Hyperbolic Smoothing Clustering Methods

Adilson Elias Xavier

PESC-COPPE-UFRJ

Rio de Janeiro, 03 Dezembro 2014

Outline of Presentation

1 - Cluster Analysis

2 - Hyperbolic Smoothing Clustering Method (HSCM)

3 - The HSCM applied for Solving the Minimum Sum-of-Squares Clustering Problem

4 - Partition into Boundary and Gravitational Regions (PBGR)

5 - HSCM + PBGR Applied to General Clustering Problem

6 - HSCM + PBGR Applied to Minimum Sum-of-Squares Clustering Problem (MSSC)

7 – HSCM Applied to MSSC ➔ Large Instances (Bagirov et al – December 2012)

8 - HSCM + PBGR Applied to MSSC ➔ Very Large Instances (March 2013)

9 - HSCM + PBGR Applied to MSSC ➔ Medium and Large Instances (April 2014)

10 - CONCLUSIONS

Cluster Analysis

In the Cluster Analysis scope, algorithms use traditionally two main strategies:

- Hierarchical clustering methods

- Partition clustering methods...

- Here we introduce a New Approach for solving the problem...

Hyperbolic Smoothing Clustering Method (HSCM)

Hyperbolic Smoothing Clustering Method - HSCM

General Problem Formulation:

$S = \{S_1, \ldots, S_m\}$ Set of m observations

$x_i \ , \ i = 1, \ldots, q$ Centroides of the clusters

$z_j = \min_{i=1,\ldots,q} \| s_j - x_i \|$ Distance of the observation j to nearest centroi

Objective function specification

$$Minimize \quad \sum_{j=1}^{m} f(z_j)$$

where $f(z_j)$ is an arbitrary monotonic increasing function of the distance z_j .

Hyperbolic Smoothing Clustering Method - HSCM

General Problem Formulation:

$$Minimize \quad \sum_{j=1}^{m} f(z_j)$$

$$z_j = \min_{i=1,\dots,q} \left\| s_j - x_i \right\|, \; j = 1,\dots,m$$

Trivial Examples of monotonic increasing functions:

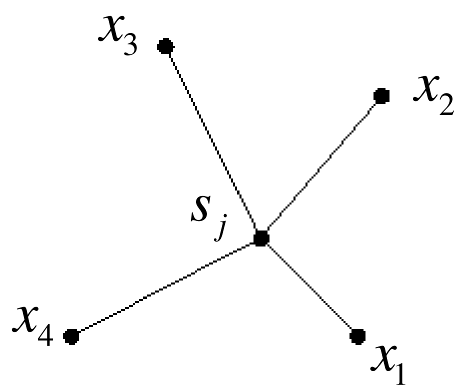
$$\sum_{j=1}^{m} f(z_j) = \sum_{j=1}^{m} Z_j^2 \qquad \sum_{j=1}^{m} f(z_j) = \sum_{j=1}^{m} Z_j$$

Possible Distance Metrics:

L_2 (Euclidean) L_1 (Manhattan) L_∞ (Chebychev) L_p (Minkowshi)

HSCM: Resolution Methodology

The General Clustering Problem



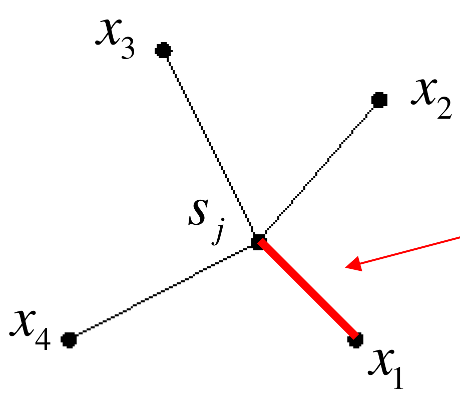
Consider a general observation point s_j.

Let be $x_i, i = 1,\ldots,q$, the centroids of clustersm where each set of these centroid coordinates will be represented by $X \in \Re^{nq}$.

HSCM: Resolution Methodology

The General Clustering Problem

Consider a general observation point s_j



$$z_j = \min_{x_i \in X} \left\| s_j - x_i \right\|$$

Let be $x_i, i = 1, \ldots, q$, the centroids of clusters where each set s_j of these centroid coordinates will be represented by $X \in \Re^{nq}$.

This procedure is strongly non differentiable.

HSCM: Resolution Methodology

Problem transformation

The general clustering problem is equivalent to the following problem:

$$\text{(P1)}\ \textit{Minimize} \quad \sum_{j=1}^{m} f(z_j)$$

$$\textit{Subject to}: \quad z_j = \min_{i=1,\ldots,q} \left\| s_j - x_i \right\| \ , \quad j = 1,\ldots,m \, .$$

HSCM: Resolution Methodology

Problem transformation

From the problem (P1), it is obtained the relaxed problem:

$$\text{(P2)} \ \ \textit{Minimize} \quad \sum_{j=1}^{m} f(z_j)$$

$$\textit{Subject to:} \quad z_j - \left\| s_j - x_i \right\| \le 0, \ \ j = 1, \ldots, m, \ \ i = 1, \ldots, q.$$

HSCM: Resolution Methodology

Problem transformation

Let us use the auxiliary function

$$\psi(y) = \max\{0, y\}$$

From the inequalities

$$z_j - \|s_j - x_i\| \leq 0\,, \quad j = 1, \ldots, m, \; i = 1, \ldots, q,$$

it follows:

$$\sum_{i=1}^{q} \Psi\left(z_j - \|s_j - x_i\|\right) = 0, \; j = 1, \cdots, m$$

HSCM: Resolution Methodology

Problem transformation

Using this set of equality constraints in place of the constraints in (P2), we obtain the following problem

$$(P3) \quad Minimize \quad \sum_{j=1}^{m} f(z_j)$$

$$Subject\ to: \sum_{i=1}^{q} \Psi\left(z_j - \left\| s_j - x_i \right\|\right) = 0,\ j = 1, \cdots, m$$

This problem corresponds also to a relaxation!!

HSCM: Resolution Methodology

Problem transformation

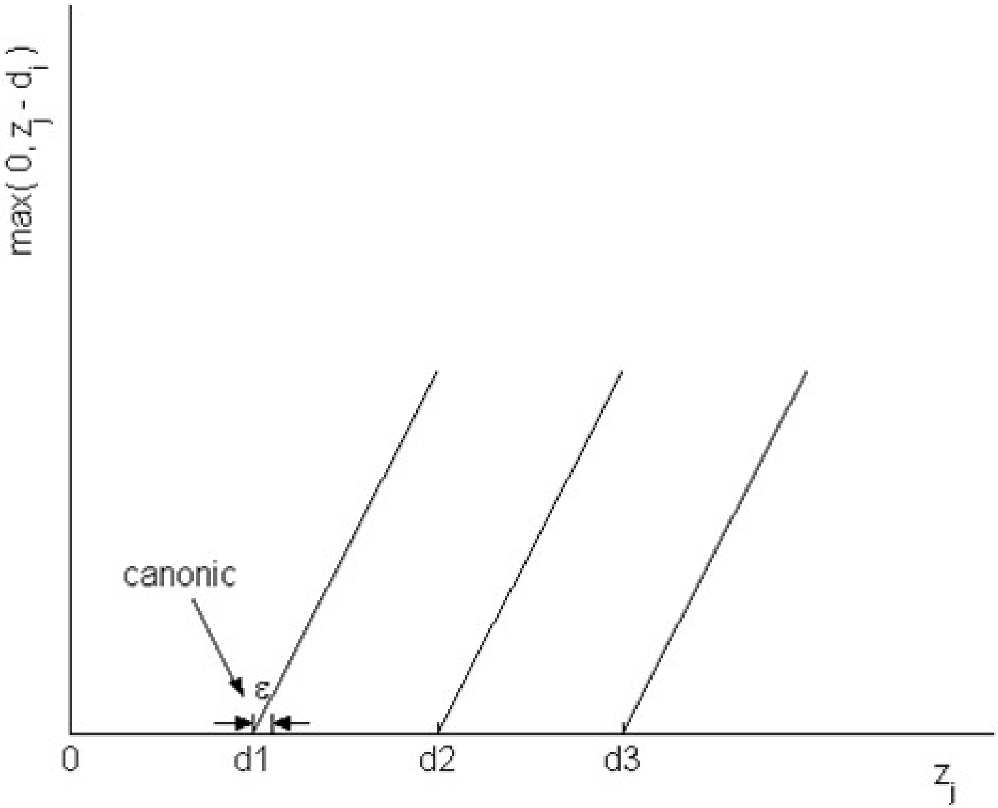
By performing a ε perturbation, it is obtained the problem:

$$(P4) \quad Minimize \quad \sum_{j=1}^{m} f(z_j)$$

$$Subject\ to: \sum_{i=1}^{q} \Psi\left(z_j - \left\| s_j - x_i \right\|\right) \geq \varepsilon,\ j = 1, \cdots, m$$

HSCM: Resolution Methodology

Problem transformation



HSCM: Resolution Methodology

Smoothing the problem

Let us define the auxiliary function

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

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

HSCM: Resolution Methodology

Smoothing the problem

Function ϕ has the following properties:

(a) $\phi(y,\tau) > \Psi(y), \forall \tau > 0$

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

(c) $\phi(y,\tau)$ is an increasing convex C^{∞} function.

HSCM: Resolution Methodology

Smoothing the problem

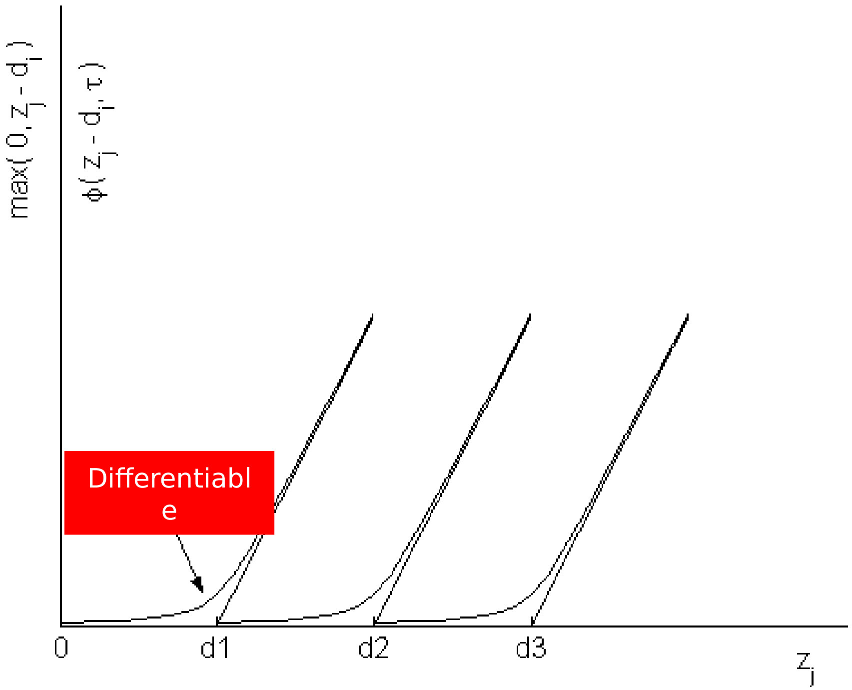
By using the Hyperbolic Smoothing approach for the problem (P4) it is obtained

$$\text{(P5)}\ \textit{Minimize} \quad \sum_{j=1}^{m} f(z_j)$$

$$\textit{Subject to}: \quad \sum_{i=1}^{q} \phi(z_j - \left\| s_j - x_i \right\|, \tau) \geq \varepsilon, \quad j = 1, \ldots, m.$$

HSCM: Resolution Methodology

Smoothing the problem



HSCM: Resolution Methodology

Smoothing the distance calculation

For the Euclidian metric, let us define the auxiliary function

$$\theta(s_j, x_i, \gamma) = \sqrt{\sum_{l=1}^{n} (s_{jl} - x_{il})^2 + \gamma^2}$$

for $\gamma > 0$.

HSCM: Resolution Methodology

Smoothing the Euclidian distance:

Function θ has the following properties:

(a) $\lim_{\gamma \to 0} \theta(s_j, x_i, \gamma) = \left\| s_j - x_i \right\|_2$;

(b) θ is a C^{∞} function.

HSCM: Resolution Methodology

Smoothing the problem

Therefore, it is now obtained the completely smooth problem:

$$\text{(P6)}\ \ \textit{Minimize} \quad \sum_{j=1}^{m} f(z_j)$$

$$\textit{Subject to:} \quad \sum_{i=1}^{q} \phi(z_j - \theta(s_j, x_i, \gamma), \tau) \geq \varepsilon, \quad j = 1, \ldots, m.$$

HSCM: Resolution Methodology

Problem resolution

By considering the KKT conditions for the problem (P6), it is possible to conclude that all inequalities will certainly be active.

$$(\text{P7})\ \textit{Minimize} \quad \sum_{j=1}^{m} f(z_j)$$

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

HSCM: Resolution Methodology

Problem resolution

Some remarks about the problem (P7):

The variable domain space has $(nq + m)$ dimensions.

All functions of this problem are C^{∞} functions in relation to the variables (x, z).

Each variable z_j appears only in one equality constraint $h(x, z_j) \quad \forall \quad j = 1, ..., m.$

HSCM: Resolution Methodology

Problem resolution

Some remarks about the problem (P7):

The parcial derivative of $h(x, z_j)$ in relation to z_j is not equal to zero for all $j = 1, \ldots, m$.

In the face of the previous remarks, it is possible to use the Implicit Function Theorem in order to calculate each component $z_j, \; j = 1, \ldots, m$, as a function of the centroids variables $x_i, \; i = 1, \ldots, q$.

HSCM: Resolution Methodology

Problem resolution

Therefore, by using the Implicit Function Theorem, it is obtained the unconstrained problem

$$\text{(P8)}\ \textit{Minimize}\ \ F(x) = \sum_{j=1}^{m} f(z_j)$$

where each z_j is obtained by the calculation of a zero of each equation

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

HSCM: Resolution Methodology

Problem resolution

Again, due to the Implicit Function Theorem, the functions $z_j(x)$ have all derivatives in relation to the variables $x_i, \; i = 1,\dots,q$. So, it is possible to calculate the gradient of the objective function of the problem (P8)

$$\nabla F(x) = \sum_{j=1}^{m} \frac{df(z_j(x))}{dz_j} \nabla z_j(x)$$

where

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

HSCM: Resolution Methodology

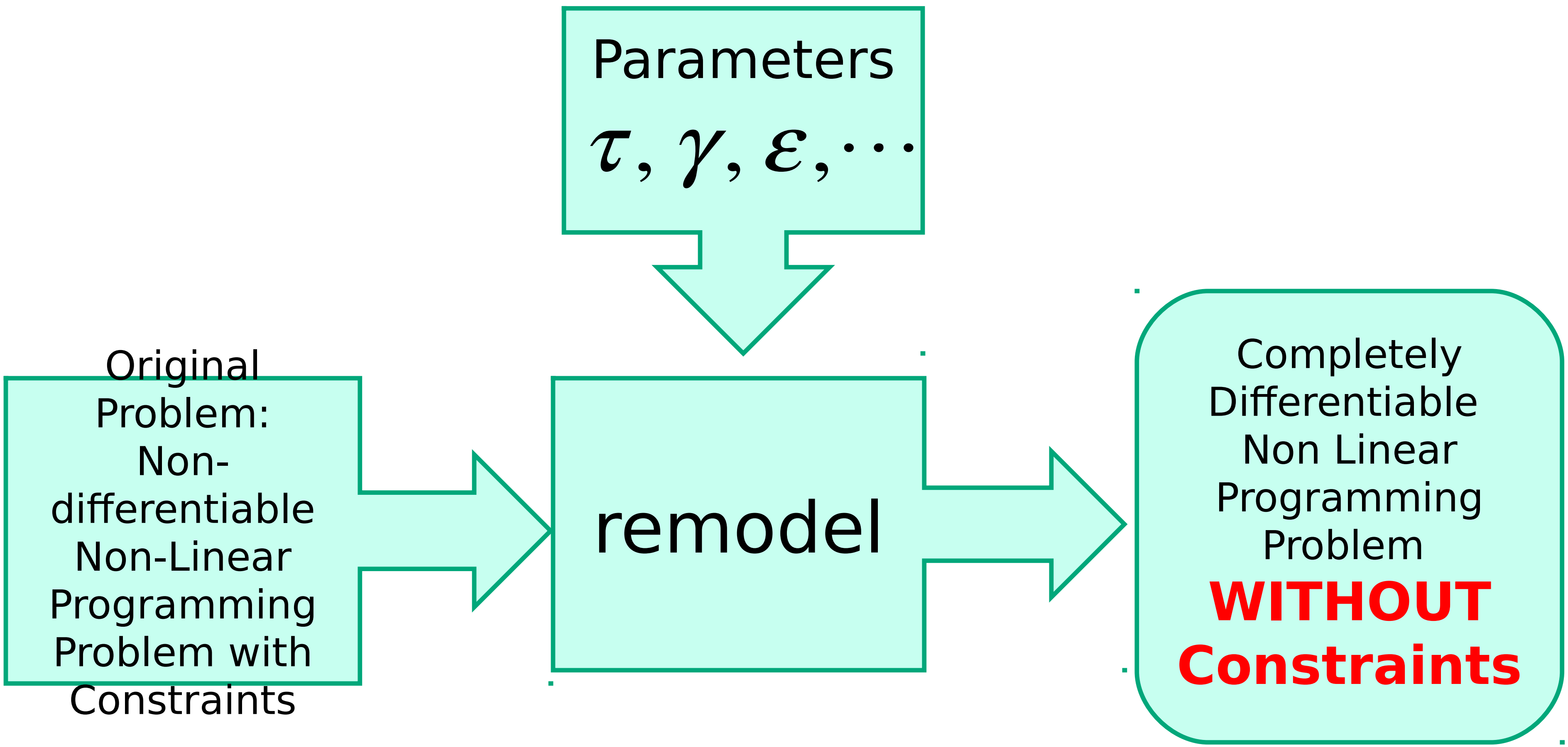
Problem resolution

Some remarks about the problem (P8):

It can be solved by making use of any method based in the first and second order derivatives information (conjugate gradient, quasi-Newton, Newton methods, etc).

It is defined in a *nq*-dimensional space, therefore a smaller dimension problem in relation to the original problem (P7).

Hyperbolic Smoothing Clustering Method:
The Main Result



Simplified HSC Algorithm

Initialization Step: Choose values $0 < \rho_1 < 1,\ 0 < \rho_2 < 1,\ 0 < \rho_3 < 1;$

Choose initial values: $x^0,\ \gamma^1,\ \tau^1,\ \varepsilon^1$; Let $k = 1$.

Main Step: Repeat indefinitely

Solve the unconstrained problem (P8)

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

with $\gamma = \gamma^k,\ \tau = \tau^k,\ \varepsilon = \varepsilon^k$, starting at the initial point x^{k-1},

and let x^k be the solution obtained.

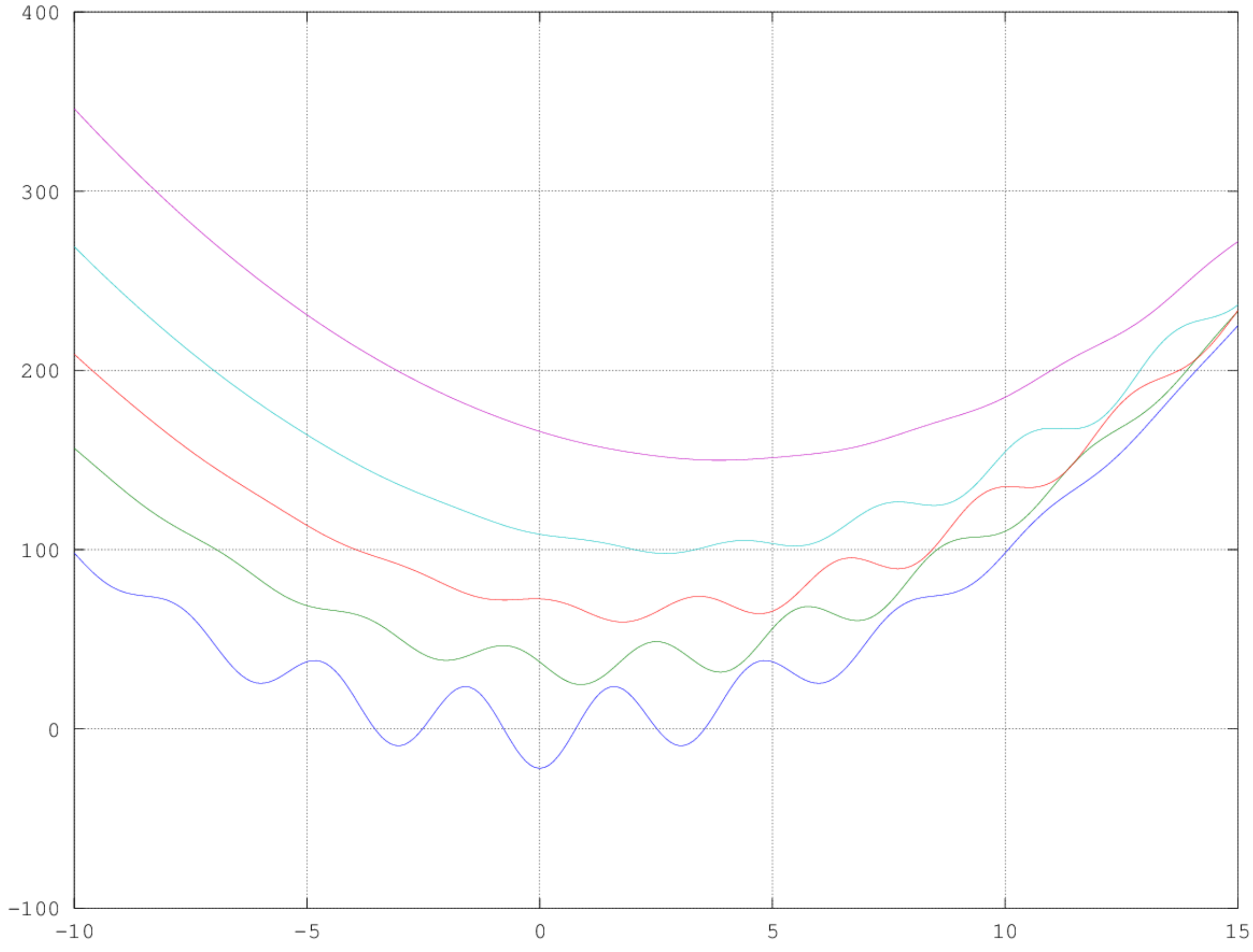
Let $\gamma^{k+1} = \rho_1 \gamma^k,\ \tau^{k+1} = \rho_2 \tau^k,\ \varepsilon^{k+1} = \rho_3 \varepsilon^k,\ k = k+1.$

Remarks about the HSCM Algorithm

The algorithm causes τ and γ approach zero, so the constraints of the subproblems it solves, given as in (P7), tend to those of (P4).

The algorithm causes ε approaches to zero, so, in a simultaneous movement, the solved problem (P4) gradually approaches to problem (P3).

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



Hyperbolic Smoothing Clustering Method

applied to:

Minimum Sum-of-Squares Clustering Problem

Pattern Recognition 43 (2010) 731–737

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

The hyperbolic smoothing clustering method

Adilson Elias Xavier

Department of Systems Engineering and Computer Science, Graduate School of Engineering (COPPE), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 20 May 2008
Received in revised form 21 October 2008
Accepted 19 June 2009

Keywords:
Cluster analysis
Min-sum-min problems
Nondifferentiable programming
Smoothing

ABSTRACT

The minimum sum-of-squares clustering problem is considered. The mathematical modeling of this problem leads to a *min-sum-min* formulation which, in addition to its intrinsic bi-level nature, has the significant characteristic of being strongly nondifferentiable. To overcome these difficulties, the resolution method proposed adopts a smoothing strategy using a special C^{∞} differentiable class function. The final solution is obtained by solving a sequence of low dimension differentiable unconstrained optimization subproblems which gradually approach the original problem. The use of this technique, called hyperbolic smoothing, allows the main difficulties presented by the original problem to be overcome. A simplified algorithm containing only the essentials of the method is presented. For the purpose of illustrating both the reliability and the efficiency of the method, a set of computational experiments was performed, making use of traditional test problems described in the literature

Problem Specification

Minimum Sum-of-Squares Clustering

Among many criteria used in Cluster Analysis, the most natural, intuitive, and frequently adopted criterion is the minimum sum-of-squares clustering (MSSC).

The Minimum Sum-of-Squares Clustering Problem

Problem definition

$$(\text{P1})\ \textit{Minimize} \quad \sum_{j=1}^{m} (z_j)^2$$

$$\textit{Subject to}: \quad z_j = \min_{i=1,\ldots,q} \left\| s_j - x_i \right\| \ , \quad j = 1,\ldots,m\,.$$

Problem Specification

The Minimum Sum-of-Squares Clustering Problem

It produces a global optimization mathematical problem.

It is a mathematical problem both non-smooth and nonconvex, with a large number of local minimizers.

It is in the class of NP-hard problems (Brucker, 1978).

Basic Literature

Methodologies for the solution of MSSC problem

Dynamic programming (Jensen, 1969).

Branch and Bound algorithms (Koontz *et alli*, 1975, Diehr, 1985, Hanjoul and Peeters, 1985, Hansen and Jaumard, 1997).

Heuristic Algorithms - The *k*-means type algorithms is the heuristic more frequently used for the MSSC problem resolution (Späth, 1980, Merle *et alli*, 2000, Hansen and Jaumard, 1997, Selim and Ismail, 1984, Hansen *et alli*, 2002).

Basic Literature

Methodologies for the solution of MSSC problem

Metaheuristic Algorithms:

Simulated annealing (Brown and Entail, 1992, Selim and Al-Sultan, 1991, Sun *et alli*, 1994).

Tabu search (Al-Sultan, 1995).

Genetic algoritms (Reeves, 1991).

Al-Sultan and Khan (1996) show that these metaheuristics spend more CPU time than the *k*-means algorithms.

Basic Literature

Methodologies for the solution of MSSC problem

Combination of local search algoritms, simulated annealing, tabu search, *k*-means and neural networks (Eikelder and Erk, 2004).

BAGIROV *et alli* (2001) present a global optimization approach.

Interior points method (du Merle *et alli*, 2000).

Basic Literature

Methodologies for the solution of MSSC problem

Variable neighborhood search algorithms, VNS, is developed by Mladenovic and Hansen (1997) and Hansen and Mladenovic (2001).

J-means algorithm, that is a modification of *k*-means algorithms, is presented by Hansen and Mladenovic (2001).

Basic Literature

Methodologies for the solution of MSSC problem

Global *k*-means heuristic, an improvement of the *k*-means algorithm, is presented by Likas *et alli* (2003) and HANSEN *et alli* (2002).

Bilinear programming techniques (MANGASARIAN, 1997).

Computational Results

- PC Intel Celeron with 2.7GHz CPU and 512MB RAM;
- Compaq Visual FORTRAN, Version 6.1;
- Quasi-Newton algorithm employing the BFGS updating formula from the Harwell Library (http://www.cse.scitech.ac.uk/nag/hsl/).

HSCM - Computational Results: TSPLIB-3038

Minimum Sum-of-Squares Clustering Formulation

Results published in Pattern Recognion 43 (2010) 731-737

q	Putative Global Minimum	Best Solution HSC Algorithm	E_{Best}	Best Solution Frequence	Start Points HSC Algorithm	E_{Mean}	CPU Time
2	0.31688E10	0.31705E10	0.05	10	10	0.05	0.60
3	0.21763E10	0.21776E10	0.06	7	10	1.08	1.18
4	0.14790E10	0.14793E10	0.02	10	10	0.02	1.81
5	0.11982E10	0.11986E10	0.03	9	10	0.06	3.09
6	0.96918E9	0.96936E9	0.02	10	10	0.02	9.56
7	0.83966E9	0.83967E9	0.001	7	10	4.66	10.65
8	0.73475E9	0.73491E9	0.02	1	10	1.35	16.24
9	0.64477E9	0.64471E9	**-0.01**	2	10	0.38	10.05
10	0.56025E9	0.56030E9	+0.01	10	10	0.01	16.16
20	0.26681E9	0.26675E9	**-0.02**	1	10	2.51	62.90
30	0.17557E9	0.17543E9	**-0.08**	3	10	0.36	169.17
40	0.12548E9	0.12541E9	**-0.06**	1	20	1.34	333.92
50	0.98400E8	0.98893E8	+0.50	1	40	1.46	662.33
60	0.82006E8	0.81115E8	**-1.09**	1	10	0.19	1049.97
80	0.61217E8	0.61025E8	**-0.31**	1	10	0.63	2038.31
100	0.48912E8	0.48470E8	**-0.90**	1	10	0.19	3499.76

HSCM: Results for the Pla85900

Minimum Sum-of-Squares Clustering Formulation

Results published in Pattern Recognion 43 (2010) 731-737

q	$f_{Calculated}$	Occur.	E_{Mean}	$Time_{Mean}$
2	0.37491E16	4	0.86	23.07
3	0.22806E16	10	0.00	47.41
4	0.15931E16	10	0.00	76.34
5	0.13397E16	1	0.80	124.32
6	0.11366E16	8	0.12	173.44
7	0.97110E15	4	0.42	254.37
8	0.83774E15	8	0.55	353.61
9	0.74660E15	3	0.68	438.71
10	0.68294E15	4	0.29	551.98

Hyperbolic Smoothing and Partition into Boundary and Gravitational Regions Applied to General Clustering Problem

The Computational Task of the HSC Algorithm

The most relevant computational task associated to HSC Algorithm is the determination of the zeros of m equations:

$$\sum_{i=1}^{q} \phi\left(z_j - \theta(s_j, x_i, \gamma), \tau \right) - \varepsilon = 0, \; j = 1, \ldots, m$$

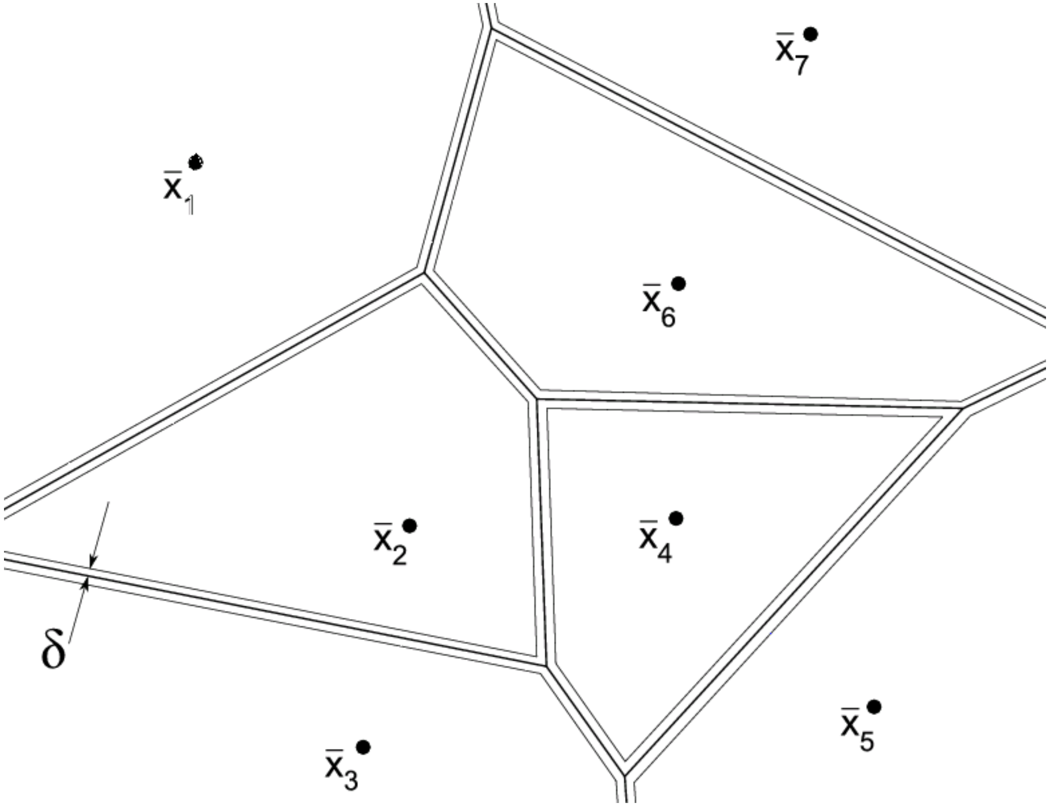
The Partition into Boundary and Gravitational Regions

The basic idea of the approach is the partition of the set of observations in two non overlapping parts:

- the first set corresponds to the observation points that are relatively close to two or more centroids;

- the second set corresponds to the observation points that are significantly close to a unique centroid in comparison with the other ones.

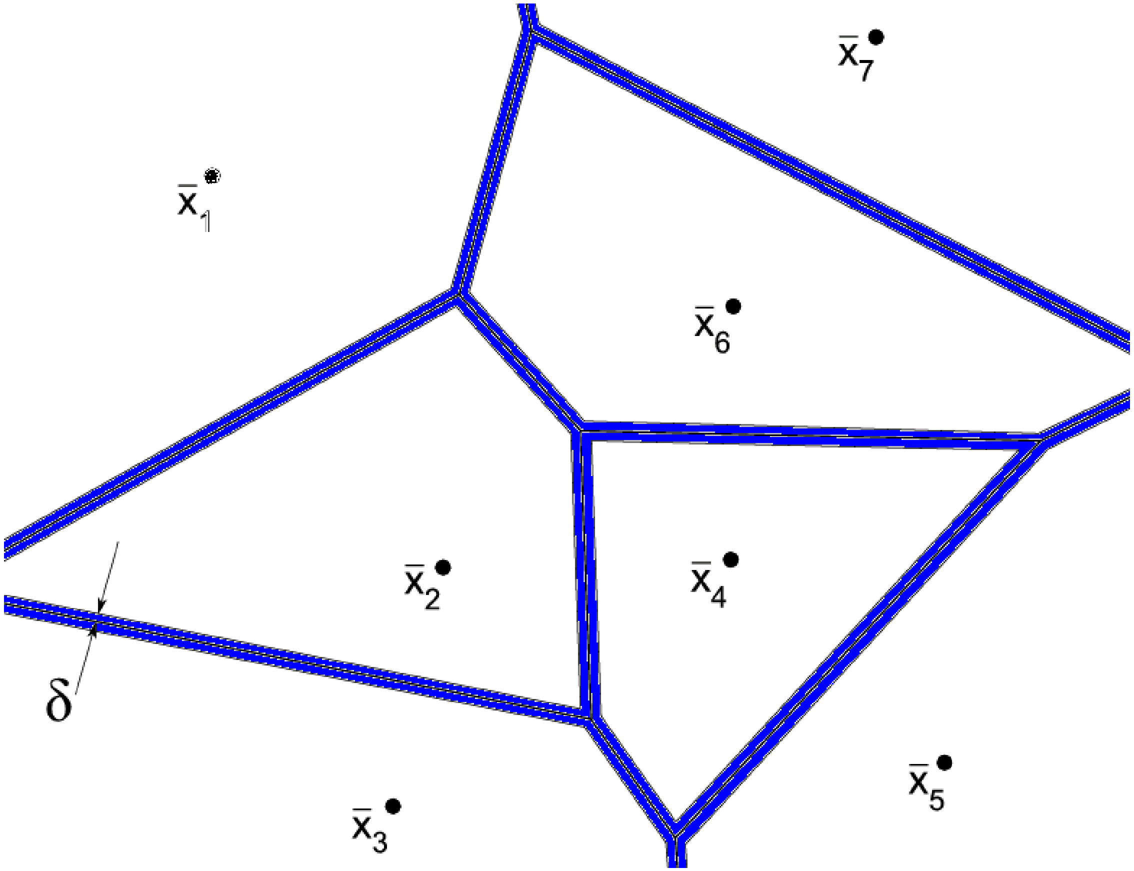
Partition of the Set of Observations



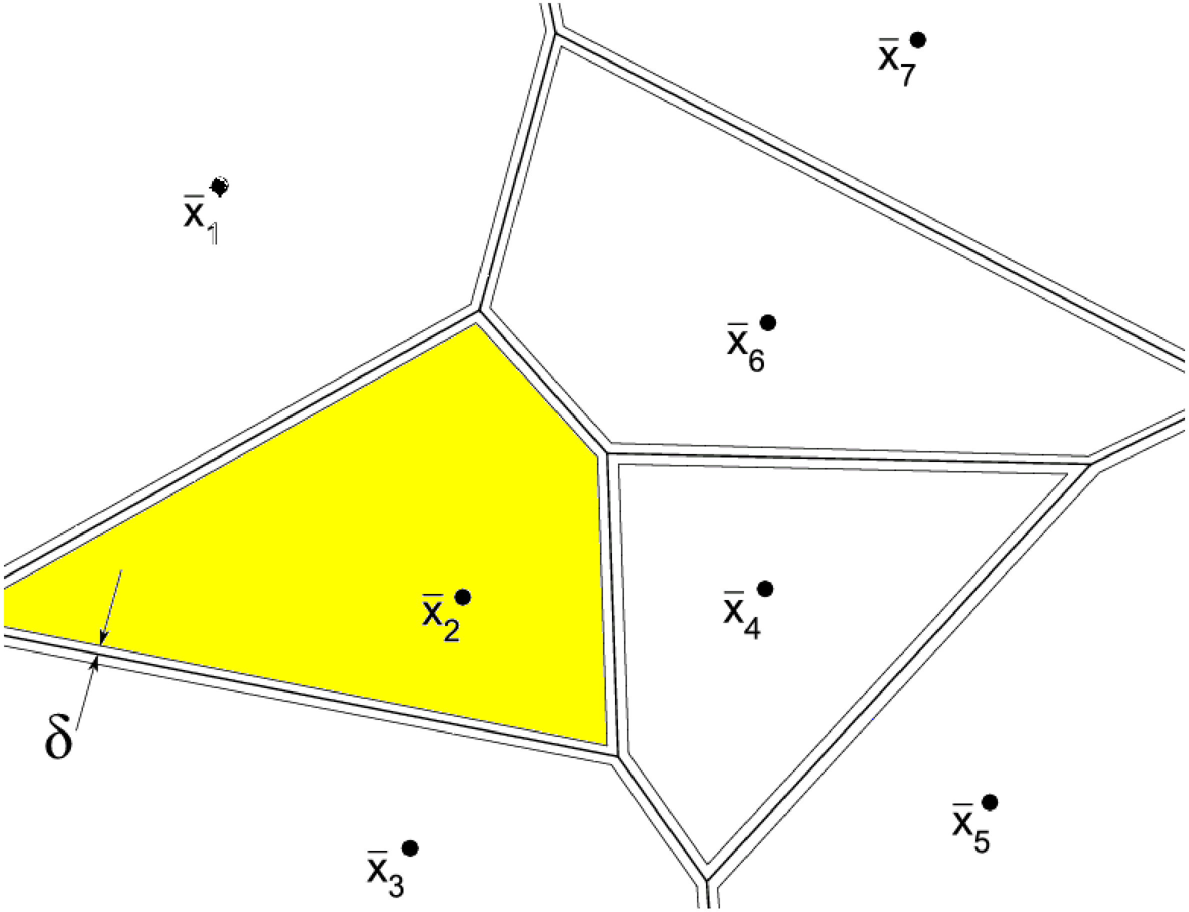
$\bar{x}$: the referencial point

δ: the band zone width

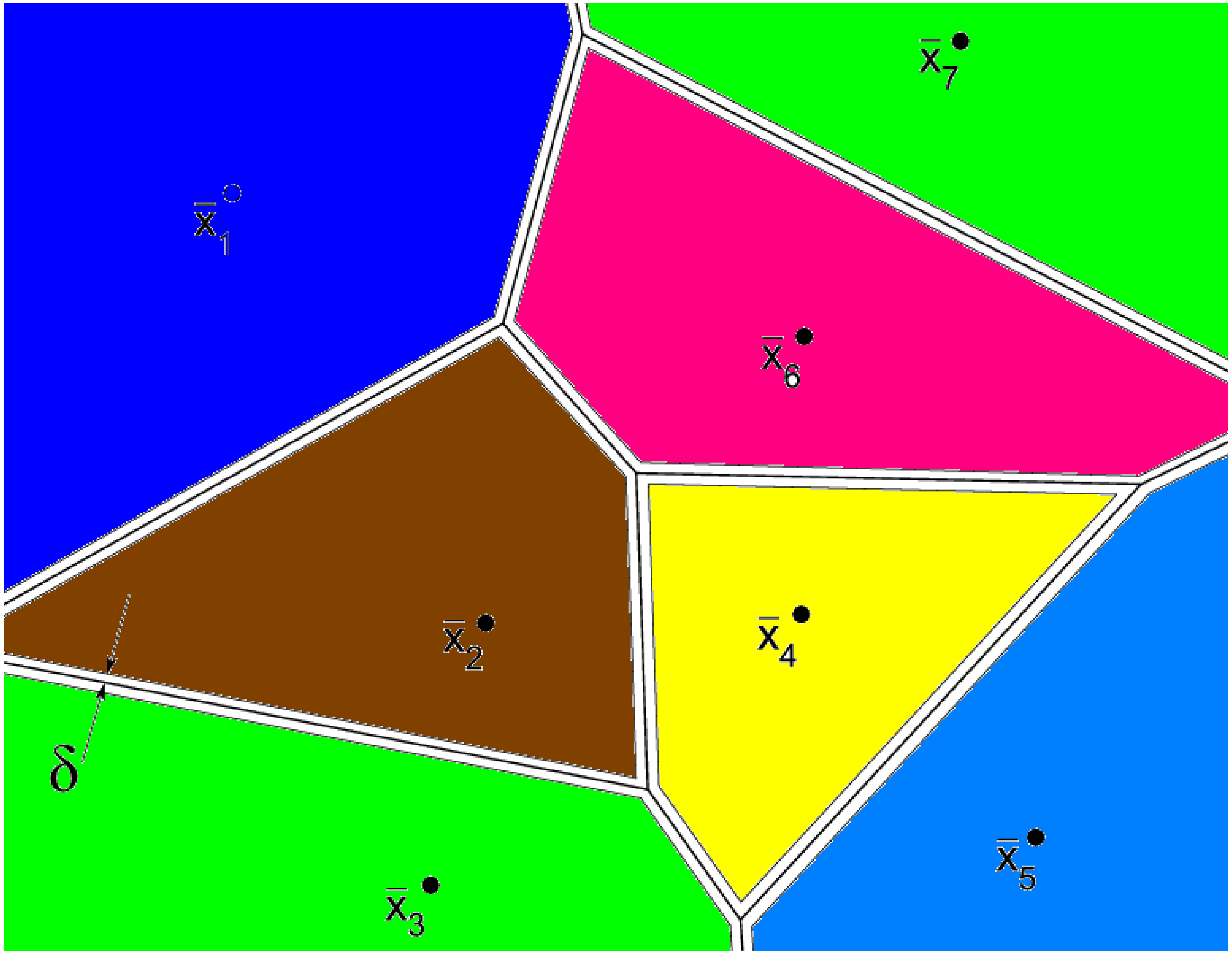
The Boundary Band Zone



Gravitational Region => $\bar{x}_2$



Gravitational Regions



The Partition of the Set of Observations

Let us define:

- J_B is the set of boundary observations;
- J_G is the set of gravitational observations.

Considering this partition, the objective function can be expressed in the following way:

$$Minimize \;\; F(x) = \sum_{j=1}^{m} f(z_j) = \sum_{j \in J_B} f(z_j) + \sum_{j \in J_G} f(z_j)$$

So, the objective function can be expressed in the following way:

$$Minimize \; F(x) = f_B(x) + f_G(x)$$

The First Component of the Objective Function

The component associated with the set of boundary observations, can be calculated by using the previous presented smoothing approach:

$$Minimize \;\; F_B(x) = \sum_{j \in J_B} f(z_j)$$

where each z_j results from the calculation of a zero of each equation:

$$h(x, z_j) = \sum_{i=1}^{q} \phi\Big(z_j - \theta(s_j, x_i, \gamma), \tau\Big) - \varepsilon = 0, \; j \in J_B$$

The Second Component of the Objective Function

The component associated with the gravitational regions, can be calculated in a simpler form:

$$Minimize \ \ F_G(x) = \sum_{i=1}^{q} \sum_{j \in J_i} f(\|s_j - x_i\|)$$

where J_i is the set of observations associated to cluster i.

Special Clustering Formulations

The Hyperbolic Smoothing and Partition into Boundary and Gravitational Regions connection presents particular and advantageous properties for:

- Minimum Sum-of-Squares-Euclidian Distances

- Minimum Sum-of-Distances Manhattan Distances

Hyperbolic Smoothing and Partition into Boundary and Gravitational Regions Applied to Minimum Sum-of-Squares Clustering Problem

Pattern Recognition 44 (2011) 70–77

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Solving the minimum sum-of-squares clustering problem by hyperbolic smoothing and partition into boundary and gravitational regions

Adilson Elias Xavier *, Vinicius Layter Xavier

Department of Systems Engineering and Computer Science, Graduate School of Engineering (COPPE), Federal University of Rio de Janeiro, P.O. Box 68511, Rio de Janeiro RJ 21941-972, Brazil

ARTICLE INFO

Article history:
Received 11 January 2010
Received in revised form
20 May 2010
Accepted 6 July 2010

Keywords:
Cluster analysis
Min-sum-min problems
Nondifferentiable programming
Smoothing

ABSTRACT

This article considers the minimum sum-of-squares clustering (MSSC) problem. The mathematical modeling of this problem leads to a *min-sum-min* formulation which, in addition to its intrinsic bi-level nature, has the significant characteristic of being strongly nondifferentiable. To overcome these difficulties, the proposed resolution method, called hyperbolic smoothing, adopts a smoothing strategy using a special C^∞ differentiable class function. The final solution is obtained by solving a sequence of low dimension differentiable unconstrained optimization subproblems which gradually approach the original problem. This paper introduces the method of partition of the set of observations into two nonoverlapping groups: "data in frontier" and "data in gravitational regions". The resulting combination of the two methodologies for the MSSC problem has interesting properties, which drastically simplify the computational tasks.

The Second Component for the MSSC

For The Minimum Sum-of-Squares Clustering formulation,
the second component, associated with the gravitational regions,
can be calculated in a direct form:

$$Minimize \;\; F_G(x) = \sum_{i=1}^{q} \sum_{j \in J_i} \left\| s_j - v_i \right\|_2^2 + \sum_{i=1}^{q} \left| J_i \right| \left\| x_i - v_i \right\|_2^2$$

where J_i is the set of observations associated to centroid i

v_i is the gravity center of cluster i.

The Gradient of the Second Component for the MSSC

The gradient of the second component, associated with the gravitational regions, can be calculated in an easy way:

$$\nabla F_G(x) = \sum_{\iota=1}^{q} 2\left|J_i\right| \left(x_i - v_i \right)$$

where J_i is the number the of observations associated to i centroid

Simplified Accelerated Algorithm: AHSCM

Initialization Step: Choose values $0 < \rho_1 < 1,\ 0 < \rho_2 < 1,\ 0 < \rho_3 < 1;$

Choose initial values: $x^0,\ \gamma^1,\ \tau^1,\ \varepsilon^1$; Let $k = 1$.

Specify the initial boundary band width: δ^1.

Main Step: Repeat indefinitely

By using $\bar{x} = x^{k-1}$ and $\delta = \delta^k$ determine partitions J_B and J_G.

Solve the problem

$$minimize\ \ f(x) = \sum_{j \in J_B} z_j(x)^2 + \sum_{j \in J_G} z_j(x)^2$$

with $\gamma = \gamma^k,\ \tau = \tau^k,\ \varepsilon = \varepsilon^k$, starting at the initial point x^{k-1},

and let x^k be the solution obtained.

Redefine the boundary band width value δ^{k+1}.

Let $\gamma^{k+1} = \rho_1 \gamma^k,\ \tau^{k+1} = \rho_2 \tau^k,\ \varepsilon^{k+1} = \rho_3 \varepsilon^k,\ k = k+1.$

Comparison of Results: TSPLIB-3038

Hyperbolic Smoothing Clustering & Accelerated HSC Methods

Results published in Pattern Recognion 44 (2011) 70-77

q	f_{opt}	Algorithm HSC		Algorithm AHSC-L2		Speed up
		Error	Time (s)	Error	Time (s)	
2	0.31688E10	0.05	0.60	0.05	0.07	**8.6**
10	0.56025E09	0.01	16.16	0.01	0.28	**57.7**
20	0.26681E09	**-0.02**	62.90	0.05	0.59	**107**
30	0.17557E09	**-0.08**	169.17	0.31	0.86	**197**
40	0.12548E09	**-0.06**	333.92	**-0.11**	1.09	**306**
50	0.98400E08	0.50	662.33	0.44	1.36	**487**
60	0.82006E08	**-1.09**	1049.97	**-0.80**	1.91	**550**
80	0.61217E08	**-0.31**	2038.31	**-0.73**	6.72	**303**
100	0.48912E08	**-0.90**	3499.76	**-0.60**	9.79	**357**

Comparison of Results: Pla85900

Hyperbolic Smoothing Clustering & Accelerated HSC Methods

Results published in Pattern Recognion 44 (2011) 70-77

q	f_{opt}	Algorithm HSC		Algorithm AHSC-L2		Speed up
		Error	Time (s)	Error	Time (s)	
2	0.37491E16	0.86	23.07	0.58	3.65	**6.3**
3	0.22806E16	0.00	47.41	0.04	4.92	**9.6**
4	0.15931E16	0.00	76.34	0.00	5.76	**13.3**
5	0.13397E16	0.80	124.32	1.35	7.78	**16.0**
6	0.11366E16	0.12	173.44	1.25	7.87	**22.0**
7	0.97110E15	0.42	254.37	0.87	9.33	**27.3**
8	0.83774E15	0.55	353.61	0.37	12.96	**27.3**
9	0.74660E15	0.68	438.71	0.25	13.00	**33.8**
10	0.68294E15	0.29	551.98	0.46	14.75	**37.4**

Results for Page-Blocks instance

Minimum Sum-of-Squares Clustering Formulation

q	f_{opt}	MGKM		AHSC-L2	
		Error	Time (s)	Error	Time (s)
2	0.57937E11	0.00	6.92	**0.00**	**0.12**
10	0.45662E10	0.00	34.09	**-72**	**0.56**
20	0.17139E10	0.19	62.09	**-1**	**1.86**
30	0.94106E09	0.00	89.42	1.21	**4.23**
40	0.62570E09	0.00	118.55	**-179**	**7.74**
50	0.42937E09	0.00	149.77	**-144**	**14.31**
60	0.31185E09	0.33	184.06	**-3**	**27.78**
80	0.20576E09	0.00	258.69	**-113**	**50.56**
100	0.14545E09	0.10	346.94	1.18	**90.31**

m = 5437 n = 10

Hyperbolic Smoothing Applied to Minimum Sum-of-Squares Clustering Problem

Bagirov et al - December 2012

An incremental clustering algorithm based on hyperbolic smoothing

A.M. Bagirov[1], B. Ordin[2], G. Ozturk[3] and A.E. Xavier[4]

[1] *School of Science, Information Technology and Engineering, University of Ballarat, Victoria, Australia,*

[2] *Department of Mathematics, Faculty of Science, Ege University, Izmir, Turkey,*

[2] *Department of Industrial Engineering, Faculty of Engineering, Anadolu University, Eskisehir, Turkey,*

[4] *Department of Systems Engineering and Computer Science, Graduate School of Engineering, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil.*

Comparison Results

General k-means, Modified General k-means, Hyperbolic Smoothing

TSPLIB3038

		GKM			MGKM			SMOOTH		
q	$\mathbf{f_{opt}}$	**E**	**α**	**t**	**E**	**α**	**t**	**E**	**α**	**t**
2	$0.31688 \cdot 10^{10}$	0.00	0.09	0.06	0.00	0.19	0.11	0.00	0.26	0.25
5	$0.11982 \cdot 10^{10}$	0.00	0.38	0.25	0.00	0.75	0.39	0.00	0.95	0.98
10	$0.56025 \cdot 10^{9}$	2.78	0.87	0.48	0.58	1.70	0.81	0.00	2.47	2.64
15	$0.35604 \cdot 10^{9}$	0.07	1.36	0.70	1.06	2.66	1.20	0.00	4.71	5.16
20	$0.26681 \cdot 10^{9}$	2.00	1.87	0.94	0.48	3.62	1.61	0.11	7.99	8.86
25	$0.21450 \cdot 10^{9}$	0.78	2.41	1.20	0.23	4.59	1.98	0.01	12.49	14.01

Comparison Results

General k-means, Modified General k-means, Hyperbolic Smoothing

Pendigit

		GKM			MGKM			SMOOTH		
q	$\mathbf{f_{opt}}$	**E**	**α**	**t**	**E**	**α**	**t**	**E**	**α**	**t**
2	$0.12812 \cdot 10^{9}$	0.39	0.12	2.56	0.00	0.24	5.25	0.00	0.20	5.37
5	$0.75304 \cdot 10^{8}$	0.00	0.49	9.73	0.00	0.97	18.72	0.00	0.83	22.60
10	$0.49302 \cdot 10^{8}$	0.00	1.10	20.45	0.00	2.18	39.31	0.00	2.21	61.71
15	$0.39067 \cdot 10^{8}$	0.00	1.71	30.79	0.00	3.40	59.19	0.00	3.99	120.46
20	$0.34123 \cdot 10^{8}$	0.00	2.33	41.25	0.17	4.62	78.37	0.16	7.12	250.80
25	$0.30038 \cdot 10^{8}$	0.24	2.95	51.87	0.24	5.85	98.47	0.00	10.75	446.37

Comparison Results

General k-means, Modified General, Hyperbolic Smoothing

D15112

		GKM			MGKM			SMOOTH		
q	f_{opt}	E	α	t	E	α	t	E	α	t
2	$0.36840 \cdot 10^{12}$	0.00	0.22	2.75	0.00	0.46	4.51	0.00	0.42	5.15
5	$0.13271 \cdot 10^{12}$	0.00	0.92	8.50	0.00	1.83	13.85	0.00	1.52	18.41
10	$0.64892 \cdot 10^{11}$	0.78	2.07	14.87	0.78	4.12	24.98	0.00	3.44	41.50
15	$0.43136 \cdot 10^{11}$	0.26	3.24	21.00	0.26	6.44	35.51	0.00	5.74	68.23
20	$0.32177 \cdot 10^{11}$	0.25	4.43	26.99	0.25	8.77	45.68	0.00	8.65	101.82
25	$0.25423 \cdot 10^{11}$	0.03	5.62	32.71	0.03	11.11	55.47	0.00	11.96	139.53

Comparison Results

General k-means, Modified General k-means, Hyperbolic Smoothing

Pla85900

		GKM			MGKM			SMOOTH		
q	**f_{opt}**	**E**	**α**	**t**	**E**	**α**	**t**	**E**	**α**	**t**
2	$0.37491 \cdot 10^{16}$	0.00	0.07	89.54	0.00	0.15	115.99	0.00	0.12	123.96
5	$0.13397 \cdot 10^{16}$	0.00	0.30	408.60	0.00	0.59	512.31	0.00	0.37	452.96
10	$0.68294 \cdot 10^{15}$	0.00	0.67	754.67	0.00	1.33	994.29	0.00	0.79	1011.17
15	$0.46029 \cdot 10^{15}$	0.51	1.04	1083.04	0.51	2.07	1448.94	0.00	1.24	1596.67
20	$0.35087 \cdot 10^{15}$	0.01	1.42	1355.85	0.00	2.82	1844.34	0.00	1.71	2210.91
25	$0.28323 \cdot 10^{15}$	0.74	1.80	1570.37	0.73	3.57	2186.20	0.00	2.21	2863.18

Very Large Clustering Instances

Hyperbolic Smoothing and Partition into Boundary and Gravitational Regions
Applied to
Minimum Sum-of-Squares Clustering Problem

Synthetic Problems – 03/2013

Synthetic Problem
5.000.000 Observations
Euclidian Space with n = 4 dimensions

q	$f_{AHSC\text{-}L2}$	Occur.	E_{Mean}	T_{Mean}
2	0.366820E7	4	1.48	10.18
3	0.283568E7	1	1.61	14.32
4	0.233063E7	1	0.18	18.68
5	0.184130E7	2	0.13	18.07
6	0.158542E7	4	0.18	22.24
7	0.132902E7	1	0.23	27.04
8	0.107989E7	1	0.23	27.40
9	0.835887E6	2	0.24	27.21
10	0.597059E6	10	0.00	24.31
c	0.600000E6	-	-	-

Synthetic Problem 5.000.000 Observations Euclidian Space with n = 5 dimensions

q	$f_{AHSC-L2}$	Occur.	E_{Mean}	T_{Mean}
2	0.381813E7	2	1.15	11.61
3	0.298569E7	1	2.64	16.51
4	0.248090E7	1	1.56	21.53
5	0.199132E7	1	0.12	20.62
6	0.173498E7	1	0.06	26.48
7	0.147899E7	2	0.17	31.93
8	0.122977E7	2	0.24	33.12
9	0.985936E6	1	0.21	33.17
10	0.747074E6	10	0.00	28.25
c	0.750000E6	-	-	-

Synthetic Problem
5.000.000 Observations
Euclidian Space with n = 6 dimensions

q	$f_{AHSC\text{-}L2}$	Occur.	E_{Mean}	T_{Mean}
2	0.396811E7	2	0.80	12.95
3	0.313868E7	3	1.12	17.69
4	0.263094E7	2	0.82	23.03
5	0.214135E7	1	0.09	22.29
6	0.188503E7	1	0.14	27.34
7	0.162983E7	2	0.18	33.93
8	0.137984E7	2	0.11	36.84
9	0.113584E7	2	0.17	34.33
10	0.897046E6	10	0.00	27.90
c	0.900000E6	-	-	-

Synthetic Problem
5.000.000 Observations
Euclidian Space with n = 8 dimensions

q	$f_{AHSC\text{-}L2}$	Occur.	E_{Mean}	T_{Mean}
2	0.426813E7	1	1.33	12.65
3	0.343566E7	2	1.11	19.52
4	0.293101E7	1	1.41	23.01
5	0.244127E7	1	0.05	24.96
6	0.218579E7	3	0.11	30.39
7	0.192898E7	1	0.14	37.74
8	0.168137E7	1	0.03	39.85
9	0.143580E7	1	0.17	35.50
10	0.119702E7	10	0.00	30.29
c	0.120000E7	-	-	-

Synthetic Problem
5.000.000 Observations
Euclidian Space with n = 10 dimensions

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

Medium and Large Clustering Instances

Hyperbolic Smoothing and Partition into Boundary and Gravitational Regions
Applied to
Minimum Sum-of-Squares Clustering Problem

Comparison with Bagirov et al

January April, 2014

Medium and Large Clustering Instances

Solving Problems of the Literature by the Accelerated Hyperbolic Smoothing Clustering Method

Vinicius Layter Xavier & Adilson Elias Xavier
Federal University of Rio de Janeiro – Brazil

EUROPT 2014 - Perpignan

Medium Size Instances
Starting Point - Loose Way

Breast Cancer (m=683 , n=9)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	1,93E 04	0	0	0	0,02	0	0,05	0,00	0,02
5	1,37E 04	2,28	0,03	1,86	0,05	0	0,61	0,55	0,04
10	1,02E 04	0,26	0,06	0,28	0,11	0	2,14	0,09	0,13
15	8,69E 03	1,02	0,08	1,07	0,16	0	9,08	0,12	0,30
20	7,65E 03	3,64	0,11	1,8	0,2	0	27,36	0,21	0,51
25	6,90E 03	5,08	0,14	0,93	0,27	0	44,65	0,65	1,00

Medium Size Instances
Starting Point - Loose Way

TSPLIB1060 (m=1060 , n=2)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	**9,83E 09**	**0**	**0**	**0**	**0,02**	**0**	**0,05**	**0,31**	**0,01**
5	**3,79E 09**	**0,01**	**0,03**	**0,01**	**0,05**	**0**	**0,17**	**0,56**	**0,02**
10	**1,75E 09**	**0,23**	**0,06**	**0,05**	**0,09**	**0**	**0,7**	**1,08**	**0,05**
15	**1,12E 09**	**0,09**	**0,09**	**0,06**	**0,16**	**0,02**	**1,25**	**1,21**	**0,09**
20	**7,92E 08**	**0,41**	**0,12**	**0,41**	**0,2**	**0,03**	**2,7**	**2,60**	**0,14**
25	**6,07E 08**	**1,8**	**0,16**	**1,8**	**0,25**	**0,02**	**4,63**	**1,92**	**0,21**

Medium Size Instances Starting Point - Loose Way

TSPLIB3038 (m=3038 , n=2)

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	3,17E 09	0	0,06	0	0,11	0	0,25	0,06	0,03
5	1,20E 09	0	0,25	0	0,39	0	0,98	0,32	0,11
10	5,60E 08	2,78	0,48	0,58	0,81	0	2,64	2,76	0,33
15	3,56E 08	0,07	0,7	1,06	1,2	0	5,16	0,03	0,66
20	2,67E 08	2	0,94	0,48	1,61	0,11	8,86	0,22	1,30
25	2,15E 08	0,78	1,2	0,23	1,98	0,01	14,01	0,55	1,54

Medium Size Instances Starting Point - Loose Way

Pendigit (m=10992 , n=16)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	1,28E08	0,39	2,56	0	5,23	0	5,37	0,00	0,37
5	7,53E07	0	9,73	0	18,72	0	22,6	0,00	1,52
10	4,93E07	0	20,45	0	39,31	0	61,71	2,40	3,19
15	3,91E07	0	30,79	0	59,19	0	120,46	1,44	6,02
20	3,41E07	0	41,25	0,17	78,37	0,16	250,8	-0,02	9,75
25	3,00E07	0,24	51,87	0,24	98,47	0	446,37	1,59	12,61

Large Size Instances Starting Point - Loose Way

D15112 **(m=15112 , n=2)**

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	3,68E11	0	2,75	0	4,51	0	5,15	0,00	0,38
5	1,33E11	0	8,5	0	13,85	0	18,41	0,00	1,24
10	6,49E10	0,78	14,87	0,78	24,98	0	41,5	0,00	2,48
15	4,31E10	0,26	21	0,26	35,51	0	68,23	0,00	4,09
20	3,22E10	0,25	26,99	0,25	45,68	0	101,82	0,30	6,41
25	2,54E10	0,03	32,71	0,03	55,47	0	139,53	-0,48	9,07

Large Size Instances Starting Point - Loose Way

Letters (m=20000 , n=16)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	**1,38E 06**	**0**	**9,63**	**0**	**17,35**	**0**	**14,4**	**0,00**	**0,49**
5	**1,08E 06**	**1,94**	**38,28**	**0,87**	**61,78**	**0**	**63,1**	**0,91**	**1,95**
10	**8,58E 05**	**0**	**79,76**	**0**	**131,7 7**	**0,21**	**163,7 1**	**0,19**	**5,45**
15	**7,45E 05**	**0,48**	**116,7**	**0**	**200,0 7**	**0,4**	**312,2 5**	**0,66**	**10,28**
20	**6,74E 05**	**0,53**	**153,2 1**	**0,34**	**265,9 7**	**0**	**509,4 5**	**0,05**	**14,22**
25	**6,23E 05**	**1,47**	**187,7 9**	**0,58**	**331,0 7**	**0**	**749,9 6**	**0,04**	**20,25**

Large Size Instances Starting Point - Loose Way

Shuttle **(m=58000 , n=9)**

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	2,13E09	0	63,2	0	123,82	0	86,74	0,00	0,14
5	7,24E08	0	251,65	0	521,81	0	351,5	0,00	0,72
10	2,83E08	0,02	581,28	0	1207,2	0	813,73	0,00	2,86
15	1,53E08	0	888,24	0	1776,88	0	1272,75	0,00	5,85
20	1,06E08	0	1195,84	0	2338,99	0	1798,88	-3,55	12,38
25	7,98E07	0,17	1512,55	0,17	2917,41	0	2396,82	-3,13	26,05

Medium Size Instances
Starting Point - Loose Way

Pla85900 (m=85900 , n=2)

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	**3,75E15**	**0**	**89,54**	**0**	**115,99**	**0**	**123,96**	**0,00**	**1,01**
5	**1,34E15**	**0**	**408,6**	**0**	**512,31**	**0**	**452,96**	**0,00**	**3,48**
10	**6,83E14**	**0**	**754,67**	**0**	**994,29**	**0**	**1011,17**	**0,00**	**9,49**
15	**4,60E14**	**0,51**	**1083,04**	**0,51**	**1448,94**	**0**	**1596,67**	**0,23**	**16,94**
20	**3,51E14**	**0,01**	**1355,85**	**0**	**1844,34**	**0**	**2210,91**	**-0,28**	**27,51**
25	**2,83E14**	**0,74**	**1570,37**	**0,73**	**2186,2**	**0**	**2863,18**	**0,27**	**39,30**

Medium Size Instances Starting Point - Tight Way

Breast Cancer (m=683 , n=9)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	1,93E08	0	0	0	0.02	0	0.05	0.00	0.04
5	1,37E08	2.28	0.03	1.86	0.05	0	0.61	0.00	0.14
10	1,02E08	0.26	0.06	0.28	0.11	0	2.14	0.16	0.37
15	8,69E07	1.02	0.08	1.07	0.16	0	9.08	-0.36	0.62
20	7,65E07	3.64	0.11	1.80	0.20	0	27.36	-0.19	1.12
25	6,90E07	5.08	0.14	0.93	0.27	0	44.65	0.11	1.53

Medium Size Instances
Starting Point - Tight Way

TSPLIB1060 **(m=1060 , n=2)**

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	**9,83E13**	**0**	**0**	**0**	**0.02**	**0**	**0.05**	**0.31**	**0.04**
5	**3,79E13**	**0.01**	**0.03**	**0.01**	**0.05**	**0**	**0.17**	**0.54**	**0.13**
10	**1,75E13**	**0.23**	**0.06**	**0.05**	**0.09**	**0**	**0.70**	**0.42**	**0.39**
15	**1,12E13**	**0.09**	**0.09**	**0.06**	**0.16**	**0.02**	**1.25**	**1.21**	**0.87**
20	**7,92E12**	**0.41**	**0.12**	**0.41**	**0.20**	**0.03**	**2.70**	**1.35**	**1.55**
25	**6,07E12**	**1.80**	**0.16**	**1.80**	**0.25**	**0.02**	**4.63**	**1.81**	**2.03**

Medium Size Instances Starting Point - Tight Way

TSPLIB3038 (m=3038 , n=2)

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	3,17E13	0	0.06	0	0.11	0	0.25	0.06	0.11
5	1,20E13	0	0.25	0	0.39	0	0.98	0.00	0.21
10	5,60E12	2.78	0.48	0.58	0.81	0	2.64	0.56	0.69
15	3,56E12	0.07	0.70	1.06	1.20	0	5.16	0.03	1.40
20	2,67E12	2.00	0.94	0.48	1.61	0.11	8.86	0.22	2.39
25	2,15E12	0.78	1.20	0.23	1.98	0.01	14.01	0.18	3.84

Medium Size Instances Starting Point - Tight Way

Pendigit (m=10992 , n=16)

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	1,28E12	0.39	2.56	0	5.23	0	5.37	0.00	0.38
5	7,53E11	0	9.73	0	18.72	0	22.60	0.00	1.76
10	4,93E11	0	20.45	0	39.31	0	61.71	0.00	4.99
15	3,91E11	0	30.79	0	59.19	0	120.46	0.00	10.33
20	3,41E11	0	41.25	0.17	78.37	0.16	250.80	-0.31	19.21
25	3,00E11	0.24	51.87	0.24	98.47	0	446.37	-0.12	33.06

Large Size Instances
Starting Point - Tight Way

D15112 (m=15112 , n=2)

		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	*E*	*t*	*E*	*t*	*E*	*t*	*E*	*t*
2	3,68E15	0	2.75	0	4.51	0	5.15	0.00	0.31
5	1,33E15	0	8.50	0	13.85	0	18.41	0.00	1.36
10	6,49E14	0.78	14.87	0.78	24.98	0	41.50	-0.62	4.21
15	4,31E14	0.26	21.00	0.26	35.51	0	68.23	0.15	9.31
20	3,22E14	0.25	26.99	0.25	45.68	0	101.82	0.62	17.43
25	2,54E14	0.03	32.71	0.03	55.47	0	139.53	-0.49	30.22

Large Size Instances
Starting Point - Tight Way

Letters (m=20000 , n=16)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	1,38E10	0	9.63	0	17.35	0	14.40	0.00	0.43
5	1,08E10	1.94	38.28	0.87	61.78	0	63.10	0.87	2.03
10	8,58E09	0	79.76	0	131.77	0.21	163.71	0.00	7.61
15	7,45E09	0.48	116.70	0	200.07	0.40	312.25	-0.08	19.89
20	6,74E09	0.53	153.21	0.34	265.97	0	509.45	0.30	26.06
25	6,23E09	1.47	187.79	0.58	331.07	0	749.96	-0.53	53.17

Large Size Instances
Starting Point - Tight Way

Shuttle (m=58000 , n=9)

		GKM		MGKM		HSCM		AHSCM	
k	***F_{opt}***	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	**2,13E13**	**0**	**63.20**	**0**	**123.82**	**0**	**86.74**	**0.00**	**0.20**
5	**7,24E12**	**0**	**251.65**	**0**	**521.81**	**0**	**351.50**	**0.00**	**1.35**
10	**2,83E12**	**0.02**	**581.28**	**0**	**1207.20**	**0**	**813.73**	**0.00**	**4.87**
15	**1,53E12**	**0**	**888.24**	**0**	**1776.88**	**0**	**1272.75**	**0.00**	**11.09**
20	**1,06E12**	**0**	**1195.84**	**0**	**2338.99**	**0**	**1798.88**	**-3.60**	**26.80**
25	**7,98E11**	**0.17**	**1512.55**	**0.17**	**2917.41**	**0**	**2396.82**	**-3.13**	**51.97**

Large Size Instances
Starting Point - Tight Way

Pla85900 (m=85900 , n=2)

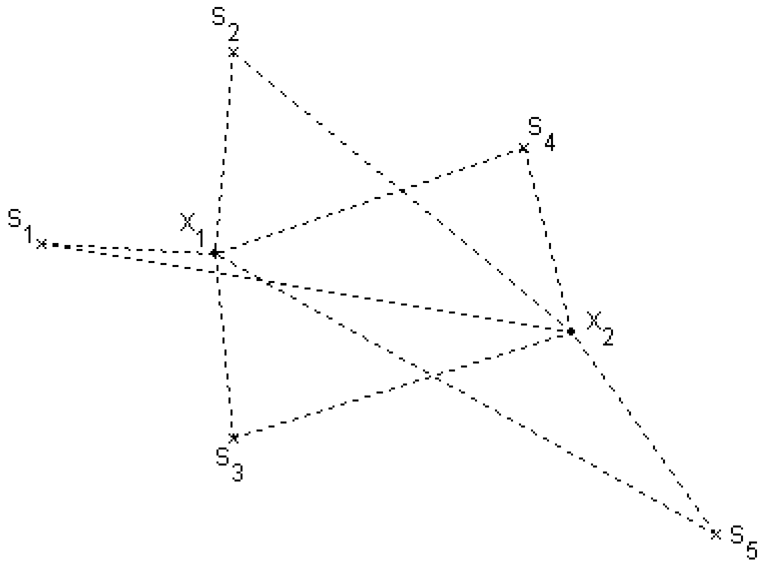
		GKM		MGKM		HSCM		AHSCM	
k	F_{opt}	***E***	***t***	***E***	***t***	***E***	***t***	***E***	***t***
2	**3,75E19**	**0**	**89.54**	**0**	**115.99**	**0**	**123.96**	**0.00**	**1.17**
5	**1,34E19**	**0**	**408.60**	**0**	**512.31**	**0**	**452.96**	**0.00**	**5.04**
10	**6,83E18**	**0**	**754.67**	**0**	**994.29**	**0**	**1011.17**	**0.00**	**15.66**
15	**4,60E18**	**0.51**	**1083.04**	**0.51**	**1448.94**	**0**	**1596.67**	**0.00**	**38.01**
20	**3,51E18**	**0.01**	**1355.85**	**0**	**1844.34**	**0**	**2210.91**	**0.00**	**73.09**
25	**2,83E18**	**0.74**	**1570.37**	**0.73**	**2186.20**	**0**	**2863.18**	**0.13**	**135.62**

Conclusions

Performance of the HSCM

The performance of HSCM can be attributed to the complete

differentiability of the approach. So, each centroid sees permanently every observation point, conversely, each observation point can permanently see every centroid and attract it



Performance of the AHSCM Algorithm

- The robustness performance of the AHSCM Algorithm can be attributed to the complete differentiability of the approach.

- The speed performance of the AHSCM Algorithm can be attributed to the partition of the set of observations in two non overlapping parts.This last approach engenders a drastic simplification of computational tasks.

Conclusions

In view of the preliminary results obtained, where the proposed methodology performed efficiently and robustly, these algorithms can represent a possible approach for dealing with the solving of clustering of problems of real applications.

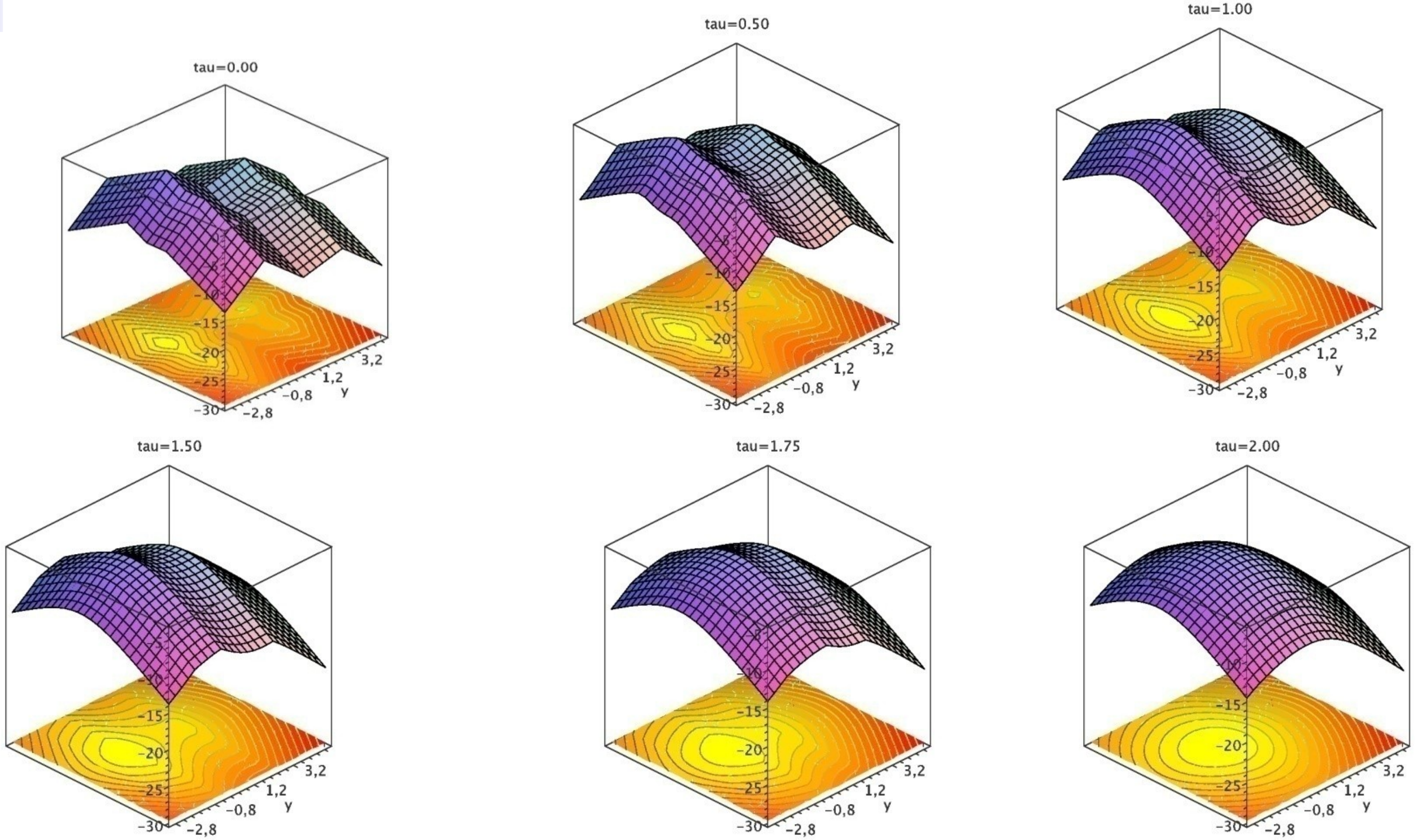
Conclusions

The last, but not the least:

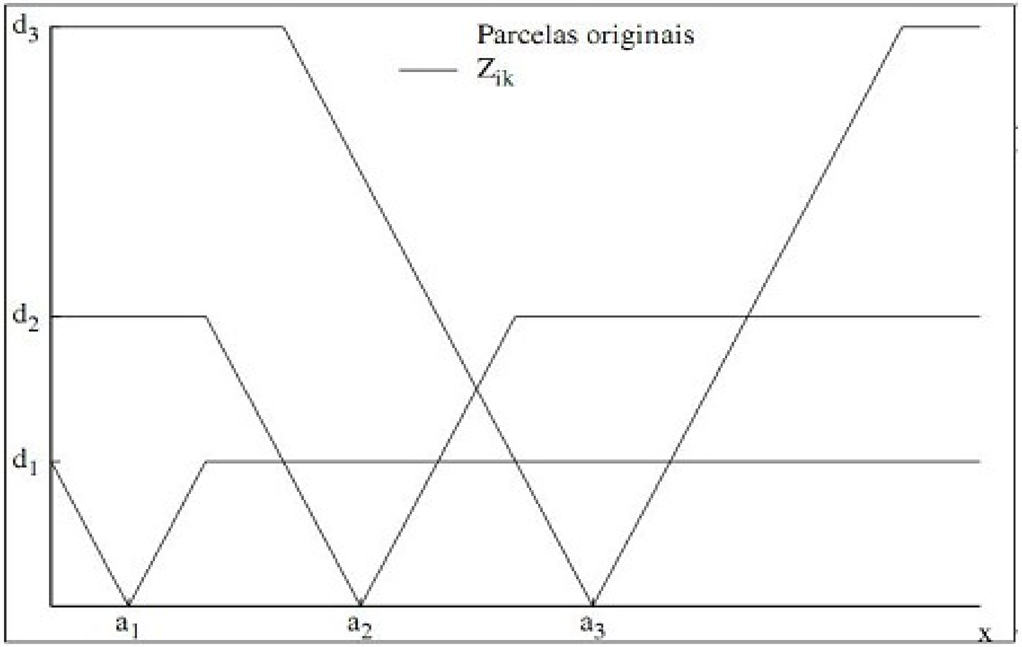
The hyperbolic smoothing approach eliminates a lot of local minima points, remaining only the deepest ones.

END

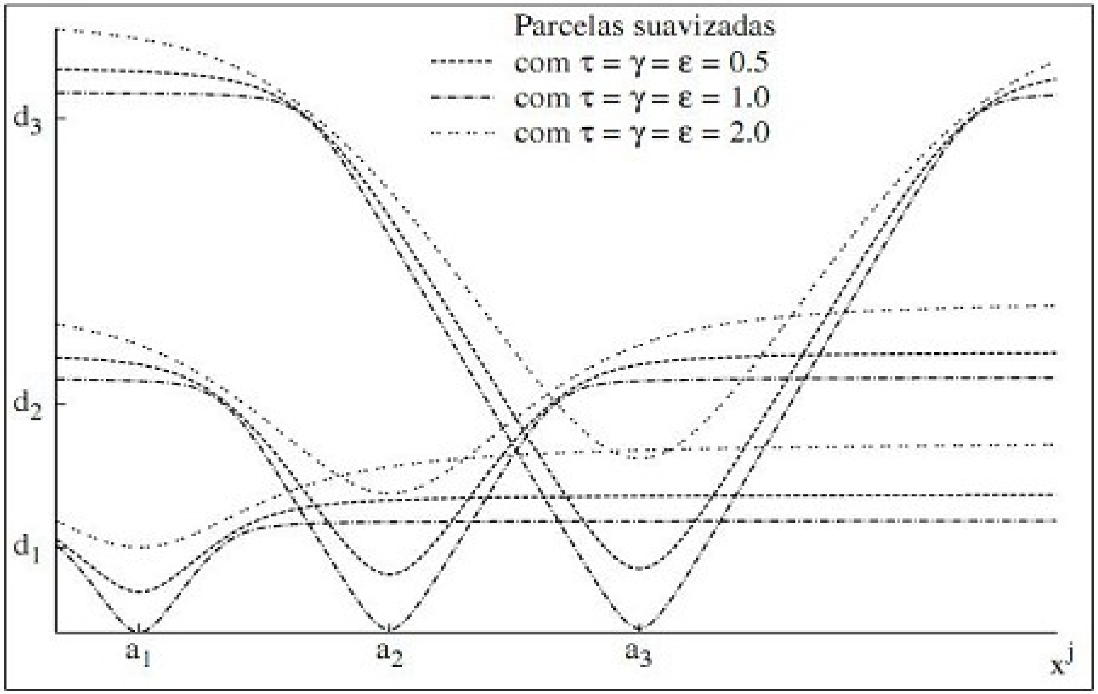
Convexification effect by the Hyperbolic Smoothing approach
Geometric distance problem



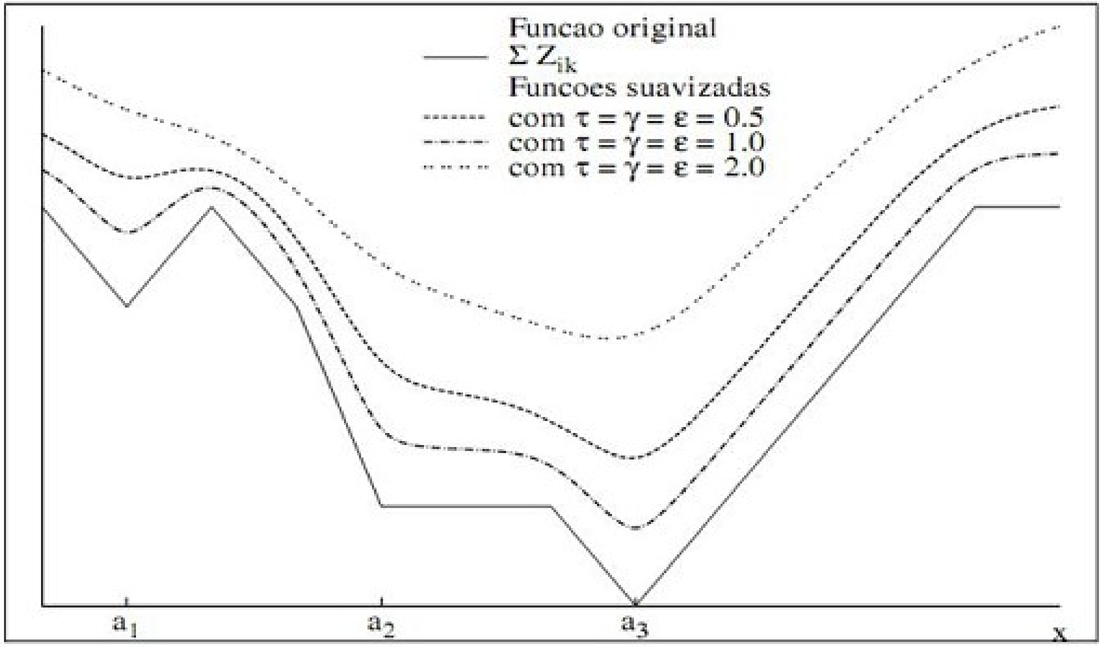
Smoothing of the objective function terms of a specific classification problem



Smoothing of the objective function terms of a specific classification problem



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