1

Mathematical Properties of the Ordered Median Function

1.1 Introduction

When talking about solving a location problem in practice at least the following phases have to be considered:

- Definition of the problem
- Identification of side constraints
- Choice of the right objective function(s)
- Data collection
- Data analysis (data mining)
- Optimization (actual resolution)
- Visualization of the results
- Discussion if the problem is solved or if the phases have to be started again.

Typically researchers in location theory have been concentrating very much on the resolution (optimization phase) of a given problem. The type of the problem and to some extent also the side constraints were motivated by classical papers in location theory (see [206, 209]). The idea of having a facility placed at a location which is in average good for each client, led to the median objective function (also called Weber or Fermat-Weber objective), see [209]. Finding a location which is even for the most remote client as good as possible, brought up the idea of the center objective (see [163]). The insight, that both points of view might be too extreme, led to the cent-dian approach (see [87]). In addition, researchers always distinguished between continuous, network and discrete location problems. Therefore, the main scope of researchers can be seen as picking a problem from the table in Figure 1.1 by selecting a row and a column, maybe adding some additional constraints to it, and then finding good solution procedures.

	decision space			
objective function		continuous	network	discrete
	median			
	center			
	cent-dian			
	...			

Fig. 1.1. A simplified classification of most of the classical location problems

Although modern location theory[1] is now more than 90 years old the focus of the researchers has been problem oriented. In this direction enormous advances have been achieved since the first analytical approaches. Today sophisticated tools from computer science, computational geometry and combinatorial optimization are applied. However, several years ago a group of researchers realized that although there is a vast literature of papers and books in location theory, a common theory is still missing (There are some exceptions, as for instance [184].) We found the question interesting and even more challenging. Thus, we started in 1996 to work on a unified framework. This monograph summarizes results we published in the last years in several papers. We want to present to the reader a neat picture of our ideas. Our aim is to have a common way for expressing most of the relevant location objectives with a single methodology for all the three major branches of locational analysis: continuous, network and discrete location problems. Or, looking at the table in Figure 1.1, we do not want to unify location problems with respect to the decision space, but with respect to the way the objective function is stated.

We will start in the next section with an illustrative example which will make our ideas transparent without going deep into the mathematics. After that, we give a rigorous mathematical foundation of the concept.

1.2 Motivating Example

Consider the following situation: Three decision makers (Mr. Optimistic, Mr. Pessimistic and Mrs. Realistic) are in a meeting to decide about a new service

[1] By modern location theory, we mean location theory with an economical background.

facility of a governmental institution. The purpose of the service facility is totally unclear but it is agreed on that its major clients live in five places ($P1, \ldots, P5$) in the city. The three decision makers have offers for four different locations from the local real estate Mr. Greedy. The task of the decision makers is now to decide which of the four locations should be used to build the new service facility. The price of the four locations is pretty much the same and the area of each of the four locations is sufficient for the planned building. In a first step a consulting company (with the name WeKnow) was hired to estimate the cost for serving the customers at the five major places from each of the four locations under consideration. The outcome is shown in the following table:

	New1	New2	New3	New4
P1	5	2	5	13
P2	6	20	4	2
P3	12	10	9	2
P4	2	2	13	1
P5	5	9	2	3

Now the discussion is focussing on how to evaluate the different alternatives. Mr. Optimistic came well prepared to the meeting. He found a book by a person called Weber ([206]) in which it is explained that a good way is to take the location where the sum of costs is minimized.

The outcome is $New4$ with an objective value of 21.

Mr. Pessimistic argues that people might not accept the new service in case some of them have to spend too much time in reaching the new institution. He also had a look at some books about locational decisions and in his opinion the maximal cost for a customer to reach the new institution has to be minimized.

The outcome is $New1$ with an objective value of 12.

Mr. Optimistic, however, does not gives up so easily. He says: "Both evaluations are in some sense too extreme. Recently, a friend of mine told me about the Cent-Dian approach which in my understanding is a compromise between our two criteria". We simply have to take the sum (or median) objective (f), the center objective (g) and estimate an α between 0 and 1. The objective function is then calculated as $\alpha f + (1 - \alpha)g$. Mr. Pessimistic agrees that this is in principle a good idea. But having still doubts he asks

- How do we estimate this α?
- How much can we change the α before the solution becomes different?

Mr. Optimistic is a little confused about these questions and also does not remember easy answers in the books he read.

Suddenly, Mrs. Realistic steps in and argues: "The ideas I heard up to now sound quite good. But in my opinion we have to be a little bit more involved with the customers' needs. Isn't it the case that we could neglect the two closest customer groups?. They will be fine of anyway. Moreover, when looking at the table, I recognized that whatever solution we will take one group will be always quite far away. We can do nothing about it, but the center objective is determined by only this far away group. Couldn't we just have an objective function which allows to leave out the k_1 closest customers as well as the k_2 furthest customers." Both, Mr. Optimistic and Mr. Pessimistic, nod in full agreement. Then Mr. Pessimistic asks the key question: "How would such an objective function look like?" Even Mr. Optimistic has to confess that he did not come across this objective. However, he claims that it might be the case, that the objective function Mrs. Realistic was talking about is nothing else than a Cent-Dian function.

Mrs. Realistic responds that this might very well be the case but that she is not a mathematician and she even wouldn't know the correct α to be chosen. She adds that this should be the job of some mathematicians and that the book of Weber is now more than 90 years old and therefore the existence of some unified framework to answer these questions should be useful.

Mr. Pessimistic adds that he anyhow would prefer to keep his original objective and asks if it is also possible that each decision maker keeps his favorite objective under such a common framework.

The three of them decided in the end to stop the planning process for the new institution, since nobody knew what it was good for. Instead they took the money allocated originally to this project and gave it to two mathematicians (associated with the company NP) to work "hard" and to develop answers to their questions.

The core tasks coming out of this story:

- Find a relatively easy way of choosing an objective function.
- Are there alternatives which can be excluded independently of the specific type of objective function?

Of course, the story could have ended in another way leaving similar questions open: The discussion continues and Mrs. Realistic asks if they really should restrict themselves to the alternatives which Mr. Greedy provided. She suggests to use the streets (as a network) and compute there the optimal location in terms of distance. Then the neighborhood of this location should be checked for appropriate ground. Mr. Pessimistic adds that he knows that some streets are about to be changed and some new ones are built. Therefore he would be in favor of using a planar model which approximates the distances typical for the city and then proceed with finding an optimal location. After that he would like to follow the same plan as Mrs. Realistic.

1.3 The Ordered Median Function

In this section we formally introduce a family of functions which fulfils all the requirements discussed in Section 1.2. The structure of these functions involves a nonlinearity in the form of an ordering operation. It is clear that this step introduces a degree of complication into the function. Nevertheless, it is a fair price to be paid in order to handle the requirements shown in the previous section. We will review some of the interesting properties of these functions in a first step to understand their behavior. Then, we give an axiomatic characterization of this objective function.

We start defining the ordered median function (OMf). This function is a weighted average of ordered elements. For any $x \in \mathbb{R}^M$ denote $x_{\leq} = (x_{(1)}, \ldots, x_{(M)})$ where $x_{(1)} \leq x_{(2)} \leq \ldots x_{(M)}$. We consider the function:

$$\begin{aligned} sort_M : \mathbb{R}^M &\longrightarrow \mathbb{R}^M \\ x &\longrightarrow x_{\leq}. \end{aligned} \tag{1.1}$$

Let $\langle \cdot, \cdot \rangle$ denote the usual scalar product in $\mathbb{R}^M$.

Definition 1.1 *The function $f_\lambda : \mathbb{R}^M \longrightarrow \mathbb{R}$ is an ordered median function, for short $f_\lambda \in$ OMf(M), if $f_\lambda(x) = \langle \lambda, sort_M(x) \rangle$ for some $\lambda = (\lambda_1, \ldots, \lambda_M) \in \mathbb{R}^M$.*

It is clear that ordered median functions are nonlinear functions. Whereas the nonlinearity is induced by the sorting. One of the consequences of this sorting is that the pseudo-linear representation given in Definition 1.1 is pointwise defined. Nevertheless, one can easily identify its linearity domains (see chapters 2 or 3). The identification of these regions provides us a subdivision of the framework space where in each of its cells the function is linear. Obviously, the topology of these regions depends on the framework space and on the lambda vector. A detailed discussion is deferred to the corresponding chapters. Different choices of lambda lead also to different functions within the same family. We start showing that the most used objective functions in location theory, namely median, center, centdian or k-centrum are among the functions covered by the ordered median functions.

Some operators related to the ordered median function have been developed by other authors independently. This is the case of the ordered weighted operators (OWA) studied by Yager [214] to aggregate semantic preferences in the context of artificial intelligence; as well as SAND functions (isotone and sublinear functions) introduced by Francis, Lowe and Tamir [78] to study aggregation errors in multifacility location models.

We start with some simple properties and remarks concerning ordered median functions. Most of them are natural questions that appear when a family of functions is considered. Partial answers are summarized in the following proposition.

Proposition 1.1 *Let $f_\lambda(x), f_\mu(x) \in$ OMf(M).*

1. *$f_\lambda(x)$ is a continuous function.*
2. *$f_\lambda(x)$ is a symmetric function, i.e. for any $x \in \mathbb{R}^M$ $f_\lambda(x) = f_\lambda(sort_M(x))$.*

Table 1.1. Different instances of ordered median functions.

λ	$f_\lambda(x)$	Meaning
$(1/M,\dots,1/M)$	$\frac{1}{M}\sum_{i=1}^{M} x_i$	mean average of x
$(0,\dots,0,1)$	$\max_{1\le i\le M} x_i$	maximum component of x
$(1,0,\dots,0,0)$	$\min_{1\le i\le M} x_i$	minimum component of x
$(\alpha,\dots,\alpha,\alpha,1)$	$\alpha\sum_{i=1}^{M} x_i + (1-\alpha)\max_{1\le i\le M} x_i$	α-centdian, $\alpha\in[0,1]$
$(0,\dots,0,1,\overset{k}{\dots},1)$	$\sum_{i=M-k+1}^{M} x_{(i)}$	k-centrum of x
$(1,\overset{k}{\dots},1,0,\dots,0)$	$\sum_{i=1}^{k} x_{(i)}$	anti-k-centrum of x
$(-1,0,\dots,0,1)$	$\max_{1\le i\le M} x_i - \min_{1\le i\le M} x_i$	range of x
$(\alpha,0,\dots,0,1-\alpha)$	$\alpha \min_{1\le i\le M} x_i +$ $(1-\alpha)\max_{1\le i\le M} x_i$	Hurwicz criterion $\alpha\in[0,1]$
$(0,\dots,0,1,\dots,1,0,\dots,0)$	$\sum_{i=k_1+1}^{M-k_2} x_{(i)}$	(k_1,k_2)-trimmed mean
$\lambda_1 \gg \lambda_2 \gg \dots \gg \lambda_M$	$\sum_{i=1}^{M} \lambda_i x_{(i)}$	lex-min of x in any bounded region
$\vdots$	$\vdots$	$\vdots$

3. *$f_\lambda(x)$ is a convex function iff $0\le\lambda_1\le\dots\le\lambda_M$.*
4. *Let $\alpha\in\mathbb{R}$. $f_\alpha(x)\in OMf(1)$ iff $f_\alpha(x)=\alpha x$.*
5. *If c_1 and c_2 are constants, then the function $c_1 f_\lambda(x)+c_2 f_\mu(x)\in OMf(M)$.*
6. *If $f_\lambda(x)\in OMf(M)$ and $f_\alpha(u)\in OMf(1)$, then the composite function is an ordered median function of x on $\mathbb{R}^M$.*
7. *If $\{f_{\lambda^n}(x)\}$ is a set of ordered median functions that pointwise converges to a function f, then $f\in OMf(M)$.*
8. *If $\{f_{\lambda^n}(x)\}$ is a set of ordered median functions, all bounded above in each point x of $\mathbb{R}^M$, then the pointwise maximum (or sup) function defined at each point x is not in general an OMf (see Example 1.1).*
9. *Let $p<M-1$ and $x^p=(x_1,\dots,x_p)$, $x^{\backslash p}=(x_{p+1},\dots,x_M)$. If $f_\lambda(x)\in OMf(M)$ then $f_{\lambda^p}(x^p)+f_{\lambda^{\backslash p}}(x^{\backslash p}) \lesseqgtr f_\lambda(x)$ (see Example 1.2).*

Proof.
The proof of (1) and (3) can be found in [179]. The proofs of items (2) and (4) to (6) are straightforward and therefore are omitted.
To prove (7) we proceed in the following way. Let $f(x)$ be the limit function. Since, $f_{\lambda^n}(x) = \langle \lambda^n, sort_M(x) \rangle$ then at x it must exist λ such that $\lim_{n\to\infty} \lambda^n = \lambda$. Hence, at the point x we have: $f(x) = \langle \lambda, sort_M(x) \rangle$. Now, since the above limit in λ^n does not depend on x then we get that $f(y) = \langle \lambda, sort_M(y) \rangle$, for all y.

Finally, counterexamples for the assertions (8) and (9) are given in the examples 1.1 and 1.2. □

Example 1.1
Let us consider the following three lambda vectors: $\lambda^1 = (-1, 0, 0, 1)$, $\lambda^2 = (0, 0, 1, 0)$, $\lambda^3 = (1, 1, 0, 0)$; and for any $x \in \mathbb{R}^4$ let

$$f_{max}(x) = \max_{i=1,2,3} \{f_{\lambda^i}(x)\}.$$

The following table shows the evaluation of the functions f_{λ^i} and f_{max} at different points.

x	$f_{\lambda^1}(x)$	$f_{\lambda^2}(x)$	$f_{\lambda^3}(x)$	$f_{max}(x)$
(1,2,3,4)	3	3	3	3
(0,1,2,4)	4	2	1	4
(0,0,1,2)	2	1	0	2
(1,3,3,3.5)	2.5	3	4	4

For the first three points the unique representation of the function f_{max} as an OMf(4) is obtained with $\lambda_{max} = (3, 0, -4, 3)$. However, for the fourth point $\hat{x} = (1, 3, 3, 3.5)$ we have

$$\langle \lambda_{max}, sort_M(\hat{x}) \rangle = 1.5 < f_{max}(\hat{x}) = 4.$$

Example 1.2
Let us take $x = (4, 1, 3)$ and $p = 1$. With these choices $x^1 = 4$ and $x^{\backslash 1} = (1, 3)$. The following table shows the possible relationships between $f_\lambda(x)$ and $f_{\lambda^1}(x^1) + f_{\lambda^{\backslash 1}}(x^{\backslash 1})$.

λ	$f_\lambda(x)$	λ^1	$f_{\lambda^1}(x^1)$	$\lambda^{\backslash p}$	$f_{\lambda^{\backslash 1}}(x^{\backslash 1})$	symbol
(1,1,1)	8	1	4	(1,1)	4	=
(1,2,3)	19	1	4	(2,3)	11	>
(4,1,2)	15	4	16	(1,2)	8	<

In order to continue the analysis of the ordered median function we need to introduce some notation that will be used in the following. We will consider in $\mathbb{R}^M_{0+}$ a particular simplex denoted by $S_M^{\leq}$ which is defined as

$$S_M^{\leq} = \left\{(\lambda_1, \dots, \lambda_M) : 0 \leq \lambda_1 \leq \dots \leq \lambda_M, \sum_{i=1}^{M} \lambda_i = 1\right\}. \tag{1.2}$$

Together with this simplex, we introduce two cones $\Lambda_M^{\leq}$ and $\Lambda_{r_1,\dots,r_k}$

$$\Lambda_M^{\leq} = \{(\lambda_1, \dots, \lambda_M) : 0 \leq \lambda_1 \leq \dots \leq \lambda_M\} \tag{1.3}$$

$$\begin{aligned} \Lambda_{r_1,\dots,r_k} = \Big\{\lambda \in \mathbb{R}^M_+ : \lambda = (\lambda^1, \dots, \lambda^k) \in \mathbb{R}^M_+,\ \lambda^i = (\lambda^i_1, \dots, \lambda^i_{r_i}) \in \mathbb{R}^{r_i}_+, \\ \max_j \lambda^i_j \leq \min_j \lambda^{i+1}_j \text{ for any } i = 1, \dots, k-1\Big\} \quad 1 \leq k \leq M, \end{aligned} \tag{1.4}$$

$$\Lambda_M = \mathbb{R}^M_+ .$$

Notice that with this notation the extreme cases are Λ_M (which corresponds to $k = 1$) and $\Lambda_{1,\dots,1}$ which is the set $\Lambda_M^{\leq}$.

In addition, let $\mathcal{P}(1 \dots M)$ be the set of all the permutations of the first M natural numbers, i.e.

$$\mathcal{P}(1 \dots M) = \{\sigma : \sigma \text{ is a permutation of } 1, \dots, M\}. \tag{1.5}$$

We write $\sigma = (\sigma(1), \dots, \sigma(M))$ or sometimes and if this is possible without causing confusion simply $\sigma = \sigma(1), \dots, \sigma(M)$ or $\sigma = (1), \dots, (M)$.

The next result, that we include for the sake of completeness, is well-known and its proof can be found in the book by Hardy et al. [98].

Lemma 1.1 *Let $x = (x_1, \dots, x_M)$ and $y = (y_1, \dots, y_M)$ be two vectors in $\mathbb{R}^M$. Suppose that $x \leq y$, then*

$$x_{\leq} = (x_{(1)}, \dots, x_{(M)}) \leq y_{\leq} = (y_{(1)}, \dots, y_{(M)}) .$$

To get more insight into the structure of ordered median functions we state some other properties concerning sorted vectors.

Theorem 1.1 *Let $x = (x_1, \dots, x_r) \in \mathbb{R}^r$ and $y = (y_1, \dots, y_s) \in \mathbb{R}^s$ be two vectors of real numbers with $r \leq s$. Let $\sigma^x \in \mathcal{P}(1 \dots r)$ and $\sigma^y \in \mathcal{P}(1 \dots s)$ such that*

$$x_{\sigma^x(1)} \leq \dots \leq x_{\sigma^x(r)}, \tag{1.6}$$

and

$$y_{\sigma^y(1)} \leq \dots \leq y_{\sigma^y(s)}. \tag{1.7}$$

Then, if

$$x_k \geq y_k, \quad \forall\, k = 1, \dots, r, \tag{1.8}$$

we have that

$$x_{\sigma^x(k)} \geq y_{\sigma^y(k)}, \quad \forall\, k = 1, \dots, r.$$

Proof.
Consider $k \in \{1, \dots, r\}$. We use the fact that because σ^x and σ^y are permutations, the set $\{\sigma^x(1), \dots, \sigma^x(k)\}$ consists of k distinct elements and the set $\{\sigma^y(1), \dots, \sigma^y(k-1)\}$ consists of $k-1$ distinct elements, where for $k = 1$ the latter set is simply taken to be the empty set. Thus, there must exist $m \in \{1, \dots, k\}$ such that $\sigma^x(m) \notin \{\sigma^y(1), \dots, \sigma^y(k-1)\}$. (If $k = 0$ simply take $m = 1$.) Now by (1.7), it must be that $y_{\sigma^y(k)} \le y_{\sigma^x(m)}$. Furthermore, by (1.8), we have $y_{\sigma^x(m)} \le x_{\sigma^x(m)}$ and by (1.6) we have $x_{\sigma^x(m)} \le x_{\sigma^x(k)}$. Hence $y_{\sigma^y(k)} \le x_{\sigma^x(k)}$, as required. □

The following lemma shows that if two sorted vectors are given, and a permutation can be found to ensure one vector is (componentwise) not greater than the other, then the former vector in its original sorted state must also be not greater (componentwise) than the latter.

Lemma 1.2 *Suppose $w, \hat{w} \in \mathbb{R}^M$ satisfy*

$$w_1 \le w_2 \le \cdots \le w_M, \tag{1.9}$$

$$\hat{w}_1 \le \hat{w}_2 \le \cdots \le \hat{w}_M, \tag{1.10}$$

and

$$\hat{w}_{\sigma(i)} \le w_i, \quad \forall\, i = 1, \dots, M \tag{1.11}$$

for some $\sigma \in \mathcal{P}(1 \dots M)$. Then

$$\hat{w}_i \le w_i, \quad \forall\, i = 1, \dots, M. \tag{1.12}$$

Proof.
We are using Theorem 1.1. Set $w'_i = \hat{w}_{\sigma(i)}$ for all $i = 1, \dots, M$, and take $r = s = M$, $p = (w_1, \dots, w_M)$, $q = (w'_1, \dots, w'_M)$, σ^x to be the identity permutation and $\sigma^y = \sigma^{-1}$, the inverse permutation of σ. Note that in this case $w_{\sigma^x(i)} = w_i$ for all $i = 1, \dots, M$, and furthermore

$$w'_{\sigma^y(i)} = w'_{\sigma^{-1}(i)} = \hat{w}_{\sigma(\sigma^{-1}(i))} = \hat{w}_i, \quad \forall\, i = 1, \dots, M. \tag{1.13}$$

It is obvious from (1.9) that the elements of p form an increasing sequence under the permutation σ^x. It is also clear from (1.10) and (1.13) that the elements of q form an increasing sequence under the permutation σ^y. Thus, the first two conditions of Theorem 1.1 are met. From (1.11) and the definition of w', we have that $w_i \ge w'_i$ for all $i = 1, \dots, M$ and the final condition of Theorem 1.1 is met. From Theorem 1.1 we thus deduce that $w_{\sigma^x(i)} \ge w'_{\sigma^y(i)}$ for all $i = 1, \dots, M$. Now for all $i = 1, \dots, M$ we have that $w_{\sigma^x(i)} = w_i$ and $w'_{\sigma^y(i)} = \hat{w}_i$, so $\hat{w}_i \le w_i$ as required. □

The next lemma shows the following: Take a sorted vector w of r real numbers that is componentwise not greater than r elements chosen from another sorted vector w' of $s \ge r$ real numbers. Then the first r entries of w' are componentwise greater or equal than w.

Lemma 1.3 *Let $w = (w_1, \dots, w_s)$ be a vector of $s \geq 1$ real numbers with*

$$w_1 \leq \dots \leq w_s.$$

Let $r \in \{1, \dots, s\}$ and let $S \in \mathbb{R}^r$ be a vector with elements no less, componentwise, than r of the elements of w, say $w' = (w'_1, \dots, w'_r)$ where for all $j = 1, \dots, r$ there exists a unique $i(j) \in \{1, \dots, s\}$ such that $w'_j \geq w_{i(j)}$, with

$$w'_1 \leq \dots \leq w'_r.$$

Then

$$w'_j \geq w_j, \quad \forall\, j = 1, \dots, r.$$

Proof.
The claim follows from Theorem 1.1, as follows. Take r and s as given. Take $y_j = w_{i(j)}$ for all $j = 1, \dots, r$ and define $y_{r+1}, \dots, y_s$ to be the components from the vector w which do not have index in $\{i(j) \ : j = 1, \dots, r\}$. Note there is a one-to-one correspondence between the elements of w and y, i.e. y is a permutation of w. Take $x_j = w'_j$ for all $j = 1, \dots, r$, so $x_j = w'_j \geq w_{i(j)} = y_j$ for all $j = 1, \dots, r$. Also take σ^x to be the identity permutation, and note that σ^y can be taken so that that $y_{\sigma^y(i)} = w_i$ for all $i = 1, \dots, s$, by the definition of w and since q is a permutation of w. Now by Theorem 1.1 it must be that for all $j = 1, \dots, r$, $x_{\sigma^x(j)} \geq y_{\sigma^y(j)} = w_j$, and so $x_j = w'_j \geq w_j$, as required.□

To understand the nature of the OMf we need a precise characterization. Exactly this will be done in the following two results using the concepts of symmetry and sublinearity.

Theorem 1.2 *A function f defined over $\mathbb{R}^M_+$ is continuous, symmetric and linear over $\Lambda^{\leq}_M$ if and only if $f \in$ OMf(M).*

Proof.
Since f is linear over $\Lambda^{\leq}_M$, there exists $\lambda = (\lambda_1, \dots, \lambda_M)$ such that for any $x \in \Lambda^{\leq}_M$ $f(x) = \langle \lambda, x \rangle$. Now, let us consider any $y \notin \Lambda^{\leq}_M$. There exists a permutation $\sigma \in \mathcal{P}(1 \dots M)$ such that $y_\sigma \in \Lambda^{\leq}_M$. By the symmetry property it holds $f(y) = f(y_\sigma)$. Moreover, for y_σ we have $f(y_\sigma) = \langle \lambda, y_\sigma \rangle$. Hence, we get that for any $x \in \mathbb{R}^M$

$$f(x) = \langle \lambda, x_{\leq} \rangle.$$

Finally, the converse is trivially true. □

There are particular instances of the λ vector that make their analysis interesting. One of them is the convex case. The convex ordered median function corresponds to the case where $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_M$. For this case, we can obtain a characterization without the explicit knowledge of a linearity region. Let $\lambda = (\lambda_1, \dots, \lambda_M)$ with $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_M$; and $\lambda_\pi = (\lambda_{\pi(1)}, \dots, \lambda_{\pi(M)})$ being $\pi \in \mathcal{P}(1 \dots M)$.

Theorem 1.3 *A function f defined over $\mathbb{R}^M$ is the support function of the set $S_\lambda = conv\{\lambda_\pi : \pi \in \mathcal{P}(1...M)\}$ if and only if f is the convex ordered median function*

$$f_\lambda(x) = \sum_{i=1}^{M} \lambda_i x_{(i)}. \tag{1.14}$$

Proof.
Let us assume that f is the support function of S_λ. Then, we have that

$$f(x) = \sup_{s \in S_\lambda} \langle s, x \rangle = \max_{\pi \in \mathcal{P}(1...M)} \langle \lambda_\pi, x \rangle = \max_{\pi \in \mathcal{P}(1...M)} \langle \lambda, x_\pi \rangle = \sum_{i=1}^{M} \lambda_i x_{(i)}.$$

Conversely, it suffices to apply Theorem 368 in Hardy, Littlewood and Polya [98] to the expression in (1.14). □

This result allows us to give global bounds on the ordered median function which do not depend on the weight λ.

Proposition 1.2

1. $f_{(1,0,\ldots,0)}(x) \leq f_\lambda(x) \leq f_{(0,\ldots,0,1)}(x) \quad \forall\, x \in X, \quad \forall\, \lambda \in S_M^{\leq}$.
2. $f_{(1/M,\ldots,1/M)}(x) \leq f_\lambda(x) \leq f_{(0,\ldots,0,1)}(x) \quad \forall\, x \in X, \quad \forall\, \lambda \in S_M^{\leq}$.

Proof.
We only prove (2.) since the proof of (1.) is similar.
It is well-known that $S_M^{\leq}$ is the convex hull of

$$\{e_M, 1/2(e_M + e_{M-1}), \ldots, 1/M \sum_{i=1}^{M} e_i\}$$

where e_i is the vector with 1 in the i-th component and 0 everywhere else (see e.g. Claessens et al. [41]). For each $x \in X$ the function $\lambda \to f_\lambda(x)$ is linear in λ. Thus, we have that the minimum has to be achieved on the extreme points of $S_M^{\leq}$, i.e.

$$\min_{\lambda \in S_M^{\leq}} f_\lambda(x) = \min \left\{ x_{(M)}, 1/2\Big(x_{(M)} + x_{(M-1)}\Big), \ldots, 1/M \sum_{i=1}^{M} x_{(i)} \right\}$$

thus

$$\min_{\lambda \in S_M^{\leq}} f_\lambda(x) = \frac{1}{M} \sum_{i=1}^{M} x_{(i)}.$$

Analogously for the maximum we obtain that

$$\max_{\lambda \in S_M^{\leq}} f_\lambda(x) = x_{(M)}.$$

□

It is important to remark that these bounds are tight. For instance, for (2.) the lower bound is realized with $\bar{\lambda} = (1/M, \dots, 1/M)$ and the upper bound with $\hat{\lambda} = (0, \dots, 0, 1)$, since both $\bar{\lambda}$ and $\hat{\lambda}$ are in $S_M^{\leq}$.

Before continuing, it is worth introducing another concept that will help us in the following. For any vector $x \in \mathbb{R}^M$ and $k = 1, ..., M$, define $r_k(x)$ to be the sum of the k largest components of x, i.e.,

$$r_k(x) = \sum_{t=M-k+1}^{M} x_{(t)}. \tag{1.15}$$

This function is usually called in the literature k-centrum and plays a very important role in the analysis of the ordered median function. The reason is easily understood because of the representation in (1.16). For any $\lambda = (\lambda_1, \dots, \lambda_M) \in \Lambda_M^{\leq}$ we observe that

$$f_\lambda(x) = \sum_{k=1}^{M} (\lambda_k - \lambda_{k-1}) r_{M-k+1}(x). \tag{1.16}$$

(For convenience we set $\lambda_0 = 0$.)

Convexity is an important property within the scope of continuous optimization. Thus, it is crucial to know the conditions that ensure this property. In the context of discrete optimization convexity cannot even be defined. Nevertheless, in this case submodularity plays a similar role. (The interested reader is referred to the Chapter by McCormick in the Handbook Discrete Optimization [130].) In the following, we also prove a submodularity property of the convex ordered median function.

Let $x = (x_i)$, $y = (y_i)$, be vectors in $\mathbb{R}^M$. Define the *meet* of x, y to be the vector $x \bigwedge y = (\min\{x_i, y_i\})$, and the *join* of x, y by $x \bigvee y = (\max\{x_i, y_i\})$. The meet and join operations define a lattice on $\mathbb{R}^M$.

Theorem 1.4 *(***Submodularity Theorem [173]***) Given $\lambda = (\lambda_1, ..., \lambda_M)$, satisfying $0 \leq \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_M$. For any $x \in \mathbb{R}^M$ define the function $f_\lambda(x) = \sum_{i=1}^{M} \lambda_i x_{(i)}$. Then, $f_\lambda(x)$ is submodular over the lattice defined by the above meet and join operations, i.e., for any pair of vectors x, y in $\mathbb{R}^M$,*

$$f_\lambda(x \bigvee y) + f_\lambda(x \bigwedge y) \leq f_\lambda(x) + f_\lambda(y).$$

Proof.
To prove the above theorem it is sufficient to prove that for any $k = 1, ..., M$,

$$r_k(x \bigvee y) + r_k(x \bigwedge y) \leq r_k(x) + r_k(y). \tag{1.17}$$

Given a pair of vectors x and y in $\mathbb{R}^M$, let $c = x \bigwedge y$ and $d = x \bigvee y$. To prove the submodularity inequality for $r_k(x)$, it will suffice to prove that for any subset C' of k components of c, and any subset D' of k components of d, there exist a subset X' of k components of x and a subset Y' of k components of y, such that the sum of the $2k$ elements in $C' \cup D'$ is smaller than or equal to the sum of the $2k$ elements in $X' \cup Y'$. Formally, we prove the following claim:

Claim: Let I and J be two subsets of $\{1, 2, .., M\}$, with $|I| = |J| = k$. There exist two subsets I' and J' of $\{1, 2, ..., M\}$, with $|I'| = |J'| = k$, such that

$$\sum_{i \in I} c_i + \sum_{j \in J} d_j \leq \sum_{s \in I'} x_s + \sum_{t \in J'} y_t.$$

Proof of Claim: Without loss of generality suppose that $x_i \neq y_i$ for all $i = 1, ..., M$. Let $I_x = \{i \in I : c_i = x_i\}$, $I_y = \{i \in I : c_i = y_i\}$, $J_x = \{j \in J : d_j = x_j\}$, and $J_y = \{j \in J : d_j = y_j\}$. Since $x_i \neq y_i$ for $i = 1, ..., M$, we obtain $|I_x| + |I_y| = |J_x| + |J_y| = k$, I_x and J_x are mutually disjoint, and I_y and J_y are mutually disjoint. Therefore, if $|I_x| + |J_x| = k$, (which in turn implies that $|I_y| + |J_y| = k$), the claim holds with equality for $I' = I_x \cup J_x$, and $J' = I_y \cup J_y$. Hence, suppose without loss of generality that $|I_x| + |J_x| > k$, and $|I_y| + |J_y| < k$. Define $I'' = I_x \cup J_x$, and $J'' = I_y \cup J_y$. Let $K = I'' \cap J''$. $|I''| > k$, and $|J''| < k$. We have

$$\sum_{i \in I} c_i + \sum_{j \in J} d_j = \sum_{i \in K} (x_i + y_i) + \sum_{s \in I'' - K} x_s + \sum_{t \in J'' - K} y_t. \tag{1.18}$$

On each side of the last equation we sum exactly k components from c, ("minimum elements"), and k components from d, ("maximum elements"). Moreover, the set of components $\{x_s : s \in I'' - K\} \cup \{y_t : t \in J'' - K\}$ contains exactly $k - |K|$ minimum elements and exactly $k - |K|$ maximum elements. In particular, the set $\{x_s : s \in I'' - K\}$ contains at most $k - |K|$ maximum elements. Therefore, the set $\{x_s : s \in I'' - K\}$ contains at least $q = |I''| - |K| - (k - |K|) = |I''| - k$ minimum elements. Let $I^* \subset I'' - K$, $|I^*| = q$, denote the index set of such a subset of minimum elements. We therefore have,

$$x_i \leq y_i, i \in I^*. \tag{1.19}$$

Note that from the construction I^* and J'' are mutually disjoint. Finally define $I' = I'' - I^*$ and $J' = J'' \cup I^*$, and use (1.18) and (1.19) to observe that the claim is satisfied for this choice of sets. □

1.4 Towards Location Problems

After having presented some general properties of the ordered median functions we will now describe the use of this concept in the location context since this is the major focus of this book.

In location problems we usually are given a set $A = \{a_1, \dots, a_M\}$ of clients and we are looking for the locations of a set $X = \{x_1, \dots, x_N\}$ of new facilities. The quality of a solution is evaluated by a function on the relation between A and X, typically written as $c(A, X) = (c(a_1, X), \dots, c(a_M, X))$ or simply $c(A, X) = (c_1(X), \dots, c_M(X))$. $c(A, X)$ may express time, distance, cost, personal preferences,... Assuming that the quality of the service provided decreases with an increase of $c(a_i, X)$, the objective function to be optimized depends on the cost vector $c(A, X)$. Thus, from the server point of view, a monotonicity principle has to be required because the larger the components of the cost vector, the lower the quality of the service provided (the reader may note that this monotonicity principle is reversed when the new facilities are obnoxious or represent any risk for the users).

On the other hand, looking at the problem from the clients' point of view, it is clear that the quality of the service obtained for the group (the entire set A) does not depend on the name given to the clients in the set A. That is to say, the service is the same if the cost vector only changes the order of its components. Thus, seeing the problem from this perspective, the symmetry principle (see Milnor [138]) must hold (recall that the symmetry principle for an objective function f states that the value given by f to a point u does not depend on the order of the components of u). Therefore, for each cost vector $c_\sigma(A, X)$ whose components are a permutation of the components of $c(A, X)$, $f(c(A, X)) = f(c_\sigma(A, X))$. These two principles have been already used in the literature of location theory and their assumption is accepted (see e.g. Buhl [29], Carrizosa et al. [33], Carrizosa et. al [32] or Puerto and Fernández [169]).

We have proved that the ordered median function is compatible with these principles. By identifying x with $c(A, X)$ we can apply the concept of ordered median functions to location problems. It means that the independent variable of the general definition given in 1.1 is replaced by the cost relation among the clients and the new facilities. The components of $c(A, X)$ are related by means of the lambda vector so that different choices will generate different location models (see Table 1.1).

Of course, the main difficulty is not simply to state the general problem, but to provide structural properties and solution procedures for the respective decision spaces (continuous, network and discrete). Exactly, this will be the content of the remaining three parts of this book. Before starting these detailed parts, we want to give three illustrative examples showing the relation to classical problems in location theory.

Example 1.3

(Three points Fermat problem.) This is one of the classical problems in metric geometry and it dates back to the XVII century. Its formulation is disarmingly simple, but is really rich and still catches the attention of researchers from different areas.

According to Kuhn [126] the original formulation of this problem is credited to Pierre de Fermat (1601-1665), a French mathematician, who addressed the problem with the following challenge: *"let he who does not approve my method attempt the solution of the following problem: given three points in the plane, find a fourth point such that the sum of its distances to the three given points is a minimum".*

Denote by $x = (x_1, x_2)$ the point realizing the minimum of the Fermat problem; and assume that the coordinates of the three given points are $a = (a_1, a_2), b = (b_1, b_2), c = (c_1, c_2)$. If we take the vector $\lambda = (1, 1, 1)$, the Fermat problem consists of finding the minimum of the following ordered median function (which is in fact the classical median function):

$$f_\lambda((d(x, a), d(x, b), d(x, c))) = d(x, a) + d(x, b) + d(x, c),$$

where for instance $d(x, a) = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2}$ is the Euclidean distance from x to a.

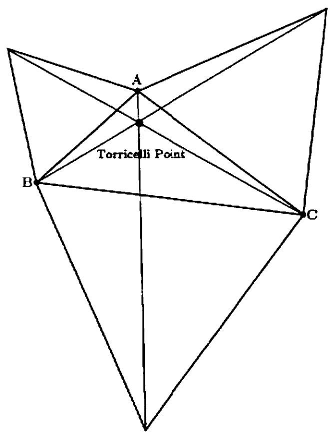

Fig. 1.2. The Torricelli point

The first geometrical solution is credited to Torricelli (1608-1647), the reader can see in Figure 1.3 the construction of the so called Torricelli point. Later in this book we will see how the same solution can be obtained using the theory of the ordered median function.

Example 1.4
In the next example we look at a realistic planning problem in the city of Kaiserslautern. We are given a part of the city map (see Figure 1.3) and we have to find a location for a take away restaurant. The model which was

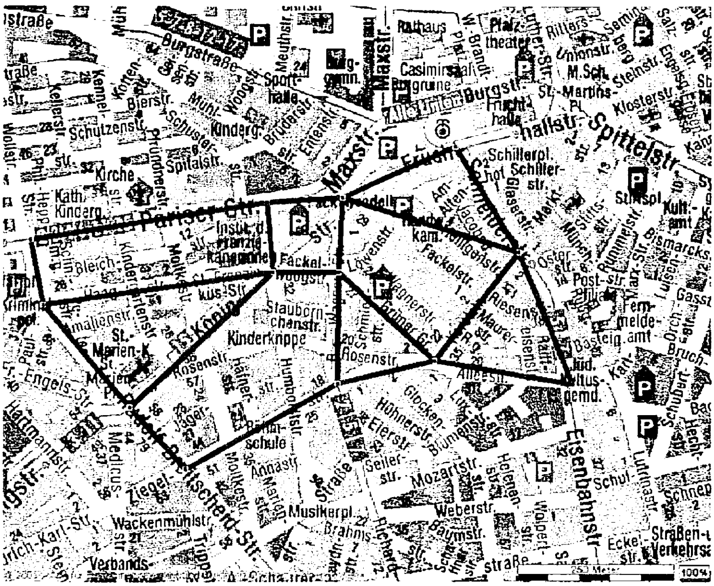

Fig. 1.3. A part of the city map of Kaiserslautern

agreed on is to use the length of the streets of the city map as measure for the distance. In addition the demand of a street is always aggregated into the next crossing. These crossings can be seen as nodes of a graph and the streets are then the corresponding edges (see Figure 1.4).

Therefore A consists of the nodes of the graph and $X = (x)$ is the location of the take away restaurant. We want to allow x to be either in a node of the graph or at any point on an edge (street). The vector $c(A, x)$ is given by the all-pair shortest path matrix (see Table 1.2).

For measuring distances among edges we simply assume that the distance to a node along an incident edge growths linear.

Moreover, it is said to be important that we do not have customers served bad. Therefore we choose a model in which we will only take into account the furthest 3 customers. All others then are remaining within a reasonable distance.

This means that we choose $\lambda = (0, \dots, 0, 1, 1, 1)$.

As a result we get, that the optimal location would be in Node 6 with an objective value of 246. If we would have decided only to take into account the two furthest customers, we would have gotten two optimal locations: One on the edge $[6, 7]$, a quarter of the edge length away from Node 6 and another one on the edge $[6, 9]$ only a fraction of 0.113 away from Node 6. The objective value is 172. We will see later in the book how the solution procedure works.

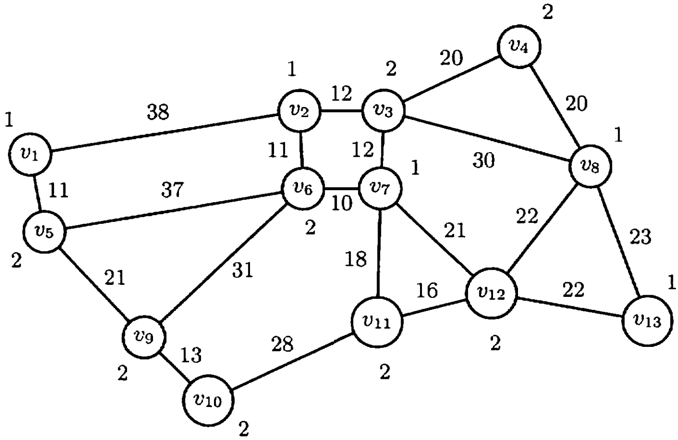

Fig. 1.4. The network model for the KL-city street map

Table 1.2. Distance matrix of Example 1.4.

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}
v_1	0	38	50	70	11	48	58	80	32	45	73	79	101
v_2	38	0	12	32	48	11	21	42	42	55	39	42	64
v_3	50	12	0	20	59	22	12	30	53	58	30	33	53
v_4	70	32	20	0	79	42	32	20	73	78	50	42	43
v_5	11	48	59	79	0	37	47	89	21	34	62	68	90
v_6	48	11	22	42	37	0	10	52	31	44	28	31	53
v_7	58	21	12	32	47	10	0	42	41	46	18	21	43
v_8	80	42	30	20	89	52	42	0	79	66	38	22	23
v_9	32	42	53	73	21	31	41	79	0	13	41	57	79
v_{10}	45	55	58	78	34	44	46	66	13	0	28	44	66
v_{11}	73	39	30	50	62	28	18	38	41	28	0	16	38
v_{12}	79	42	33	42	68	31	21	22	57	44	16	0	22
v_{13}	101	64	53	43	90	53	43	23	79	66	38	22	0

Example 1.5
We are coming back to the discussion between Mr. Optimistic, Mr. Pessimistic and Mrs. Realistic. The situation has changed since the local government decided to have two of these service facilities (this happened since they could not justify one). The rest of the data remains the same and is reprinted below:

	New_1	New_2	New_3	New_4
P_1	5	2	5	13
P_2	6	20	4	2
P_3	12	10	9	2
P_4	2	2	13	1
P_5	5	9	2	3

So, the task now is to select two out of the four possible locations. It is further assumed (by Mrs. Realistic) that a client always goes to the closest of the two service facilities (since one gets the same service in both locations). It is Mr. Pessimistic's turn and therefore he insists on only taking the largest costs into account. This can be modeled in the ordered median framework by setting

$$\lambda = (0, 0, 0, 0, 1) \ .$$

Moreover, $c(P_i, \{New_k, New_l\}) = \min\{c(P_i, New_k), c(P_i, New_l)\}$. We are therefore looking for a solution of the so called 2-Center problem. The possible solutions are $\{New_1, New_2\}$, $\{New_1, New_3\}$, $\{New_1, New_4\}$, $\{New_2, New_3\}$, $\{New_2, New_4\}$, $\{New_3, New_4\}$.

The computation works as follows:

$$c(A, \{New_1, New_2\}) = (2, 6, 10, 2, 5)$$

and

$$\langle \lambda, c(A, \{New_1, New_2\})_{\leq} \rangle = \langle (0, 0, 0, 0, 1), (2, 2, 5, 6, 10) \rangle = 10 \ .$$

The remaining five possibilities are computed analogously:

- $c(A, \{New_1, New_3\}) = (5, 4, 9, 2, 2)$ and the maximum is 9.
- $c(A, \{New_1, New_4\}) = (5, 2, 2, 1, 3)$ and the maximum is 5.
- $c(A, \{New_2, New_3\}) = (2, 4, 9, 2, 2)$ and the maximum is 9.
- $c(A, \{New_2, New_4\}) = (2, 2, 2, 1, 3)$ and the maximum is 3.
- $c(A, \{New_3, New_4\}) = (5, 2, 2, 1, 2)$ and the maximum is 5.

As a result Mr. Pessimistic voted for the solution $\{New_2, New_4\}$.

Springer

http://www.springer.com/978-3-540-24321-2

Location Theory
A Unified Approach
Nickel, S.; Puerto, J.
2005, XXII, 437 p., Hardcover
ISBN: 978-3-540-24321-2