Empirical convergence speed of inclusion functions for facility location problems ⋆

B. Tóth and J. Fernández

University of Murcia, Dpt. Statistics and Operations Research, Murcia, Spain

T. Csendes

University of Szeged, Institute of Informatics, Szeged, Hungary

Abstract

One of the key points in interval global optimization is the selection of a suitable inclusion function which allows to solve the problem efficiently. Usually, the tighter the inclusions provided by the inclusion function, the better, because this will make the accelerating devices used in the algorithm more effective at discarding boxes. On the other hand, whereas more sophisticated inclusion functions may give tighter inclusions, they require more computational effort than other providing larger overestimations. In an earlier paper the *empirical convergence speed* of inclusion functions was defined and studied, and it was shown to be a good indicator of the inclusion precision. If the empirical convergence speed is analyzed for a given type of functions, then one can select the appropriate inclusion function to be used when dealing with those type of functions. In this paper we present such a study, dealing with functions used in competitive facility location problems.

Key words: Inclusion function, convergence speed, facility location
AMS Subject Classification: 65G30, 65G40, 65K05, 90C30, 90B85

1 Empirical convergence speed

The following notation will be use throughout. $X = [\underline{X}, \overline{X}] \subseteq \mathbb{R}$ is a real interval, $\mathbb{I}$ is the set of real intervals, $X = (X_1, \ldots, X_n) \in \mathbb{I}^n$ is an n-dimensional

⋆ This work has been supported by the Grants OTKA T 034350, T 048377, OMFB D-30/2000, OMFB E-24/2001 and by the Ministry of Science and Technology of Spain under the research projects BEC2002-01026 and HH2004-0014. B. Tóth was on leave from Research Group on Artificial Intelligence of the Hungarian Academy of Sciences and the University of Szeged, Szeged, Hungary.

Email address: `{boglarka,josefdez}@um.es` (B. Tóth and J. Fernández).

interval (also called *box*), $w(X) = \overline{X} - \underline{X}$ is the *width* of $X \in \mathbb{I}$ and $w_{rel}(X) = w(X)/\max\{\min_{x\in X}|x|, 1\}$ its relative width. The width of a box $X \in \mathbb{I}^n$ is to be understood as $w(X) = \max_{i=1,\ldots,n} w(X_i)$. The midpoint of the interval X will be denoted by $mid(X)$.

The standard global optimization problem can be defined as

$$\min_{x\in X} f(x), \tag{1}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously differentiable function and $X \in \mathbb{I}^n$ is a box. This paper is related with the solution of (1) via interval branch-and-bound methods [13]. In particular, we are interested in finding the most suitable inclusion function for solving (1).

Definition 1 *Let f be a continuous function. A function $F : \mathbb{I}^n \rightarrow \mathbb{I}$ is an* inclusion function *of f, if for every $X \in \mathbb{I}^n$ and $x \in X$, $f(x) \in F(X)$, i.e. $f(X) = \{f(x) \mid x \in X\} \subseteq F(X)$.*

Usually, a surrogate for this problem is to find the inclusion function providing tighter inclusions, since in this way the different accelerating devices used in interval B&B methods are more efficient at discarding boxes, thus the problem is solved faster. The classical measure for the quality of an inclusion function is the convergence order.

Definition 2 *Let F be an inclusion function of f and Y be a box, where f is defined. Then the* convergence order *of F is at least α, if there exists a positive constant c such that the inequality*

$$w(F(X)) - w(f(X)) \leq c \cdot w(X)^{\alpha} \tag{2}$$

holds for every $X \subseteq Y$.

However, this definition implies a *worst case* analysis. Instead of that, in a recent paper [15], the *empirical convergence speed* of F has been defined with the aim of measuring the *average behaviour*. The empirical convergence speed is obtained by approximating the values of α and c in the equation obtained from (2) by changing the '$\leq$' sign to '$=$'. To do it, the equation is transformed into

$$\log_{10}\left(w\left(F(X)\right) - w\left(f(X)\right)\right) = \log_{10}(c) + \alpha \log_{10}\left(w(X)\right).$$

Then, the actual values of $w(F(X))$, $w(f(X))$ and $w(X)$ are computed for a given set of boxes, and the α and c constants that best fit the computed widths are determined with linear regression. This is in accordance to the definition of the theoretical convergence order, except that in order to get the best approximation, the inequality is changed to equality. In this way we will obtain information about the *average behaviour* of the inclusion function, instead of the *worst case* analysis. This change fits exactly in what is needed for algorithmic development, since when used in an algorithm, the average behaviour of inclusion functions for realistic size intervals is what matters.

For a given test function f we can use different sets of boxes to determine α and c. In this paper two different types of sets have been considered:

– sequences of embedded boxes, all containing a global minimizer point, and
– sequences of random intervals converging to a point.

As explained before, for every box in a sequence we compute the inclusion function, an approximation of the range over the box, and the width of the box. The hardest one to produce is the approximation of the range. We obtain it by computing an enclosure of the minimum and the maximum of the function in the given interval with a Moore-Skelboe type branch-and-bound global optimization algorithm [13], using the termination criterion $w_{rel}([\tilde{f}, \overline{F}(X)]) \leq 10^{-8}$, where $\tilde{f}$ is the best upper bound found so far by the algorithm.

In particular, in this paper we discuss the application of empirical convergence speed values to some facility location problems. The inclusion functions examined in the study were

natural interval extension: denoted by $F_n(X)$,
centered form[9]: $F_c(X) = F(c) + F'(X)(X - c)$ where $c = mid(X)$ and $F'(X)$ is the inclusion function of the gradient (obtained by automatic differentiation),
Baumann's optimal centered form [1]: $F_b(X) = F(b) + F'(X)(X - b)$, where $b = (U\underline{X} - L\overline{X})/(U - L)$ and $F'(X) = [L, U]$ is the inclusion of the gradient given by automatic differentiation,
slope arithmetic form [3]: $F_s(X) = F(c) + s_f(c, X)(X - c)$, where $c = mid(X)$ and $s_f(c, X)$ is the inclusion of the slopes to any point in X from c given by the slope arithmetic, and
affine arithmetic form [11]: $(F_a(X))$ where a variable x is represented by a first-degree polynomial $\hat{x} = x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + \ldots + x_n\varepsilon_n$, x_i-s are finite floating-point numbers, and the ε_i-s are symbolic real variables in [-1,1].

In what follows, we denote the above inclusion functions by $F_i(X), i \in I = \{n, c, b, s, a\}$. Some of those inclusion functions provide more useful information, which can be used in global optimization algorithms. For instance, with the centered and Baumann forms an inclusion of the gradient is available, which allows us to use the monotonicity test. Another pruning test described in [10] can be applied to Baumann, centered and slope arithmetic forms. This fact has also to be taken into account when selecting an inclusion function for solving a problem (in the results in Table 2, we used this information when solving the problems with the different inclusion functions).

2 The examined facility location problem

In the present work we consider a special type of objective functions: those belonging to competitive facility location problems [5]. Competitive location deals with the problem of locating facilities to provide a service (or goods) to the customers (or consumers) of a given geographical area where other competing facilities offering the same service are already present (or will enter to the market in the near future). Many models on the subject are available in the literature that vary in the ingredients which form the model (location space, number of facilities to be located, patronising behaviour, attraction of customers to the facilities, ...). Many references can be found in [7,12]. Next, we present a very general model for locating a new single facility in the plane.

Assume that a chain of warehouses or shops wants to set up a new facility in a planar market, where similar facilities of competitors, and possibly of its own chain, are already present. Fixed demand points split their demand in a probabilistic way over all facilities in the market, proportionally with their attraction to each facility through a gravitational type model. The demand is deterministic and inelastic. The objective is the maximization of the profit obtained by the chain. Both the location and the quality (design) of the new facility are to be found. The notation used in the model is the following.

x,α	location and quality of the new facility,
n	number of demand points,
p_i	demand points ($i = 1, \dots, n$),
ω_i	demand (or buying power) at p_i,
m	number of existing facilities,
f_j	existing facilities ($j = 1, \dots, m$),
k	number of existing facilities belonging to the chain,
d_{ij}	distance between demand point p_i and facility f_j,
d_{ix}	distance between demand point p_i and the new facility x,
α_{ij}	quality of facility f_j as perceived by demand point p_i,
$g_i(\cdot)$	a non-negative non-decreasing function,
γ_i	weight for the quality of x as perceived by demand point p_i,
$\alpha_{ij}/g_i(d_{ij})$	attraction that demand point p_i feels for facility f_j,
$\gamma_i\alpha/g_i(d_{ix})$	attraction that demand point p_i feels for the new facility x.

The total market share attracted by the chain is given by

$$M(x,\alpha) = \sum_{i=1}^{n} \omega_i \frac{\dfrac{\gamma_i\alpha}{g_i(d_{ix})} + \displaystyle\sum_{j=1}^{k} \frac{\alpha_{ij}}{g_i(d_{ij})}}{\dfrac{\gamma_i\alpha}{g_i(d_{ix})} + \displaystyle\sum_{j=1}^{m} \frac{\alpha_{ij}}{g_i(d_{ij})}},$$

and the profit it gets (to be maximized) is given by

$$\Pi(x,\alpha) = F(M(x,\alpha)) - G(x,\alpha), \tag{3}$$

where $F(\cdot)$ is an increasing function which transforms the market share into expected sales and $G(x,\alpha)$ gives the operating costs of the new facility. It has been suggested in [5] to choose $F(M(x,\alpha)) = c \cdot M(x,\alpha), c > 0$, and $G(x,\alpha) = G_1(x) + G_2(\alpha)$ with $G_1(x) = \sum_{i=1}^{n} \omega_i \frac{1}{(d_{ix})^2+\phi_i}$ ($\phi_i > 0$ are given parameters) and $G_2(\alpha) = e^{\frac{\alpha}{\alpha_0}+\alpha_1} - e^{\alpha_1}$ (with $\alpha_0 > 0$ and α_1 given values). For more details on the model see [5]. Notice that Π, the objective function of our location problem, is a complex, highly nonlinear function. We are interesting in finding the most suitable inclusion function for Π, which allows us to solve the corresponding location problem efficiently.

3 Numerical Results and Conclusion

For the experiments the C-XSC-2.0 library [3] was used with the automatic differentiation tool of CToolbox-2.0 software [6,8]. For the affine arithmetic form the freely available libaa library was used [4]. When an inclusion was accurate (compared to our range approximation), and hence the overestimation was zero, its logarithm was set to -30.

Table 1
The determined empirical convergence speed $(\alpha, \log_{10} c)$ values for the studied inclusion functions, using all the examined facility location problems. The determined α values are followed in brackets by the correlation coefficient of the regression.

	α				
Sequence	Natural	Centered	Baumann	Affine	Slope
Opt	1.026 (.99)	2.667 (.89)	3.555 (.80)	2.662 (.85)	1.997 (.98)
Rand	1.067 (.99)	2.368 (.94)	4.546 (.81)	2.901 (.87)	1.993 (.99)
All	1.060 (.99)	2.422 (.93)	4.300 (.77)	2.853 (.86)	1.993 (.99)
	$\log_{10} c$				
Opt	2.608	3.937	3.936	3.461	2.052
Rand	2.712	3.450	-3.379	3.245	2.010
All	2.697	3.518	-2.161	3.288	2.017

We examined 12 objective functions of the form (3). In 6 of them we considered 50 demand points and 5 existing facilities (generated in a random way, as well as the rest of the parameters), and in the other functions 100 demand points and 10 existing facilities. The results obtained were very similar for all the cases but one. In Table 1 we can see the approximated α and c values for the inclusion functions studied, when all the computed widths of all the examined functions are considered.

We can see that for both the optimum following sequences (Opt) and the random sequences of boxes (Rand) the Baumann form is the inclusion function with the highest convergence speed, as well as when we consider the computed widths of all the sequences together (All). This is due to the fact that on monotonous boxes the Baumann form is accurate. The second best inclusion is the affine arithmetic form, and in the only analyzed case in which the Baumann was not the best function, it was the affine arithmetic form the one providing the best values. The centered and slope forms follow in this ranking, and finally, as expected, we have the natural interval extension, with an α value very close to the theoretical convergence order of 1.

Figure 1 illustrates the relation of the investigated empirical convergence speed values when we considered the 12 investigated functions and all the sequences.

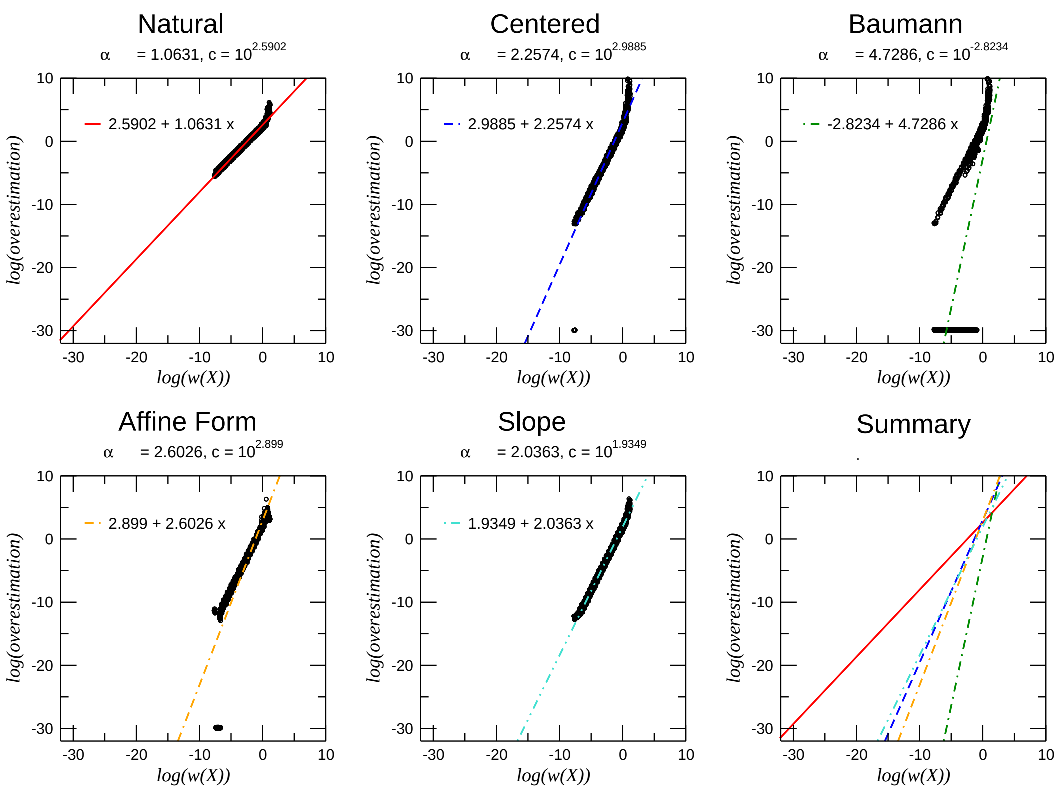

Fig. 1. The results of the linear regression for all the data on log-log scales indicating that the Baumann form and the natural interval extension are the best choice depending on the width of the box.

From the picture we can conclude that for the examined functions the empirical convergence speed suggests to use the Baumann form for every box with width smaller than 19.56 and the natural interval extension if the width is bigger (see the 'Summary' picture in Fig. 1, in which we can see the breaking point at which it is recommended to change of inclusion function).

Algorithm 1 Adaptive multi-inclusion B&B algorithm

Input: $(X, I(w(Y)))$ $I(w(Y)) \to J(\subseteq I)$*: indexes of inclusion functions*
$\mathcal{L}_\mathcal{W} \leftarrow X,\ \mathcal{L}_\mathcal{S} \leftarrow \emptyset$
while ($\mathcal{L}_\mathcal{W} \neq \emptyset$)
  Select an interval Y from $\mathcal{L}_\mathcal{W}$ — *Selection Rule*
  Evaluate $F(Y) = \bigcap_{i \in I(w(Y))} F_i(Y)$, Update $\tilde{f}$ — *Cut-Off Test*
  Discard $Z \in \mathcal{L}_\mathcal{S}, \mathcal{L}_\mathcal{W}$ if $\underline{F}(Z) > \tilde{f}$
  if (Y cannot be discarded or reduced) — *Discarding Tests*
    Divide Y into subintervals Y_1, Y_2 — *Division Rule*
    if (Y_i satisfies the termination criterion) — *Termination Rule*
      Store Y_i in $\mathcal{L}_\mathcal{S}$
    else Store Y_i in $\mathcal{L}_\mathcal{W}$

To check in practice the usefulness of the empirical convergence speed, we applied an adaptive multi-inclusion algorithm (Algorithm 1), and measured the time needed for solving a location problem when using different inclusion functions. The obtained results can be seen in Table 2. We give the average time needed for solving all the problems. The first four columns refer to the cases when only one inclusion function was used (the results for the natural extension are not given, since the program lasted more than two hours). In the next three columns we have used the natural interval extension in addition to the given inclusion function (notice that the natural interval extension is given as a by-product when evaluating the gradients or slopes with the used libraries). The last column contains the result when we switch from one inclusion to another depending on the width of the box to be evaluated. It can be seen from the table that the results are in correspondence with the suggestion from the empirical convergence speed.

Table 2
The running times for the solution of the facility location problems in seconds. In the rows we give the results for accuracy 10^{-2}, 10^{-4} and 10^{-8}, respectively.

Accuracy	Only one inclusion function				With natural extension			Switching
	Cent.	Slope	Affine	Baum.	Cent.	Slope	Baum.	na/na-ba-af
10^{-2}	362.2	2683.6	492.7	294.6	346.0	2685.7	281.8	
10^{-4}	480.1	3695.1	642.6	385.6	457.6	3721.0	368.6	
10^{-8}	990.7	—	1653.2	840.9	943.5	—	806.8	922.3

In future research, we plan to examine different inclusion functions, e.g. the Taylor model [2]. We also plan to consider the effect of other aspects related to the inclusion functions when used in a global optimization algorithm, e.g., other discarding tests.

References

[1] E. Baumann, Optimal centered forms. BIT, 28(1988) 80–87.

[2] M. Berz and G. Hofstäster, Computation and application of Taylor polynomials with interval remainder bounds. Reliable Computing, 4(1998) 83–97.

[3] M. Bräuer, W. Hofschuster, and W. Krämer, Steigungsarithmetiken in C-XSC. Preprint 2001/3, Universität Wuppertal, 2001.

[4] O. Gay, Libaa: Une librarie C++ d' Affine Arithmetic. Technical Report, École Polytechnique Fédérale de Lausanne, 2003.

[5] J. Fernández, B. Pelegrín, F. Plastria, and B. Tóth, Solving a Huff-like competitive location and design model for profit maximization in the plane, European Journal of Operational Research, to appear, available at `http://www.inf.u-szeged.hu/~boglarka/compet.ps.gz`

[6] R. Hammer, M. Hocks, U. Kulisch, and D. Ratz, C++ Toolbox for Verified Computing I: Basic Numerical Problems: Theory, Algorithms, and Programs. Springer-Verlag, Berlin, 1995.

[7] M. Kilkenny and J.F. Thisse, Economics of location: a selective survey, Computers and Operations Research, 26(1999) 1369–1394.

[8] R. Klatte, U. Kulisch, A. Wiethoff, C. Lawo, and M. Rauch, C–XSC, A C++ Class Library for Extended Scientific Computing. Springer-Verlag, Berlin, 1993.

[9] R. Krawczyk and K. Nickel, The centered form in interval arithmetics: Quadratic convergence and inclusion isotonicity. Computing 28(1982) 117–137.

[10] J.A. Martínez, L.G. Casado, I. García, Ya.D. Sergeyev, and B. Tóth, On an efficient use of gradient information for accelerating interval global optimization algorithms. Numerical Algorithms, 37(2004) 61–69.

[11] F. Messine, Extensions of Affine Arithmetic: Application to Global Optimization, Journal of Universal Computer Science, 8(1992) 992-1015.

[12] F. Plastria, Static competitive facility location: an overview of optimisation approaches, European Journal of Operational Research, 129(2001) 461–470.

[13] H. Ratschek and J. Rokne, New Computer Methods for Global Optimization. Ellis Horwood Ltd., Chichester, England, 1988.

[14] D. Ratz, Automatic Slope Computation and its Application in Nonsmooth Global Optimization. Shaker Verlag, Aachen, Germany, 1998.

[15] B. Tóth and T. Csendes, Empirical investigation of the convergence speed of inclusion functions in a global optimization context, Reliable Computing, 11(2005) 253–273.