Flying Elephants: Método para Resolver Problemas Não-Diferenciáveis

Adilson Elias Xavier

Vinicius Layter Xavier

Federal University of Rio de Janeiro - BRAZIL

Seminários PESC

Rio de Janeiro, 27 Novembro 2015.

Outline of Presentation

1 - Introduction

2 - Fundamental Smoothing Procedures

3 - Geometry Distance Problem

4 - Covering Problems

5 - Clustering Problems

6 - Fermat-Weber Problem

7 - Hub location Problems

8 - Conclusions

Many do not believe, but the elephants really fly!!

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).

Introduction

The new name of the methodology, Flying Elephants, is definitively 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.

Introduction

The Flying feature is directly derived from the complete differentiability property of the method, which has the necessary power to make the flight of the heavy elephant feasible.

Moreover, it permits intergalactic trips into spaces with large number of dimensions, differently from the short local searches associated to traditional heuristic algorithms.

On the other side, the convexification feature also associated to the FE method is analogous to the local action of the Elephant landing, eliminating a lot of local minima points.

Fundamental

Smoothing

Procedures

Fundamental Smoothing Procedures

There are only two
Fundamental Smoothing Procedures

The main principle is always to perform transformations on the original formulation to make possible to use these two fundamental procedures in order to obtain a succedaneous problem completely differentiable .

This is the idea!!

Smoothing of the absolute value function

To smooth the absolute value function $|u|$
we use the function:

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

Fundamental Smoothing Procedures

Function θ has the following properties:

(a) $\lim\limits_{\gamma \to 0} \theta(u,\gamma) = |u|$

(b) θ is a C^{∞} function.

(c) $\theta'(u,\gamma) = u/(u^2+\gamma^2)^{1/2}$

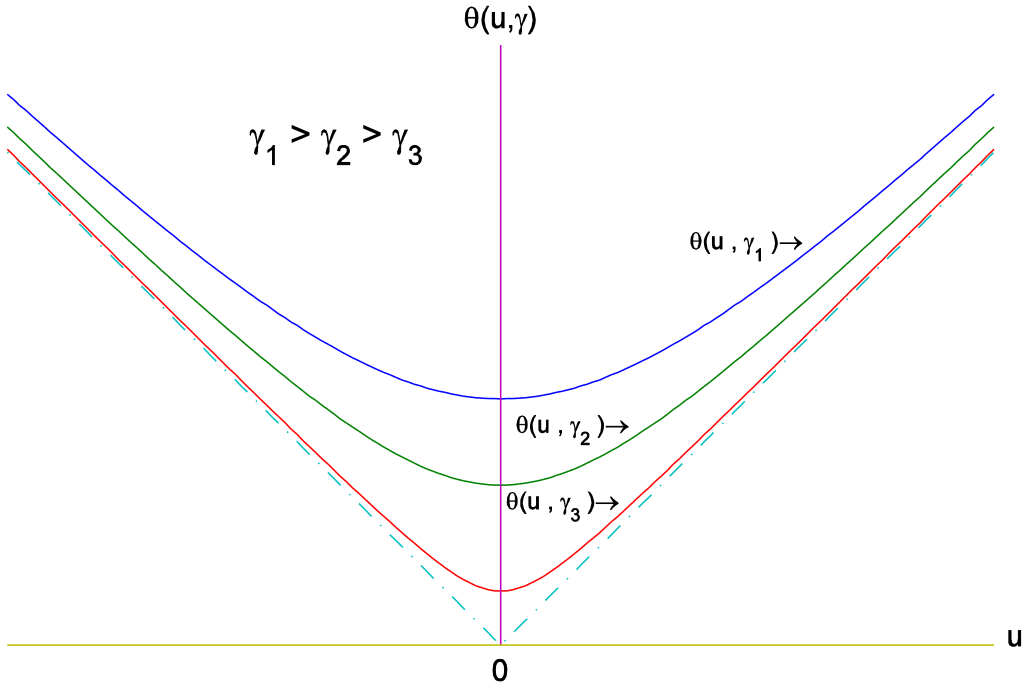
(d) $\theta''(u,\gamma) = \gamma^2/(u^2+\gamma^2)^{3/2}$

(e) $\theta''(0,\gamma) = 1/\gamma$

(f) $\lim\limits_{\gamma \to 0} \theta''(0,\gamma) \to \infty$

Fundamental Smoothing Procedures

Smoothing of the absolute value function



Fundamental Smoothing Procedures

Now, we will present the smoothing procedure of the function $\psi(y,\lambda) = \lambda \max(0, y)$. For this purpose, let us define the function

$$\phi(y,\lambda,\tau) = (\lambda y + \sqrt{\lambda^2 y^2 + \tau^2})/2$$

for $y \in \Re$ and $\tau > 0$.

Fundamental Smoothing Procedures

Function ϕ has the following properties:

(a) $\phi(y,\lambda,\tau) > \psi(y,\lambda), \quad \forall \tau > 0$

(b) $\lim_{\tau \to 0} \phi(y,\lambda,\tau) = \psi(y,\lambda)$

(c) $\phi(y,\lambda,\tau)$ is na increasing convex C^{∞} function in y

(d) $\phi'(y,\lambda,\tau) = \lambda + \lambda^2 y / (\lambda^2 y^2 + \tau^2)^{1/2}$

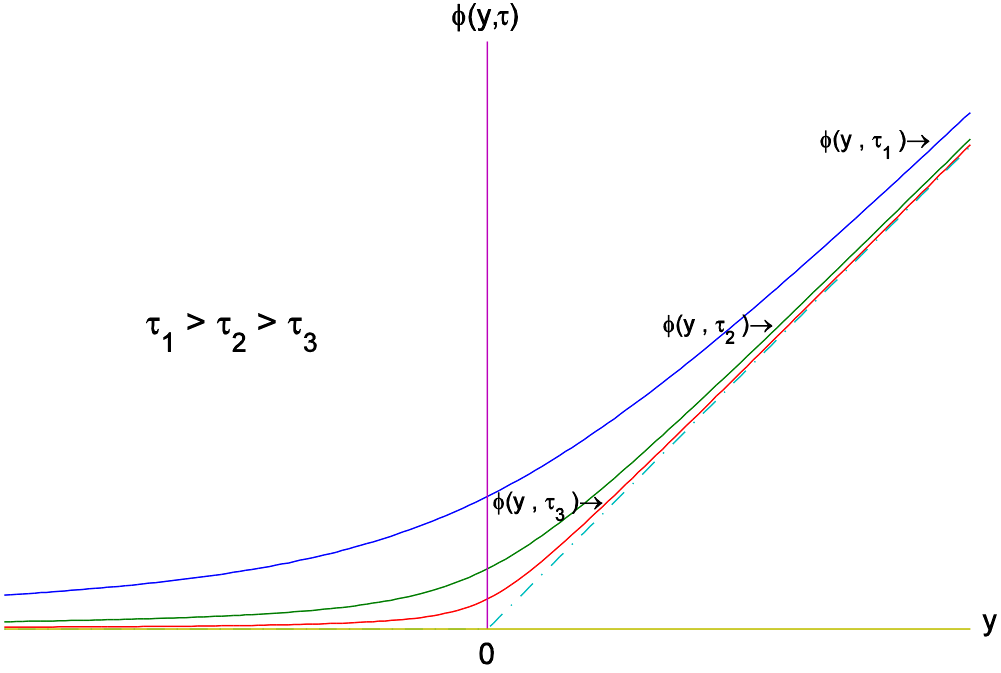
(e) $\phi''(y,\lambda,\tau) = \lambda^2 \tau^2 / (\lambda^2 y^2 + \tau^2)^{3/2}$

(f) $\phi''(0,\lambda,\tau) = \lambda^2 / \tau$

(g) $\lim_{\tau \to 0} \phi''(0,\lambda,\tau) \to \infty$

Fundamental Smoothing Procedures

Smoothing of the function ψ



Geometry

Distance

Problem

Publications

Michael de Souza, Adilson Elias Xavier, Carlile Lavor and Nelson Maculan, "*Hyperbolic Smoothing and Penalty Tecniques Applied to Molecular Structure Determination*", **Operations Research Letters, Vol. 39, pp. 461-465, 2011,** doi:10.1016/j.orl.2011.07.007

Adilson Elias Xavier and Helder Manoel Venceslau, "Solving the Geometric Distance Problem by the Hyperbolic Smoothing Approach", DGA 2013 - Workshop on Distance Geometry and Applications, Manaus, June 24-27, 2013.

Geometry Distance Problem

Let $G = (V, E)$ denote a graph, in which for each $\text{arc}(i, j) \in E$, 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}

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

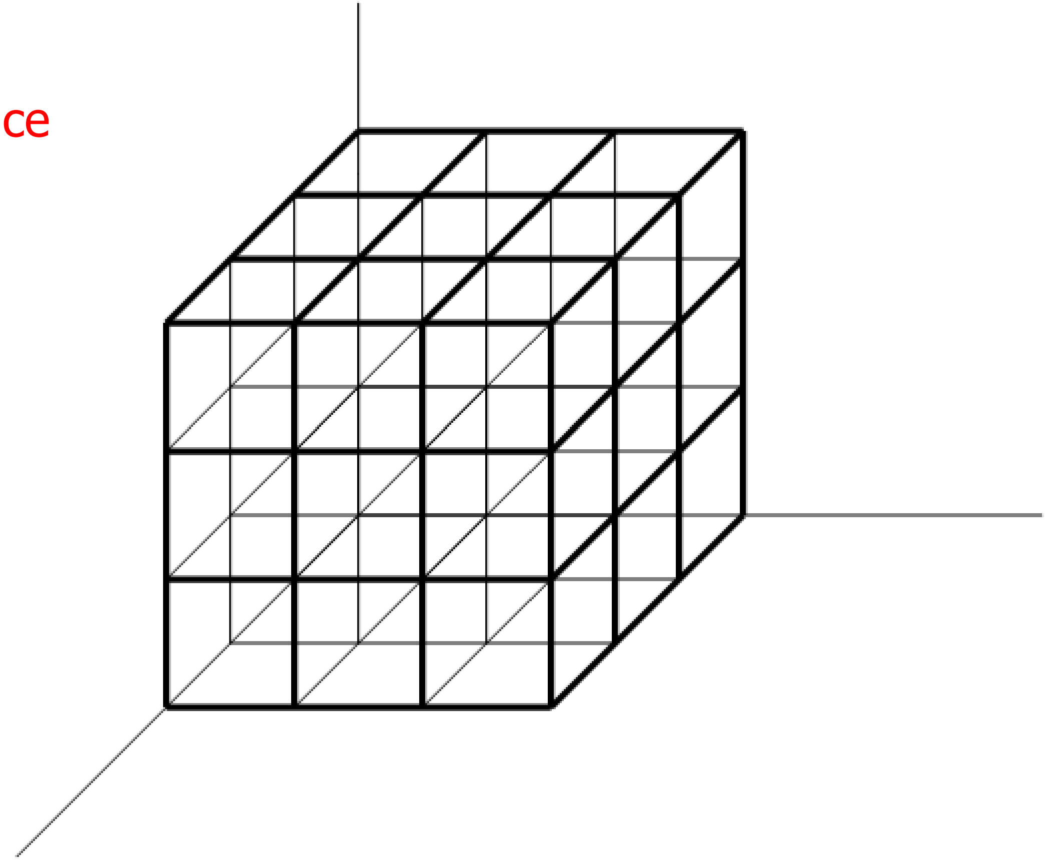
Geometry Distance Problem

For solving the previous problem by using the Flying Elephant technique it is only necessary to use the function and $\theta(u,\gamma)$ to define $u = \|x_i - x_j\|$:

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

Geometry Distance Problem

Moré-Wu Instance
s=4



Results of FE Technique Applied to Moré-Wu Instance

s	$m = s^3$	$n = 3s^3$	p	$Occur.$	f_{Med}	$Time$	$Occ.n.s$
3	27	81	198	0	-	0.1	8
4	64	192	888	6	0.27E-6	0.7	5
5	125	375	2800	8	0.29E-5	2.8	7
6	216	648	7110	8	0.19E-4	7.6	4
7	343	1029	15582	5	0.16E-4	19	5
8	512	1536	30688	8	0.29E-3	45	3
9	729	2187	55728	6	0.86E-3	97	0
10	1000	3000	94950	7	0.95E-3	45	1
11	1331	3993	153670	6	0.17E-2	81	0

Results of FE Technique applied to Moré-Wu Instance

s	$m = s^3$	$n = 3s^3$	p	$Occur.$	f_{Med}	$Time$	$Occ.n.s$
12	1728	5184	238392	8	0.15E-1	143	0
13	2197	6591	356928	7	0.32E-1	222	-
14	2744	8232	518518	8	0.18E-1	380	-
15	3375	10125	733950	6	0.65E-1	543	-
16	4096	12288	1015680	7	0.42E-1	835	-
17	4913	14739	1377952	6	0.16E0	1270	-
18	5832	17496	1836918	7	0.21E0	1853	-
19	6859	20577	2410758	8	0.24E0	2335	-
20	8000	24000	3119800	8	0.59E0	3187	-

Covering

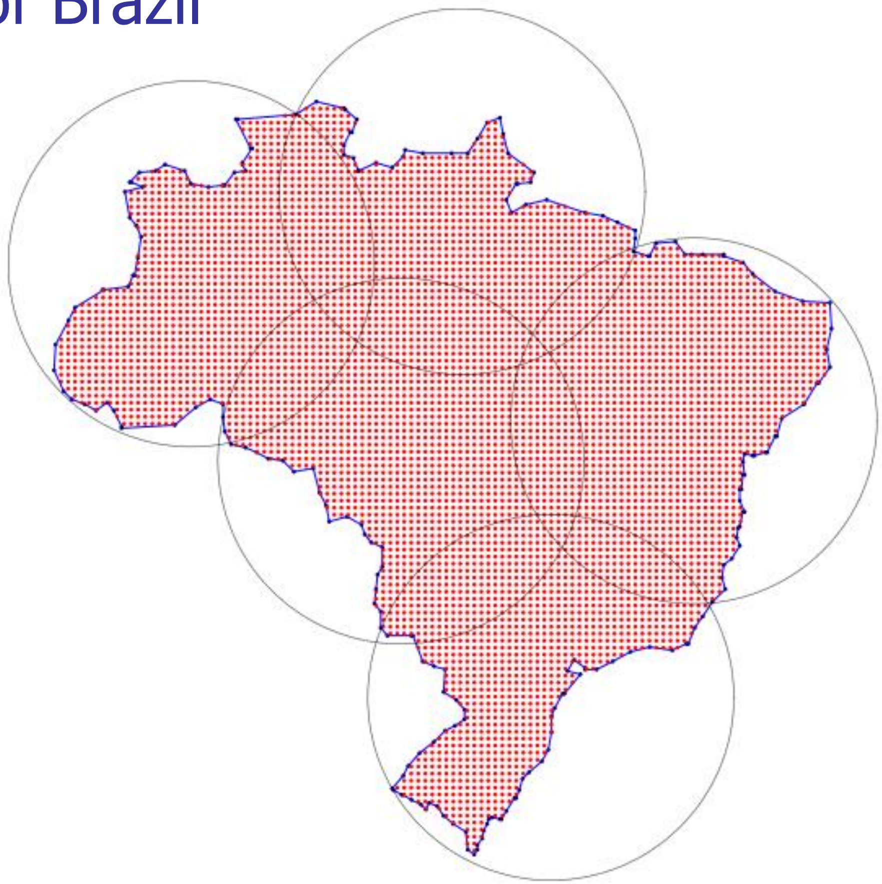
Problems

Publications

A. E. Xavier e A. A. F. Oliveira (2005), “Optimum Covering of Plane Domains by Circles Via Hyperbolic Smoothing Method”, **Journal of Global Optimization, Volume 31 Number 3, March 2005, Pages 493-504 , Springer**, http://dx.doi.org/10.1007/s10898-004-0737-8

Covering Problem Conceptualization

Coverages of Brazil

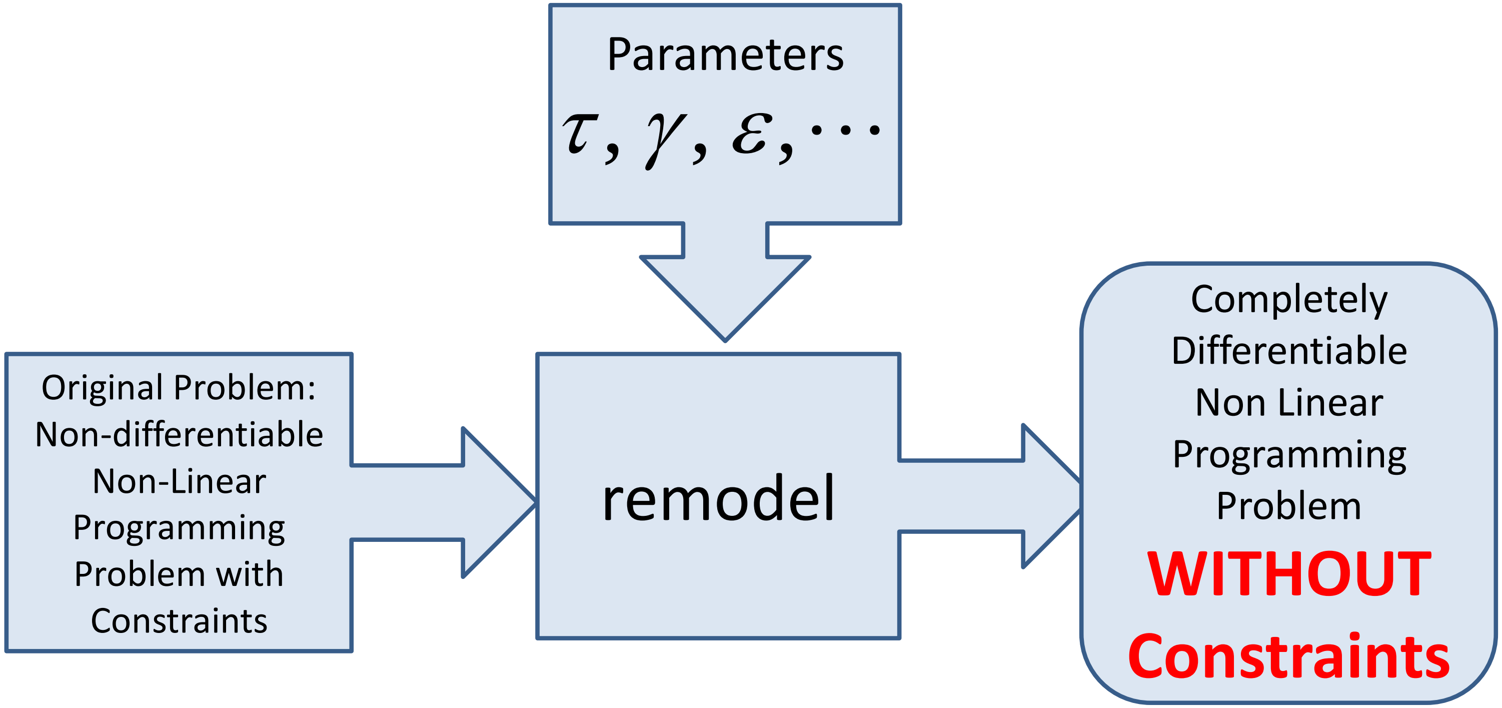


Covering Problems

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 centres of the circles that must cover this set of points

$$X^* = \arg\min_{X \in \Re^{2q}} \max_{\mathrm{j}=1,\cdots,\mathrm{m}} \min_{i=1,\cdots,q} \left\| \mathrm{s}_\mathrm{j} - \mathrm{x}_\mathrm{i} \right\|_2$$

Flying Elephants Transformations



Covering Problems

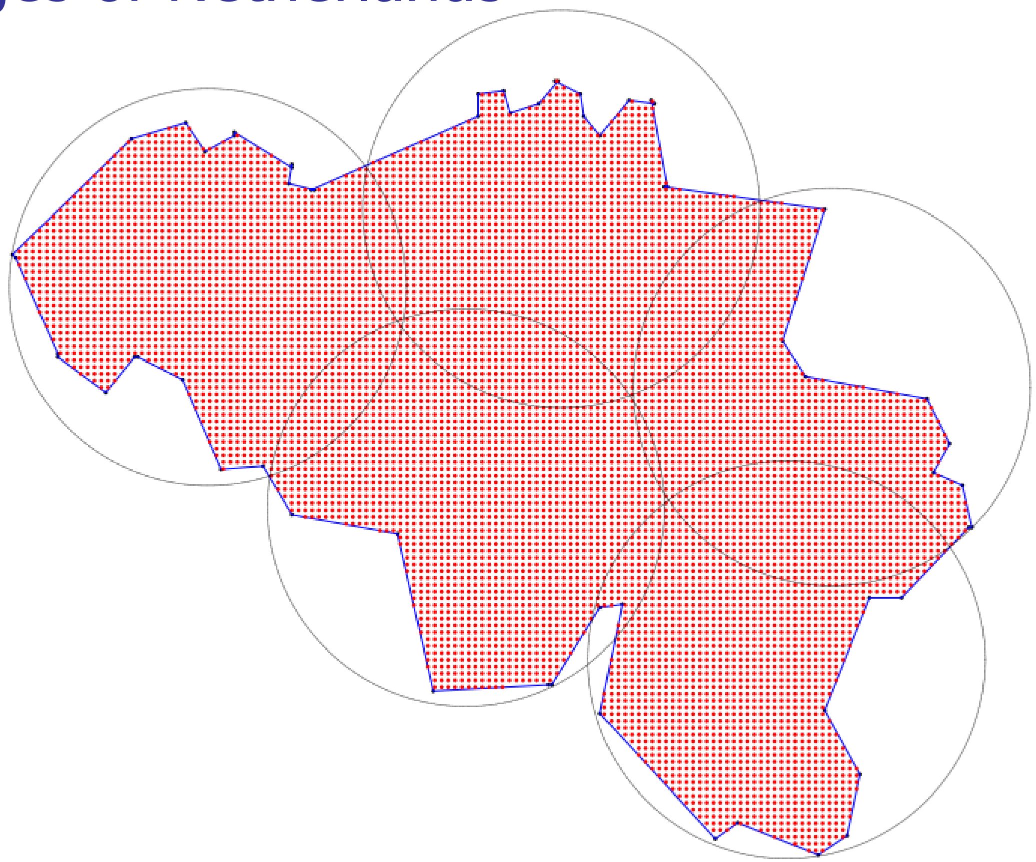
By performing an ε perturbation and by using the FE approach, the three-level strongly nondifferentiable $\min - \max - \min$ problem can be transformed in a one-level completely smooth one:

$$\text{minimize} \quad z$$

$$\text{subject to:} \quad \sum_{i=1}^{q} \phi\left(z - \left\| s_j - x_i \right\|_2, \tau\right) \geq \varepsilon, \quad j = 1, \cdots, m$$

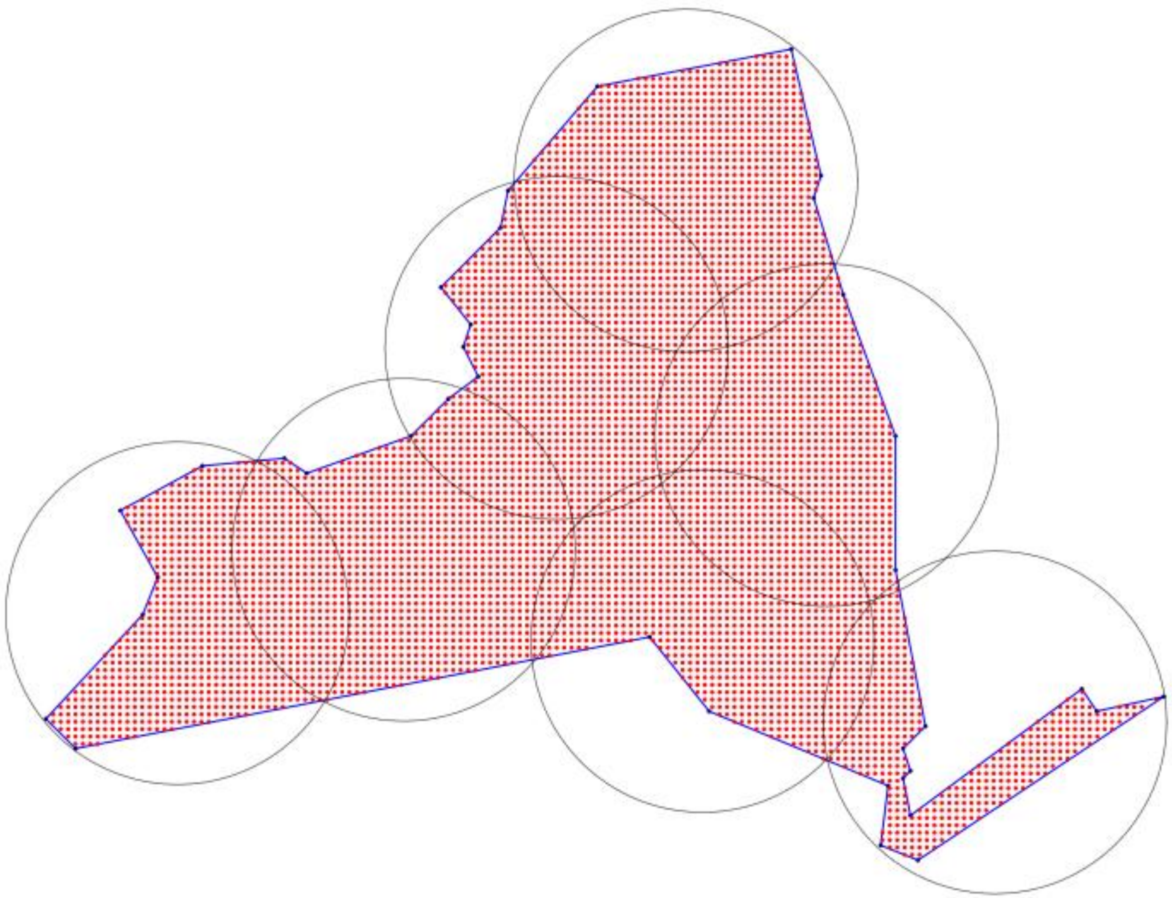
Covering Problems

Coverages of Netherlands



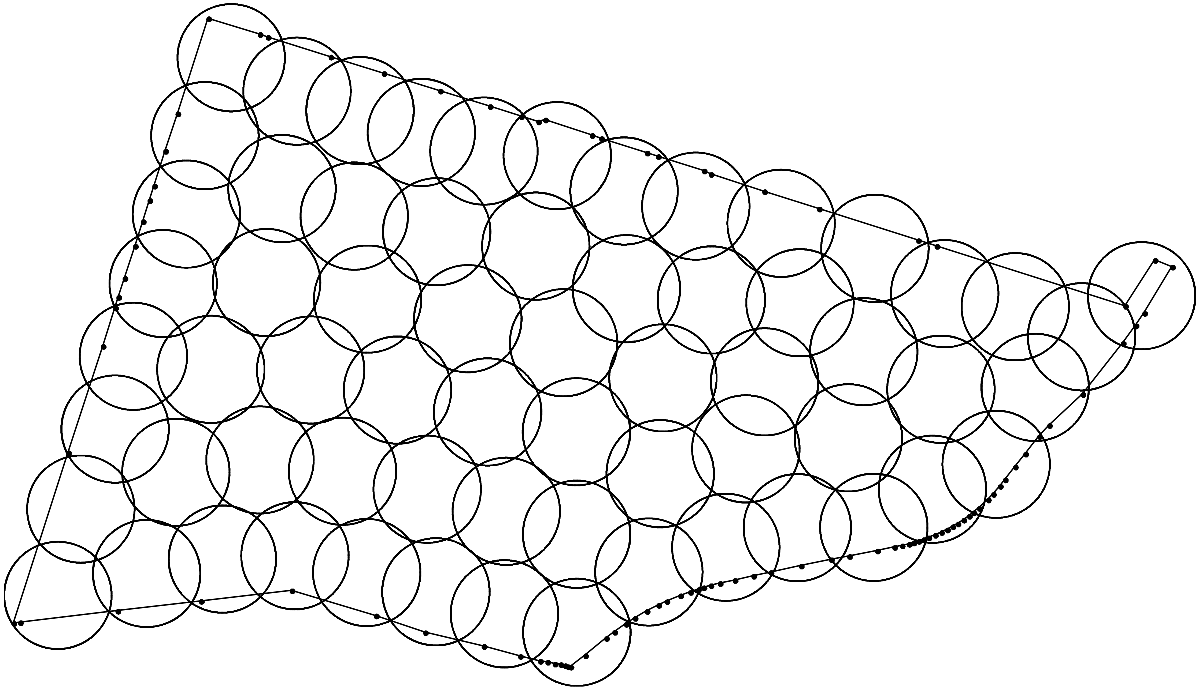
Covering Problems

Coverages of the state of New York



Covering Problems

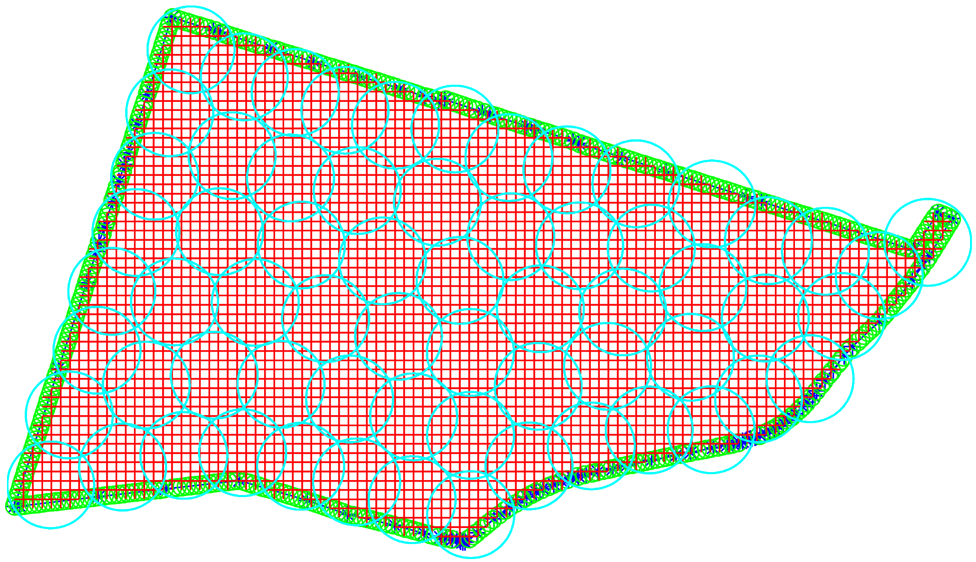
Coverages of Dionisio Torres District – Fortaleza - Brazil



64 circles

Covering Problems

Coverages of Dionisio Torres District – Fortaleza - Brazil



64 circles

Clustering

Problems

Publications – Part 1

Adilson Elias Xavier, “The Hyperbolic Smoothing Clustering Method”, **Pattern Recognition, Vol. 43, pp. 731-737, 2010** doi:10.1016/j.patcog.2009.06.018

Adilson Elias Xavier and Vinicius Layter Xavier, "Solving the Minimun Sum-of-Squares Clustering Problem by Hyperbolic Smoothing and Partition into Boundary and Gravitational Region ", **Pattern Recognition, Vol. 44, pp. 70-77, 2011.** doi:10.1016/j.patcog.2010.07.004

Adilson Elias Xavier and Vinicius Layter Xavier “Solving Very Large Problems by the Accelerated Hyperbolic Smoothing Clustering Method”, EURO-2013, Rome, June 30 - July 4, 2013.

Publications – Part 2

Adilson Elias Xavier and Vinicius Layter Xavie& Sergio Barbosa Villas-Boas "*Solving the Minimum of L1 Distances Clustering Problem by the Hyperbolic Smoothing Approach and Partition into Boundary and Gravitational Regions*", **Studies in Classification, Data Analysis and Knowledge Organization, Editors-in-chief: Bock, H.-H., Gaul, W.A., Vichi, M., Weihs, C.** Series Editors: Baier, D., Critchley, F., Decker, R., Diday, E., Greenacre, M., Lauro, C.N., Meulman, J., Monari, P., Nishisato, S., Ohsumi, N., Opitz, O., Ritter, G., Schader, M., Springer-Verlag GmbH, Heidelberg, 2013, http://www.springer.com/series/1564.

Publications – Part 3

A.M. Bagirov, B. Ordin, G. Ozturk and A.E. Xavier, *"An incremental clustering algorithm based on hyperbolic smoothing"*, **submitted to Computational Optimization and Applications.**

Adilson E. Xavier , Luiz A. A. Oliveira , Jose F. M. Pessanha , Vinicius L. Xavier, "Application of Accelerated Hyperbolic Smoothing Clustering Method to Determining Load Profiles", **submitted to EJOR, European Journal of Operational Research, 2013.**

Clustering Problems

Let S denote a set of m patterns or observations from an Euclidean n-space, to be clustered into a given number q of disjoint clusters. Let $x_i, i = 1, \cdots, q$ be the centroids of the clusters, where each $x_i \in \Re^n$. Given a point s_j of S, we initially calculate the Euclidean distance from s_j to the nearest center. This is given by $z_j = \min_{i=1,\cdots,q} \left\| \mathrm{s_j - x_i} \right\|_2$. The most frequent measurement of the quality of a clustering associated to a specific position of q centroids is provided by the minimum sum of the squares (MSSC) of these distances:

Clustering Problems

The minimum sum of the squares (MSSC) of these distances:

$$\text{minimize} \quad \sum_{j=1}^{m} z_j^2$$

$$\text{subject to:} \quad z_j = \min_{i=1,\cdots,q} \left\| s_j - x_i \right\|_2, \quad j = 1, \cdots, m$$

Clustering Problems

By using FE approach, it is possible to use the Implicit Function Theorem to calculate each component $z_j,\ j = 1, \cdots, m$ as a function of the centroid variables $x_i, i = 1, \cdots, q$. In this way, the unconstrained problem

$$\text{minimize} \qquad f(x) = \sum_{j=1}^{m} z_j(x)^2$$

where each $z_j(x)$ is obtained by the calculation of a zero of

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

Clustering 5000000 Synthetic Observations with n = 10 Dimensions

q	$f_{AHSC\text{-}L2}$	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	-	-	-

Fermat-Weber Problem

(Multisource Weber Problem)

(continuous p-center Problem)

Publications

Xavier, V.L.: Resolução do Problema de Agrupamento segundo o Critério de Minimização da Soma de Distâncias, **M.Sc. thesis - COPPE - UFRJ, Rio de Janeiro, 2012**

Xavier, V.L., França, F.M.G., Xavier, A.E. and Lima, P.M.V., *"A Hyperbolic Smoothing Approach to the Multisource Weber Problem"*, **accepted for publication on the Special Issue of Journal of Global Optimization dedicated to EURO XXV 2012, Vilnius, Lithuania.**

Fermat-Weber Problem

Let $S = \{s_1, \cdots, s_m\}$ denote a set of m cities or locations in an Euclidean planar space $\Re^2$, with a corresponding set of demands $W = \{w_1, \cdots, w_m\}$ to be attended by q a given number of facilities. To formulate the Fermat-Weber 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, $x_i \in \Re^2$. Given a point $s_j \in S$, we initially calculate the Euclidean distance from s_j to the nearest centroid: $z_j = \min_{i=1,\cdots,q} \left\| \mathrm{s_j - x_i} \right\|_2$.

Fermat-Weber Problem

The Fermat-Weber problem considers the placing of q facilities in order to minimize the total transportation cost:

$$\text{minimize} \quad \sum_{j=1}^{m} w_j z_j$$

$$\text{subject to:} \quad z_j = \min_{i=1,\cdots,q} \left\| s_j - x_i \right\|_2, \quad j = 1, \cdots, m$$

Fermat-Weber Problem

By using FE approach, it is possible to use the Implicit Function Theorem to calculate each component $z_j, j=1,\cdots,m$ as a function of the centroid variables $x_i, i=1,\cdots,q$. In this way, the unconstrained problem is obtained

$$\text{minimize} \qquad f(x)=\sum_{j=1}^{m} w_j z_j(x)$$

Fermat-Weber Problem

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 :

$$h_j(z_j, x) = \sum_{j=1}^{m} \phi(z_j - \theta(s_j, x_i, \gamma), \tau) - \varepsilon = 0, \qquad j = 1, \cdots, m$$

Fermat-Weber Problem – Pla85900

q	$f_{FE_{Best}}$	$occur.$	E_{Mean}	$Time_{Mean}$
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

Continuous Hub Location Problem

(multiple allocation p-Hub median problem)

Publications

Claudio Joaquim Martagão Gesteira Neto - "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.

Adilson Elias Xavier, Claudio Martagão Gesteira e Vinicius Layter Xavier, *"Solving the Continuous Multiple Allocation p-Hub Median Problem by the Hyperbolic Smoothing Approach"*, **accepted for publication in Optimization: A Journal of Mathematical Programming and Operations Research Optimization, Taylor & Francis, Special Issue of Optimization dedicated to EURO XXV 2012, Vilnius, Lithuania**

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.

The assumption is that each pair of cities is directly connected by the shortest distance route between them.

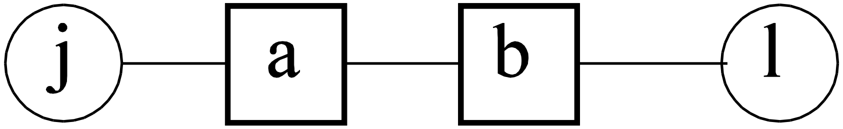
Hub location Problems

The connections between each pair of cities j and l, have always three parts:

1 - from the origin city j to a first hub a;

2 - from hub a to a second hub b;

3 - from hub b to destination city l.

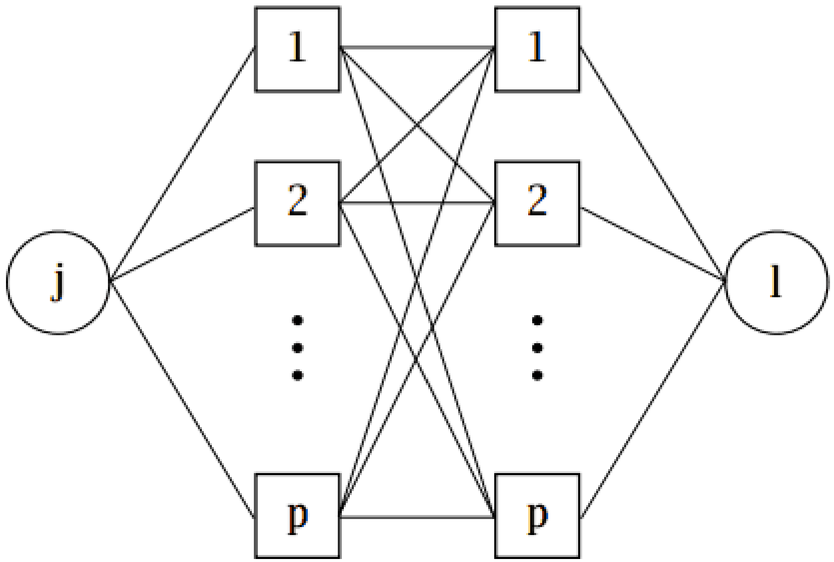


Hub location Problems

Let $S = \{s_1, \cdots, s_m\}$ denote a set of m cities or consumer points in a planar region. Let w_{jl} be the demand between two points j and l . Let $x_i, i = 1, \cdots, p$ be the hubs, where each $x_i \in \Re^2$.

Hub location Problems

The set of possible connections between city j and city l.



Multiple allocation is permitted!

Hub location Problems

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

$$\text{minimize} \quad \sum_{j=1}^{m} \sum_{l=1}^{m} w_{jl} z_{jl}$$

$$\text{subject to:} \quad \mathrm{z}_{\mathrm{jl}} = \min_{\mathrm{a,b=1,\cdots,p}} z_{jabl}, \quad j = 1, \cdots, m$$

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

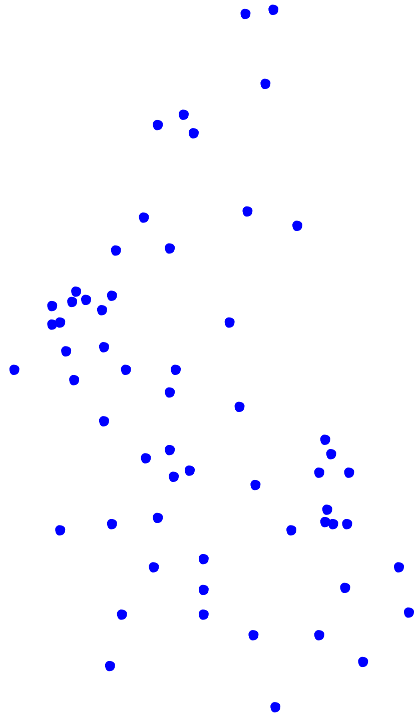
Hub location Problems

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

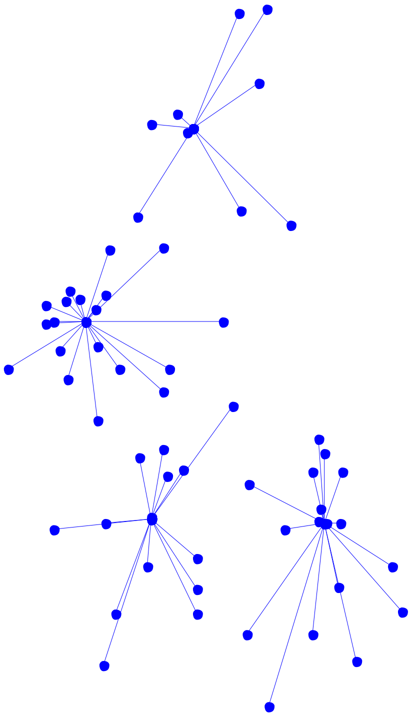
$$\text{minimize} \qquad f(x) = \sum_{j=1}^{m}\sum_{l=1}^{m} w_{jl} z_{jl}(x)$$

Hub location Problems

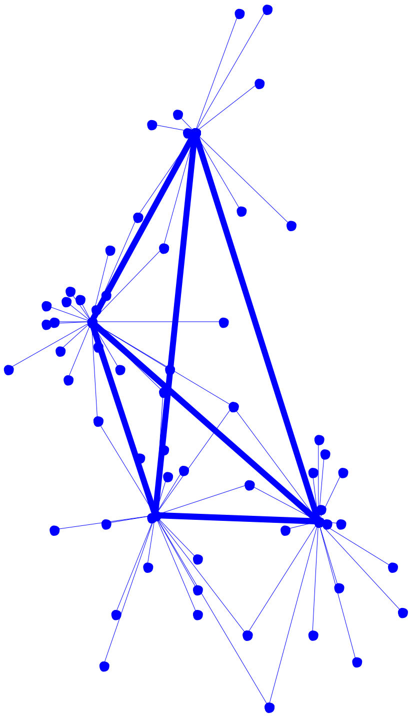
German Towns: coordinates of 59 towns (Späth, 1980)



Hub location Problems
Alpha=0. => Fermat Weber problem

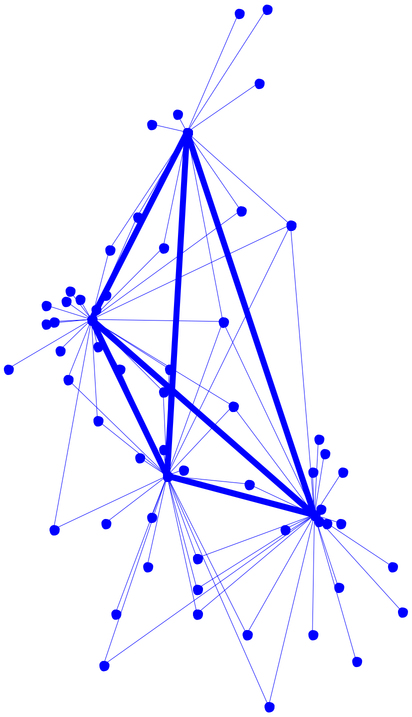


Hub location Problems
Alpha=0.25



Hub location Problems

Alpha=0.5



Hub location Problems

Hub Location Problem - dsj1000 TSPLIB instance $\alpha = 0.5$

p	$f_{\mathrm{FE}_{\mathrm{Best}}}$	$occur.$	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

Conclusions

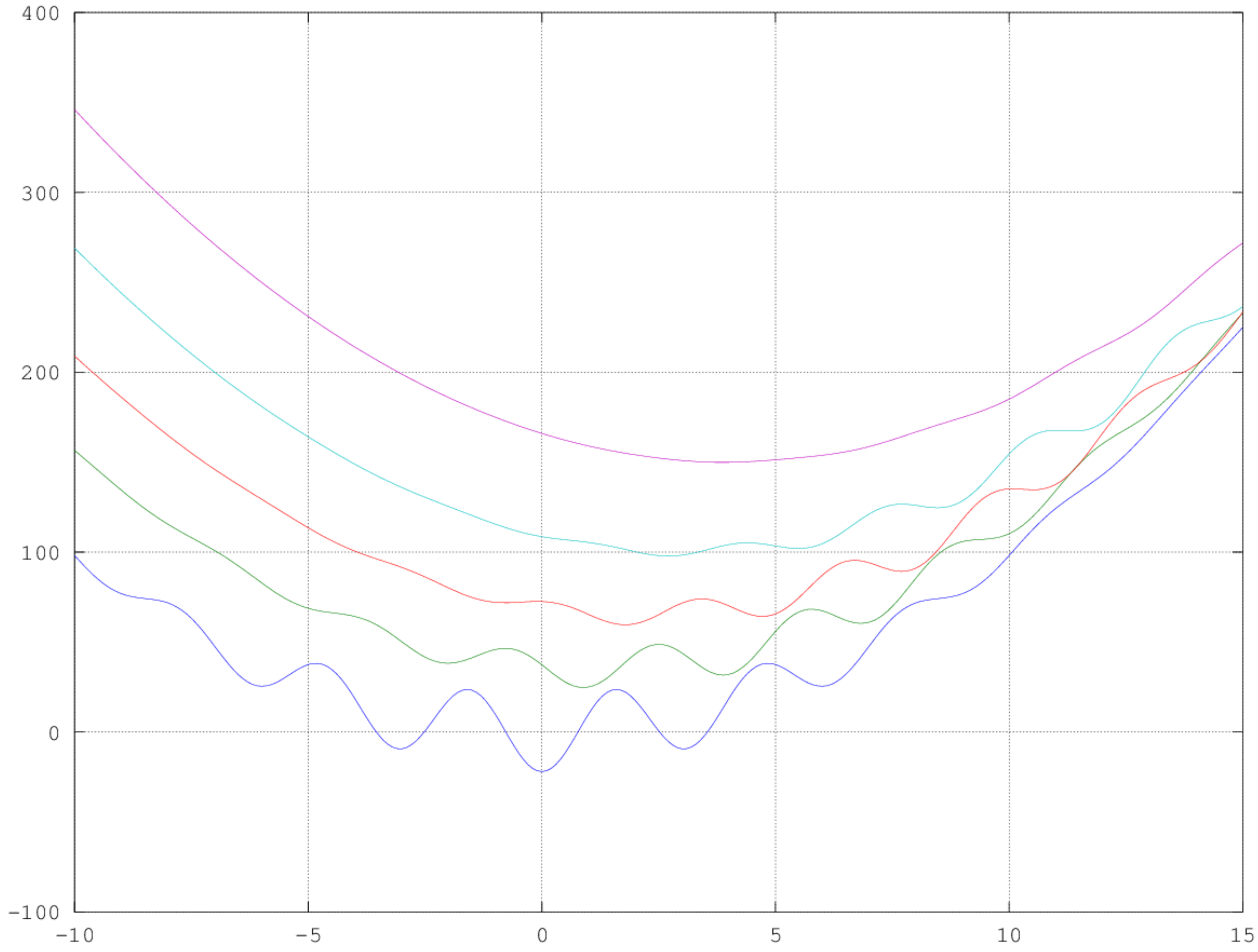
Performance of the **Flying Elephants** Method

The performance of Flying Elephants Method can be attributed to the complete differentiability of the approach.

So, the succedaneum formulation can be comfortably solved by using the most powerful and efficient algorithms, such as conjugate gradient or quasi-Newton algorithms.

Computational experiments for all 5 related problems obtained unprecedented results, which exhibits a high level performance according to the different criteria of consistency, robustness and efficiency.

Additional Effect Produced by the Smoothing Procedures: Elimination of Local Minimum Points



Max-Cut Problem

(repto lançado pelo
Prof. Manoel Campelo –UFC - Fortaleza)

Max-Cut Problem

The max-cut problem specification:

Temos um grafo G=(V,E), onde V é o conjunto de n nós
E é o conjunto de arcos

O problema é particionar o conjunto de nós V em duas partes V1 e V2, de maneira que a Soma dos Pesos dos Arcos entre V1 e V2 seja máxima.

Max-Cut Problem

A cada nó i é associada uma variável x_i que é igual a 1 se o nó pertence à partição V1 e igual a 0 em caso contrário.

$$\text{maximize} \quad \sum_{(i,j)} c_{ij} \max(x_i - x_j, x_j - x_i)$$

$$\text{maximize} \quad \sum_{(i,j)} c_{ij} \left\| x_i - x_j \right\|$$

Max-Cut Problem

Problema Suavizado

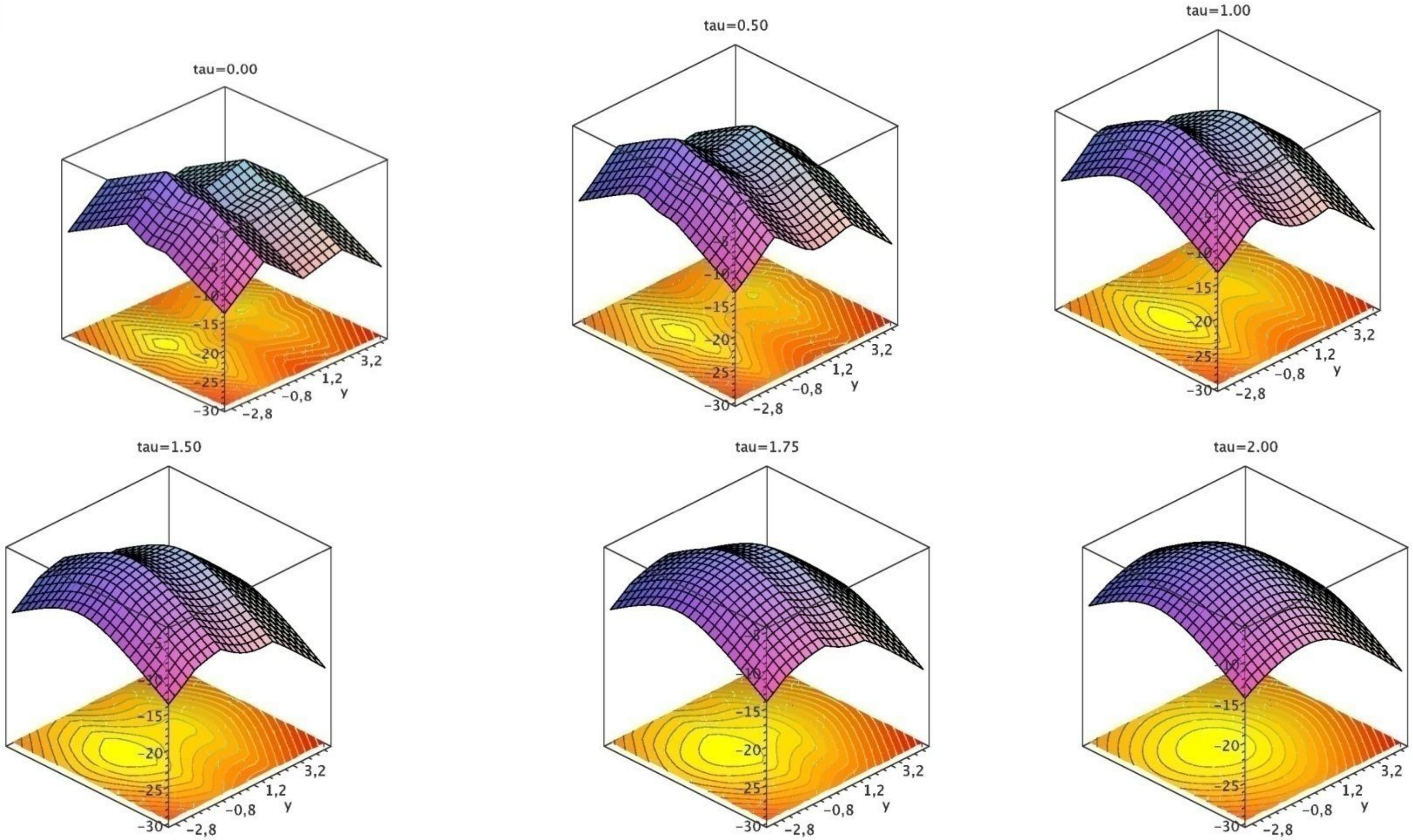
x_i

$$\text{maximize} \quad \sum_{(i,j)} c_{ij}\,\theta\left(x_i, x_j, \gamma \right)$$

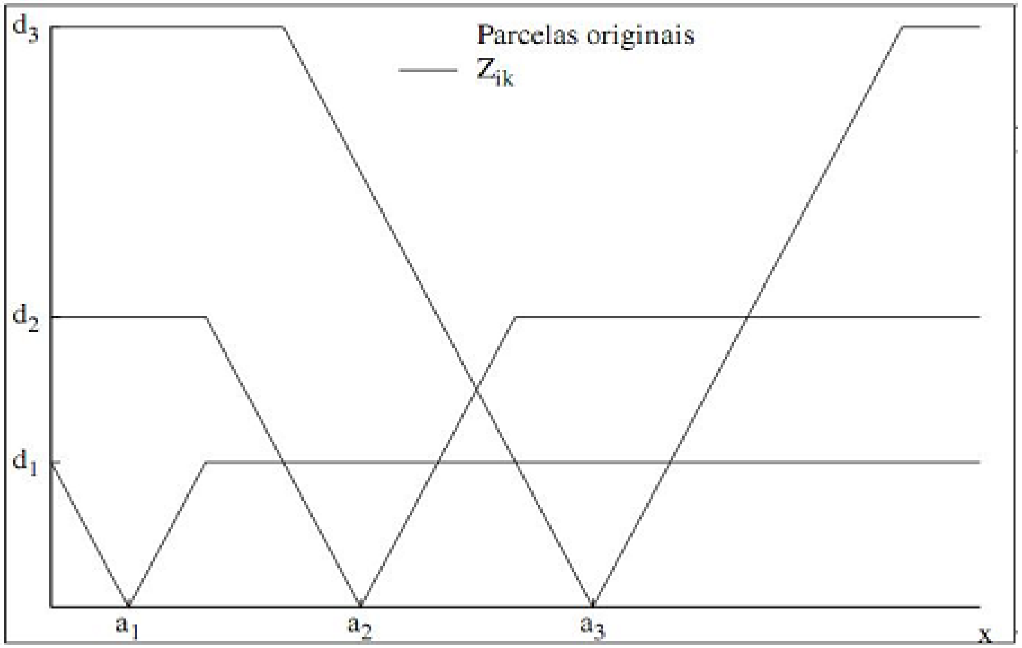
$$1 > x_i > 0$$

END

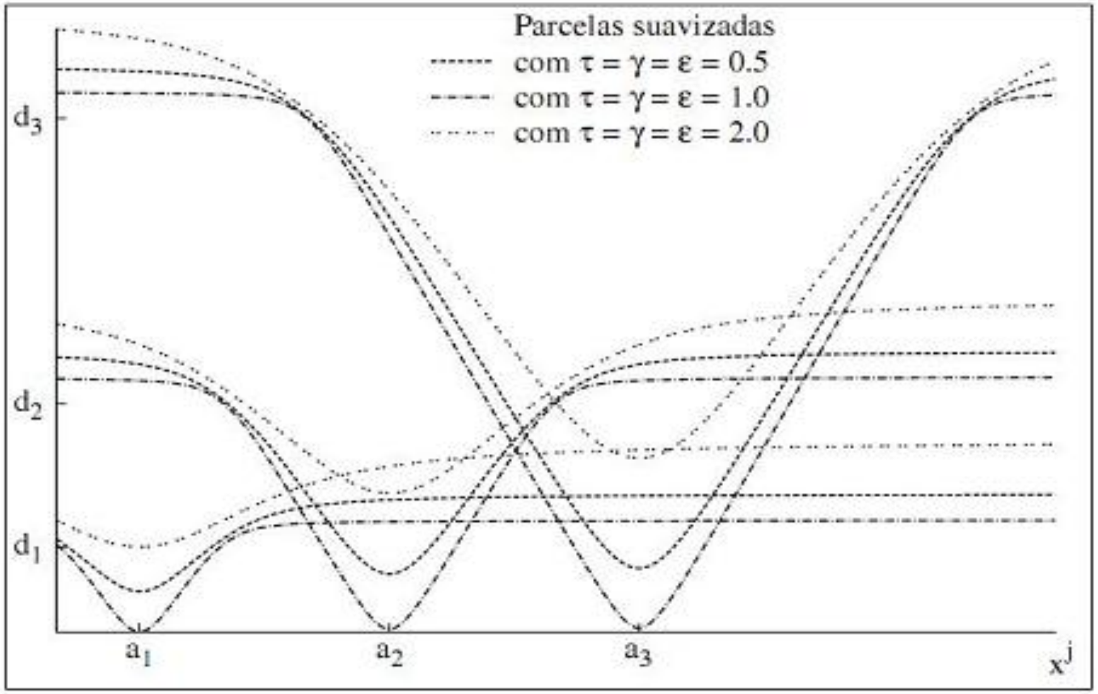
Convexification effect by the Hyperbolic Smoothing approach
Geometric distance problem



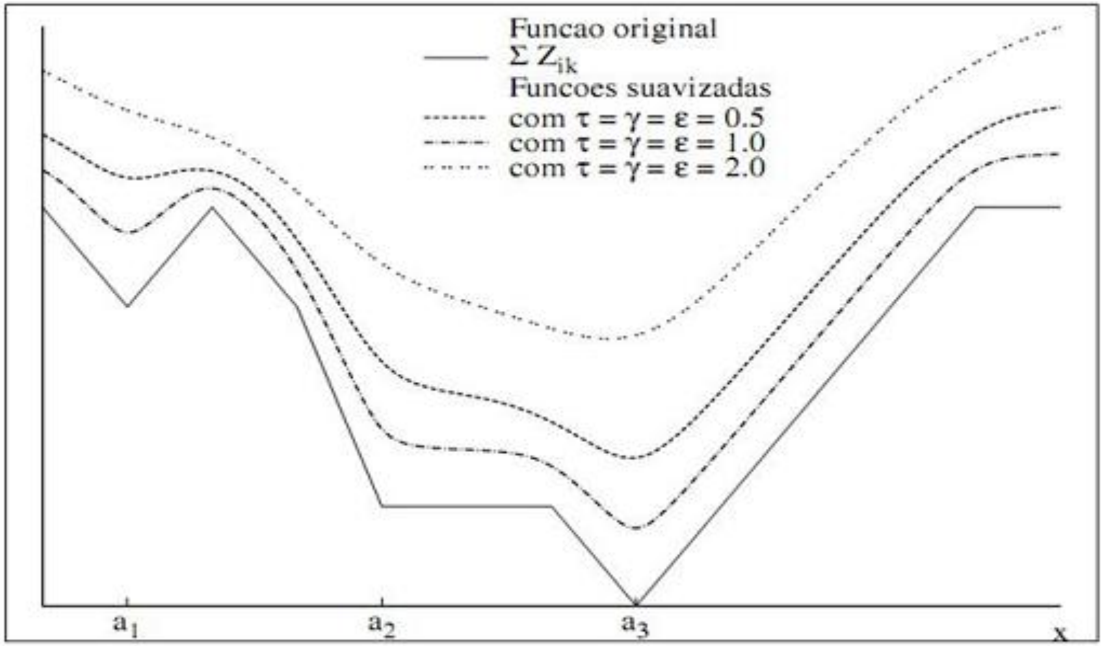
Smoothing of the objective function terms of a specific classification problem



Smoothing of the objective function terms of a specific classification problem



Smoothing of the objective function terms of a specific classification problem: Global Effect on the Objective Function



Quasi-Convexification effect by the Hyperbolic Smoothing approach associated to the Covering of a Region with Equal Circles Problem