$F(X) := W\big(f(c)+(x-c)^n \cdot h_0, X\big)+(X-c)^n \cdot \big(h(X)-h_0\big)$$

where $(X-c)^n$ is computed as the product of n intervals all equal to $X-c$. Set

$$g(x) := f(c)+(x-c)^n h_0, \quad R(X) := (X-c)^n \cdot \big(h(X)-h_0\big).$$

Then $F(X)$ is a *remainder form of f on X* in the sense of Cornelius-Lohner [2], Theorem 4. See also [7], Chapter 6.7. Therefore, since $h_0 \in h(X)$ we have by Theorem 4 in [2] that

$$\begin{aligned} q\big(W(f,X), F(X)\big) &\leq 2\,|R(X)| \leq 2\,|X-c|^n \cdot |h(X)-h_0| \\ &\leq 2\big(d(X)\big)^n \cdot d\big(h(X)\big) \\ &\leq 2\sigma\big(d(X)\big)^{n+1}. \end{aligned}$$

On the other hand it holds that

$$\begin{aligned} f(X) &= f(c)+W\big((x-c)^n, X\big) \cdot h(X) \\ &= f(c)+W\big((x-c)^n, X\big) \cdot \big(h_0+h(X)-h_0\big) \\ &\subseteq W\big(f(c)+(x-c)^n \cdot h_0, X\big)+W\big((x-c)^n, X\big) \cdot \big(h(X)-h_0\big) \\ &\subseteq W\big(f(c)+(x-c)^n \cdot h_0, X\big)+(X-c)^n \big(h(X)-h_0\big) \\ &= F(X). \end{aligned}$$

Therefore

$$q\big(W(f,X), f(X)\big) \leq q\big(W(f,X), F(X)\big) \leq 2\sigma\big(d(X)\big)^{n+1}.$$

5. Appendix: Proof of the Lemma

We consider first the case that $h(w) \geq 0$ and that *n is an even integer*. We then have

$$f(c)+W\big((x-c)^n, X\big) \cdot h(w) = [f(c), f(c)+\max\big((x_1-c)^n, (x_2-c)^n\big) \cdot h(w)]$$

and

$$W(f,X) = [f(c), f(c)+(v-c)^n \cdot h(v)]$$

with some $v \in X$ for which

$$(v-c)^n h(v) \geq (x-c)^n h(x) \text{ for all } x \in X.$$

We only have to show that

$$(v-c)^n h(v) \geq \max\big((x_1-c)^n, (x_2-c)^n\big) \cdot h(w).$$

The validity of this inequality can be seen as follows.

By the definition of v it holds that

$$(v-c)^n h(v) \geq \max\left((x_1-c)^n h(x_1), (x_2-c)^n h(x_2)\right).$$

By the definition of $h(w)$ and since n is even the inequalities

$$(x_1-c)^n \cdot h(x_1) \geq (x_1-c)^n \cdot h(w)$$

and

$$(x_2-c)^n \cdot h(x_2) \geq (x_2-c)^n \cdot h(w)$$

hold. Therefore

$$\max\left((x_1-c)^n h(x_1), (x_2-c)^n h(x_2)\right) \geq \max\left((x_1-c)^n \cdot h(w), (x_2-c)^n \cdot h(w)\right) = \\ = \max\left((x_1-c)^n, (x_2-c)^n\right) \cdot h(w)$$

and the assertion is proved.

Consider now the case that $h(w) \geq 0$ and that *n is an odd integer*. In this case we have

$$f(c) + W\left((x-c)^n, X\right) \cdot h(w) = \\ = [f(c) + (x_1-c)^n h(w), f(c) + (x_2-c)^n h(w)]$$

and

$$W(f, X) = [f(c) + (u-c)^n h(u), f(c) + (v-c)^n \cdot h(v)]$$

where $u, v \in X$.

We first show that

$$f(c) + (x_1-c)^n h(w) \geq f(c) + (u-c)^n h(u).$$

By definition of u it holds that

$$(u-c)^n h(u) \leq (x_1-c)^n h(x_1).$$

Since $h(w) \leq h(x_1)$ and $(x_1-c)^n \leq 0$ it follows that

$$(x_1-c)^n h(w) \geq (x_1-c)^n h(x_1)$$

and therefore the assertion holds.

To complete the proof of (14) under our present assumption we have to show that

$$f(c) + (x_2-c)^n h(w) \leq f(c) + (v-c)^n h(v).$$

By definition of v it holds that

$$(v-c)^n h(v) \geq (x_2-c)^n h(x_2).$$

Since $h(w) \leq h(x_2)$ and $(x_2-c)^n \geq 0$ it follows that

$$(x_2-c)^n h(w) \leq (x_2-c)^n h(x_2)$$

and therefore the assertion holds.

The final part of the proof consists in considering *the case* $h(w) < 0$. Under this assumption the proof can again be performed by considering the subcases in which n is an even and odd integer, respectively. The details proceed similarly as in the case $h(w) \geq 0$. We omit the details. □

References

[1] Alefeld, G., Herzberger, J.: Introduction to Interval Computations. New York: Academic Press 1983.

[2] Cornelius, H., Lohner, R.: Computing the range of values of real functions with accuracy higher than second order. Computing *33*, 331 – 347 (1984).

[3] Hansen, E. R.: The centered form. In: Topics in Interval Analysis (Hansen, E., ed.). Oxford: 1968.

[4] Herzberger, J.: Zur Approximation des Wertebereichs reeller Funktionen durch Intervallausdrücke. Computing, Suppl. *1*, 57 – 64 (1977).

[5] Moore, R. E.: Interval Analysis. Englewood Cliffs, N.J.: Prentice-Hall 1966.

[6] Moore, R. E.: Methods and Applications of Interval Analysis. Philadelphia: SIAM 1979.

[7] Ratschek, H., Rokne, J.: Computer Methods for the Range of Functions. Chichester: Ellis Horwood 1984.

G. Alefeld and R. Lohner
Institut für Angewandte Mathematik
Universität Karlsruhe
Kaiserstrasse 12
D-7500 Karlsruhe
Federal Republic of Germany