UNIVERSIDAD NACIONAL DE CÓRDOBA

FACULTAD DE MATEMÁTICA, ASTRONOMÍA Y FÍSICA

SERIE "A"

TRABAJOS DE MATEMÁTICA

Nº 95/09

A projected Weiszfeld's algorithm for the box-constrained Weber location problem

Elvio A. Pilotta - Germán A. Torres

Editores: Jorge R. Lauret – Jorge G. Adrover

CIUDAD UNIVERSITARIA – 5000 CÓRDOBA
REPÚBLICA ARGENTINA

A projected Weiszfeld's algorithm for the box-constrained Weber location problem

Elvio A. Pilotta *   Germán A. Torres †

July 30, 2009

Abstract

The Fermat-Weber problem consists in finding a point in $\mathbb{R}^n$ that minimizes the weighted sum of distances from m points in $\mathbb{R}^n$ that are not collinear. An application that motivated this problem is the optimal location of industries in the 2-dimensional case. The Weber problem is a generalization of the well-known Fermat problem. An usual method to solve the Weber problem, proposed by Weiszfeld in 1937, is based in a fixed-point iteration. In recent years there has been a growing interest in formalizing properties of well-definition of the method and proving convergence results.

In this work we generalize the Weber location problem considering box-constrained restrictions. We propose a fixed-point iteration with projections on the restrictions and demonstrate descending properties. It is also proved that the limit of the sequence generated by the method is a feasible point and satisfies the KKT optimality conditions. Numerical experiments are presented to illustrate the theoretical results.

Keywords: Weber problem - box constraints - fixed-point iteration

1 Introduction.

Let $a_1, ..., a_m$ be m distinct points in the space $\mathbb{R}^n$, called vertices, and positive numbers $w_1, ..., w_m$, called weights. The Fermat-Weber problem is to find a point in $\mathbb{R}^n$ that minimizes the weighted sum of Euclidean distances from the m given points, that is, we have to find:

$$\begin{aligned} \text{argmin} \quad & f(x) && (1)\\ \text{s.t.} \quad & x \in \mathbb{R}^n, && (2) \end{aligned}$$

*Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, CIEM (CONICET), Medina Allende s/n, Ciudad Universitaria (5000), Córdoba, Argentina (e-mail: pilotta@famaf.unc.edu.ar)

†Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, CIEM (CONICET), Medina Allende s/n, Ciudad Universitaria (5000), Córdoba, Argentina (e-mail: torres@famaf.unc.edu.ar)

Work sponsored by CONICET and SECYT-UNC

where f is called the Weber function and it is defined by:

$$f(x) = \sum_{j=1}^{m} w_j \|x - a_j\|, \qquad w_j > 0, \qquad j = 1, ..., m. \tag{3}$$

It is a well-known fact that this function is strictly convex if the vertices are not collinear (we will assume this hypothesis from now on).

Pierre de Fermat (1601 – 1665) (see [1]) formulated the Fermat problem: "Given three points in a plane, find a fourth point such that the sum of its distances to the three given points is a minimum!". Several solutions, based on geometrical arguments, were proposed by E. Torricelli (1608 – 1647) and T. Simpson (1710 – 1761). In [2] historical details and geometric aspects are offered.

In [3] A. Weber formulated the problem (1)-(2) from an economical point of view. The vertices represent customers or demands, the solution to the problem denotes the location of a new facility, and the weights are costs between the new facility and the customers.

Among several schemes to solve the Weber location problem (see [6, 7, 8, 9]), one of the most popular methods was presented by E. V. Weiszfeld in [4]. The Weiszfeld algorithm is an iterative method based on the first-order necessary conditions for a stationary point of the objective function:

$$x^{(k+1)} = T(x^{(k)}), \qquad k \in \mathbb{N} \cup \{0\}, \tag{4}$$

where

$$T(x) = \begin{cases} \dfrac{\displaystyle\sum_{j=1}^{m} \frac{w_j a_j}{\|x - a_j\|}}{\displaystyle\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}}, & x \neq a_1, ..., a_m, \\ \\ a_i, & x = a_i \text{ for some } i = 1, ..., m. \end{cases} \tag{5}$$

This iteration function is derived from nullifying the gradient of the Weber function (3):

$$\nabla f(x) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} (x - a_j), \qquad x \neq a_1, ..., a_m. \tag{6}$$

Despite of the simplicity of the algorithm (4), the proof of convergence of the algorithm is a difficult issue. H. Kuhn showed a counterexample (see [5]) where he pointed out an error in a Weiszfeld convergence statement, because it is possible that, for some k, the point $x^{(k)}$ is a vertex. Besides that, Kuhn proved that the algorithm (4) converges to the unique optimum for all but a denumerable number of starting points $x^{(0)}$. However, Chandrasekaran and Tamir (see [10]) detected a flaw in the Kuhn's statement and showed that the system $T(x) = a_i$ may have a continuum set of solutions even when the vertices a_1, ..., a_m are not collinear. They conjectured that if the convex hull of the set of vertices is of full dimension, then the set of initial points for which the sequence generated by the Weiszfeld algorithm yields in a vertex is denumerable. In [11], J. Brimberg proved the conjecture of Chandrasekaran and Tamir, but in [12] Cánovas et al. showed counterexamples to the proofs in Brimberg's paper. Finally, in [13], R. Cañavate claims that the Weiszfeld algorithm converges for all points but a set of measure zero.

Generalizations and new thecniques for the Fermat-Weber location problem have been developed in recent years. In [14] the authors study the so called Regional Weber Problem, which allows the demand not to be concentrated onto a finite set of points, but follows an arbitrary probability measure. Hamacher and Klamroth (see [15]) consider distances defined by block norms. A modified Weiszfeld algorithm that is monotonically convergent is presented in [16]. S.-D. Lee (see [17]) generalizes conventional P-median location problems by considering the unreliability of facilities. In [18] the authors consider the problem of finding the location of a single facility in a region divided by a straight line, where the distances are measured differently, on different sides of this boundary line. Pfeiffer and Klamroth (see [19]) exploit similarities of continuous location problems and network location problems. In [20] barriers are defined and the solution of locating facilities in the plane in the presence of barriers is generalized. Location of facilities in planar networks are investigated in [21], and in [22] heuristics is used to generate initial solutions for the capacitated multi-source Weber problem.

In this paper we propose an algorithm in order to find a point x^* that is solution of the following problem:

$$\begin{aligned} \text{argmin} \quad & f(x) && (7) \\ \text{s.t.} \quad & x \in [l, u], && (8) \end{aligned}$$

where $l = (l_1, ..., l_n)$, $u = (u_1, ..., u_n)$ and $[l, u] = \{x \in \mathbb{R}^n : l_i \leq x_i \leq u_i, i = 1, ..., n\}$. In practical terms, we are interested in locating a facility minimizing the sum of the Euclidean distances but restricting the location to a previously determined area, that in our case is a box in $\mathbb{R}^n$. Since the function f is strictly convex and the set $[l, u]$ is convex, the problem (7)-(8) has a unique solution. Besides that, the solution x^* holds the KKT conditions, that is, x^* is solution if and only if x^* is a KKT point.

The paper is structured as follows: in Section 2 we define the proposed algorithm, as well as auxiliary functions and multipliers. In Section 3 we prove propositions that are needed in Section 4, which is devoted to present the main results about convergence and minimization of the algorithm. Section 5 refers to the numerical experiments to see the agreement between experiments and theory. Finally, Section 6 is dedicated to the conclusions.

2 Some definitions.

This section is devoted to define entities that will be necessary for the convergence proof of the proposed algorithm.

Let us define a function $P : \mathbb{R}^n \to \mathbb{R}^n$ defined as:

$$(P(x))_i = \begin{cases} l_i & \text{if} \quad x_i < l_i, \\ x_i & \text{if} \quad x_i \in [l_i, u_i], \\ u_i & \text{if} \quad x_i > u_i, \end{cases} \qquad i = 1, ..., n \tag{9}$$

The efect of the function P is to map any point in the space $\mathbb{R}^n$ to the box $[l,u]$. We then propose a modified Weiszfeld algorithm in which we come back to the box in each step:

$$x^{(k+1)} = \tilde{T}(x^{(k)}), \qquad k = 0, 1, ..., \qquad \tilde{T} = P \circ T. \tag{10}$$

We expect that the limit of the sequence (10) be the solution of the problem (7)-(8) and we will prove it throughout this paper. From now on we will assume that each $x^{(k)}$ defined in (10) is different from the vertices a_1, ..., a_m.

Remark 2.1 *The functional $\tilde{T}$ is continuous in $\mathbb{R}^n - \{a_1, ..., a_m\}$ and the sequence $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ is contained in the box $[l,u]$.*

Proof. Since functions P and T are continuous in $\mathbb{R}^n - \{a_1, ..., a_m\}$, the functional $\tilde{T}$ is continuous. Besides that, if x is a vector in $\mathbb{R}^n$ then $P(x) \in [l,u]$. This proves that the sequence $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ is contained in $[l,u]$.

■

We define the following multipliers associated to a point x and to the problem (7)-(8):

Definition 2.2 *Let x be a point in $\mathbb{R}^n$ different from the vertices a_1, a_2, ..., a_m. We define:*

$$\alpha_k(x) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[l_k - (a_j)_k\right], \qquad k = 1, ..., n, \tag{11}$$

$$\beta_k(x) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[(a_j)_k - u_k\right], \qquad k = 1, ..., n. \tag{12}$$

These coefficients $\alpha_k(x)$, $\beta_k(x)$, $k = 1, ..., n$, assume a well determined sign according to $T(x)$:

Remark 2.3 *Let x be a point in $\mathbb{R}^n - \{a_1, ..., a_m\}$. Then for each $k = 1, ..., n$:*

(a) $(T(x))_k < l_k$ *if and only if* $\alpha_k > 0$.

(b) $(T(x))_k > u_k$ *if and only if* $\beta_k > 0$.

(c) $(T(x))_k \in [l_k, u_k]$ *if and only if* $\alpha_k \le 0$ *and* $\beta_k \le 0$.

Proof. The listed properties are deduced rewriting the inequalities. Therefore, for the (a) case we have:

$$(T(x))_k < l_k \iff \frac{\displaystyle\sum_{j=1}^{m} \frac{w_j (a_j)_k}{\|x - a_j\|}}{\displaystyle\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} < l_k \iff \sum_{j=1}^{m} \frac{w_j (a_j)_k}{\|x - a_j\|} < \sum_{j=1}^{m} \frac{w_j l_k}{\|x - a_j\|} \iff$$

$$\iff \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[l_k - (a_j)_k\right] > 0 \iff \alpha_k(x) > 0.$$

The cases (b) and (c) are similar to (a) and are left to the reader.

■

It is also necessary to identify the indices associated to the inequalities of the previous remark. Then, we define the following sets:

Definition 2.4 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$. We define the following set of indices:*

$$\begin{aligned} L(x) &= \{k \in \mathbb{N} : 1 \leq k \leq n,\ (T(x))_k < l_k\}, & (13)\\ U(x) &= \{k \in \mathbb{N} : 1 \leq k \leq n,\ (T(x))_k > u_k\}. & (14)\\ I(x) &= \{k \in \mathbb{N} : 1 \leq k \leq n,\ (T(x))_k \in [l_k, u_k]\}, & (15)\end{aligned}$$

Notice that:

- $k \in L(x)$ if and only if $\alpha_k(x) > 0$,
- $k \in U(x)$ if and only if $\beta_k(x) > 0$,
- $k \in I(x)$ if and only if $\alpha_k(x) \leq 0$ and $\beta_k(x) \leq 0$.

We will define auxiliary functions that take into account the projection (9) in order to generalize the Weiszfeld algorithm.

Definition 2.5 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$. We define:*

(a) $E_x : \mathbb{R}^n \to \mathbb{R}^n$, *where:*

$$E_x(y) = \begin{cases} l_k, & k \in L(x), \\ y_k, & k \in I(x), \\ u_k, & k \in U(x), \end{cases} \qquad k = 1, ..., n. \tag{16}$$

(b) $g_x : \mathbb{R}^n - \{a_1, ..., a_m\} \to \mathbb{R}$, *where:*

$$g_x(y) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \|E_x(y) - a_j\|^2. \tag{17}$$

(c) If $I(x) = \{i_1, ..., i_r\} \neq \emptyset$, *define* $I_x : \mathbb{R}^r \to \mathbb{R}^n$, *where:*

$$(I_x(z))_k = \begin{cases} l_k, & k \in L(x), \\ z_s, & k = i_s \in I(x), \\ u_k, & k \in U(x), \end{cases} \qquad k = 1, ..., n. \tag{18}$$

(d) If $I(x) = \{i_1, ..., i_r\} \neq \emptyset$, *define* $P_x : \mathbb{R}^n \to \mathbb{R}^r$, *where:*

$$(P_x(y))_k = y_{i_k}, \qquad k = 1, ..., r. \tag{19}$$

A useful property of the functions defined above is pointed out in the following remark.

Remark 2.6 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $I(x) = \{i_1, ..., i_r\} \neq \emptyset$ and $r = \#I(x)$. Then $E_x \circ I_x = I_x$.*

Proof. Let z and k be such that $z \in \mathbb{R}^n$ and $1 \leq k \leq n$. If $k \in I(x)$ then $(E_x \circ I_x(z))_k = (I_x(z))_k$ according to (16). If $k \in L(x)$, then $(E_x \circ I_x(z))_k = l_k = (I_x(z))_k$ due to (16) and (18). If $k \in U(x)$, then $(E_x \circ I_x(z))_k = u_k = (I_x(z))_k$ for the same reason.

■

3 Properties.

This section is dedicated to prove auxiliary results involving the functions defined in the previous section. These results will be helpful to demonstrate the main theorems in the next section.

First of all we can realise that the function g_x is related to the Weber function f through a formula that will allow us to deduce an inequality.

Proposition 3.1 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$. Then,*

$$g_x(x) = f(x) + 2 \sum_{i \in L(x)} (l_i - x_i)\, \alpha_i(x) - 2 \sum_{i \in U(x)} (u_i - x_i)\, \beta_i(x) + \gamma \tag{20}$$

where

$$\gamma = - \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[\sum_{i \in L(x)} (l_i - x_i)^2 + \sum_{i \in U(x)} (u_i - x_i)^2 \right] \leq 0. \tag{21}$$

Proof. Using algebraic steps we treat some parts of the g_x formula. In this way, if $i \in L(x)$ then:

$$\begin{aligned}
\left[l_i - (a_j)_i\right]^2 &= \left(l_i - x_i + x_i - (a_j)_i\right)^2 = \\
&= \left(x_i - (a_j)_i\right)^2 + (l_i - x_i)^2 + 2\,(l_i - x_i)\left(x_i - (a_j)_i\right) = \\
&= \left(x_i - (a_j)_i\right)^2 + (l_i - x_i)^2 + 2\,(l_i - x_i)\left(x_i - l_i + l_i - (a_j)_i\right) = \\
&= \left(x_i - (a_j)_i\right)^2 - (l_i - x_i)^2 + 2\,(l_i - x_i)\left(l_i - (a_j)_i\right),
\end{aligned} \tag{22}$$

and if $i \in U(x)$ then:

$$\begin{aligned}
\left[u_i - (a_j)_i\right]^2 &= \left(u_i - x_i + x_i - (a_j)_i\right)^2 = \\
&= \left(x_i - (a_j)_i\right)^2 + (u_i - x_i)^2 + 2\,(u_i - x_i)\left(x_i - (a_j)_i\right) = \\
&= \left(x_i - (a_j)_i\right)^2 + (u_i - x_i)^2 + 2\,(u_i - x_i)\left(x_i - u_i + u_i - (a_j)_i\right) = \\
&= \left(x_i - (a_j)_i\right)^2 - (u_i - x_i)^2 - 2\,(u_i - x_i)\left((a_j)_i - u_i\right).
\end{aligned} \tag{23}$$

Now, from the definition of g_x and E_x (see (16) and (17)), it is clear that:

$$g_x(x) = \sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \left\{ \sum_{i\in I(x)} \left[x_i - (a_j)_i\right]^2 + \sum_{i\in L(x)} \left[l_i - (a_j)_i\right]^2 + \sum_{i\in U(x)} \left[u_i - (a_j)_i\right]^2 \right\}. \quad (24)$$

Replacing (22) and (23) in (24) we get the following identity:

$$\begin{aligned} g_x(x) \;&=\; \sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \left\{ \sum_{i\in I(x)} \left[x_i - (a_j)_i\right]^2 + \sum_{i\in L(x)} \left[x_i - (a_j)_i\right]^2 + \sum_{i\in U(x)} \left[x_i - (a_j)_i\right]^2 \right\} + \\ &+\; 2 \sum_{i\in L(x)} (l_i - x_i) \sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \left[l_i - (a_j)_i\right] - \\ &-\; 2 \sum_{i\in U(x)} (u_i - x_i) \sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \left[(a_j)_i - u_i\right] - \\ &-\; \sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \left[\sum_{i\in L(x)} \left(l_i - (a_j)_i\right)^2 + \sum_{i\in U(x)} \left(u_i - (a_j)_i\right)^2 \right]. \end{aligned} \quad (25)$$

Notice that the first term of equation (25) is the Weber function, whereas the last term is the number γ as in (21). In the second and third term of (25) we can replace by $\alpha_k(x)$ and $\beta_k(x)$ using (11) and (12). Therefore, we get equation (20) and this concludes the proof.

■

Notice that if the point x were in the box $[l, u]$, due to the properties of the multipliers $\alpha_k(x)$ and $\beta_k(x)$, we would obtain an inequality between $g_x(x)$ and $f(x)$, that is, $g_x(x) \leq f(x)$. We would like to have a strict inequality in order that the sequence $\left\{f\left(x^{(k)}\right)\right\}$ is strictly descendent.

We investigate some properties of the function g_x and realise that is a strictly convex function in a space of dimension $r < n$, where $r = \#I(x)$ and considering $I(x)$ different from the empty set.

Proposition 3.2 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $I(x) = \{i_1, ..., i_r\} \neq \emptyset$ and $r = \#I(x)$. Then, the function $\tilde{g}_x = g_x \circ I_x$ is strictly convex.*

Proof. It can be seen that:

$$\nabla \tilde{g}_x(z) = 2 \sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \begin{pmatrix} z_1 - (a_j)_{i_1} \\ \vdots \\ z_r - (a_j)_{i_r} \end{pmatrix} \quad (26)$$

and

$$\nabla^2 \tilde{g}_x(z) = \begin{pmatrix} 2\sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2\sum_{j=1}^{m} \frac{w_j}{\|x-a_j\|} \end{pmatrix} \quad (27)$$

Thus, $\nabla^2 \tilde{g}_x$ is symetric and positive definite. Therefore, the function $\tilde{g}_x$ is strictly convex. ■

Consequently, the optimality conditions of first order for the function $\tilde{g}_x$ are necessary and sufficient. We can find a unique point where $\tilde{g}_x$ has a global minimum.

Proposition 3.3 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $I(x) = \{i_1, ..., i_r\} \neq \emptyset$ and $r = \#I(x)$. Then, the function $\tilde{g}_x$ has a unique global minimum point in $P_x \circ P \circ T(x)$.*

Proof. We have to look for a root of $\nabla \tilde{g}_x$.

$$\nabla \tilde{g}_x(z) = 0 \iff 2\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \begin{pmatrix} z_1 - (a_j)_{i_1} \\ \vdots \\ z_r - (a_j)_{i_r} \end{pmatrix} = 0 \tag{28}$$

$$\iff \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} z_k = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} (a_j)_{i_k}, \quad k = 1, ..., r, \tag{29}$$

$$\iff z_k = \frac{\displaystyle\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} (a_j)_{i_k}}{\displaystyle\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}}, \quad k = 1, ..., r. \tag{30}$$

$$\iff z = P_x \circ P \circ T(x) \tag{31}$$

■

Now we have to find a relationship between the functions $\tilde{g}$ with g_x.

Remark 3.4 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $I(x) = \{i_1, ..., i_r\} \neq \emptyset$ and $r = \#I(x)$. Then:*

$$g_x(y) = \tilde{g}_x \circ P_x(y). \tag{32}$$

Proof. Starting from the definition of g_x in (17) we have that:

$$g_x(y) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left\{ \sum_{k=1}^{r} \left[y_{i_k} - (a_j)_{i_k} \right]^2 + \sum_{i \in L(x)} \left[l_i - (a_j)_i \right]^2 + \sum_{i \in U(x)} \left[u_i - (a_j)_i \right]^2 \right\}. \tag{33}$$

Using the definition of the functions I_x and P_x as in (18)-(19) and replacing in the first term of (33):

$$\begin{aligned} g_x(y) \;=\; & \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left\{ \sum_{i \in I(x)} \left[(I_x \circ P_x(y))_i - (a_j)_i \right]^2 + \right. \\ & \left. + \sum_{i \in L(x)} \left[l_i - (a_j)_i \right]^2 + \sum_{i \in U(x)} \left[u_i - (a_j)_i \right]^2 \right\}. \end{aligned} \tag{34}$$

Looking at the definition of I_x once again, we have that:

$$g_x(y) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \|I_x \circ P_x(y) - a_j\|^2. \tag{35}$$

Finally, by Remark 2.6 and (17) we obtain:

$$g_x(y) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \|E_x\left(I_x \circ P_x(y)\right) - a_j\|^2 = g_x \circ I_x \circ P_x(y) = \tilde{g}_x \circ P_x(y). \tag{36}$$

■

The next two propositions have the objective of determining a strict inequality between the function g_x at the point where it reaches the global minimum and $f(x)$. First of all, we have to prove the following:

Proposition 3.5 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$. Then $g_x(P \circ T(x)) \leq g_x(x)$. If $I(x) \neq \emptyset$ and $P_x(P \circ T(x)) \neq P_x(x)$ then $g_x(P \circ T(x)) < g_x(x)$.*

Proof. If $I(x) = \emptyset$ then the function g_x is constant, therefore $g_x(P \circ T(x)) = g_x(x)$. Let us suppose now that $I(x) \neq \emptyset$ and $P_x(P \circ T(x)) = P_x(x)$. Then $g_x(P \circ T(x)) = \tilde{g}_x(P_x(P \circ T(x))) = \tilde{g}_x(P_x(x)) = g_x(x)$. Finally, if $I(x) \neq \emptyset$ and $P_x(P \circ T(x)) \neq P_x(x)$ then $g_x(P \circ T(x)) = \tilde{g}_x(P_x(P \circ T(x))) < \tilde{g}_x(P_x(x)) = g_x(x)$ due to Proposition 3.3.

■

Proposition 3.6 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $x \in [l, u]$ and $x \neq P \circ T(x)$. Then $g_x(P \circ T(x)) < f(x)$.*

Proof. Suppose first that $I(x) = \emptyset$. Then $L(x) \neq \emptyset$ or $U(x) \neq \emptyset$. This implies that there exists $i \in L(x) \cup U(x)$ such that $x_i \neq l_i$ or $x_i \neq u_i$. From Proposition 3.1 and Proposition 3.5 we have that:

$$g_x(P \circ T(x)) \leq g_x(x) = f(x) + 2 \sum_{i \in L(x)} \left(l_i - x_i\right) \alpha_i(x) - 2 \sum_{i \in U(x)} \left(u_i - x_i\right) \beta_i(x) + \gamma < f(x). \tag{37}$$

Now suppose that $I(x) \neq \emptyset$ and $P_x(x) = P_x(P \circ T(x))$. Since $x \neq P \circ T(x)$ there exists $i \in L(x) \cup U(x)$ such that $x_i \neq \left(P \circ T(x)\right)_i$. Then there exists $i \in L(x) \cup U(x)$ such that $x_i \neq l_i$ or $x_i \neq u_i$. The same argument as above applies.

Finally, let us assume that $I(x) \neq \emptyset$ and $P_x(x) \neq P_x(P \circ T(x))$. Using Proposition (3.5) we have that $g_x(P \circ T(x)) < g_x(x) = f(x) + 2\sum_{i \in L(x)} \left(l_i - x_i\right) \alpha_i(x) - 2\sum_{i \in U(x)} \left(u_i - x_i\right) \beta_i(x) + \gamma \leq f(x)$.

■

The next lemma states an equality that relates the Weber function and g_x at the point $P \circ T(x)$. Besides that, this result is crucial in the next section.

Lemma 3.7 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $x \in [l, u]$ and $x \neq P \circ T(x)$. Then*

$$g_x(P \circ T(x)) = f(x) + 2\left(f(P \circ T(x) - f(x))\right) + \delta, \qquad \delta \geq 0. \tag{38}$$

Proof. By definition 2.5 we have:

$$g_x\left(P \circ T(x)\right) = \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \|E_x(P \circ T(x)) - a_j\|^2 = \tag{39}$$

Adding and substracting $\|x - a_j\|$ we have:

$$\begin{aligned} g_x\left(P \circ T(x)\right) &= \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[\|x - a_j\| + \left(\|E_x(P \circ T(x)) - a_j\| - \|x - a_j\|\right)\right]^2 = \\ &= \sum_{j=1}^{m} w_j \|x - a_j\| + 2 \sum_{j=1}^{m} w_j \left(\|E_x(P \circ T(x)) - a_j\| - \|x - a_j\|\right) + \\ &+ \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left(\|E_x(P \circ T(x)) - a_j\| - \|x - a_j\|\right)^2. \end{aligned} \tag{40}$$

Notice that the first term of (40) is the Weber function, and the last term is a non-negative number, so we will define it as δ. Thus:

$$g_x\left(P \circ T(x)\right) = f(x) + 2\left(\sum_{j=1}^{m} w_j \|E_x(P \circ T(x)) - a_j\| - \sum_{j=1}^{m} w_j \|x - a_j\|\right) + \gamma = \tag{41}$$

It is easy to see that $E_x \circ P \circ T = P \circ T$, therefore:

$$\begin{aligned} g_x\left(P \circ T(x)\right) &= f(x) + 2\left(\sum_{j=1}^{m} w_j \|P \circ T(x) - a_j\| - \sum_{j=1}^{m} w_j \|x - a_j\|\right) + \delta = \\ &= f(x) + 2\left(f(P \circ T(x)) - f(x)\right) + \delta. \end{aligned} \tag{42}$$

This concludes the proof.

■

4 Main results.

This section states the main results about convergence of the sequence (10) but mainly that the limit of it is a solution of the problem (7)-(8), that is, that the limit of the sequence satisfies the Karush-Kuhn-Tucker optimality conditions.

The next theorem states that if a point x is in the box $[l, u]$, then the next iterated point produces a decrease in the objective function.

Theorem 4.1 *Let x be in $\mathbb{R}^n - \{a_1, ..., a_m\}$, $x \in [l, u]$ and $x \neq P \circ T(x)$. Then*

$$f(P \circ T(x)) < f(x). \tag{43}$$

Proof. Using Proposition 3.6 and Lemma 3.7 we have that:

$$g_x(P \circ T(x)) < f(x) \iff f(x) + 2\left(f(P \circ T(x)) - f(x)\right) + \delta < f(x) \iff \tag{44}$$

$$\iff f(P \circ T(x)) - f(x) + \frac{\delta}{2} < 0 \iff f(P \circ T(x)) + \frac{\delta}{2} < f(x). \tag{45}$$

Then:

$$f(P \circ T(x)) \leq f(P \circ T(x)) + \frac{\delta}{2} < f(x). \tag{46}$$

■

As a direct consequence of Theorem 4.1 we have:

Theorem 4.2 *Let $x^{(0)}$ be an arbitrary point in $\mathbb{R}^n$ such that $x^{(0)} \in [l, u]$. If the sequence (10) is well defined (that is, $x^{(k)} \neq a_1, ..., a_m$) then the sequence $\left\{f(x^{(k)})\right\}_{k=0}^{\infty}$ is strictly descendent.*

We can assure that at least we will have a subsequence of (10) convergent to some point x^*.

Theorem 4.3 *If the sequence $\left\{x^{(k)}\right\}_{k=0}^{\infty}$ can be generated, then there exists a subsequence convergent to a point $x^* \in [l, u]$.*

Proof. By Remark 2.1 the sequence is contained in the box $[l, u]$. By compactness, there exists a limit point x^*.

■

Corollary 4.4 *If $x^{(k)}$ converges to x^* when k tends to infinity, then x^* is a fixed point of $\tilde{T} = P \circ T$.*

Proof. This a consequence of Remark 2.1.

■

The KKT optimality conditions of first order for the problem (7)-(8) are the following:

$$\begin{aligned}
\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}(x - a_j) - \sum_{i=1}^{n} \mu_i e_i + \sum_{i=1}^{n} \mu_{i+n} e_i &= 0, && (47)\\
\mu_i (l_i - x_i) &= 0, \quad i = 1, ..., n, && (48)\\
\mu_{i+n} (x_i - u_i) &= 0, \quad i = 1, ..., n, && (49)\\
\mu_i &\geq 0, \quad i = 1, ..., 2n, && (50)\\
l_i - x_i &\leq 0, \quad i = 1, ..., n, && (51)\\
x_i - u_i &\leq 0, \quad i = 1, ..., n, && (52)
\end{aligned}$$

where the vectors e_i are the canonical vectors.

Now, we would like to prove that any fixed point of $\tilde{T}$ is the solution to the problem (7)-(8), and finally make the conexion between the modified Weiszfeld algorithm and a minimization problem.

Theorem 4.5 *The following are equivalent:*

(a) x is a KKT point,

(b) x is a solution of the problem (7)-(8)

(c) x is a fixed point of $\tilde{T}$.

Proof. Because the function f is strictly convex and the set $[l, u]$ is convex, it holds that (a) is equivalent to (b).

Now, we will prove that (a) implies (c). Let us suppose that x is a KKT point, that is, x is a minimum point of f subject to $[l, u]$. If x were not a fixed point of $\tilde{T}$, we would have that $x \neq P \circ T(x)$, and then $f(P \circ T(x)) < f(x)$ by Proposition 4.1. This contradicts the hypothesis.

In order to demonstrate that (c) implies (a), we assume that x is a fixed point of $P \circ T$, that is, $x = P \circ T(x)$. We will prove that x holds equations (47)-(52). We will define the following multipliers:

$$\mu_i = \begin{cases} \alpha_i, & i \in L(x) \\ 0, & i \notin L(x), \end{cases} \qquad i = 1, ..., n, \tag{53}$$

$$\mu_{i+n} = \begin{cases} \beta_i, & i \in U(x), \\ 0, & i \notin U(x), \end{cases} \qquad i = 1, ..., n. \tag{54}$$

By construction, the equation (50) holds. Due to the fact that $x = P \circ T(x)$ and by equation (9), the point x is in the box $[l, u]$, so equations (51) and (52) hold.

Let i be such that $1 \leq i \leq n$:

- If $i \notin L(x)$ then $\mu_i = 0$.
- If $i \in L(x)$ then $(T(x))_i < l_i$, and therefore $(P \circ T(x))_i = l_i$. So, $x_i - l_i = 0$.

This implies equation (48).

Similarly,

- If $i \notin U(x)$ then $\mu_{i+n} = 0$.
- If $i \in U(x)$ then $(T(x))_i > u_i$, and therefore $(P \circ T(x))_i = u_i$. So, $x_i - u_i = 0$.

This implies equation (49).

Let us check equation (47) for each component k, $k = 1, ..., n$:

- If $k \in I(x)$, then $(T(x))_k \in [l_k, u_k]$. This means that $x_k = (P \circ T(x)) = (T(x))_k$. Besides that $\mu_k = \mu_{k+n} = 0$. Now, the left-hand side of equation (47) says that:

$$\frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] - \mu_k + \mu_{k+n}}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = \frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right]}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = \tag{55}$$

$$= \quad x_k - \frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} (a_j)_k}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = x_k - (T(x))_k = 0. \tag{56}$$

- If $k \in L(x)$, then $(T(x))_k < l_k$. This means that $x_k = (P \circ T(x))_k = l_k$. Besides that $\mu_k = \alpha_k$ and $\mu_{k+n} = 0$. The left-hand side of (47) says that:

$$\frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] - \mu_k + \mu_{k+n}}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = \frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] - \alpha_k}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = \tag{57}$$

$$= \quad \frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] - \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[l_k - (a_j)_k\right]}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = x_k - l_k = 0. \tag{58}$$

- If $k \in U(x)$, then $(T(x))_k > u_k$. This means that $x_k = (P \circ T(x))_k = u_k$. Besides that $\mu_k = 0$ and $\mu_{k+n} = \beta_k$. The left-hand side of (47) says that:

$$\frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] - \mu_k + \mu_{k+n}}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = \frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] + \beta_k}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = \tag{59}$$

$$= \quad \frac{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[x_k - (a_j)_k\right] + \sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|} \left[(a_j)_k - u_k\right]}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} = x_k - u_k = 0. \tag{60}$$

Therefore, this concludes the proof.

■

5 Numerical experiments.

We implemented the algorithm (10) in standard Fortran 90 language using Intel Fortran Compiler 10.1. The code was compiled and executed in a PC running Linux OS, Intel(R) Core(RM)2 Duo, T5750, 2.00 GHz.

In order to test the examples, the problem (7)-(8) was also solved with an Optimization Toolbox implemented in Matlab 7.5: function *fmincon* for the constrained problem, and function *fminunc* for the unconstrained problem. Details on the algorithms behind this M-functions can be seen in the Matlab documentation (see [24] and [23] and the bibliography cited there).

The following two examples show interesting aspects of the algorithm. In both cases we take as vertices and their associated weights:

$$x_1 = (1,0)\,, \quad w_1 = 5, \tag{61}$$
$$x_2 = (0,0)\,, \quad w_2 = 3, \tag{62}$$
$$x_3 = (0,1)\,, \quad w_3 = 2. \tag{63}$$
$$x_4 = (1,4)\,, \quad w_3 = 3. \tag{64}$$

The next table presents the settings of the two examples:

	Example 1	Example 2
Initial point	$(1.0, 3.5)$	$(0.5, 1.0)$
Box	$[0,1] \times [1.5, 3.5]$	$[0.25, 0.75] \times [0,1]$
Tolerance	10^{-5}	10^{-5}

In the first example (see Figure 1) the algorithm took 10 iterations to reach the solution for the determined tolerance. Notice that the solution of the unrestricted problem is outside the box, and the solution of (7)-(8) is not simply a projection of the solution without restrictions. The results are the following:

- Without restrictions (using Matlab): $(0.65394, 0.29279)$
- With restrictions (using Matlab): $(0.47294, 1.50000)$
- With restrictions (using (10)): $(0.47293, 1.50000)$

In the second example (see Figure 2), 37 iterations were needed to reach the solution. In this case the global minimum is inside the box, therefore the solutions of the unrestricted problem and the solution of (7)-(8) are the same. The results are the following:

- Without restrictions (using Matlab): $(0.65394, 0.29279)$
- With restrictions (using Matlab): $(0.65386, 0.29282)$
- With restrictions (using (10)): $(0.65391, 0.29281)$

In both cases the sequence generated by the algorithm is convergent to the solution, thus the theoretical results correspond to the experiments.

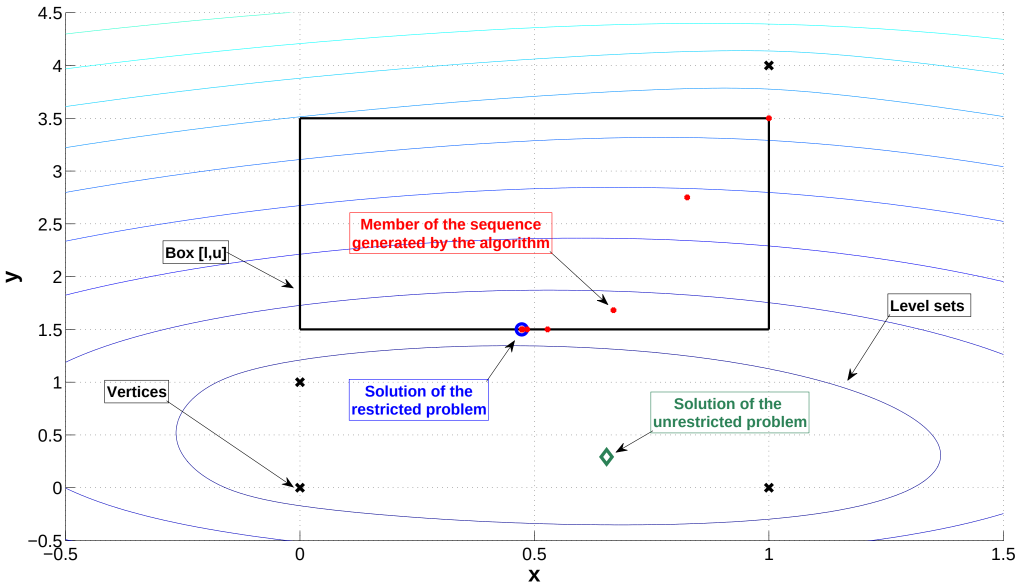

Figure 1: Example 1

6 Conclusions.

This paper proposes a modified Weiszfeld algorithm consisting in two stages: first, iterate using the Weiszfeld iteration function (4), and second, project onto the box using (9). We realise that the limit of the sequence generated by the algorithm (10) is in fact the solution of a minimization problem with box constraints (7)-(8). We prove that the sequence $\left\{f(x^{(k)})\right\}$ is descendent and that a point x^* is a fixed-point of the iteration function if and only if x^* is a KKT point. That property allows us to relate the algorithm and the minimization problem.

Numerical experiments are performed in order to confirm the theoretical results. We show an example where the solution is outside the box, and another example where the solution is outside the box. In the last case, the iteration is the same as the Weiszfeld iteration. In both cases it can be seen that the sequence generated by the proposed algorithm converges to the solution of the restricted problem.

References

[1] **G. O. Wesolowsky**, *The Weber problem: history and perspectives*, J. Location Sci., 1, pp. 5–23, 1993.

[2] **Y. S. Kupitz** and **H. Martini**, *Geometric aspects of the generalized Fermat-Torricelli problem*, Intuitive Geometry, Mathematical Studies, Vol. 6, Bolyai Society, pp. 55–127, 1997.

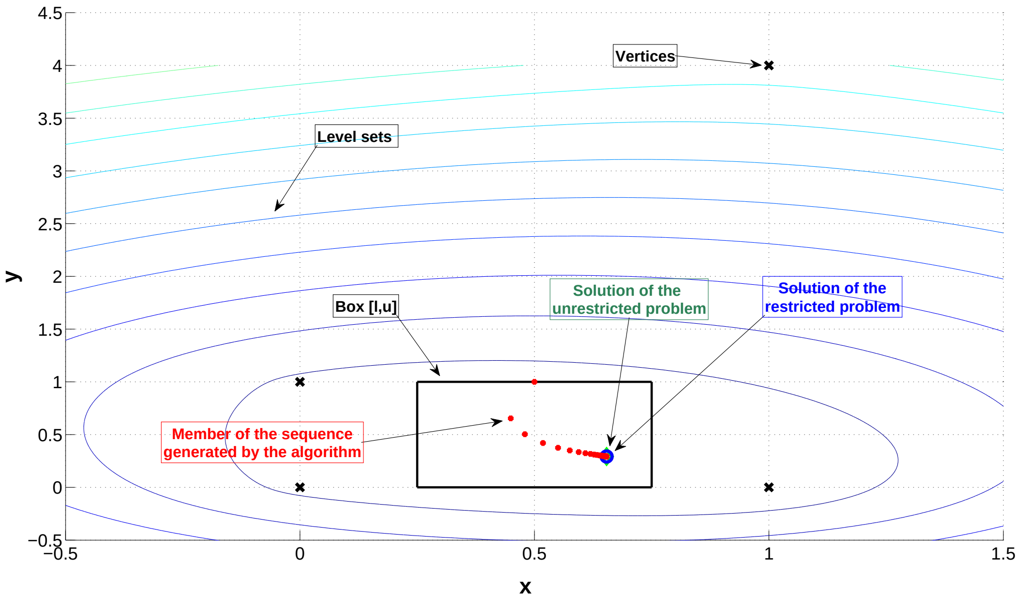

Figure 2: Example 2

[3] **A. Weber**, *Über den Standort der Industrien*, Tübingen. (Translated by C. J. Friederich (1929): "Theory of the location of industries". Chicago, University of Chicago Press.).

[4] **E. Weiszfeld**, *Sur le point par lequel la somme des distances de n points donnés est minimum*, Tohoku Mathematics Journal 43, pp. 355–386, 1937.

[5] **H. W. Kuhn**, *A note on Fermat's problem*, Mathematical Programming 4, pp. 98–107, 1973.

[6] **J. A. Chatelon**, **D. W. Hearn** and **T. J. Lowe**, *A subgradient algorithm for certain minimax and minisum problems*, Mathematical Programming 15, pp. 130–145, 1978.

[7] **J. W. Eyster**, **J. A. White** and **W. W. Wierwille**, *On solving multifacility location problems using a hyperboloid approximation procedure*, AIEE Transactions 5, pp. 1–6, 1973.

[8] **H. W. Kuhn**, *On a pair of dual nonlinear programs*, in: J. Abadie, ed., Nonlinear Programming (North-Holland, Amsterdam, 1967), pp. 38–54.

[9] **M. L. Overton**, *A quadratically convergence method for minimizing a sum of Euclidean norms*, Mathematical Programming 27, pp. 34–63, 1983.

[10] **R. Chandrasekaran** and **A. Tamir**, *Open questions concerning Weiszfeld's algorithm for the Fermat-Weber location problem*, Mathematical Programming 44, pp. 293–295, 1989.

[11] **J. Brimberg**, *The Fermat-Weber location problem revisited*, Mathematical Programming 71, pp. 71–76, 1995.

[12] **L. Cánovas**, **R. Cañavate** and **Alfredo Marín**, *On the convergence of the Weiszfeld algorithm*, Mathematical Programming, pp. 327–330, 2002.

[13] **R. Cañavate**, *Algunas cuestiones teóricas sobre la validez del algoritmo de Weiszfeld para el problema de Weber*, Recta, ASEPUMA (Asociación Española de Profesores Universitarios de Matemáticas para la Economía y la Empresa), 2005.

[14] **E. Carrizosa**, **M. Muñoz-Márquez** and **J. Puerto**, *The Weber problem with regional demand*, European Journal of Operational Research 104, pp. 358–365, 1998.

[15] **H. W. Hamacher** and **K. Klamroth**, *Planar Weber location problems with barriers and block norms*, Annals of Operations Research 96, pp. 191–208, 2000.

[16] **Y. Vardi** and **C.-H. Zhang**, *A modified Weiszfeld algorithm for the Fermat-Weber location problem*, Mathematical Programming, Ser. A 90, pp. 559–566, 2001.

[17] **S.-D. Lee**, *On solving unreliable planar location problems*, Computers & Operations Research 28, pp. 329–344, 2001.

[18] **M. Zaferanieh**, **H. Taghizadeh Kakhki**, **J. Brimberg** and **G. O. Wesolowsky**, *A BSSS algorithm for the single facility location problem in two regions with different norms*, European Journal of Operational Research 190, pp. 79–89, 2008.

[19] **B. Pfeiffer** and **K. Klamroth**, *A unified model for Weber problems with continuous and network distances*, Computers & Operations Research 35, pp. 312–326, 2008.

[20] **M. Bischoff**, **T. Fleischmann** and **K. Klamroth**, *The multi-facility location-allocation problem with polyhedral barriers*, Computers & Operations Research 36, pp. 1376–1392, 2009.

[21] **T. Drezner**, **Z. Drezner** and **C. Scott**, *Location of a facility minimizing nuisance to or from a planar network*, Computers & Operations Research 36, pp. 135–148, 2009.

[22] **M. Luis**, **S. Salhi** and **G. Nagy**, *Region-rejection based heuristics for the capacitated multi-source Weber problem*, Computers & Operations Research 36, pp. 2007–2017, 2009.

[23] http://www.mathworks.com/access/helpdesk/help/toolbox/optim/ug/fminunc.html

[24] http://www.mathworks.com/access/helpdesk/help/toolbox/optim/ug/fmincon.html