Solving Non-Differentiable Problems by the Flying Elephants Approach

Adilson Elias Xavier
Vinicius Layter Xavier
COPPE - Federal University of Rio de Janeiro
{adilson,vinicius}@cos.ufrj.br

ABSTRACT:

Flying Elephants (FE) is a generalization and a new interpretation of the Hyperbolic Smoothing approach. The article introduces the fundamental smoothing procedures. It presents a general overview of successful applications of the approach for solving a select set of five important problems, namely: distance geometry, covering, clustering, Fermat-Weber and hub location. For each problem it is presented the original non-smooth formulation and the succedaneous completely differentiable one. Computational experiments for all related problems obtained results, which exhibited a high level of performance according to the different criteria of consistency, robustness and efficiency. For each problem, some results to illustrate the performance of FE are also presented.

KEYWORDS. Non-smooth Programming, Smoothing, Distance Geometry, Clustering, Location.

1. Introduction

The core idea of the Flying Elephants method is the smoothing of a given non-differentiable problem. In a sense, the process whereby this is achieved is a generalization and a new interpretation of a smoothing scheme, called Hyperbolic Smoothing (HS), presented in Santos (1997) for non-differentiable problems in general. This technique was developed through an adaptation of the hyperbolic penalty method originally introduced by Xavier (1982) in order to solve the general non-linear programming problem.

By smoothing we fundamentally mean the substitution of an intrinsically non-differentiable two-level problem by a C^{∞} differentiable single-level alternative. This is achieved through the solution of a sequence of differentiable sub-problems which gradually approaches the original problem. Each sub-problem, owing to its being indefinitely differentiable, can be comfortably solved by using the most powerful and efficient algorithms, such as conjugate gradient, quasi-Newton or Newton methods.

First, the FE method incorporates any C^{∞} smoothing scheme, for instance the hyperbolic smoothing approach. The HS approach has been applied for solving a set of mathematical hard problems. Despite these problems present a non-differentiable and a non-convex structure with a large number of local minimizers, the HS method has produced efficiently and reliably very deep local minima. The paper presents a survey of successful applications of the HS approach for solving a select set of important problems, namely: distance geometry (Macambira ((2003), Xavier (2003), Souza(2010) and Souza et al. (2011)), covering (Xavier & Oliveira (2005)), clustering (Xavier (2010), Xavier &-Xavier (2011) and Bagirov et al. (2012b)), Fermat-Weber (Xavier (2012) and Xavier et al. (2012)) and hub location (Gesteira (2012), and Xavier, Gesteira & Xavier (2012)). There are other successful applications which are not presented in this survey, such as: minimax (Chaves(1997) and Bagirov et al. (2012a)) and packing problems.

The new name of the methodology, Flying Elephants, is definitely not associated to any analogy with the biology area. It is only a metaphor, but this name is fundamentally associated with properties of the method. The Flying feature is directly derived from the C^{∞} differentiability property of the method, which has the necessary power to make the fly of the heavy elephant feasible. Moreover, it permits intergalactic trips into spaces with large number of dimensions, differently of the short local searches associated to traditional heuristic algorithms. On the other side, the convexification feature also associated to the HS method is analogous to the local action of the Elephant landing, eliminating a lot of local minima points.

2. The Fundamental Smoothing Procedures

The Flying Elephants method is based on the hyperbolic smoothing of the non-differentiable functions belonging to the optimization problem formulation. We will present the two basic smoothing procedures. First, we will consider the smoothing of the absolute absolute value function $|u|$ where $u \in \Re$. For this purpose, $y > 0$, let us define the function:

$$\theta(u,\gamma) = \sqrt{u^2 + \gamma^2} \qquad (1)$$

Function θ has the following properties:

(a) $\lim_{\gamma \to 0} \theta(u,\gamma) = |u|$;

(b) θ *is* C^{∞} *function.*

For smoothing the function $\psi(u) = \max(0,u)$ we use:

$$\phi(u,\tau) = \left(u + \sqrt{u^2 + \tau^2}\right)/2 \qquad (2)$$

Function ψ has the following properties:

(a) $\lim_{\gamma \to 0} \phi(u,\gamma) = \psi(u)$

(b) $\phi(u,\tau)$ is an increasing convex C^{∞} function in variable u.

3. Distance Geometry Problem

Let $G = (V, E)$ denote a graph, in which for each arc $(i, j) \in E$ it is associated a measure $a_{ij} > 0$. The problem consists of associating a vector $x_i \in \Re^n$ for each knot $i \in V$, basically addressed to represent the position of this knot into a n-dimensional space, so that Euclidean distances between knots, $\|x_i - x_j\|$, corresponds appropriately to the given measures a_{ij}:

$$minimize \;\; f(x) = \sum_{(i,j) \in E} \left(\|x_i - x_j\| - a_{ij}\right)^2 \qquad (3)$$

This formulation presents the non-differentiable property due the presence of the Euclidean norm term. Moreover, the objective function is non-convex, so the problem has innumerable local minima. For solving the problem (3) by using the FE technique is only necessary to use the function $\theta(u,\gamma)$ and to define $u = \|x_i - x_j\|$:

$$minimize\ \ f(x) = \sum_{(i,j)\in E} \left(\theta\left(\|x_i - x_j\| - a_{ij}, \gamma\right) - a_{ij}\right)^2 \qquad (4)$$

Beside its smoothing properties the function θ also has the important convexification power. Xavier (2003) shows the following theoretical result:

Proposition 1: There is a value $\bar{\gamma}$ such as, for all values $\gamma > \bar{\gamma}$ the Hessian matrix $\nabla^2 F(x)$ will be positive definite.

Souza et al (2011) extend the theoretical result of Proposition 1 to a more general distance geometry problem. The presented computational results for classical instances clearly show both the robustness and the efficiency of the FE method.

In order to illustrate the computational properties of the Flying Elephants method, we took the traditional lattice problem originally proposed by Moré and Wu (1995). This instance is a synthetic problem, where the knots are located on the intersection of s planes that cut a cube in the three principal directions in equal intervals. The numerical experiments have been carried out on a Intel Core i7-2620M Windows Notebook with 2.70GHz and 8 GB RAM. The columns of Table 1 show the number of splits of the cube (s), the number of variables of the problem $(m = 3s^3)$, the occurrences of correct solutions obtained in 10 tentative solutions *(Occur.)*, the average value of the correct solutions f_{FE} and the mean CPU time (T_{Mean}) given in seconds associated to 10 tentative solutions.

s	$n = 3s^3$	Occur.	$f_{FE_{Aver}}$	T_{Mean}		s	$n = 3s^3$	Occur.	$f_{FE_{Aver}}$	T_{Mean}
3	81	0	-	0.1		12	5184	8	0.15E-1	143
4	192	6	0.27E-6	0.7		13	6591	7	0.32E-1	222
5	375	8	0.29E-5	2.8		14	8232	8	0.18E-1	380
6	648	8	0.19E-4	7.6		15	10125	6	0.65E-1	543
7	1029	5	0.16E-4	19		16	12288	7	0.42E-1	835
8	1536	8	0.29E-3	45		17	14739	6	0.16E0	1270
9	2187	6	0.86E-3	97		18	17496	7	0.21E0	1853
10	3000	7	0.95E-3	45		19	20577	8	0.24E0	2335
11	3993	6	0.17E-2	81		20	24000	8	0.59E0	3187

Table 1: Distance Geometry Problem - Moré-Wu Lattice Instance

This considered lattice instance is a synthetic problem which of global solution f^* assumes a value equal zero, $f^*=0$. So, the small values exhibited in the column $f_{FE_{Aver}}$ of Table 1 indicate the robustness of the FE approach, from cases s=4 to s=20, the last one with 3119800 arcs. An and Tao (2000) proposed an approach for solving this problem based on the difference of convex functions optimization algorithms and exhibits similar results, but with up to s=16.

4. Covering Problems

Let S be a finite region in $\Re^2$. A set of q figures constitutes a covering of order *1* of S; if each point $s \in S$ belongs to at least one figure. Coverages of a higher order can be defined in a similar manner. Problems inherent to the covering of $\Re^2$ regions by circles, of $\Re^3$ regions by spheres, and even regions in higher dimensional spaces have been the object of research for many decades. We consider the special case of covering a finite plane domain S optimally by a given number q of circles. We first discretize the domain S; into a finite set of m points s_j, $j = 1, \cdots, m$. Let x_i, $i = 1, \cdots, q$ be the centers of the circles that must cover this set of points.

The optimal placing of the centers must provide the best-quality covering, that is, it must minimize the most critical covering. If X^* denotes an optimal placement and X is the set of all placements, the covering problem assumes a *min-max-min* form:

$$X^* = \underset{X \in \Re^{2q}}{\arg\min} \; \max_{j=1,\dots,m} \; \min_{i=1,\dots,q} \left\| s_j - x_i \right\|_2 , \tag{5}$$

By performing an ε perturbation and by using the FE approach the three-level strongly non-differentiable *min-max-min* problem can be transformed in a one-level completely smooth one:

$$\begin{aligned} & \textit{minimize} \;\; z \\ & \textit{subject to}: \sum_{i=1}^{q} \phi \left(z - \left\| s_j - x_i \right\|_2 , \tau \right) \geq \varepsilon, \quad j = 1, \dots, m \end{aligned} \tag{6}$$

In order to show the computational properties of the Flying Elephants method, Figure 1 shows the results obtained in the solution of three covering problems: Brazil (5 circles), The Netherlands (5 circles) and state of New York (7 circles). The number of discretization points were, respectively, 6620, 9220 and 7225. It is possible to find very few works presenting computational results with similar quality to those obtained by FE approach, we cite Wei (2008).

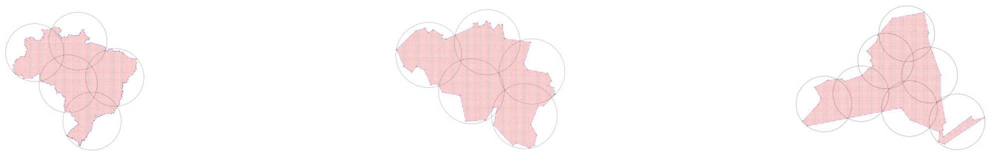

Figure 1: Coverages of Brazil, Netherlands and state of New York

5. Clustering Problems

Let $S = \{s_1, s_2, \dots, s_m\}$ denote a set of m cities or locations in an Euclidean n-dimensional space $\Re^n$, to be clustered into a given number q of disjoint clusters. To formulate the original clustering problem as a $\min - sum - \min$ problem, we proceed as follows. Let x_i, $i = 1, \cdots, q$ be the locations of facilities or centroids. The set of these centroids coordinates will be represented by $X \in \Re^{nq}$. Given a point $s_j \in S$, we initially calculate the Euclidian distance from s_j to the nearest centroid:

$$z_j = \min_{i=1,\cdots,q} \left\| s_j - x_i \right\|_2$$

The most frequent measurement of the quality of a clustering associated to a position of q centroids is provided by the minimum sum of squares (MSSC) of these distances.

$$\begin{aligned} & minimize \ \sum_{j=1}^{m} z_j^2 \\ & subject\ to: \ z_j = \min_{i=1,\ldots,q} \left\| s_j - x_i \right\|_2, \quad j = 1,\ldots,m \end{aligned} \tag{7}$$

As the partial derivative of $h_j = (x, z_j)$ with respect to $z_j, j = 1,\ldots,m$ is not equal to zero, it is possible to use the Implicit Function Theorem, see by example: Jittorntrum (1978), to calculate each component $z_j, j = 1,\ldots,m$ as a function of the centroid variables $x_i, i = 1, \ldots, q$. In this way, the unconstrained problem

$$minimize \ \ f(x) = \sum_{j=1}^{m} z_j(x)^2 , \tag{8}$$

is obtained, where each $z_j(x)$ results from the calculation of the single zero of each equation below, since each term ϕ above strictly increases together with variable z_j:

$$h_j(x, z_j) = \sum_{i=1}^{q} \phi(z_j - \theta(s_j, x_i, \gamma), \tau) - \varepsilon = 0, \quad j = 1,\ldots,m \tag{9}$$

Again, due to the Implicit Function Theorem, the functions $z_j(x)$ have all derivatives with respect to the variables $x_i, i = 1, \ldots, q$, and therefore it is possible to calculate the gradient of the objective function of problem (8),

$$\nabla f(x) = \sum_{j=1}^{m} 2\, z_j(x)\, \nabla z_j(x). \tag{10}$$

where

$$\nabla z_j(x) = -\nabla h(x, z_j) / \frac{\partial h(x, z_j)}{\partial z_j}, \tag{11}$$

while $\nabla h(x, z_j)$ and $\partial h(x, z_j) / \partial z_j$ are obtained from equations (9) and from definitions of function $\phi(y, \tau)$ and $\theta(s_j, x_i, \gamma)$.

Xavier (2010) introduces the use of the FE approach for solving the MSSC problem. The presented computational results exhibit a performance with robustness, efficiency and consistency, as well as with the capacity to solve large instances. Xavier & Xavier (2011) present a pruning scheme which speeds up the performance of the FE approach up to 500 times maintaining the same robustness. Bagirov et al (2012b) propose an algorithm which is based on the combination of FE approach, without the pruning scheme, and a incremental approach to get a good starting point. It is considered the solution of the largest clustering instances presented in the literature. A comparison with other two top algorithms demonstrates that the proposed algorithm is more accurate.

Below we present a new computational experiment in order to exhibit the performance of the FE method and, in a particular way, to show its capacity for solving very large clustering problems. We generate a synthetic data sets with m=5000000 observations in space with n=10 dimensions. The observations were generated as random perturbations of 10 known centers. The numerical experiments have been carried out on a Intel Core i7-2620M Windows Notebook with 2.70GHz and 8 GB RAM.

Table 2 presents a synthesis of the computational results. We vary the number of clusters q=2, and for each number of clusters, ten different randomly chosen starting points were used. The columns show the number of clusters (q), the best solution produced ($f_{FE_{Best}}$) by the FE approach, the number of occurrences of the best solution (Occur.), the average deviation of the 10 solutions (E_{Mean}) in relation to the best solution obtained and CPU mean time given in seconds (T_{Mean}) associated to 10 tentative solutions. The last row of each table, represented by character c of center, informs the sum of variances of the 10 synthetic groups, which is greater than the obtained value by the FE approach for the case q=10. This result demonstrates unequivocally the robustness of new method.

q	$f_{FE_{Best}}$	Occur.	E_{Mean}	T_{Mean}
2	0.456807E7	3	0.94	16.12
3	0.373567E7	1	1.21	24.69
4	0.323058E7	1	0.91	32.90
5	0.274135E7	1	0.09	26.06
6	0.248541E7	1	0.04	36.55
7	0.222897E7	1	0.19	43.24
8	0.197977E7	2	0.12	45.38
9	0.173581E7	2	0.10	42.78
10	0.149703E7	10	0.00	32.98
c	0.150000E7	-	-	-

Table 2: Clustering 5000000 Synthetic Observation with $n = 10$ Dimensions

6. The Fermat-Weber Problem

Let $S = \{s_1, s_2, \dots, s_m\}$ denote a set of m cities or locations in an Euclidean planar space $\Re^2$, with a corresponding set of demands $W = \{w_1, w_2, \dots, w_m\}$ to be attended by q, a given number of facilities. To formulate the Fermat-Weber problem as a $\min - sum - \min$, we proceed as follows. Let $x_i,\ i = 1, \cdots, q$ be the locations of facilities or centroids. The set of these centroid coordinates will be represented by $X \in \Re^{2q}$. Given a point $s_j \in S$, we initially calculate the Euclidian distance from s_j to the nearest centroid: $z_j = \min_{i=1,\cdots,q} \|s_j - x_i\|_2$. The Fermat-Weber problem considers the placing of q facilities in order to minimize the total transportation cost:

$$\begin{aligned} & \textit{minimize} \ \sum_{j=1}^{m} w_j z_j \\ & \textit{subject to}: \ z_j = \min_{i=1,\dots,q} \left\| s_j - x_i \right\|_2, \quad j = 1,\dots,m \end{aligned} \tag{12}$$

As the partial derivative of $h_j = (x, z_j)$ with respect to z_j, $j = 1,\dots,m$ is not equal to zero, it is possible to use the Implicit Function Theorem to calculate each component z_j, $j = 1,\dots,m$ as a function of the centroid variables x_i, $i = 1, \dots, q$. This way, the unconstrained problem:

$$\textit{minimize} \ \ f(x) = \sum_{j=1}^{m} w_j z_j(x) \ , \tag{13}$$

is obtained, where each $z_j(x)$ results from the calculation of the single zero of each equation below, since each term ϕ above strictly increases together with variable z_j:

$$h_j(x, z_j) = \sum_{i=1}^{q} \phi(z_j - \theta(s_j, x_i, \gamma), \tau) - \varepsilon = 0, \quad j = 1,\dots,m \tag{14}$$

Again, due to the Implicit Function Theorem, the functions $z_j(x)$ have all derivatives with respect to the variables x_i, $i = 1, \dots, q$, and therefore it is possible to calculate the gradient of the objective function of problem (13),

$$\nabla f(x) = \sum_{j=1}^{m} w_j \nabla z_j(x). \tag{15}$$

where

$$\nabla z_j(x) = -\nabla h(x, z_j) / \frac{\partial h(x, z_j)}{\partial z_j} . \tag{16}$$

This way, it is easy to solve problem (13) by making use of any method based on first order derivative information. Finally, it must be emphasized that problem (13) is defined on a ($2q$)-dimensional space, so it is a small problem, since the number of clusters, q; is, in general, very small for real applications.

Xavier (2012) introduces the use of the FE approach to solve the Fermat-Weber problem. The computational experiments with the new approach show a performance similar to the top algorithms for solving problems up to 1060 cities, the previous largest instance, see Brimberg et al (2000) and Plastino et al (2011). Xavier (2012) and Xavier et al (2012) present also results for problems never considered in the literature, with up to 85900 cities, a new superior bound size about 80 times larger.

q	$f_{FE_{Best}}$	Occur.	E_{Mean}	T_{Mean}
2	0.163625E11	6	0.27	25.33
3	0.127835E11	10	0.00	50.91
4	0.108063E11	10	0.00	74.62
5	0.984539E10	7	0.11	121.02
6	0.902515E10	10	0.00	156.63
7	0.836416E10	3	0.18	206.71
8	0.778239E10	10	0.00	260.89
9	0.737264E10	9	0.09	317.09
10	0.704126E10	1	0.19	381.33
15	0.576935E10	10	0.00	937.84
20	0.502191E10	1	0.13	1690.06
30	0.411982E10	2	0.08	4062.92
40	0.358238E10	1	0.11	8169.64

Table 3: Fermat-Weber Problem - Pla85900 Instance

We reproduce in Table 3 the results of the experiment for the new largest instance Pla85900. Numerical experiments have been carried out on a PC Intel Pentium T4300, 2.1GHz CPU with 4GB RAM, Windows 7, 32 bits.

Table 3 contains: the number of facilities q; the best objective function value produced by the FE method $f_{FE_{Best}}$ by using a set of random starting points; the number of occurrences of the best solution Occur.; the perceptual deviation value E_{Mean} of the set of 10 solutions related to the best value produced by the FE method and the mean CPU time given in seconds T_{Mean}. The low values in the column E_{Mean} show unequivocally the consistence of the FE algorithm. As there is not any recorded result for this instance the obtained values for the objective function and for the CPU time are a challenge for future research.

7. Hub location Problems

The continuous p-hub median problem is a location problem which requires finding a set of p hubs in a planar region, in order to minimize a particular transportation cost function. To formulate this problem, we proceed as follows. Let $S=\{s_1,s_2,\ldots,s_m\}$ denote a set of m cities or locations in an Euclidean planar space $\Re^2$. Let w_{jl} be the demand between two points j and l. Let x_i, $i=1,\cdots,p$ be the hubs, where each $x_i \in \Re^2$.

Concerning the hub-and-spoke problem under consideration, the connections between each pair of points j and l; as depicted by Figure 1, have always three parts: from the origin point j to a first hub a, from a to a second hub b and from b to destination point l. Multiple allocations is permitted, meaning that any given point can be served by one or more hubs. The first and the second hubs can be coincident (i.e., a=b), meaning that a unique hub is used to connect the origin point j and the destination point l. Figure (2) shows the p^2 possible connections between two cities.

The p-hub median problem corresponds to minimizing the total cost between all pairs of cities taking the unitary cost value for all connections:

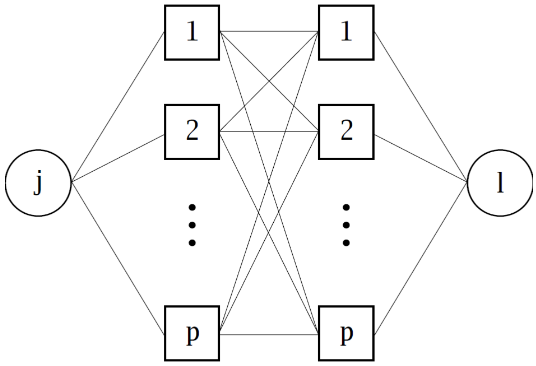

Figure 2: The set of connections between point j and point l

$$\begin{aligned} & minimize \sum_{j=1}^{m} \sum_{l=1}^{m} w_{jl} z_{jl} \\ & subject\ to: \ z_{jl} = \min_{a,b=1,\dots,p} z_{jabl}, \quad j,l = 1,\dots,m \end{aligned} \tag{17}$$

where $z_{jabl} = \left\| s_j - x_a \right\|_2 + \alpha \left\| x_a - x_b \right\|_2 + \left\| x_b - s_l \right\|_2$ and α is the reduction factor: $0 \le \alpha \le 1$.

By using FE approach, it is possible to use once more the Implicit Function Theorem to calculate each component $z_{jl}, \ j,l = 1,\dots,m$ as a function of the centroid variables $x_i, \ i = 1,\cdots,p$. So, the unconstrained problem

$$minimize \ \ f(x) = \sum_{j=1}^{m} \sum_{l=1}^{m} w_{jl} z_{jl}(x) \ , \tag{18}$$

is obtained, where each $z_{jl}(x)$ results from the calculation of the single zero of each equation:

$$\begin{aligned} h_{jl}(x, z_{jl}) = \sum_{a=1}^{p} \sum_{b=1}^{p} [\, & \phi(z_{jl} - (\theta(s_j, x_a, \gamma), \tau) + \\ & \alpha\, \theta(s_j, x_a, \gamma), \tau) + \theta(s_j, x_a, \gamma), \tau)) - \varepsilon = 0, \quad j,l = 1,\dots,m \end{aligned} \tag{19}$$

Again, due to the Implicit Function Theorem, the functions $z_{jl}(x)$ have all derivatives with respect to the variables $x_i, \ i = 1,\cdots,p$ and therefore it is possible to calculate the gradient of the objective function (18):

$$\nabla f(x) = \sum_{j=1}^{m} \sum_{l=1}^{m} w_{jl} \nabla z_{jl}(x). \tag{20}$$

where

$$\nabla z_{jl}(x) = -\nabla h(x, z_{jl}) / \frac{\partial h(x, z_{jl})}{\partial z_{jl}}, \tag{21}$$

while $\nabla h(x, z_{jl})$ and $\partial h(x, z_{jl}) / \partial z_{jl}$ are directly obtained from equations (1), (2) and (19).

This way, it is easy to solve problem (18) by making use of any method based on first order derivative information. Finally, it must be emphasized that problem (18) is defined on a *(2p)*-dimensional space, so it is a small problem, since the number of hubs, p, is small, in general, for real world applications.

Gesteira (2012) and Xavier et al (2012) introduce the use of the FE approach to solve the continuous hub location problem. It is presented computational results for instances up to 1000 cities or about 500000 different origin-destination pairs. The number of hubs reaches the value p=5, which implies 12.5 million different path connections *jabl,* see Figure 2. Contreras et al (2011) consider a discrete hub location problem and solve problems up to 500 cities. To the best knowledge of these authors: *the new instances are by far the largest and most difficult ever solved for any type of hub location problem.*

Below in Table 4 we reproduce the computational results obtained for solving the largest instance. The numerical experiments have been carried out on a PC Intel Celeron with a 2.7GHz CPU and 512MB RAM. The first column presents the specified number of hubs (p). The second column presents the best objective function value $f_{FE_{Best}}$ produced by the FE method in 10 tentatives. The next three columns present the number of occurrences of the best solution (Occur.), the percentage average deviation of the T solutions E_{Mean} in relation to the best solution obtained $f_{FE_{Best}}$ and the CPU mean time given in seconds $Time_{\mathrm{Mean}}$. The low values in the column E_{Mean} show unequivocally the consistence of the FE algorithm. As there is not any recorded result for this instance the obtained values for objective function and CPU time are a challenge for future research.

q	$f_{FE_{Best}}$	Occur.	E_{Mean}	$Time_{\mathrm{Mean}}$
2	0.342083E12	10	0.00	376.66
3	0.285747E12	10	0.00	1296.32
4	0.263992E12	9	0.07	3754.33
5	0.248652E12	4	0.35	8234.88

Table 4: Hub Location Problem - dsj1000 TSPLIB instance ($\alpha = 0.5$)

8. Conclusions

This article presents a general review of successful applications of the FE approach for solving a select set of five important problems. For each problem, the performance of the FE approach can be attributed to the complete differentiable formulation, therefore powerful optimization methods can be applied for solve them. Furthermore, the FE has the convexification feature that eliminate a lot of local minima points.

In short, computational experiments for all related problems obtained results, which exhibited a high level of performance of the FE approach according to the different criteria of consistency, robustness and efficiency. The robustness, consistency and efficiency performances can be attributed to the complete differentiability of the approach. Based on the success of these previous experiences, we believe that the FE methodology can also be used for solving a broad class of non-smooth problems with similar characteristics, like those contemplated in the seminal survey written by Rubinov (2006).

References

An, Hoai L. and Pham D. Tao (2000), Large-Scale Molecular Conformation Via the Exact Distance Geometry Problem. Lecture Notes in Economics and Mathematical Systems, 481: 260-277.

Bagirov, A.M., Al-Nuaimat A. and Sultanova, N. (2012), Hyperbolic smoothing function method for minimax problems, Optimization, No. 11, 1-24

Bagirov, A.M., Ordin, B., Ozturk, G. and Xavier, A.E. (2012), An incremental clustering algorithm based on smoothing techniques. submitted to Computational Optimization and Applications.

Brimberg, J.; Hansen, P.; Mladenovic, N. and Taillard, E. D. (2000) Improvements and Comparison of Heuristics for Solving the Multisource Weber Problem. Operations Research, 48: 129-135.

Chaves, A.M.V. (1997) Resolução do problema minimax via suavização. M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.

Contreras, I.; Jean-Franois Cordeau, and Gilbert Laporte (2011), Benders Decomposition for Large- Scale Uncapacitated Hub Location. Operations Research, 59: 1477-1490.

Gesteira, Claudio (2012), Resolvendo Problemas de Localização de Hubs com Alocação Múltipla numa Modelagem Contínua Tipo P-Medianas Usando a Abordagem de Suavização Hiperbólica, M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro, 2012.

Jittorntrum, K. (1978) An Implicit Function Theorem. *Journal of Optimization Theory and Applications*, 25(4): 575-577.

Macambira, A.F.U. (2003) Determinação de estruturas de proteínas via suavização e penalização hiperbólica. M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.

Moré, J.J. and Wu., Z. (1995) ε-optimal solutions to distance geometry problems via global continuation. Mathematics and Computer Science Division, Argonne Lab., Preprint MCS-P520-0595.

Plastino, A.; Fuchshuber, R.; Martins, S.L.; Freitas, A.A. and Salhi, S. (2011), A Hybrid Data Mining Metaheuristic for the p-Median Problem, Statistical Analysis and Data Mining, Vol. 4, pp. 313-335.

Rubinov, A. (2006), Methods for Global Optimization of Nonsmooth Functions with Applications. Applied and Computational Mathematics, 5: 3-15.

Santos, A.B.A. (1997) Problemas de programação não-diferenciável: Uma metodologia de suavização. M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.

Souza, M.F. (2010) Suavização hiperbólica aplicada à otimização de geometria molecular. D.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.

Souza, M.F., Xavier, A.E., Lavor, C. and Maculan, N. (2011) Hyperbolic smoothing and penalty techniques applied to molecular structure determination. Operations Research Letters, 39: 461--465.

Wei, Hu (2008) Solving continuous space location problems. Ph.D. Thesis - The Ohio State University.

Xavier, A.E. (1982) Penalização hiperbólica: Um novo método para resolução de problemas de otimização. M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.

Xavier, A.E. (2003) Convexificação do problema de distância geométrica através da técnica de suavização hiperbólica. Workshop em Biociências COPPE UFRJ.

Xavier, A.E. (2010) The hyperbolic smoothing clustering method. Pattern Recognition, 43: 731--737.

Xavier, A.E., Gesteira, C.M. and Xavier, V.L. (2012) The continuous multiple allocation p-hub median problem solving by the hyperbolic smoothing approach: Computational performance, accepted to Optimization.

Xavier, A.E. and Oliveira, AA.F. (2005) Optimal covering of plane domains by circles via hyperbolic smoothing. Journal of Global Optimization, 31: 493-504.

Xavier, A.E. (2011) and Vinicius Layter Xavier. Solving the minimum sum-of-squares clustering problem by hyperbolic smoothing and partition into boundary and gravitational regions. Pattern Recognition, 44: 70-77.

Xavier, Vinicius Layter (2012), Resolução do Problema de Agrupamento segundo o Critério de Minimização da Soma de Distâncias, M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.

Xavier, V.L., França, F.M.G., Xavier, A.E. and Lima, P.M.V. (2012) A hyperbolic smoothing approach to the multisource weber problem, submitted to Journal of Global Optimization.