Operations Research Letters 38 (2010) 207–214

Contents lists available at ScienceDirect

Operations Research Letters

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

A generalized Weiszfeld method for the multi-facility location problem

Cem Iyigun [a,*], Adi Ben-Israel [b]

[a] Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey
[b] RUTCOR—Rutgers Center for Operations Research, Rutgers University, NJ, USA

ARTICLE INFO

Article history:
Received 15 May 2009
Accepted 15 October 2009
Available online 26 November 2009

Keywords:
Fermat–Weber location problem
Multi-facility location problem
Decomposition method
Probabilistic assignments
Weiszfeld method

ABSTRACT

An iterative method is proposed for the K facilities location problem. The problem is relaxed using probabilistic assignments, depending on the distances to the facilities. The probabilities, that decompose the problem into K single-facility location problems, are updated at each iteration together with the facility locations. The proposed method is a natural generalization of the Weiszfeld method to several facilities.

1. Introduction

The **Fermat–Weber location problem** (also **single-facility location problem**) is to locate a **facility** that will serve optimally a set of **customers**, given by their **locations** and **weights**, in the sense of minimizing the weighted sum of distances traveled by the customers. A well known method for solving the problem is the **Weiszfeld method** [30], a gradient method that expresses and updates the sought center as a convex combination of the data points.

The **multi-facility location problem** (**MFLP**) is to locate a (given) number of facilities to serve the customers as above. Each customer is assigned to a single facility, and the problem (also called the **location–allocation problem**) is to determine the optimal locations of the facilities, as well as the optimal assignments of customers (assignment is absent in the single-facility case.)

MFLP is NP hard, [26]. We relax it by replacing rigid assignments with probabilistic assignments, as in [4,3] and [8]. This allows a decomposition of MFLP into single-facility location problems, coupled by the membership probabilities that are updated at each iteration.

2. The problem

Notation. $\overline{1,K} := \{1, 2, \ldots, K\}$. $\|\mathbf{x}\|$ denotes the **Euclidean norm** of a vector $\mathbf{x} \in \mathbb{R}^n$. The **Euclidean distance** $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ is used throughout.

Let $\mathbf{X} := \{\mathbf{x}_i : i \in \overline{1,N}\}$ be a set of N **data points** in $\mathbb{R}^n$, with given **weights** $\{w_i > 0 : i \in \overline{1,N}\}$. Typically, the points $\{\mathbf{x}_i\}$ are the **locations** of **customers**, the weights $\{w_i\}$ are their **demands**.

Given an integer $1 \le K < N$, the **MFLP** is to locate K facilities, and assign each customer to one facility, so as to minimize the sum of weighted distances

$$\min_{\mathbf{c}_1,\mathbf{c}_2,\ldots,\mathbf{c}_K} \sum_{k=1}^{K} \sum_{\mathbf{x}_i \in \mathcal{C}_k} w_i\, d(\mathbf{x}_i, \mathbf{c}_k) \tag{L.K}$$

where $\{\mathbf{c}_k\}$ are the **locations** (or **centers**) of the **facilities**, and $\mathcal{C}_k$ is the **cluster** of customers that are assigned to the kth facility.

For $K = 1$, one gets the **Fermat–Weber location problem**: given $\mathbf{X}$ and $\{w_i : i \in \overline{1,N}\}$ as above, find a point $\mathbf{c} \in \mathbb{R}^n$ minimizing the sum of weighted distances,

$$\min_{\mathbf{c} \in \mathbb{R}^n} \sum_{i=1}^{N} w_i\, d(\mathbf{x}_i, \mathbf{c}), \tag{L.1}$$

see [12,23,24,31] and their references.

If the points $\{\mathbf{x}_i\}$ are not collinear, as is assumed throughout, the objective function of (L.1)

$$f(\mathbf{c}) = \sum_{i=1}^{N} w_i\, d(\mathbf{x}_i, \mathbf{c}) \tag{1}$$

is strictly convex, and (L.1) has a unique optimal solution.

The gradient of (1) is undefined if $\mathbf{c}$ coincides with one of the data points $\{\mathbf{x}_i\}$. For $\mathbf{c} \notin \mathbf{X}$,

$$\nabla f(\mathbf{c}) = -\sum_{i=1}^{N} w_i \frac{\mathbf{x}_i - \mathbf{c}}{\|\mathbf{x}_i - \mathbf{c}\|}, \tag{2}$$

* Corresponding author.
E-mail addresses: iyigun@ie.metu.edu.tr (C. Iyigun), adi.benisrael@gmail.com (A. Ben-Israel).

0167-6377/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2009.11.005

and the optimal center $\mathbf{c}^*$, if not in $\mathbf{X}$, is characterized by $\nabla f(\mathbf{c}^*) = \mathbf{0}$, expressing it as a convex combination of the points $\mathbf{x}_i$,

$$\mathbf{c}^* = \sum_{i=1}^{N} \lambda_i \mathbf{x}_i, \text{ with weights } \lambda_i = \frac{w_i/\|\mathbf{x}_i - \mathbf{c}^*\|}{\sum_{m=1}^{N} w_m/\|\mathbf{x}_m - \mathbf{c}^*\|} \text{ that depend on } \mathbf{c}^*.$$

This circular result gives rise to the **Weiszfeld iteration**, [30],

$$\mathbf{c}_+ := T(\mathbf{c}) \tag{3}$$

where $\mathbf{c}_+$ is the updated center, $\mathbf{c}$ is the current center, and

$$T(\mathbf{c}) := \begin{cases} \displaystyle\sum_{i=1}^{N} \left(\frac{w_i/\|\mathbf{x}_i - \mathbf{c}\|}{\sum_{m=1}^{N} w_m/\|\mathbf{x}_m - \mathbf{c}\|} \right) \mathbf{x}_i, & \text{if } \mathbf{c} \notin \mathbf{X}; \\ \mathbf{c}, & \text{if } \mathbf{c} \in \mathbf{X}. \end{cases} \tag{4}$$

In order to extend $\nabla f(\mathbf{c})$ to all $\mathbf{c}$, Kuhn [21] modified it as follows: $\nabla f(\mathbf{c}) := -R(\mathbf{c})$, where

$$R(\mathbf{c}) := \begin{cases} -\nabla f(\mathbf{c}), & \text{if } \mathbf{c} \notin \mathbf{X}; \\ \max\{0, \|R^j\| - w_j\} \dfrac{R^j}{\|R^j\|}, & \text{if } \mathbf{c} = \mathbf{x}_j \in \mathbf{X}, \end{cases} \tag{5}$$

$$\text{where} \quad R^j := \sum_{i \neq j} \frac{w_i}{\|\mathbf{x}_i - \mathbf{x}_j\|} (\mathbf{x}_i - \mathbf{x}_j) \tag{6}$$

is the **resultant force** of $N-1$ forces of magnitude w_i and direction $\mathbf{x}_i - \mathbf{x}_j$, $i \neq j$. The following properties of the mappings $R(\cdot)$, $T(\cdot)$, the optimal center $\mathbf{c}^*$ and any point $\mathbf{x}_j \in \mathbf{X}$ were proved by Kuhn [21]:

$$\mathbf{c} = \mathbf{c}^* \iff R(\mathbf{c}) = \mathbf{0}. \tag{7a}$$

$$\mathbf{c}^* \in \operatorname{conv} \mathbf{X} \text{ (the convex hull of } \mathbf{X}). \tag{7b}$$

If $\mathbf{c} = \mathbf{c}^*$ then $T(\mathbf{c}) = \mathbf{c}$. Conversely, if $\mathbf{c} \notin \mathbf{X}$,

$$T(\mathbf{c}) = \mathbf{c} \text{ then } \mathbf{c} = \mathbf{c}^*. \tag{7c}$$

$$\text{If } T(\mathbf{c}) \neq \mathbf{c} \text{ then } f(T(\mathbf{c})) < f(\mathbf{c}). \tag{7d}$$

$$\mathbf{x}_j = \mathbf{c}^* \iff w_j \geq \|R^j\|. \tag{7e}$$

If $\mathbf{x}_j \neq \mathbf{c}^*$,

$$\text{the direction of steepest descent of } f \text{ at } \mathbf{x}_j \text{ is } R^j/\|R^j\|. \tag{7f}$$

If $\mathbf{x}_j \neq \mathbf{c}^*$ there exists $\delta > 0$ such that

$$0 < \|\mathbf{c} - \mathbf{x}_j\| \implies \|T^s(\mathbf{c}) - \mathbf{x}_j\| > \delta \text{ for some } s. \tag{7g}$$

$$\lim_{\mathbf{c} \to \mathbf{x}_j} \frac{\|T(\mathbf{c}) - \mathbf{x}_j\|}{\|\mathbf{c} - \mathbf{x}_j\|} = \frac{\|R^j\|}{w_j}. \tag{7h}$$

$$\text{For any } \mathbf{c}_0, \text{ if no } \mathbf{c}_r := T^r(\mathbf{c}_0) \in \mathbf{X}, \text{ then } \lim_{r \to \infty} \mathbf{c}_r = \mathbf{c}^*. \tag{7i}$$

These results are generalized in Theorem 1 to the case of several facilities.

Remark 1. Another claim in [21], that

$T^r(\mathbf{c}_0) \to \mathbf{c}^*$

for all but a denumerable number of initial centers $\mathbf{c}_0$,

was refuted by Chandrasekaran and Tamir [6]. Convergence can be assured by modifying the algorithm (3) and (4) at a non-optimal center that coincides with a data point $\mathbf{x}_j$. Balas and Yu [1] proposed moving from $\mathbf{x}_j$ in the direction R^j (6) of steepest descent, assuring a decrease of the objective, and non-return to $\mathbf{x}_j$ by (7c) and (7d). Vardi and Zhang [29] guaranteed exit from $\mathbf{x}_j$ by augmenting the objective function with a quadratic of the distances to the other data points. Convergence was also addressed by Ostresh [27], Eckhardt [13], Drezner [11], Brimberg [5], Beck et al. [2], and others.

3. Probabilistic assignments

For $1 < K < N$, the problem (L.K) is NP hard, [26]. It can be solved polynomially in N for $K = 2$, see [10], and possibly for other given K.

We relax the problem by using **probabilistic** (or soft) assignments, with **cluster membership probabilities**,

$$p_k(\mathbf{x}) := \operatorname{Prob}\{\mathbf{x} \in \mathcal{C}_k\}, \quad k \in \overline{1, K},$$

assumed to depend only on the distances $\{d(\mathbf{x}, \mathbf{c}_k) : k \in \overline{1, K}\}$ of the point $\mathbf{x}$ from the K centers. A reasonable assumption is

ASSIGNMENT TO A FACILITY IS MORE PROBABLE THE CLOSER IT IS (A)

and a simple way to model it,

$$p_k(\mathbf{x})\, d(\mathbf{x}, \mathbf{c}_k) = \frac{1}{w} D(\mathbf{x}), \quad k \in \overline{1, K}, \tag{8}$$

where w is the weight of $\mathbf{x}$, and $D(\cdot)$ is a function of $\mathbf{x}$, that does not depend on k. There are other ways to model assumption (A), but (8) works well enough for our purposes.

Model (8) expresses probabilities in terms of distances, displaying neutrality among facilities in the sense of the Choice Axiom of Luce, [25, Axiom 1], see [19, Appendix A]. Other issues such as attractiveness, introduced in the Huff model [15,16], see also [9], are ignored.

Using the fact that probabilities add to one, we get from (8),

$$p_k(\mathbf{x}) = \frac{1/d(\mathbf{x}, \mathbf{c}_k)}{\sum_{\ell=1}^{K} 1/d(\mathbf{x}, \mathbf{c}_\ell)} = \frac{\prod_{j \neq k} d(\mathbf{x}, \mathbf{c}_j)}{\sum_{\ell=1}^{K} \prod_{m \neq \ell} d(\mathbf{x}, \mathbf{c}_m)}, \quad k \in \overline{1, K}, \tag{9}$$

interpreted as $p_k(\mathbf{x}) = 1$ if $d(\mathbf{x}, \mathbf{c}_k) = 0$, i.e., $\mathbf{x} = \mathbf{c}_k$. In the special case $K = 2$,

$$\begin{aligned} p_1(\mathbf{x}) &= \frac{d(\mathbf{x}, \mathbf{c}_2)}{d(\mathbf{x}, \mathbf{c}_1) + d(\mathbf{x}, \mathbf{c}_2)}, \\ p_2(\mathbf{x}) &= \frac{d(\mathbf{x}, \mathbf{c}_1)}{d(\mathbf{x}, \mathbf{c}_1) + d(\mathbf{x}, \mathbf{c}_2)}. \end{aligned} \tag{10}$$

From (8), we similarly get

$$\frac{D(\mathbf{x})}{w} = \frac{\prod_{j=1}^{K} d(\mathbf{x}, \mathbf{c}_j)}{\sum_{\ell=1}^{K} \prod_{m \neq \ell} d(\mathbf{x}, \mathbf{c}_m)}, \tag{11}$$

which is (up to a constant) the **harmonic mean** of the distances $\{d(\mathbf{x}, \mathbf{c}_j) : j \in \overline{1, K}\}$. In particular,

$$D(\mathbf{x}) = w\, \frac{d(\mathbf{x}, \mathbf{c}_1) d(\mathbf{x}, \mathbf{c}_2)}{d(\mathbf{x}, \mathbf{c}_1) + d(\mathbf{x}, \mathbf{c}_2)}, \quad \text{for } K = 2. \tag{12}$$

The function (11) is called the **joint distance function** (**JDF**) at $\mathbf{x}$.

Abbreviating $p_k(\mathbf{x})$ by p_k, Eq. (8) is an optimality condition for the extremum problem

$$\min \left\{ w \sum_{k=1}^{K} p_k^2\, d(\mathbf{x}, \mathbf{c}_k) : \sum_{k=1}^{K} p_k = 1,\ p_k \geq 0,\ k \in \overline{1, K} \right\} \tag{13}$$

with variables $\{p_k\}$. The squares of probabilities in (13) are explained as a device for smoothing the underlying objective, min $\{\|\mathbf{x} - \mathbf{c}_k\| : k \in \overline{1, K}\}$, see the seminal article by Teboulle [28].