Chapter 2 Location estimation based on the minimum distance approach

In this chapter, we determine the facility location using the minimum distance approach. Note that we use "location estimation" in the title since we use the statistical approach (i.e. weighted median) a lot in this chapter, where the "location estimation" corresponds to the "location determination" in operations research.

2.1 Objective functions based on the distances

In mathematics, a distance function is a generalization of the concept of physical distance. In physics or everyday usage, the distance may refer to a physical length of how far apart objects are. Virtually all continuous facility location models are concerned with the distance between points (their coordinates). In this chapter, we briefly introduce the general distance formulation and investigate what distance measures are usually applied in the facility location problem.

2.1.1 Distance approach

The distance between two points is usually measured by L^p norm. Let d be the dimensional number. Given a point X=(x_1, x_2, …, x_d) and another point Y= (y_1, y_2, …, y_d), the L^p-norm distance, also known as Minkowski distance, is defined as:

$$L^p = \left(\sum_{j=1}^{d} \left|x_j - y_j\right|^p\right)^{\frac{1}{p}} \tag{2.1}$$

Two popular distances are the Manhattan distance (L^1 norm) and Euclidean distance (L^2 norm). To provide more flexibility of the above L^p-norm distance, two parameters (i.e., p and q) are used and the distance function is defined as

$$L^{(p,q)} = \left(\sum_{j=1}^{d} \left|x_j - y_j\right|^p\right)^{q} \tag{2.2}$$

It is obvious that $L^{(p=1,q=1)}$ is Manhattan distance (L^1 norm) and $L^{(p=2,q=\frac{1}{2})}$ is Euclidean distance (L^2 norm). In addition, the squared Euclidean distance (squared L^2 norm) is also famous in previous studies, which is $L^{(p=2,q=1)}$. As an illustration, we provide three different distances in a two-dimensional plane in Fig. 2.1. Specifically, the Manhattan distance (L^1 norm), Euclidean distance (L^2 norm), and squared Euclidean distance (squared L^2 norm) are presented as follows.

(1) Manhattan distance

As shown in Fig. 2.1(a), the Manhattan distance (L^1 norm) is applied in the urban area of interest due to the street configuration (Chiu and Chen 2009; Gao et al. 2019). For instance, the determination of the bicycle-sharing stations or the ambulance stations needs to consider the Manhattan distance. The travel distance $L^{(1,1)}(A,F)$ from point A to point F is the same as the sum of each street distance (i.e., A to B, B to C, C to D, D to E), as shown in Fig. 2.1(a). It is obvious that the $L^{(1,1)}(A,F)$ from the lower-left corner A to the upper-right corner F only allows moving rightward and upward among the number of meshes. Then the Manhattan distance $L^{(1,1)}(A,F)$ is given by

$$L^{(1,1)}(A,F) = L^{(1,1)}(A,B) + L^{(1,1)}(B,C) + L^{(1,1)}(C,D) + L^{(1,1)}(D,E) + L^{(1,1)}(E,F) \tag{2.3}$$

(2) Euclidean distance

Fig. 2.1(b) shows an illustration of the Euclidean distance (L^2 norm) that is the straight-line distance between two points, which is widely applied in the rural area of interest, long-distance cases, and signal propagation cases (Esnaf and Küçükdeniz 2009; Chanta et al. 2014). For instance, refer to the location determination of 4/5G stations or plants within a large-scale area. The travel distance $L^{(2,0.5)}(A,F)$ from the lower-left corner A to the upper-right corner F is the sum of lengths of the straight-line distances connecting those two points via points B, C, D, and E that are straight on this line. It is obvious that the Euclidean distance from A to F only allows moving along the straight line to have the same distance value. Then the Euclidean distance $L^{(2,0.5)}(A,F)$ is given by

$$L^{(2,\frac{1}{2})}(A,F) = L^{(2,\frac{1}{2})}(A,B) + L^{(2,\frac{1}{2})}(B,C) + L^{(2,\frac{1}{2})}(C,D) + L^{(2,\frac{1}{2})}(D,E) + L^{(2,\frac{1}{2})}(E,F) \tag{2.4}$$

(3) Squared Euclidean distance

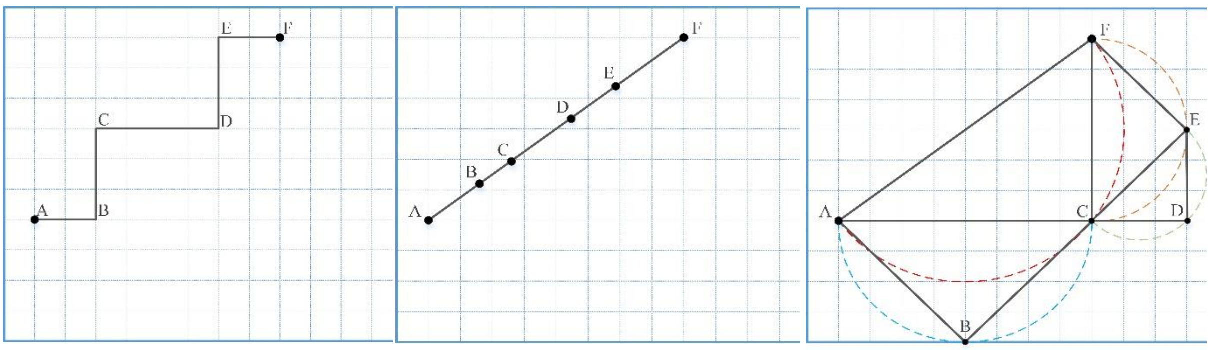
As shown in Fig. 2.1(c), $L^{(2,1)}(A,F)$ is the *squared Euclidean* distance (squared L^2 norm) from points A to F, it can be considered as the squared length of the straight-line distance connecting those two points, which is given by

$$L^{(2,1)}(A,F) = L^{(2,1/2)}(A,F)^2 \tag{2.5}$$

Note that $L^{(2,1)}(A,F)$ can be replaced by $L^{(2,1)}(A,C) + L^{(2,1)}(C,F)$ due to the Pythagorean Theorem. Similarily, $L^{(2,1)}(A,C)$ can be replaced by $L^{(2,1)}(A,B) + L^{(2,1)}(B,C)$ and $L^{(2,1)}(C,F)$ can be replaced by $L^{(2,1)}(C,E) + L^{(2,1)}(E,F)$, where $L^{(2,1)}(C,E)$ is the summation of $L^{(2,1)}(C,D)$ and $L^{(2,1)}(D,E)$. Thus, we can rewrite $L^{(2,1)}(A,F)$ as

$$\begin{aligned} L^{(2,1)}(A,F) &= L^{(2,1)}(A,C) + L^{(2,1)}(C,F) \\ &= L^{(2,1)}(A,B) + L^{(2,1)}(B,C) + L^{(2,1)}(C,E) + L^{(2,1)}(E,F) \\ &= L^{(2,1)}(A,B) + L^{(2,1)}(B,C) + L^{(2,1)}(C,D) + L^{(2,1)}(D,E) + L^{(2,1)}(E,F) \\ &= \cdots \end{aligned} \tag{2.6}$$

It is obvious that the travel distance would be formulated by the squared Euclidean distance. However, as shown in Fig. 2.1(c), it forms unreasonable tracks from points A to F via points B, C, D, and E.



(a) Manhattan distance (b) Euclidean distance (c) Squared Euclidean distance

Fig. 2.1. Illustration of three distances

2.1.2 Metric

In mathematics, a distance function defines a measure between each pair of elements of a set. A metric is a distance satisfying all the conditions (i)-(iv) below. A set with a metric is called a metric space, for more details, see Section 1.4 in Rudin (1987). To check the validity of different distance functions, we need to see if they satisfy the metric properties. According to the distance metrics from Parnas and Ron (2003), when $\mathfrak{x}$, $\mathfrak{y}$, and $\mathfrak{z}$ are the points within the feasible region, we say that a metric should hold the following four conditions:

(i) Non-negativity: $d(\mathfrak{x}, \mathfrak{y}) \geq 0$.

(ii) Identity of indiscernible: $d(\mathfrak{x}, \mathfrak{y}) = 0 \Leftrightarrow \mathfrak{x} = \mathfrak{y}$.

(iii) Symmetry: $d(\mathfrak{x}, \mathfrak{y}) = d(\mathfrak{y}, \mathfrak{x})$.

(iv) Triangle inequality: $d(\mathfrak{x}, \mathfrak{y}) \leq d(\mathfrak{x}, \mathfrak{z}) + d(\mathfrak{z}, \mathfrak{y})$.

It is easily seen that Manhattan and Euclidean distances hold for all of the above conditions. However, the squared Euclidean distance only holds for the conditions (i), (ii), and (iii). There is always the following common equation for the squared Euclidean distance:

$$L^{(2,1)}(\mathfrak{x}, \mathfrak{y}) = L^{(2,1)}(\mathfrak{x}, \mathfrak{z}) + L^{(2,1)}(\mathfrak{z}, \mathfrak{y}) - 2 \times \cos\theta \times \sqrt{L^{(2,1)}(\mathfrak{z}, \mathfrak{y}) \times L^{(2,1)}(\mathfrak{x}, \mathfrak{z})} \tag{2.7}$$

where θ is the angle of two lines $(\mathfrak{z}, \mathfrak{y})$ and $(\mathfrak{x}, \mathfrak{z})$. While $90^{\circ} < \theta < 180^{\circ}$, it is easy to have the following inequality:

$$L^{(2,1)}(\mathfrak{x}, \mathfrak{y}) > L^{(2,1)}(\mathfrak{x}, \mathfrak{z}) + L^{(2,1)}(\mathfrak{z}, \mathfrak{y}) \tag{2.8}$$

The above inequality in (2.8) conflicts with the above condition (iv). Then we can conclude that the squared Euclidean distance is not a metric. It should be noted that, in statistics, many statistical distance functions satisfy only conditions (i) and (ii). For more details, see Chapter 2 of Basu et al. (2011).

2.1.3 Objective functions

Based on the above three distances, we summarize their general objective functions (minimization of the total weighted distance) in a two-dimensional plane. Here, two common assumptions need to be declared:

(i) The demand points with different priority levels (weights) and locations are known.

(ii) The linear relationship between travel cost and distance.

The notations used in the proposed models are given as follows.

Parameters

n Number of demand points, ($i = 1,2, \dots, n$).

x_i x-coordinate of demand point i.

y_i y-coordinate of demand point i.

z_i z-coordinate of demand point i.

$\mathcal{w}_i$ Weight (Priority level) of demand point i.

Decision variables

u x-coordinate of the facility.

v y-coordinate of the facility.

w z-coordinate of the facility.

Generally, the objective function needs to be minimized so that the total weighted distance-related travel time or cost is optimized. Before the objectives are presented, we normalize the weight using the following equation

$$\omega_i = \frac{\mathcal{w}_i}{\sum_{i=1}^{n} \mathcal{w}_i} \tag{2.9}$$

where ω_i is the normalized weight of demand point i.

- *Manhattan distance-based objective function*

When the Manhattan distance is considered in the two-dimensional single-facility location problem, the objective is to minimize the total weighted Manhattan distance Obj_1, which is given by

$$\mathrm{Obj}_1 = \sum_{i=1}^{n} \omega_i(|x_i - u| + |y_i - v|) \tag{2.10}$$

where $|x_i - u| + |y_i - v|$ is the distance measured along the axis at right angles, which is usually applied in the urban area of interest (Chiu and Chen 2009; Gao et al. 2019). It should be noted that the normalized weight influences the objective function value, but it would not affect its final solution. Similarly, the same situation happens in other objective functions.

- *Euclidean distance-based objective function*

When the Euclidean distance is considered in the two-dimensional single-facility location problem, the objective is to minimize the total weighted Euclidean distance Obj_2, which is given by

$$\mathrm{Obj}_2 = \sum_{i=1}^{n} \omega_i\sqrt{(x_i - u)^2 + (y_i - v)^2} \tag{2.11}$$

where $\sqrt{(x_i - u)^2 + (y_i - v)^2}$ is the straight-line distance between two points, which is widely applied in the rural area of interest and long-distance cases (Esnaf and Küçükdeniz 2009; Chanta et al. 2014).

- *Squared Euclidean distance-based objective function*

Next, we consider the squared Euclidean distance for the two-dimensional single-facility location problem. The objective is to minimize the weighted squared Euclidean distance Obj_3, which is given by

$$\mathrm{Obj}_3 = \sum_{i=1}^{n} \omega_i[(x_i - u)^2 + (y_i - v)^2] \tag{2.12}$$

where $(x_i - u)^2 + (y_i - v)^2$ is the squared value of the straight-line distance between two points.

- *General objective function*

Based on the $L^{(p,q)}$ distance function defined in (2.2) of Section 2.1.1, we propose the general objective function for the two-dimensional single-facility location problem, which is given by

$$\mathrm{Obj}_{(p,q)} = \sum_{i=1}^{n} \omega_i (|x_i - u|^p + |y_i - v|^p)^q \tag{2.13}$$

where p and q are nonnegative values. Then the above objective functions Obj_1 in (2.10), Obj_2 in (2.11), and Obj_3 in (2.12) are the special cases of $\mathrm{Obj}_{(p,q)}$.

2.2 Minimum distance approach

In this chapter, we provide the two-dimensional facility location methods for the three different distances introduced in Section 2.1 and evaluate their performances using the robustness properly, relative efficiency, and optimized objective function values.

2.2.1 Minimization of the total weighted Manhattan distance

Previously, some studies have investigated the minimizer of the Obj_1 in (2.10) using different methods (Francis et al. 1992; Sule 2001; Heragu 2008). However, the intrinsic principles and properties of the minimizer have received insufficient attention. Therefore, we determine the minimizer of the Obj_1 in (2.10) from the perspective of the statistical viewpoint and investigate the properties of the minimizer, which will be detailed later. With the objective function of Obj_1 in (2.10), let $(\tilde{u}, \tilde{v})$ be the optimal facility location. Then we have

$$(\tilde{u}, \tilde{v}) = \underset{(u,v)}{\operatorname{argmin}}(\mathrm{Obj}_1) = \underset{(u,v)}{\operatorname{argmin}}\left[\sum_{i=1}^{n} \omega_i(|x_i - u| + |y_i - v|)\right] \tag{2.14}$$

The minimizer of Obj_1 in (2.10), which is denoted by $(\tilde{u}, \tilde{v})$, can be obtained by the following estimating equations that need to be solved for both u and v according to Section 1.3 of Hettmansperger and McKean (2010).

$$\frac{\partial \mathrm{Obj}_1}{\partial u} = \sum_{i=1}^{n} \omega_i \operatorname{sgn}(x_i - u) = 0 \tag{2.15}$$

$$\frac{\partial \mathrm{Obj}_1}{\partial v} = \sum_{i=1}^{n} \omega_i \operatorname{sgn}(y_i - v) = 0 \tag{2.16}$$

Then the optimal values $\tilde{u}$ and $\tilde{v}$ can be calculated separately. It is easily seen that $\tilde{u}$ is the weighted median of the x-axis observations and $\tilde{v}$ is the weighted median of the y-axis observations, which will be detailed later. Note that the weighted median was first suggested by Edgeworth (1888) and since then it has been widely used in many applications (Vazler et al. 2012). As an illustration, we briefly introduce the conventional median. Then we introduce the weighted median. As the values $\tilde{u}$ and $\tilde{v}$ can be obtained separately, we consider the weighted median for the x-axis observations. Then the weighted median for the y-axis observations is easily obtained using the same method. Because the median can be obtained by using the empirical cumulative distribution function, we briefly introduce the definition of the empirical cumulative distribution function which was introduced in Definition 2.1 of Owen (2001). Then we propose a new definition that can consider the weights.

Definition 2.1. Given a set of observations $x_1, x_2, \ldots, x_n$, the empirical cumulative distribution function F_n is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{I}(x_i \leq x), \qquad x \in \mathbb{R} \tag{2.17}$$

where $\mathrm{I}(A)$ represents the indicator function defined as

$$\mathrm{I}(A) = \begin{cases} 1 & \text{if } A \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

Definition 2.2. Csorgo (1983) defined the left sample quantile function (inverse cumulative distribution function) as below.

$$F_{n,L}{}^{-1}(p) = \inf\{x: F_n(x) \geq p\} \tag{2.19}$$

Using this, the conventional median $F_n{}^{-1}(1/2)$ is obtained as:

$$F_{n,L}{}^{-1}\left(\frac{1}{2}\right) = \inf\left\{x: F_n(x) \geq \frac{1}{2}\right\} = x_{(k)} \qquad \text{if } \frac{k-1}{n} < \frac{1}{2} \leq \frac{k}{n}, k = 1,2,\ldots,n \tag{2.20}$$

where $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$.

Definition 2.3. Wasserman (2013) defined the right sample quantile function (inverse cumulative distribution function), which is given by

$$F_{n,R}{}^{-1}(p) = \inf\{x: F_n(x) > p\} \tag{2.21}$$

Note that the definition by Wasserman (2013) is slightly different from that by Csorgo (1983). It is easily seen that

$$\inf\{x: F_n(x) \geq p\} \leq \inf\{x: F_n(x) > p\} \tag{2.22}$$

Thus, the sample quantile by Csorgo (1983) is called the left quantile, while the sample quantile by Wasserman (2013) is the right quantile. Rychlik (2001) and Hosseini (2010) showed that

$$F_{n,R}{}^{-1}(p) = \inf\{x: F_n(x) > p\} = \sup\{x: F_n(x) \leq p\} \tag{2.23}$$

Based on the quantile function by Wasserman (2013), we have the corresponding median $F_n{}^{-1}(1/2)$, which is obtained as:

$$F_{n,R}{}^{-1}\left(\frac{1}{2}\right) = \inf\left\{x: F_n(x) > \frac{1}{2}\right\} = x_{(k+1)} \qquad \text{if } \frac{k-1}{n} < \frac{1}{2} \leq \frac{k}{n}, k = 1,2,\dots,n \tag{2.24}$$

where $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$.

It is obvious that the difference between (2.19) and (2.21) is the case when $F_n(1/2) = 0.5$ with even n. It is obvious that (2.19) takes the left bound value and (2.21) takes the right bound value. And both of them are the minimizers (medians) to the total Manhattan distance of the x-axis observations.

The above definitions on the empirical distribution and the sample quantile do not consider the weights of the observations. Thus, we suggest the new definitions below.

Definition 2.4: Given a set of observations x_1, x_2, …, x_n with corresponding positive weights ω_1, ω_2, …, ω_n such that $\sum_{i=1}^{n} \omega_i = 1$, we have the empirical cumulative distribution function $G_n(x)$ with weights, which is defined as

$$G_n(x) = \sum_{i=1}^{n} \omega_i \mathrm{I}(x_i \leq x) \tag{2.25}$$

Note that the above G_n includes the conventional empirical cumulative distribution function F_n in (2.4), as a special case when $\omega_i = 1/n$. Similar to the definition of the sample quantile function in Csorgo (1983), we define the sample quantile function with weights.

Definition 2.5: Given a set of observations x_1, x_2, …, x_n with corresponding positive weights ω_1, ω_2, …, ω_n such that $\sum_{i=1}^{n} \omega_i = 1$, the sample left and right quantiles with weights are given by

$$G_{n,L}{}^{-1}(p) = \inf\{x: G_n(x) \geq p\} \tag{2.26}$$

$$G_{n,R}{}^{-1}(p) = \inf\{x: G_n(x) > p\} \tag{2.27}$$

Next, our goal is to obtain the minimizer of the Obj_1 in (2.10). As we can calculate the weighted medians for x-axis and y-axis separately, we only focus on the weighted median of x-axis observations, which is given by

$$G_{n,L}{}^{-1}\left(\frac{1}{2}\right) = \inf\left\{x: G_n(x) \geq \frac{1}{2}\right\} = x_{(k)} \qquad \text{if} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} \leq \sum_{j=1}^{k} \omega_{(j)}, k = 1,2,\ldots,n \tag{2.28}$$

$$G_{n,R}{}^{-1}\left(\frac{1}{2}\right) = \inf\left\{x: G_n(x) > \frac{1}{2}\right\} = x_{(k+1)} \qquad \text{if} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} \leq \sum_{j=1}^{k} \omega_{(j)}, k = 1,2,\ldots,n \tag{2.29}$$

where $\omega_{(j)}$ is the weight for $x_{(j)}$.

According to the equations in (2.23) and (2.29), we can deduce the following equation. Then we have

$$G_{n,R}{}^{-1}\left(\frac{1}{2}\right) = \inf\left\{x: G_n(x) > \frac{1}{2}\right\} = \sup\left\{x: G_n(x) \leq \frac{1}{2}\right\} = x_{(k+1)} \tag{2.30}$$

It is obvious that the above $x_{(k)}$ minimizes the weighted L^1 norm (Manhattan distance). However, it is not a unique minimizer when there is a tied value at $\sum_{j=1}^{k} \omega_{(j)} = 1/2$. Thus, we consider two cases:

(i) $\sum_{j=1}^{k-1} \omega_{(j)} < 1/2$ and $1/2 < \sum_{j=1}^{k} \omega_{(j)}$

(ii) $\sum_{j=1}^{k-1} \omega_{(j)} < 1/2$ and $1/2 = \sum_{j=1}^{k} \omega_{(j)}$

To illustrate how to obtain the weighted median, we use four examples. In Examples 2.1 and 2.2, we consider five observations but with different weights. Then we consider six observations with different weights in Examples 2.3 and 2.4. The observations and the weights of these four examples are provided in Table 2.1. After sorting the observations with weights for each of the examples, it is easily seen that Examples 2.1 and 2.3 belong to Case (i) and Examples 2.2 and 2.4 belong to Case (ii).

Table 2.1 Datasets of four examples

ID	Example 2.1		Example 2.2		Example 2.3		Example 2.4	
	x_i	ω_i	x_i	ω_i	x_i	ω_i	x_i	ω_i
$i = 1$	1.5	0.1	1.5	0.1	0.5	0.1	0.5	0.1
$i = 2$	2.2	0.2	2.2	0.2	3.4	0.1	4.5	0.1
$i = 3$	4.4	0.2	3.6	0.2	5.6	0.1	5.6	0.1
$i = 4$	5.6	0.2	5.6	0.3	1.2	0.2	1.2	0.2
$i = 5$	3.6	0.3	4.4	0.2	4.5	0.2	2.6	0.2
$i = 6$	--	--	--	--	2.6	0.3	3.4	0.3

Example 2.1

As presented in Table 2.2, the sorted observations $x_{(1)}$, $x_{(2)}$, …, $x_{(5)}$ with their corresponding weights $\omega_{(1)}$, $\omega_{(2)}$, …, $\omega_{(5)}$ such that $\sum_{j=1}^{5} \omega_{(j)} = 1$ are provided based on Example 2.1.

Table 2.2 Sorted observations with weights in Example 2.1

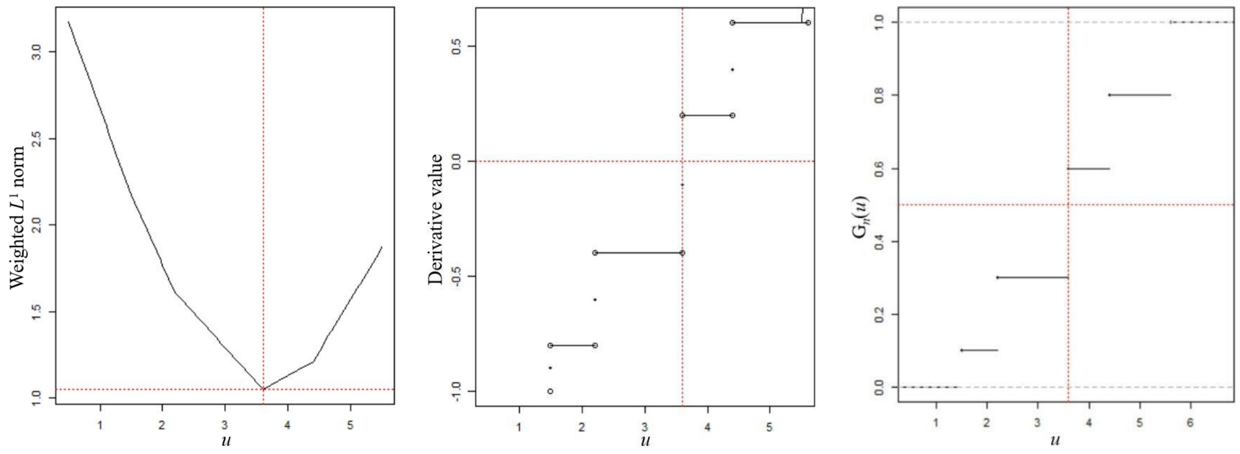
ID	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
$x_{(j)}$	1.5	2.2	3.6	4.4	5.6
$\omega_{(j)}$	0.1	0.2	0.3	0.2	0.2

Then we have the following sample quantile function:

$$G_{n,L}{}^{-1}\left(\frac{1}{2}\right) = G_{n,R}{}^{-1}\left(\frac{1}{2}\right) = x_{(k)} \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} < \sum_{j=1}^{k} \omega_{(j)} \tag{2.31}$$

where $n = 5$ and $k = 3$. Then the weighted median is given by $\tilde{u} = x_{(3)} = 3.6$.

To provide a clear visual illustration of the weighted median for Example 2.1, we present the relationship between u and the weighted L^1 norm (Manhattan distance) in Fig. 2.2(a). The weighted L^1 norm gets the minimum value when $\tilde{u} = x_{(3)} = 3.6$. The same result can be obtained with the method developed in Francis et al. (1992). Besides, we calculate the derivative value of the weighted L^1 norm respect to u in Fig. 2.2(b). It shows that the changing point is the weighted median ($\tilde{u}$=3.6) when the derivative value changes from negative to positive. Besides, Fig. 2.2(c) also shows the location of the weighted median $\tilde{u}$ associating with the empirical cumulative distribution function. Accordingly, we obtain the weighted median $\tilde{u} = x_{(3)}$=3.6 through the above equation (2.31), which is considered as the optimal facility location on the x-axis.



(a) (b) (c)

Fig. 2.2. Weighted median of x-axis observations in Example 2.1

In this case, we have the general formulation to calculate the weighted median, which is given by

$$\tilde{u} = \text{weighted median}(x_1, x_2, \ldots, x_n) = x_{(k)} \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} < \sum_{j=1}^{k} \omega_{(j)} \tag{2.32}$$

Example 2.2

As presented in Table 2.3, the sorted observations $x_{(1)}$, $x_{(2)}$, …, $x_{(5)}$ with their corresponding weights $\omega_{(1)}$, $\omega_{(2)}$, …, $\omega_{(5)}$ such that $\sum_{j=1}^{5} \omega_{(j)} = 1$ are provided based on Example 2.2.

Table 2.3 Sorted observations with weights in Example 2.2

ID	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
$x_{(j)}$	1.5	2.2	3.6	4.4	5.6
$\omega_{(j)}$	0.1	0.2	0.2	0.2	0.3

Then we have the following sample quantile function:

$$G_{n,L}{}^{-1}\left(\frac{1}{2}\right) < G_{n,R}{}^{-1}\left(\frac{1}{2}\right) \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.33}$$

Specifically, we have

$$G_{n,L}{}^{-1}\left(\frac{1}{2}\right) = x_{(k)} \quad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.34}$$

$$G_{n,R}{}^{-1}\left(\frac{1}{2}\right) = x_{(k+1)} \quad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.35}$$

where $n = 5$ and $k = 3$.

It is obvious there is a tied value at $\sum_{j=1}^{k} \omega_{(j)} = 1/2$ even though the number of observations is odd, which is different from the conventional median because there is only one value for the conventional median when the number of observations is odd. In this case, $x_{(3)}$ is not the unique minimizer of the weighted L^1 norm. To illustrate the weighted median, we present the relationship between u and the weighted L^1 norm of the x-axis in Fig. 2.3(a). As shown in Fig. 2.3(a), there is an infinite number of minimizers of the weighted L^1 norm. Any values between the left quantile $x_{(k)}$ and right quantile $x_{(k+1)}$ can minimize the weighted L^1 norm. Then the weighted median is given by

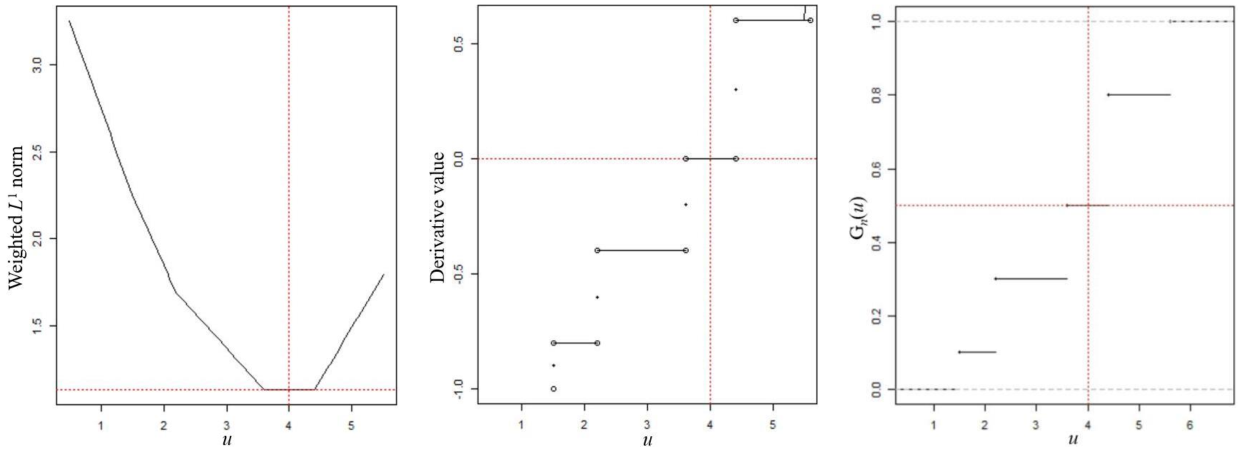
$$\tilde{u} = \lambda x_{(k)} + (1-\lambda) x_{(k+1)} \quad \text{while } 0 \leq \lambda \leq 1 \text{ and } \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.36}$$

In this thesis, we suggest that $\tilde{u}$ is the average value of the left and right quantiles, which is given by

$$\tilde{u} = \frac{x_{(k)} + x_{(k+1)}}{2} = \frac{3.6 + 4.4}{2} = 4.0 \tag{2.37}$$

Note that λ is not necessarily 0.5 in general, but it is reasonable to choose $\lambda = 0.5$ for the case of the quantile of 0.5.

Besides, as shown in Fig. 2.3 (b), the derivative value also has tied values when it changes from negative to positive. Both Figs. 2.3 (b) and (c) validate the suggested method in (2.37) to obtain the weighted median. Thus, the weighted median $\tilde{u} = 4.0$ is obtained that is considered as the optimal location of the facility on the x-axis.



(a) (b) (c)

Fig. 2.3. Weighted median of x-axis observations in Example 2.2

In this case, we have the general formulation to calculate the weighted median, which is given by

$$\tilde{u} = \text{weighted median}(x_1, x_2, \dots, x_n) = \frac{x_{(k)} + x_{(k+1)}}{2} \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.38}$$

Example 2.3

As presented in Table 2.4, the sorted observations $x_{(1)}$, $x_{(2)}$, …, $x_{(6)}$ with their corresponding weights $\omega_{(1)}$, $\omega_{(2)}$, …, $\omega_{(6)}$ such that $\sum_{j=1}^{6} \omega_{(j)} = 1$ are provided for Example 2.3.

Table 2.4 Sorted observations with weights in Example 2.3

ID	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$x_{(j)}$	0.5	1.2	2.6	3.4	4.5	5.6
$\omega_{(j)}$	0.1	0.2	0.3	0.1	0.2	0.1

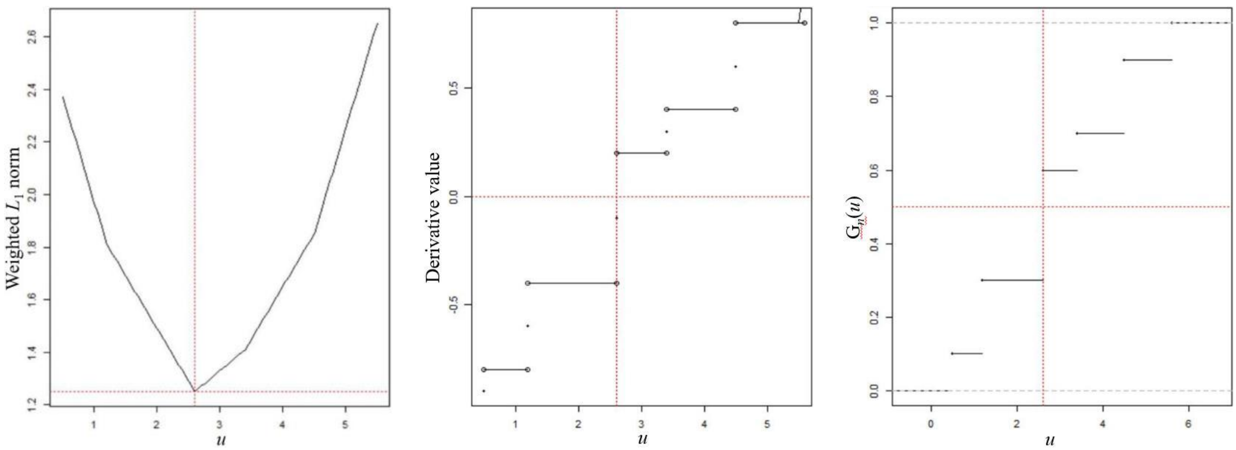
Then we have the following sample quantile function:

$$G_{n,L}{}^{-1}\left(\frac{1}{2}\right) = G_{n,R}{}^{-1}\left(\frac{1}{2}\right) = x_{(k)} \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} < \sum_{j=1}^{k} \omega_{(j)} \tag{2.39}$$

where $n = 6$ and $k = 3$. This result is $x_{(3)}$=2.6.

To provide a clear visual illustration of the weighted median for Example 2.3, we present the relationship between u and the weighted L^1 norm (Manhattan distance) in Fig. 2.4(a). The weighted L^1

norm gets the minimum value when $\tilde{u} = x_{(3)}$=2.6. It should be noted that there is only one minimizer when the number of observations is even, while there is an infinite number of values for the conventional median when the number of observations is even. Besides, we calculate the derivative value of the weighted L^1 norm respect to u in Fig. 2.4(b). It shows that the changing point is the weighted median ($\tilde{u}$=2.6) when the derivative value changes from negative to positive. Besides, Fig. 2.4(c) also shows the location of the weighted median $\tilde{u}$ associating with the empirical cumulative distribution function. Accordingly, we obtain the weighted median $\tilde{u} = x_{(3)}$=2.6 through the above equation (2.39), which is considered as the optimal facility location on the x-axis. We have the same solution using (2.32).



(a) (b) (c)

Fig. 2.4. Weighted median of x-axis observations in Example 2.3

Example 2.4

As presented in Table 2.5, the sorted observations $x_{(1)}$, $x_{(2)}$, …, $x_{(6)}$ with their corresponding weights $\omega_{(1)}$, $\omega_{(2)}$, …, $\omega_{(6)}$ such that $\sum_{j=1}^{6} \omega_{(j)} = 1$ are provided for Example 2.4.

Table 2.5 Sorted observations with weights in Example 2.4

ID	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$x_{(j)}$	0.5	1.2	2.6	3.4	4.5	5.6
$\omega_{(j)}$	0.1	0.2	0.2	0.3	0.1	0.1

Based on the formulations from (2.33) to (2.35), we have the following sample quantile function:

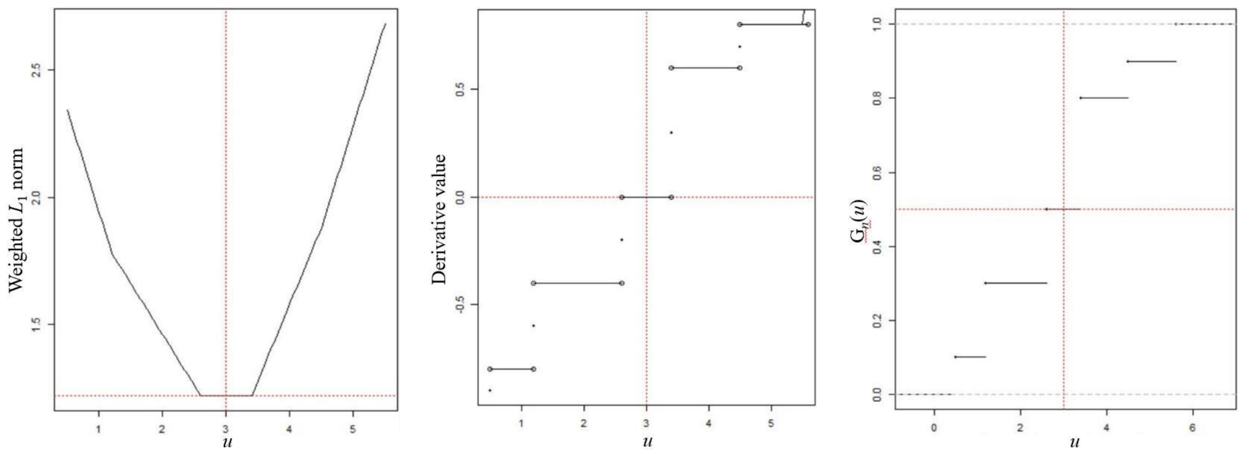
$$G_{n,L}{}^{-1}\left(\frac{1}{2}\right) = x_{(k)} < G_{n,R}{}^{-1}\left(\frac{1}{2}\right) = x_{(k+1)} \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.40}$$

where $n = 6$ and $k = 3$.

According to the suggested formulation in (2.38), we have the final weighted median, which is given by

$$\tilde{u} = \frac{x_{(k)} + x_{(k+1)}}{2} = \frac{x_{(3)} + x_{(4)}}{2} = 3.0 \qquad \text{while} \sum_{j=1}^{k-1} \omega_{(j)} < \frac{1}{2} = \sum_{j=1}^{k} \omega_{(j)} \tag{2.41}$$

It is obvious that there is a tied value at $\sum_{j=1}^{k} \omega_{(j)} = 1/2$. We present the relationship between u and the weighted L^1 norm of the x-axis in Fig. 2.5(a). There is an infinite number of minimizers of the weighted L^1 norm. Any values between the left quantile $x_{(3)}$ and right quantile $x_{(4)}$ can minimize the weighted L^1 norm. Besides, we calculate the derivative value of the weighted L^1 norm respect to u in Fig. 2.5(b). It shows that the changing point is the weighted median ($\tilde{u}$=3.0) when the derivative value changes from negative to positive. Besides, Fig. 2.5(c) also shows the location of the weighted median $\tilde{u}$ associating with the empirical cumulative distribution function.



(a) (b) (c)

Fig. 2.5. Weighted median of x-axis observations in Example 2.4

As seen in Examples 2.1, 2.2, 2.3, and 2.4, the suggested formula in (2.36) can be used for all cases: (i) $\sum_{j=1}^{k-1} \omega_{(j)} < 0.5$ and $\sum_{j=1}^{k-1} \omega_{(j)} = 0.5$. In this thesis, the minimizer (optimal location) of the objective functions can be calculated through the formula in (2.36). This statistical approach is much easier to determine the weighted median than the previous methods. Besides, with the weighted median calculated in (2.36), it is possible to investigate the properties of the minimizers, which will be detailed later.

2.2.2 Minimization of the total weighted Euclidean distance

When the Euclidean distance (L^2 norm) is considered in this model, the goal is to minimize Obj_2 in (2.11) so that the optimal location $(\tilde{u}, \tilde{v})$ is given by the minimizer of Obj_2 in (2.11), which is denoted by

$$(\tilde{u}, \tilde{v}) = \underset{(u,v)}{\operatorname{argmin}}(\mathrm{Obj}_2) = \underset{(u,v)}{\operatorname{argmin}}\left[\sum_{i=1}^{n} \omega_i(|x_i - u|^2 + |y_i - v|^2)^{\frac{1}{2}}\right] \tag{2.42}$$

Unlike the Manhattan distance, the optimal location based on the Obj_2 in (2.11) is not in explicit form. Thus, we have to obtain it by using numerical methods. Here, we use the quasi-Newton method that is also called the BFGS method (Broyden 1972; Fletcher 1970; Goldfarb 1970; Shanno 1970) to determine the optimal facility location so that the total weighted Euclidean distance can be minimized. For specific details on the BFGS method, the reader can be referred to Byrd et al. (1995). Note that the quasi-Newton method can easily find the optimal facility location due to the convexity of Obj_2 in (2.11) and it has been implemented in the R language developed by R Core Team (2019). Then we provide the proof of the convexity of Obj_2.

Lemma 2.1. *The eigenvalues of the Hessian matrix* $\mathbf{H}_i$ *are given by* 0 *and* $1/f_i$, *where* $\mathbf{H}_i$ *is constructed based on the* $f_i(u, v) = (|x_i - u|^2 + |y_i - v|^2)^{1/2}$. *The Hessian matrix* $\mathbf{H}_i$ *is given by*

$$\mathbf{H}_i = \frac{1}{f_i{}^3}\begin{bmatrix} \delta_y{}^2 & -\delta_x\delta_y \\ -\delta_y\delta_x & \delta_x{}^2 \end{bmatrix} \tag{2.43}$$

where $\delta_x = (x_i - u)$ and $\delta_y = (y_i - v)$.

Proof

To obtain the eigenvalues, we need to construct the following equation

$$\det(\mathbf{H}_i - \lambda^*\mathbf{I}_2) = \frac{1}{f_i{}^6}\det\begin{pmatrix} \delta_y{}^2 - \lambda & -\delta_x\delta_y \\ -\delta_y\delta_x & \delta_x{}^2 - \lambda \end{pmatrix} = \frac{\lambda}{f_i{}^6}\left(\lambda - \delta_x{}^2 - \delta_y{}^2\right) = 0 \tag{2.44}$$

where $\lambda^* = \lambda/f_i{}^3$.

Obviously, we obtain the eigenvalues, which are given by

$$\lambda_2{}^* = 0, \lambda_1{}^* = \frac{\delta_x{}^2 + \delta_y{}^2}{f_i{}^3} = \frac{1}{f_i} \tag{2.45}$$

Theorem 2.2. *The Obj_2 in (2.11) is convex*

Proof

As obtained in the above **Lemma 2.1**, it is obvious that all the eigenvalues for the Hessian matrix $\mathbf{H}_i$ are nonnegative numbers. Thus, the Hessian matrix $\mathbf{H}_i$ is positive semidefinite, which results in the convexity of $f_i(u, v)$. Thus, Obj_2 in (2.11) is convex. For more details, see Section 3.4 of Boyd and Vandenberghe (2004) and Theorem 2.38 of Beck (2014). Then we complete the proof.

Based on the above **Theorem 2.2**, we can obtain the minimizer of Obj_2 in (2.11), denoted by $(\tilde{u}, \tilde{v})$, through solving the following system of equations:

$$\frac{\partial \mathrm{Obj}_2}{\partial u} = \sum_{i=1}^{n} \frac{\omega_i (u - x_i)}{\sqrt{(x_i - u)^2 + (y_i - v)^2}} = 0, \qquad u \neq x_i, v \neq y_i \tag{2.46}$$

$$\frac{\partial \mathrm{Obj}_2}{\partial v} = \sum_{i=1}^{n} \frac{\omega_i (v - y_i)}{\sqrt{(x_i - u)^2 + (y_i - v)^2}} = 0, \qquad u \neq x_i, v \neq y_i \tag{2.47}$$

Then the optimal location $(\tilde{u}, \tilde{v})$ should be calculated using numerical methods. In this thesis, we use the quasi-Newton method to determine the facility location.

2.2.3 Minimization of the total weighted squared Euclidean distance

Next, we consider the squared Euclidean distance (squared L^2 norm) for the facility location problem. The goal is to minimize Obj_3 in (2.12) so that the optimal location $(\tilde{u}, \tilde{v})$ is given by the minimizer of Obj_3, which is denoted by

$$(\tilde{u}, \tilde{v}) = \underset{(u,v)}{\mathrm{argmin}}(\mathrm{Obj}_3) = \underset{(u,v)}{\mathrm{argmin}} \left[\sum_{i=1}^{n} \omega_i (|x_i - u|^2 + |y_i - v|^2) \right] \tag{2.48}$$

After taking the derivative of the Obj_3 in (2.12) with respect to u and v, we obtain the following system of equations:

$$\frac{\partial \mathrm{Obj}_3}{\partial u} = \sum_{i=1}^{n} 2\omega_i (u - x_i) = 0 \tag{2.49}$$

$$\frac{\partial \mathrm{Obj}_3}{\partial v} = \sum_{i=1}^{n} 2\omega_i (v - y_i) = 0 \tag{2.50}$$

The above system of equations is easily solved so that we obtain the optimal facility location, which is explicitly given by

$$\tilde{u} = \text{weighted mean}(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} \omega_i x_i \tag{2.51}$$

$$\tilde{v} = \text{weighted mean}(y_1, y_2, \dots, y_n) = \sum_{i=1}^{n} \omega_i y_i, \tag{2.52}$$

Note that the optimal facility location based on the weighted mean is also known as the CG because $(\tilde{u}, \tilde{v})$ is the location of fulcrum when weights are located at the demand points. This CG method is widely used in previous studies due to its simple calculation.

2.2.4 Minimization of the total weighted general distance

Finally, we consider the general distance for the facility location problem. Our goal is to minimize $\mathrm{Obj}_{(p,q)}$ in (2.13) and determine the minimizer $(\tilde{u}, \tilde{v})$, which is obtained as

$$(\tilde{u}, \tilde{v}) = \underset{(u,v)}{\operatorname{argmin}}\left(\mathrm{Obj}_{(p,q)}\right) = \underset{(u,v)}{\operatorname{argmin}}\left[\sum_{i=1}^{n} \omega_i \left(|x_i - u|^p + |y_i - v|^p\right)^q\right] \tag{2.53}$$

After taking the derivative of $\mathrm{Obj}_{(p,q)}$ in (2.13) with respect to u and v, we have the following system of equations:

$$\frac{\partial \mathrm{Obj}_{(p,q)}}{\partial u} = \sum_{i=1}^{n} \omega_i pq \left(|x_i - u|^p + |y_i - v|^p\right)^{q-1} |x_i - u|^{p-1} \mathrm{sgn}(x_i - u) = 0 \tag{2.54}$$

$$\frac{\partial \mathrm{Obj}_{(p,q)}}{\partial v} = \sum_{i=1}^{n} \omega_i pq(|x_i - u|^p + |y_i - v|^p)^{q-1}|y_i - v|^{p-1}\mathrm{sgn}(y_i - v) = 0 \tag{2.55}$$

The above system of equations needs to be solved so that we can obtain the optimal facility location. Obviously, the optimal facility location based on the $\mathrm{Obj}_{(p,q)}$ in (2.13) is not in explicit form for all p and q values. As a consequence, we need to obtain it by using numerical methods.