Project Description for the course *Optimization 1*: Location Analysis

Elisabeth Köbis*

January 26, 2022

1 Introduction

Many important problems from industry, engineering-technical or social area, logistics and the energy industry lead to location problems. Examples of this are the planning of new industrial plants, production facilities or warehouses, of hospitals, rescue stations and other public facilities, to landscaping or the placement of sensors on technical components. For example, in order to avoid hospital spacial inefficiency, it is important to learn how to optimize the spaces such that relevant services such as surgical equipment or emergency services are located at the best points of access to other stations. As another example, for a decision-maker deciding the best location to build a warehouse relative to factories, it is important to find the best place to locate this warehouse in order to minimize costs of travel and transportations. Thus, location problems appear in many variants and with different constraints depending on the practical application. Other examples of location analysis appear for instance in the following areas:

- Urban and Regional Planning (e.g. locations for emergency facilities),
- Technology (e.g. placement of sensors on technical components),
- Economy (e.g. planning new production facilities),
- Geography (e.g. landscape design),
- Environment-Oriented Project Management (e.g. development of mining landscapes),
- Engineering.

This project is intended to provide the mathematical tools to model and solve location problems.

The main goal in location analysis is to find one or more new locations with minimal distance to a set of known locations.

2 Problem Description

Let us consider m given points in the plane

$$a^i = (a_1^i, a_2^i)^\top \in \mathbb{R}^2, \quad i = 1, \dots, m.$$

$A = \{a^1, \dots, a^m\}$ is the set of all given locations. Let $\mathcal{M} := \{1, \dots, m\}$. We denote the new location, which we want to determine, by

$$x = (x_1, x_2)^\top \in \mathbb{R}^2.$$

When modeling a location problem, distance measures as well as the selected objective function play a key role. We consider the following distances:

- $d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|$ (Manhattan or Taxicab distance)
- $d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ (Euclidean distance)
- $d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$ (Maximum distance)

*Norwegian University of Science and Technology, Norway, email: elisabeth.kobis@ntnu.no

The individual distances between the unknown x and a given a^i are set differently according to the concrete application. The selection of the distance measure determines whether the distance should be measured by the linear distance (Euclidean) or by a blocknorm (Maximum or Manhattan). The Maximum distance is suitable if movement in both directions is possible and only the larger one of the two determines the length of the movement.

For the objective function, the following ones are possible:

- $\max_{a\in A} d(a,x)$ (Center problem)
- $\sum_{a\in A} d(a,x)$ (Median function)

Thus, we consider the following problems

$$\min_{x\in\mathbb{R}^2} \max_{a\in A} d(a,x) \quad \text{(Center problem)} \tag{2.1}$$

$$\min_{x\in\mathbb{R}^2} \sum_{a\in A} d(a,x) \quad \text{(Median problem, also called Fermat-Weber problem)} \tag{2.2}$$

2.1 Metrics and Norms

Mathematically, the distances between two points in the plane can be measured using an appropriate metric.

Definition 2.1. Let Y be a non-empty set of $\mathbb{R}^2$. A function $d: Y \times Y \to \mathbb{R}$ is called *metric* on Y if d fulfills the following conditions for all $x, y, z, \in Y$:

1. $d(x,y) = 0 \iff x = y$ (definiteness).
2. $d(x,y) = d(y,x)$ (symmetry).
3. $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality).

The real number $d(x,y)$ represents the distance between the points x and y.

Definition 2.2. A function $||\cdot|| : \mathbb{R}^2 \to \mathbb{R}$ is called **norm** on $\mathbb{R}^2$ if $||\cdot||$ fulfills the following conditions for all $x, y \in \mathbb{R}^2$ and for all $\alpha \in \mathbb{R}$:

1. $||x|| = 0 \iff x = 0$ (definiteness).
2. $||\alpha x|| = |\alpha| \cdot ||x||$ (positive homogeneity).
3. $||x + y|| \leq ||x|| + ||y||$ (triangle inequality).

If $||\cdot|| : \mathbb{R}^2 \to \mathbb{R}$ is a norm, then it is possible to define a metric on $\mathbb{R}^2$ that is induced by the norm $||\cdot||$ in the following way:

$$d(x,y) := ||x - y||$$

for all $x, y \in \mathbb{R}^2$.

2.2 Project Tasks - Theoretical Work

1. Show that d_1, d_2 and d_∞ are metrics. [5 points]
2. Plot the unit balls $B_i(0,1) := \{x \in \mathbb{R}^2 \mid ||x||_i \leq 1\}$ $(i \in \{1, 2, \infty\})$ of the so-called Manhattan norm, the Euclidean norm and of the maximum norm. [5 points]
3. Show that every norm is a convex function. [5 points]
4. Show that the objective function of problems (2.1) and (2.2) is convex. [5 points]
5. Give a geometric interpretation of solving problem (2.1) with Euclidean distance function. [5 points]
6. Provide a solution approach for problem (2.2) with Manhattan distance function (Hint: Note that the objective function can be separated (so-called separability). Just describe the idea of such an approach.). [5 points]

7. Consider the Median problem with **squared** Euclidean distance function

$$\min_{x\in\mathbb{R}^2} \sum_{a\in A} (d_2(a,x))^2.$$

Show that the objective function is convex, and analytically compute the uniquely determined minimizer of this problem. [10 points]

8. Consider the Median problem with Euclidean distance function

$$\min_{x\in\mathbb{R}^2} \sum_{a\in A} d_2(a,x).$$

Derive necessary optimality conditions for minimizers of this problem (Hint: Note that the partial derivatives are not defined globally). [5 points]

9. Give an example of a set of existing locations, where the set of global minimizers for the problem $\min_{x\in\mathbb{R}^2} \sum_{a\in A} d_2(a,x)$ and $\min_{x\in\mathbb{R}^2} \sum_{a\in A}(d_2(a,x))^2$ do not coincide. Explain the solution. [5 points]

3 Weiszfeld Algorithm

The above problem in task 7. was easy to solve, as the objective function is differentiable and one is able to derive the stationary points by rearranging the formula appearing in the optimality condition. If we consider the Median problem with Euclidean distance function that is **not** squared, an according procedure is not as simple anymore, as one cannot solve for a minimizer directly. Instead, one uses an iterative procedure to compute approximations of a minimizer. Here, we consider the Median problem with **weighted** Euclidean distances as follows

$$\min_{x\in\mathbb{R}^2} \sum_{i\in\mathcal{M}} v^i d_2(a^i,x), \tag{P_{d_2}}$$

where the weights v^i, $i \in \mathcal{M} = \{1,\ldots,m\}$, are real-valued nonnegative numbers.

The problem (P_{d_2}) is one of the most fundamental location problems. It consists of finding a point that minimizes the sum of its weighted distances to a given finite set of anchor points. The problem is credited to the well-known French mathematician Pierre de Fermat, who at the beginning of the seventeenth century posed the following question:

Given three points in a plane, find a fourth point such that the sum of its distances to the three given points is as small as possible.

The Italian physicist and mathematician Evangelista Torricelli (mostly known for inventing the barometer) found a construction method of this point by ruler and compass, and it is therefore also called “the Toricelli point” (see reference 2. in Section 4). At the beginning of the twentieth century, the German economist Alfred Weber incorporated weights, and was able to treat facility location problems with more than 3 facilities, and the problem was consequently called “the Fermat–Weber problem”. Other names for the problem are “the Fermat problem”, “the Weber problem”, “the Fermat–Toricelli problem”, “the Steiner problem”, and many more variants.

Before presenting an algorithm for solving problem (P_{d_2}), we give the following result.

Theorem 1 (see: Love, Morris Wesolowsky, 1988, Property 2.2)**.** *If for*

$$Test_k := \left[\Big(\sum_{i\in\mathcal{M}\setminus\{k\}} v^i \frac{a_1^k - a_1^i}{d_2(a^k,a^i)} \Big)^2 + \Big(\sum_{i\in\mathcal{M}\setminus\{k\}} v^i \frac{a_2^k - a_2^i}{d_2(a^k,a^i)} \Big)^2 \right]^{\frac{1}{2}},$$

it holds that

$$\boxed{Test_k \le v^k}$$

for one $k \in \mathcal{M}$, then $x^* = a^k$ is a global minimizer.

The so-called Weiszfeld algorithm is an iterative method based on the first-order necessary conditions for a stationary point of the objective function.

The Weiszfeld algorithm consists of the following steps:

Weiszfeld algorithm for solving (P_{d_2})

Input. Existing locations $a^i = (a_1^i, a_2^i)^T \in \mathbb{R}^2$ (that are pairwise different from each other), weights $v^i > 0$, $i \in \{1,\ldots,m\}$, error bound $\epsilon > 0$.

1. Check if the minimum is attained at an existing location $a^i, i \in \{1, \ldots, m\}$. If for

$$\text{Test}_k := \Big[\Big(\sum_{i \in \mathcal{M} \setminus \{k\}} v^i \frac{a_1^k - a_1^i}{d_2(a^k, a^i)} \Big)^2 + \Big(\sum_{i \in \mathcal{M} \setminus \{k\}} v^i \frac{a_2^k - a_2^i}{d_2(a^k, a^i)} \Big)^2 \Big]^{\frac{1}{2}},$$

   it holds that

$$\boxed{\text{Test}_k \leq v^k}$$

   for one $k \in \mathcal{M} \implies$ Set $x^* = a^k$ and END. Otherwise go to Step 2.

2. Choose a starting point $x = (x_1, x_2)$. This point can be found, for example, by solving the median problem using the squared Euclidean norm.

3. Set for $j = 1, 2$: $x_j^{\text{new}} = \dfrac{\sum_{i=1}^m v^i \frac{a_j^i}{d_2(a^i,x)}}{\sum_{i=1}^m v^i \frac{1}{d_2(a^i,x)}}$.

4. If $x^{\text{new}} = (x_1^{\text{new}}, x_2^{\text{new}})$ satisfies a stopping criterion w.r.t. the given ϵ (see Section 3.1 below) $\implies$ Set $x^* = x^{\text{new}}$. Otherwise: Set $x = x^{\text{new}}$ and go to Step 3.

Output. Approximation of the minimizer $x^* = (x_1^*, x_2^*)$.

The main iteration scheme in the algorithm is

$$\begin{aligned} x_1^{new} &:= \frac{\sum_{i=1}^m v^i \frac{a_1^i}{d_2(a^i,x)}}{\sum_{i=1}^m v^i \frac{1}{d_2(a^i,x)}} \\ x_2^{new} &:= \frac{\sum_{i=1}^m v^i \frac{a_2^i}{d_2(a^i,x)}}{\sum_{i=1}^m v^i \frac{1}{d_2(a^i,x)}}. \end{aligned} \tag{3.1}$$

3.1 Termination Criterion

Let us define $f_{d_2}(x) := \sum_{i \in \mathcal{M}} v^i d_2(a^i, x)$. Let $x \in \mathbb{R}^2$ be an iteration point that was obtained during the Weiszfeld algorithm. We define

$$\sigma(x) := \max \{ d_2(x, y) : \; y \in \text{conv}\{a^1, \ldots, a^m\} \}$$

with $\text{conv}\{a^1, \ldots, a^m\} := \{x \in \mathbb{R}^2 : \; \exists \lambda^1 \geq 0, \ldots, \lambda^m \geq 0, \; \sum_{i=1}^m \lambda^i = 1, x = \sum_{i=1}^m \lambda^i a^i\}$ being the *convex hull*.

Theorem 2. *Let $x \in \mathbb{R}^2$ be an iteration point that was obtained during the Weiszfeld algorithm. Then*

$$UB(x) := ||\nabla f_{d_2}(x)|| \cdot \sigma(x)$$

is an upper bound *for a possible improvement of the objective function $f_{d_2}(x)$, where $||\cdot||$ is the Euclidean norm in $\mathbb{R}^2$.*

Proof. Because f_{d_2} is convex, we have for an iteration point x obtained during the Weiszfeld algorithm

$$\begin{aligned} f_{d_2}(x) - f_{d_2}(x^*) &\leq \langle \nabla f_{d_2}(x), x - x^* \rangle &(3.2)\\ &\overset{Cauchy-Schwartz}{\leq} ||\nabla f_{d_2}(x)|| \cdot ||x - x^*|| &(3.3)\\ &\leq ||\nabla f_{d_2}(x)|| \cdot \sigma(x). &(3.4) \end{aligned}$$

□

Theorem 3. *Let x^* denote the global minimizer of (P_{d_2}). A* lower bound *for $f_{d_2}(x^*)$ is given by*

$$LB(x) := f_{d_2}(x) - ||\nabla f_{d_2}(x)|| \cdot \sigma(x).$$

Proof. We have from Theorem 2

$$f_{d_2}(x) - f_{d_2}(x^*) \leq ||\nabla f_{d_2}(x)|| \cdot \sigma(x),$$

and therefore

$$f_{d_2}(x) - ||\nabla f_{d_2}(x)|| \cdot \sigma(x) \leq f_{d_2}(x^*).$$

□

Based on these two results, show that the following consequence holds:

Theorem 4. *If for the lower bound, we have $LB(x) > 0$ and for a given $\epsilon > 0$*

$$\frac{UB(x)}{LB(x)} = \frac{||\nabla f_{d_2}(x)|| \cdot \sigma(x)}{f_{d_2}(x) - ||\nabla f_{d_2}(x)|| \cdot \sigma(x)} < \epsilon$$

holds true, then the current iterate has at least a relative accuracy[1] *of ϵ.*

3.2 Project Tasks - Weiszfeld Algorithm

1. Explain why the main iteration scheme is chosen according to (3.1) [3 points].
2. Prove Theorem 4 above using Theorems 2 and 3 [5 points].
3. Derive a termination criterion for the Weiszfeld algorithm based on Theorem 4 [2 points].
4. Implement the Weiszfeld algorithm using Python3. Give the pseudo-code and a test example [20 points].
5. Implement the gradient descent method with backtracking and replace steps 2.-4. in the above algorithm with it. Compare with the Weiszfeld algorithm and discuss the difference between the two algorithms and their performance. Which algorithm would you suggest, and why? [20 points]

*Note: Please submit the project by **March 13th, 2022**, via your groups in Blackboard. When submitting your project, please specify who in your group contributed to which project points (or, if this is the case, whether everyone in the group contributed to all project points).*

4 Suggested Reading Material / Additional Resources

1. R. F. Love, J. G. Morris and G. O. Wesolowsky: Facility Location: Models and Methods, North Holland, New York, 1988. The relevant pages of this reference can be found here: `http://web.tecnico.ulisboa.pt/mcasquilho/compute/_scicomp/_location/LoveMorrisWesolowsky.pdf`
2. A. Beck, S. Sabach, S. Weiszfeld's Method: Old and New Results. J Optim Theory Appl 164, 1–40, 2015. `https://doi.org/10.1007/s10957-014-0586-7`
3. Z. Drezner, H. W. Hamacher (Editors): Facility Location. Applications and Theory. Springer-Verlag Berlin Heidelberg, 2002.
4. Reza Zanjirani Farahani, Masoud Hekmatfar (Editors): Facility Location: Concepts, Models, Algorithms and Case Studies. Springer Dordrecht Heidelberg London New York, 2009.
5. `https://project-flo.de` (Project Facility Location Optimizer: A MATLAB-based software tool for solving location problems).

[1]The *relative accuracy* of the current iterate x is defined as $\left|\frac{f_{d_2}(x) - f_{d_2}(x^*)}{f_{d_2}(x^*)}\right|$.