WHERE TO BISECT A BOX? A THEORETICAL EXPLANATION OF THE EXPERIMENTAL RESULTS

R. Baker Kearfott
Department of Mathematics
University of Southwestern Louisiana
U.S.L. Box 4-1010
Lafayette, LA 70504-1010, USA
email rbk@usl.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email vladik@cs.utep.edu

I. INTRODUCTION

A. Why Bisect a Box?

Several important numerical optimization methods, especially methods with automatic results verification, contain the following *bisection* step (see, e.g., a survey [2]): at any given moment of time, we have a set of "boxes" (i.e., sets of the type $[a_1, b_1] \times \ldots \times [a_n, b_n]$) that contain potential optima of the given function $\varphi(x_1, \ldots, x_n)$. To proceed, we must divide one of these boxes into two smaller ones.

B. Where to Bisect a Box?

The Problem

We can bisect a given box in n different ways, depending on which of n sides we decided to halve. So, the natural question appears: which side should we cut? i.e., where to bisect a given box?

Traditional Bisection

Historically the first idea was to cut the *longest* side (for which $b_i - a_i \to \max$).

Ratz's Bisection is Better

Ratz has shown (in [3, 4]) that much better results are achieved if we choose a side i for which $|d_i|(b_i - a_i) \to \max$, where d_i is the known approximation for the partial derivative

$$\frac{\partial \varphi}{\partial x_i}.$$

Is Ratz's Bisection Optimal?

This is a purely experimental result, without a theoretical explanation. So, the natural question is:

- is Ratz's box-splitting strategy really the *best* one, or
- is it simply *better* than the ones known before, but an even better bisection strategy is possible?

In this paper, we show that (under certain reasonable assumptions) natural conditions really lead to Ratz's bisection.

II. INFORMAL (HEURISTIC) JUSTIFICATION OF RATZ'S BISECTION

Before we start a formal analysis of this problem, let us given an informal (heuristic) justification of Ratz's method.

One of the natural goals of bisection is to minimize the range of values of the optimized function φ. Most optimization methods are *first order methods* in the sense that at any given moment of time, we have estimates f for the value of the optimized function φ and d_i for the values of its derivatives. Based on these estimates, we can estimate the range of values $\varphi([a_1, b_1], \ldots, [a_n, b_n])$: namely, if we assume that all the intervals are narrow, we can approximate the function φ by a few first terms in its Taylor series. Since we only know the first derivatives, it is natural to use a linear approximation:

$$\varphi(x_1, \ldots, x_n) \approx \varphi_{\text{approx}}(x_1, \ldots, x_n),$$

where

$$\varphi_{\text{approx}}(x_1, \ldots, x_n) = f + \sum_{i=1}^{n} d_i(x_i - c_i), \qquad (1)$$

and $c_i = (a_i + b_i)/2$ is a midpoint of i-th side.

When $x_i \in [a_i, b_i]$, the largest possible value of $\varphi_{\text{approx}}(x_1, \ldots, x_n)$ is attained when each $x_i = b_i$ for $d_i > 0$ and $x_i = a_i$ for $d_i < 0$. Hence, this largest values is equal to $\sum |d_i|(b_i - a_i)/2$. Similarly, the smallest possible value of $\varphi_{\rm approx}(x_1,\ldots,x_n)$ is equal to $-\sum |d_i|(b_i-a_i)/2$. Hence, the range of φ is approximately equal to

$$\sum_{i=1}^n |d_i|(b_i-a_i).$$

Our goal is to choose a bisection after which this range becomes the smallest possible. When we bisect i–th side, the i–th term in this sum halves. So, to get the smallest possible range after bisection, we must choose a side i for which the corresponding term is the largest possible. This is exactly Ratz's bisection strategy.

Comment. This explanation is only applicable when the box is already narrow. But the whole idea is to find the best bisection for the cases when the boxes are wide. So, this explanation is a heuristic argument in favor of Ratz's method, but it does not answer the question of whether this method is indeed the best for boxes of arbitrary size. To answer that question, we need a more formal analysis of this problem.

III. OUR MAIN IDEA

We want to design a method that, based on the estimate f for the function, on the estimates d_i for its derivatives, and on the coordinates a_i, b_i of the box, chooses a side to be bisected.

There are some cases when the choice is clear: e.g., when all estimates d_i are the same ($d_1=d_2=\ldots=d_n$), then it makes sense to bisect along the longest side ($b_i-a_i\to\min$).

To choose a method in the general case, we will apply the following idea: The main objective of numerical algorithms is to solve real-life problems. In real-life problems, the numerical values of the variables $x_1,\ldots,x_n$ depend on the choice of the measuring units:

- If, e.g., x_1 is length, and we originally measured length in inches, then we could switch to centimeters and get new numerical values $x_1'=\lambda x_1$ (where $\lambda = 1$ in/1 cm).
- If, e.g., x_3 is a spatial coordinate, then we can define it as going in an opposite direction, thus changing x_3 into $x_3'=-x_3$.

In general, the numerical value of a physical quantity x_i is defined modulo an arbitrary re-scaling

$$x_i\to k_i x_i. \qquad (2)$$

The optimal optimization algorithm should not depend on what exactly units we use, because otherwise, if it works optimally for inches, and differently for centimeters, it will thus not be optimal for the centimeters! As a result, the optimal method of choosing a bisection should not change when we apply arbitrary linear transformations.

Under transformations (2), the values of the function remain unchanged, the coordinates of the box change accordingly, and the partial derivatives change as

$$\frac{\partial\varphi}{\partial x_i}\to\frac{1}{k_i}\cdot\frac{\partial\varphi}{\partial x_i}.$$

It is therefore natural to require that the estimate d_i for the partial derivative change as $d_i\to d_i/k_i$.

Now, we are ready to formalize this idea.

IV. FORMALIZATION OF OUR IDEA AND THE MAIN RESULT

Denotation. *Let I denote the set of all (closed) intervals on the real line R. For an interval $[a,b]$ and a real number k, the product $k\cdot[a,b]$ is defined as $[ka,kb]$ if $k>0$ and $[kb,ka]$ if $k<0$.*

Definitions. *Let an integer $n\ge 1$ be fixed.*

- *A function*

  $$S(f,d_1,\ldots,d_n,[a_1,b_1],\ldots,[a_n,b_n])$$

  from $R^{n+1}\times I^n$ to $\{1,\ldots,n\}$ is called a bisection strategy if for $d_1=\ldots=d_n$, its value is equal to i for which the width b_i-a_i of the interval $[a_i,b_i]$ is the largest.

- *By a re-scaling, we mean a transformation*

  $$f\to f,\ d_i\to\frac{d_i}{k_i},\ [a_i,b_i]\to k_i\cdot[a_i,b_i],$$

  where $k_i\ne 0$ are real numbers.

- *A bisection strategy is called invariant if the value of the function S does not change under an arbitrary re-scaling; in other words, if*

  $$S(f,\frac{d_1}{k_1},\ldots,\frac{d_n}{k_n},k_1\cdot[a_1,b_1],\ldots,k_n\cdot[a_n,b_n])=$$

  $$S(f,d_1,\ldots,d_n,[a_1,b_1],\ldots,[a_n,b_n])$$

  for all possible numbers $f,d_1,\ldots,d_n$, for all possible intervals $[a_1,b_1],\ldots,[a_n,b_n]$, and for all possible $k_1,\ldots,k_n$.

PROPOSITION. *The only invariant bisection strategy is choosing i for which $|d_i|(b_i - a_i) \to \max$.*

So, we get a justification of Ratz's bisection strategy.

Idea of the Proof

It is easy to check that Ratz's strategy is indeed invariant.

Let us now show that every invariant strategy has the desired form. Let an arbitrary tuple $(f, d_1, \ldots, d_n, [a_1, b_1], \ldots, [a_n, b_n])$ be given. To find the value of the invariant bisection strategy S for this tuple, let us take $k_i = d_i$. Then, from invariance of S, we conclude that the original choice coincides with choice for the new tuple

$$(f', d'_1, \ldots, d'_n, [a'_1, b'_1], \ldots, [a'_n, b'_n]) =$$

$$(f, \frac{d_1}{k_1}, \ldots, \frac{d_n}{k_n}, d_1 \cdot [a_1, b_1], \ldots, d_n \cdot [a_n, b_n]) =$$

$$(f, 1, \ldots, 1, d_1 \cdot [a_1, b_1], \ldots, d_n \cdot [a_n, b_n]).$$

For this new tuple, $d'_1 = \ldots = d'_n$ $(= 1)$, therefore, according to the definition of the bisection strategy, we must choose i for which the interval $[a'_i, b'_i] = d_i \cdot [a_i, b_i]$ is the widest. The width of i–th interval $[a'_i, b'_i] = d_i \cdot [a_i, b_i]$ is equal to $|d_i|(b_i - a_i)$, so, we choose i for which $|d_i|(b_i - a_i) \to \max$. Q.E.D.

ACKNOWLEDGMENTS

This work was supported in part by NASA grants NAG 9-757 and NCCW-0089, by NSF grants DUE-9750858 and EEC-9322370, and by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518.

The authors are thankful to Bruno Lang for valuable discussions, and to the anonymous referees for their help.

REFERENCES

[1] Csendes, T.; Ratz, D.; *Subdivision direction selection in interval methods for global optimization*, SIAM J. Numer. Anal., 1997.

[2] Kearfott, R.B.; *A Review of Techniques in the Verified Solution of Constrained Global Optimization Problems*, In: Kearfott, R.B; Kreinovich, V.; *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 23–59.

[3] Ratz, D.; *Automatische Ergebnisverifikation bei globalen Optimierungsproblemen*, Ph.D. dissertation, Universität Karlsruhe, 1992.

[4] Ratz, D.; *Box-Splitting Strategies for the Interval Gauss–Seidel Step in a Global Optimization Method*, Computing, 1994, Vol. 53, pp. 337–354.