ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT
Discrete Optimization 1 (2004) 23–39
DISCRETE OPTIMIZATION
www.elsevier.com/locate/disopt

Inverse median problems☆

Rainer E. Burkard[a], Carmen Pleschiutschnig[a], Jianzhong Zhang[b]

[a] Institute of Optimization, Dynamical Systems and Discrete Mathematics, Graz University of Technology, Steyrergasse 30, A-8010 Graz, Austria
[b] Department of Mathematics, City University of Hong Kong, Hong Kong

Received 31 October 2003; received in revised form 20 February 2004; accepted 11 March 2004

Abstract

The inverse p-median problem consists in changing the weights of the customers of a p-median location problem at minimum cost such that a set of p prespecified suppliers becomes the p-median. The cost is proportional to the increase or decrease of the corresponding weight. We show that the discrete version of an inverse p-median problem can be formulated as a linear program. Therefore, it is polynomially solvable for fixed p even in the case of mixed positive and negative customer weights. In the case of trees with nonnegative vertex weights, the inverse 1-median problem is solvable in a greedy-like fashion. In the plane, the inverse 1-median problem can be solved in $O(n \log n)$ time, provided the distances are measured in l_1- or l_∞-norm, but this is not any more true in $\mathbb{R}^3$ endowed with the Manhattan metric.

MSC: 90B80; 90B85; 90C27; 90C05

Keywords: Location problem; 1-Median; p-Median; Inverse optimization; Greedy algorithm

1. Introduction and problem statement

Inverse optimization problems have recently generated a considerable interest. In 1992, Burton and Toint [3] introduced the inverse shortest path problem with an interesting application to geological sciences. Given a network, they change the edge lengths as little as possible such that a given path becomes the shortest path. In the same year, Berman et al. [1] published a paper on how a transportation network can be modified in an efficient way in order to improve the known location of the facilities. Later, the same authors [2] considered the analogous problem for the minimax objective. In both papers they improve the network by changing the length of the arcs and by introducing new arcs. A similar question was treated by Zhang et al. [12]. They present a strongly polynomial algorithm for shortening the lengths in a tree network within a given budget such that the longest distance from a given facility to all other nodes becomes minimum. On the other hand, Cai et al. [4] proved that the inverse center location problem is $\mathcal{NP}$-hard, though the underlying center location problem is polynomially solvable. For further results on inverse optimization including network and location models we refer the interested reader to the excellent survey on this topic compiled recently by Heuberger [9].

In this paper we introduce inverse p-median problems. We show that the discrete inverse p-median problem can be solved in polynomial time provided p is fixed and not an input parameter. In particular, we develop a greedy-like algorithm

☆ The first and the second authors acknowledge financial support by the Spezialforschungsbereich F 003 “Optimierung und Kontrolle”, Projektbereich Diskrete Optimierung. The first and third author acknowledge partial support of Hong Kong University Grant Council under the grant 9040883 (CITYU 103003).

E-mail addresses: burkard@tugraz.at (R.E. Burkard), pleschiutschnig@opt.math.tu-graz.ac.at (C. Pleschiutschnig), mazhang@cityu.hk.edu (J. Zhang).

1572-5286/$ - see front matter
doi:10.1016/j.disopt.2004.03.003

for the inverse 1-median problem in trees with positive weights. Further, we consider inverse 1-median problems in the plane and derive particularly simple solution algorithms in the case that the distances are measured in l_1- or l_∞-norm. An example will show that the greedy method does not work in $\mathbb{R}^n, n \geqslant 3$ endowed with the Manhattan metric.

The p-median problem can be stated as follows: let (X, d) be a metric space with distance function d. Let n points $P_1, P_2, \dots, P_n$ (called *customers*) be given. Let $w_1, w_2, \dots, w_n$ be weights for the customers. Find p new points $s_1, s_2, \dots, s_p$ in X, called suppliers, such that

$$\sum_{i=1}^{n} w_i \min_{k=1,\dots,p} d(P_i, s_k)$$

becomes minimum. In the classical case the weights $w_i, i = 1, 2, \dots, n$, are positive. Recently, obnoxious facilities have also been considered which are modelled by negative weights w_i, see e.g. the surveys of Cappanera [5] and Plastria [11]. We may distinguish between *location problems in space*, say e.g., in the plane, where different distance functions may be used, and *discrete location problems*. In discrete location problems either the distances between all customers and suppliers are explicitly given or customers and suppliers are interpreted as vertices of a finite graph with positive edge lengths. In this case the distances are given by the lengths of shortest paths in the graph.

Now let us define the *inverse p-median problem*: we specify p points $s_1, s_2, \dots, s_p$ of the metric space X and want to modify the customer weights such that the set of these points becomes the p-median. Suppose that we incur the nonnegative cost c_i, if the weight w_i is increased by one unit, and we incur the nonnegative cost d_i, if we decrease the weight w_i by one unit. We assume that it is not possible to increase or decrease the customer weights arbitrarily. Namely, the customer weights have to obey the bounds $0 \leqslant \underline{w}_i \leqslant w_i \leqslant \bar{w}_i, 1 \leqslant i \leqslant n$. Now we can state the inverse p-median problem as follows:

Inverse p-median problem. *Find new customer weights $w_i^*, 1 \leqslant i \leqslant n$, such that the set of points $\{s_1, s_2, \dots, s_p\}$ is a p-median with respect to the new weights w_i^*, the new weights lie within given bounds $\underline{w}_i \leqslant w_i^* \leqslant \bar{w}_i$ for all $i, 1 \leqslant i \leqslant n$, and the total cost for changing the weights becomes minimum.*

In the discrete case we can describe the inverse p-median problem by a linear program. Let $U = \{u_1, u_2, \dots, u_n\}$ be a finite set whose elements are called customers. Every customer u_i has a weight w_i. Moreover, let $V = \{v_1, v_2, \dots, v_m\}$ be a finite set whose elements are possible suppliers and let $d(u, v)$ be the distance between customer u and supplier v. Moreover, let $\mathscr{S}$ denote the class of all subsets $S \subseteq V$ of cardinality $|S| = p$. For $\{s_1, s_2, \dots, s_p\}$ being a p-median with respect to the customer weights w_i, $i = 1, \dots, n$, it is necessary and sufficient that

$$\sum_{i=1}^{n} w_i \min_{k=1,\dots,p} d(u_i, s_k) \leqslant \sum_{i=1}^{n} w_i \min_{k \in S} d(u_i, v_k)$$

holds for all $S \in \mathscr{S}$. Let p_i denote the amount by which the customer weight w_i is increased. Similarly, let q_i denote the amount by which the customer weight w_i is decreased. Then the inverse p-median problem can be written as

$$\begin{aligned}
\min \quad & \sum_{i=1}^{n} (c_i p_i + d_i q_i) \\
\text{s.t.} \quad & \sum_{i=1}^{n} (w_i + p_i - q_i) a_{iS} \leqslant 0 \quad \text{for all } S \in \mathscr{S}, \\
& w_i - q_i \geqslant \underline{w}_i, \quad i = 1, 2, \dots, n, \\
& w_i + p_i \leqslant \bar{w}_i, \quad i = 1, 2, \dots, n, \\
& p_i \geqslant 0, \quad i = 1, 2, \dots, n, \\
& q_i \geqslant 0, \quad i = 1, 2, \dots, n,
\end{aligned}$$

where

$$a_{iS} := \min(d(u_i, s_1), d(u_i, s_2), \dots, d(u_i, s_p)) - \min_{k \in S} d(u_i, v_k).$$

Thus discrete inverse p-median problems are solvable in polynomial time provided p is given and is not an input parameter. The linear program has $O(\binom{p}{m})$ constraints. Therefore, it is tractable in practice only for a rather small p-median set. Note that due to the nonnegativity assumptions on c_i and d_i we can always assume that in the optimal solution one of the values p_i or q_i is equal to 0. Since the above considerations also apply to the case of negative customer weights and to the case of mixed positive and negative customer weights, the discrete inverse obnoxious p-median problem as well as the discrete inverse pos/neg weighted p-median problem are also solvable in polynomial time. Summarizing we have

Proposition 1.1. *The discrete inverse p-median problem, where the customers may have any real weight, is solvable in polynomial time provided p is fixed and not an input parameter.*

In the next section we deal with inverse 1-median problems in trees and derive a greedy algorithm in the case of nonnegative weights. Then we address the inverse 1-median problem in the plane, in particular, if the distances are measured in the l_1-norm. It is well known that the 1-median problem with Manhattan distance in $\mathbb{R}^n (n \geqslant 2)$ can be decomposed into n 1-median problems on a line which correspond to problems in a tree. This decomposition, however, is not possible for the inverse 1-median problem in $\mathbb{R}^n (n \geqslant 2)$ endowed with the Manhattan metric: for example, if we change a weight in order to solve the problem in x-direction, then the weight change also influences the problem in y-direction. An analysis of the situation in the plane leads, however, again to a simple algorithm for solving planar inverse 1-median problems. The planar problem where the distances are measured in the l_∞ norm can be reduced to the former case. An example shows, however, that the greedy algorithm cannot be extended to $\mathbb{R}^3$ endowed with the Manhattan metric.

2. The inverse 1-median problem in trees

In the following, we consider the case that customers and supplier correspond to the vertices of a tree graph $G = (V, E)$ with vertex set V and edge set E. All vertex weights are nonnegative. As has been shown by Hua et al. [10] and, one year later, by Goldman [7] the 1-median problem on trees has the interesting property that the solution is completely independent of the (positive) edge lengths and only depends on the weights of the vertices. Let

$$W := \sum_{i=1}^{n} w_i$$

be the sum of all vertex weights of the tree. Now, let v be an arbitrary vertex of the tree of degree k and let $v_1, v_2, \ldots, v_k$ be its immediate neighbors. If we root the tree in v, we get subtrees $T_1, T_2, \ldots, T_k$ which are rooted in $v_1, v_2, \ldots, v_k$, respectively. We denote by $w(T_i)$ the sum of all vertex weights of subtree T_i. A consequence of the considerations of Hua et al. and Goldman is the following optimality criterion.

Lemma 2.1 (Optimality criterion; [7,10]). *A vertex v is a 1-median of the given tree, if and only if the weights of all its subtrees $w(T_i)$ are not larger than $W/2$:*

$$\max_{1 \leqslant i \leqslant k} w(T_i) \leqslant \frac{W}{2}. \tag{1}$$

Now let m be a 1-median and let $s, s \neq m$ not be a 1-median of the graph with respect to the given weights. Let us assume that we want vertex s to become a 1-median with respect to modified vertex weights. If we root the tree in vertex s, we get subtrees $T_1, \ldots, T_k$, where k is the degree of vertex s. One of these subtrees contains vertex m, say T_k. If s is not a 1-median for the weight w, then we know that $w(T_i) < W/2$ for all $i = 1, \ldots, k-1$ and $w(T_k) > W/2$. This follows from the fact that m is a 1-median and s lies in one of the subtrees of weight $\leqslant W/2$ which are rooted in m. Since $T_1, \ldots, T_{k-1}$ also lie in the same subtree rooted in m, they must have a weight $< W/2$.

Let W_0 denote the current weight of vertex s which is supposed to become the 1-median and let W_i denote the sum of the current weights of the vertices in the subtrees T_i, $i = 1, \ldots, k$. Moreover, we define $H := \sum_{i=0}^{k} W_i/2$. During the following algorithm we shall change the weights $w_i, i = 1, 2, \ldots, n$, and therefore also the amounts $W_i, i = 0, 1, \ldots, k$. In our considerations the *optimality gap* $D := W_k - H$ will play an important role.

Lemma 2.2. *Let $W_1 \leqslant W/2, \ldots, W_{k-1} \leqslant W/2$. The vertex s is a 1-median, if and only if the optimality gap $D = 0$.*

Proof. $D = 0$ if and only if $W_k = W/2$. It follows directly from Lemma 2.1 that in this case s is a 1-median. □

If we change the weight of s or of a vertex in the subtrees $T_1, \ldots, T_{k-1}$ by δ, then D changes by $-\delta/2$. Otherwise, if we change the weight of a vertex in the subtree T_k by the amount δ, then the optimality gap changes by $\delta/2$, since $W_k + \delta - (H + \delta/2) = D + \delta/2$. Thus, we decrease D if we either increase the weight of s or of a vertex in T_i, $i = 1, 2, \ldots, k-1$, or if we decrease the weight of a vertex in T_k. On the other hand, any operation which worsens D must be "neutralized" by a change which decreases D by the same amount. Since both operations (the increase of D and the decrease thereafter) cause nonnegative cost and, in addition the solution without increase of D and neutralization thereafter is feasible, we can restrict ourselves to increasing the weight of s or of a vertex in T_i, $i = 1, 2, \ldots, k-1$, and decreasing the weight of vertices in T_k.

Therefore, we consider the cost coefficients c for the root and the vertices in the first $k-1$ subtrees and the cost coefficients d for the vertices in the last subtree. We order these n cost coefficients increasingly in a list

$$r_{\varphi(1)} \leqslant r_{\varphi(2)} \leqslant \cdots \leqslant r_{\varphi(n)},$$

where φ is a permutation of the set $\{1, 2, \ldots, n\}$. If vertex $v_{\varphi(l)}$ is either s or belongs to one of the first $k-1$ subtrees, then $r_{\varphi(l)}$ is the coefficient $c_{\varphi(l)}$. If, however, $v_{\varphi(l)}$ is a vertex of subtree T_k, then $r_{\varphi(l)}$ is the coefficient $d_{\varphi(l)}$.

Let us assume that $r_{\varphi(1)} = c_1$. In order to decrease the optimality gap, we can increase the weight w_1 by p_1. We distinguish two cases:

Case A: By increasing the weight w_1, we can re-establish the optimality, namely $W_k = p_1/2 + H$. Thus, p_1 becomes

$$p_1 = 2(W_k - H) = W_k - \sum_{i=0}^{k-1} W_i.$$

In this case we are done. Note that in this case the subtree T_k has exactly half the weight of the whole tree, i.e., $D = 0$. Therefore any other subtree $T_1, T_2, \ldots, T_{k-1}$ has a weight not larger than half of the total weight. Thus the optimality criterion is fulfilled.

Case B: $w_1^* := w_1 + p_1$ attains the upper bound of vertex v_1. This means that $p_1 = \bar{w}_1 - w_1$. Since p_1 cannot be increased further, we delete $r_{\varphi(1)}$ from our list and proceed with the next smallest element $r_{\varphi(2)}$.

Now, if the first element $r_{\varphi(1)}$ in the list is d_1, then v_1 lies in the subtree T_k and we can reduce the current weight of this vertex by q_1. Again we distinguish the following two cases:

Case A: By decreasing the weight w_1, we can re-establish optimality, namely

$$W_k - q_1 = \left(\sum_{i=0}^{k-1} W_i + W_k - q_1 \right) \Big/ 2.$$

Thus q_1 becomes $W_k - \sum_{i=0}^{k-1} W_i$. Note again that in this case the subtree T_k has exactly half the weight of the whole tree. Therefore the first $k-1$ subtrees have a weight smaller than $H - q_1/2$. Thus the optimality criterion is fulfilled.

Case B: $w_1^* := w_1 - q_1$ attains the lower bound of vertex v_1. This means that $q_1 = w_1 - \underline{w}_1$. Since w_1 cannot be decreased any more, we delete $r_{\varphi(1)}$ from our list and proceed with the next smallest element $r_{\varphi(2)}$.

These considerations show that we can proceed in a greedy fashion in order to solve the inverse 1-median problem on a tree: first, we generate the ordered list $r_{\varphi(1)} \leqslant r_{\varphi(2)} \leqslant \cdots \leqslant r_{\varphi(n)}$. Then we choose the first element of this list and decrease or increase the corresponding weight according to the above rules. If an upper or lower bound is met, we delete this element and proceed to the next element. As soon as Case A applies for the first time we stop. In this case vertex s has become a 1-median of G. Due to the ordering of the cost coefficients any other distribution of the weights which makes s the 1-median of G would have an equal or larger cost. Therefore the algorithm yields an optimal solution. Summarizing we get

Algorithm 1. Solves the inverse 1-median problem in a tree, vertex s being the new 1-median

1: Compute the total weight $W := \sum_{i=1}^{n} w_i$.
2: Compute the weight of the subtrees rooted in s.
    If all subtrees have a weight $\leqslant W/2$, **then** stop: the given weights are already optimal,
    else let T_0 denote the subtree with largest weight W_0.
3: **For** $i := 1, 2, \ldots, n$, **if** vertex v_i is s or belongs not to T_0, **then** $r_i := c_i$, **else** $r_i := d_i$.
4: **Set** $D := W_0 - W/2; R := \{r_1, r_2, \ldots, r_n\}$.
5: **If** $R \neq \emptyset$ find a minimum element r_i of R,
    else stop: there does not exist a feasible solution.
6: **If** $r_i = c_i$, **then** $p_i := \min(2D, \bar{w}_i - w_i); D := D - p_i/2$;
    else $q_i := \min(2D, w_i - \underline{w_i}); D := D - q_i/2$.
7: **If** $D = 0$, **then** stop: an optimal solution is reached,
    else return to Step 5.

Proposition 2.3. *The inverse* 1*-median problem in a tree can be solved in* $O(n \log n)$ *time by a greedy-type algorithm.*

Since the procedure for establishing the optimality of a vertex s uses only comparisons and additions, it also works if the vertex weights of G are nonnegative elements of an arbitrary totally ordered semigroup $(H, *, \leqslant)$ with semigroup operation $*$ and order relation $\leqslant$. An element $a \in H$ is called *nonnegative*, if we have $a * c \geqslant c$ for all $c \in H$.

Note that the optimality criterion of Lemma 2.1 only holds for trees with nonnegative weights. Therefore, it can neither be used in the case of inverse obnoxious 1-median problems on trees nor for 1-median problems in general graphs. In the latter case the edge lengths cannot be neglected. In general graphs we have to guarantee that the vertex s is the 1-median in the shortest path tree rooted in s (which could easily be achieved by the algorithm outlined above), and in addition that no other shortest path tree leads to a smaller value for the function $\sum_{i=1}^{n} w_i d(v_i, s)$, i.e., that s is indeed the 1-median of the graph. To give an example, consider a cycle with three vertices a, b and c, edge lengths $[a,b]=3, [b,c]=1, [a,c]=2$ and vertex weights $w_a = w_b = w_c = 1$. It is easy to see that vertex c is the 1-median of the cycle. Let vertex a be our candidate for the 1-median. This vertex is already the 1-median in the shortest path tree rooted in a but not the 1-median of the cycle. In order that vertex a becomes a 1-median of the cycle we have, e.g., to change the weight w_a to 2.

3. Inverse 1-median problems in the plane

The continuous 1-median problem in the plane can be stated as follows: let n points $P_1, P_2, \ldots, P_n$ be given in the plane. Every point P_i has some weight w_i. We assume throughout this section that all given weights are positive. Moreover, an appropriate distance measure $d(P_i, P_j)$ is fixed, e.g., the Euclidean distance, the Manhattan distance and so on. Find a new point P_0 (the 1-median) in the plane such that

$$\sum_{i=1}^{n} w_i d(P_i, P_0)$$

becomes minimum. In the inverse version of this problem, we are given $n+1$ points $P_0, P_1, P_2, \ldots, P_n$ in the plane with weights $w_i, i = 0, 1, \ldots, n$. By changing the weights at minimum cost we want the point P_0 to become a 1-median of these points. For this reason point P_0 must lie in the convex hull of the points P_i, $i = 1, \ldots, n$, since otherwise there will be no feasible solution. Every weight w_i can only be changed between a lower bound $\underline{w}_i \geqslant 0$ and an upper bound $\bar{w}_i$. The cost for increasing the weight w_i by one unit is c_i and the cost for decreasing this weight by one unit is d_i. The total cost is measured by the function

$$\sum_{i=0}^{n} c_i p_i + d_i q_i, \tag{2}$$

where p_i is the amount by which w_i is increased and q_i is the amount by which the weight w_i is decreased. We call a solution (p_i, q_i), $i = 0, 1, 2, \ldots, n$, feasible, if it guarantees that P_0 is a 1-median and if it fulfills the bound constraints. The cost of a feasible solution is given by (2).

In the following we shall deal with the Manhattan distance. It is well known that the 1-median problem with Manhattan distance in $\mathbb{R}^n (n \geqslant 2)$ can be split into n 1-median problems on a line which can be solved in a straightforward way. Note, however, that this decomposition is not possible for the inverse 1-median problem in $\mathbb{R}^n (n \geqslant 2)$ with Manhattan distance: if we change a weight in order to solve the problem in x-direction, then the weight change also influences the problem in y-direction. In the following we will first investigate the connection between weight changes in x- and y-directions. By analyzing the situation in the plane we will show that the weight changes must fulfill a system of three inequalities. Finding the cheapest solution amounts in solving a linear program with these three inequalities and the weight bounds as a constraint set. Though in general it is not possible to solve this LP by a greedy method (an example will be given), it can be shown that due to the special right-hand side of the LP arising from the problem in the plane it can be solved in a greedy fashion. Thus, the algorithm for solving the inverse 1-median problem in the plane under the assumption that distances are measured by the Manhattan metric turns out to be a nontrivial generalization of the greedy method used to solve the inverse 1-median problem in trees and cannot be further generalized to solve problems in higher dimensions.

First, we are going to show that the actual distances between the points do not play any role. It will only be important to know whether the points lie left or right and below or above the point P_0.

Let (x_i, y_i) be the coordinates of point P_i, $i = 0, 1, \ldots, n$. In the following we always assume that P_0 should become the 1-median. From now on we make use of the following notation:

$$X_\sim := \{i \mid x_i \sim x_0\}, \qquad WX_\sim := \sum_{i \in X_\sim} w_i,$$

$$Y_\sim := \{i \mid y_i \sim y_0\}, \qquad WY_\sim := \sum_{i \in Y_\sim} w_i,$$

where $\sim$ is any of the relations $=, <, \leqslant, >, \geqslant$. We say a point $P_i = (x_i, y_i)$ fulfills $X_\sim$ ($Y_\sim$), if $x_i \sim x_0$ ($y_i \sim y_0$) holds. Using this notation the well-known optimality conditions for the 1-median problem in the plane equipped with the Manhattan distance can be expressed as follows:

Lemma 3.1 (Optimality condition). *P_0 is a 1-median if and only if*

$$WX_< - WX_\geqslant \leqslant 0, \tag{3}$$

$$WX_> - WX_\leqslant \leqslant 0, \tag{4}$$

$$WY_< - WY_\geqslant \leqslant 0, \tag{5}$$

$$WY_> - WY_\leqslant \leqslant 0. \tag{6}$$

Remark. Note that the distances between the points do not play any role, only their weights do.

We give a short proof of this lemma, see e.g. [8], since it shows why the distances do not play any role. An analogous proof applies in the case of the space $\mathbb{R}^n (n \geqslant 3)$ equipped with the l_1-norm. Therefore the optimality conditions of Lemma 3.1 can easily be generalized for 1-median problems in $\mathbb{R}^n (n \geqslant 3)$.

Proof. Let w_i^* be those weights for the given points such that P_0 is a 1-median of the given points $P_1, \dots, P_n$ under the l_1-norm. This means that the function

$$f(\xi, \eta) = \sum_{i=0}^{n} w_i^* |\xi - x_i| + \sum_{i=0}^{n} w_i^* |\eta - y_i|$$

takes its minimum in P_0. Since the function $f(\xi, \eta)$ is separable we have

$$\min_{\xi,\eta} f(\xi, \eta) = \min_{\xi} \sum_{i=0}^{n} w_i^* |\xi - x_i| + \min_{\eta} \sum_{i=0}^{n} w_i^* |\eta - y_i|.$$

The functions $f_1(\xi) := \sum_{i=0}^{n} w_i^* |\xi - x_i|$ and $f_2(\eta) := \sum_{i=0}^{n} w_i^* |\eta - y_i|$ are piecewise linear and convex functions with breakpoints in $\xi = x_i$ and $\eta = y_i, i = 0, 1, \dots, n$, respectively. f_1 has a minimum at x_0 if and only if the unilateral derivatives $f_1'(x_0+)$ and $f_1'(x_0-)$ in point x_0 fulfill:

$$f_1'(x_0-) \leqslant 0 \quad \text{and} \quad f_1'(x_0+) \geqslant 0.$$

It is easy to see that

$$f_1'(x_0-) = WX_< - WX_\geqslant,$$

$$f_1'(x_0+) = WX_\leqslant - WX_>.$$

The same holds for the unilateral derivatives of $f_2(\eta)$ and the sets $WY_\sim$. From this the optimality conditions follow in a direct way. □

Let

$$D_X := \max(WX_< - WX_\geqslant, WX_> - WX_\leqslant) = |WX_< - WX_>| - WX_=,$$

$$D_Y := \max(WY_< - WY_\geqslant, WY_> - WY_\leqslant) = |WY_< - WY_>| - WY_=.$$

We introduce the *optimality gap*

$$D := \max\{0, D_X, D_Y\}.$$

According to Lemma 3.1 a solution is optimal, if and only if $D = 0$.

For the following considerations we shall always assume that D is positive, since otherwise P_0 would already be optimal. Let D be attained by $WX_< - WX_\geqslant$. We are going to change the weights at minimum cost such that inequalities (3)–(6) hold for the modified weights. Since we shall see later in the remark after Lemma 3.3 that the optimal solution found will fulfill the relation of the form (4) anyhow, we consider in the sequel only relations (3), (5) and (6).

Let (p_i, q_i), $i = 0, 1, \ldots, n$, be a feasible solution with total cost C' of the inverse 1-median problem where $\sum_{i=0}^{n}(p_i + q_i) > D$. We are going to show that there exists a feasible solution $(\hat{p_i}, \hat{q_i})$, $i = 0, 1, \ldots, n$, with

$$\sum_{i=0}^{n}(\hat{p_i} + \hat{q_i}) = D, \qquad \hat{p_i} \geqslant 0, \qquad \hat{q_i} \geqslant 0$$

and cost $C \leqslant C'$. We partition the plane into 6 regions, namely

$$R_0 := X_{\geqslant} \cap Y_{=},$$

$$R_1 := X_{\geqslant} \cap Y_{<},$$

$$R_2 := X_{\geqslant} \cap Y_{>},$$

$$R_3 := X_{<} \cap Y_{=},$$

$$R_4 := X_{<} \cap Y_{<},$$

$$R_5 := X_{<} \cap Y_{>}.$$

We now consider the points in a fixed region R_k. In a feasible solution the weight of these points is either increased by p_i or decreased by q_i. Let

$$r_k := \sum_{P_i \in R_k} (p_i - q_i), \qquad \text{for } k = 0, 1, \ldots, 5.$$

Due to the bounds for the weights we retain the feasibility of the solution and do not increase the cost if we can replace the values p_i and q_i by smaller values while not changing the sums $WX_{<}, WX_{\geqslant}$ and $WY_{<}, WY_{>}, WY_{=}$.

First we notice that there is a feasible solution with the following property: in any of the 6 regions $R_0, \ldots, R_5$ there are only positive values for p_i or only positive values for q_i, but not for both at the same time, i.e., $p_i \cdot q_j = 0$ holds for any pair of points $P_i, P_j \in R_k$. Namely, if in the region R_k there are positive p_i and q_j, we could increase or decrease the weights less without changing the value of r_k. For example, if $p_i = 2$ and $q_j = 3$, we get the same value for r_k if we do not change the weight of point P_i (i.e., $p_i = 0$) and set $q_j = 1$. According to the above remark this would result in another feasible solution without increasing the cost.

The feasibility of the solution implies

$$WX_{<} + r_3 + r_4 + r_5 - (WX_{\geqslant} + r_0 + r_1 + r_2) \leqslant 0 \tag{7}$$

and further that

$$WY_{<} + r_1 + r_4 - (WY_{\geqslant} + r_0 + r_2 + r_3 + r_5) \leqslant 0,$$

$$WY_{>} + r_2 + r_5 - (WY_{\leqslant} + r_0 + r_1 + r_3 + r_4) \leqslant 0.$$

Let us define

$$a_1 := r_0,$$

$$a_2 := r_1 - r_5,$$

$$a_3 := r_2 - r_4,$$

$$a_4 := -r_3.$$

Then, setting

$$D_1 := WY_{<} - WY_{\geqslant},$$

$$D_2 := WY_{>} - WY_{\leqslant},$$

we get

$$D \leqslant a_1 + a_2 + a_3 + a_4,$$

$$D_1 \leqslant a_1 - a_2 + a_3 - a_4,$$

$$D_2 \leqslant a_1 + a_2 - a_3 - a_4. \tag{8}$$

Thus we get

Lemma 3.2. *A feasible solution of the planar inverse* 1*-median problem where the distances are measured with the Manhattan metric fulfills the inequality system* (8).

In order to find an optimal solution, we show first that we can assume that inequality (7) holds as equation, i.e. that $D = a_1 + a_2 + a_3 + a_4$. This is an important observation (which is not any more true in $\mathbb{R}^3$; see the example in Section 3.2).

Lemma 3.3. *If the problem is feasible, then there exists an optimal solution with*

$$a_1 + a_2 + a_3 + a_4 = D. \tag{9}$$

Lemma 3.3 can be proven by analyzing the complimentary slackness conditions of the linear program with constraint set (8) and the bounds on the weights. Since the proof is rather technical, it is transferred to the appendix.

Remark. Note that due to (9) optimality condition (4) is also fullfilled. Namely, let $\tilde{W}X_>$ and $\tilde{W}X_\leqslant$ denote the transformed weights in $X_>$ and $X_\leqslant$. Then we have

$$\tilde{W}X_> - \tilde{W}X_\leqslant \leqslant \tilde{W}X_\geqslant - \tilde{W}X_<$$

$$= WX_\geqslant - WX_< + r_0 + r_1 + r_2 - r_3 - r_4 - r_5 = -D + a_1 + a_2 + a_3 + a_4 = 0.$$

By subtracting the second and third inequalities from the first equation in

$$D \leqslant a_1 + a_2 + a_3 + a_4,$$

$$D_1 \leqslant a_1 - a_2 + a_3 - a_4,$$

$$D_2 \leqslant a_1 + a_2 - a_3 - a_4$$

we get

$$a_2 + a_4 \leqslant 1/2(D - D_1) =: b_1, \tag{10}$$

$$a_3 + a_4 \leqslant 1/2(D - D_2) =: b_2. \tag{11}$$

In the following, we shall deal with the system

$$a_1 + a_2 + a_3 + a_4 = D, \tag{12}$$

$$a_2 \qquad + a_4 \leqslant b_1, \tag{13}$$

$$a_3 + a_4 \leqslant b_2. \tag{14}$$

First, we show that it can be assumed that all values a_1, a_2, a_3 and a_4 are nonnegative. This is the statement of the following lemma.

Lemma 3.4. *If the problem is feasible, then there exists an optimal solution where the weights of some points in* $X_\geqslant$ *are increased and the weights of some points in* $X_<$ *are decreased.*

Proof. For the proof, we have to show that there is an optimal solution where all values r_0, r_1, r_2 are nonnegative and r_3, r_4, r_5 are nonpositive. We start by proving that a_1, a_2, a_3 and a_4 can be considered as nonnegative. Assume that a_1 is negative. Due to Eq. (12) we get $a_2 + a_3 + a_4 = D + |a_1|$. Therefore, a_1 can be set to 0 by reducing some of the positive summands in a_2, a_3 and a_4 by an amount which keeps them nonnegative, but sums up to $|a_1|$. Thus the feasibility of the solution with respect to the weight bounds as well as with respect to the feasibility conditions (12)–(14) is retained. Since the weight changes are reduced, the cost does not increase.

Now let us assume that $a_2 < 0$. If $a_4 > 0$, we can reduce a_4 by $\delta := \min(a_4, |a_2|)$ by setting $a_2 := a_2 + \delta, a_4 := a_4 - \delta$. This transformation keeps the solution feasible and does not increase the cost. If after this transformation a_2 is still negative, we can reduce other positive summands a_1, a_3, a_4 arbitrarily until $a_2 = 0$ is reached, but keeping the summands nonnegative. Again, feasibility is retained and the cost is not increased. An analogous argument applies to negative values of a_3.

Thus, we now can assume that a_1, a_2 and a_3 are nonnegative. Let $a_4 < 0$. In this case we get

$$a_1 + a_2 + a_3 = D + |a_4|, \tag{15}$$

$$a_2 \leqslant b_1 + |a_4|, \tag{16}$$

$$a_3 \leqslant b_2 + |a_4|. \tag{17}$$

First we can reduce a_2 by $\delta_1 := \min(a_2, |a_4|, b_2 + |a_4| - a_3)$, i.e. $a_2 := a_2 - \delta_1$ and $a_4 := a_4 + \delta_1$. If after this transformation $a_4 = 0$, we are finished, since the feasibility has been retained. If, however, after the transformation the new value of a_4 is still negative, we have to consider several cases:

- *Case* 1: $a_2 = 0$. In this case we first reduce a_3 by $\delta_2 := \min(a_3, |a_4|)$ and, if after this transformation a_4 is still negative, the coefficient a_1. This renders inequalities (16) and (17) valid and reduces the positive values of a_1 and a_3, thus retaining the feasibility.
- *Case* 2: $a_2 > 0, \delta_1 = b_2 + |a_4| - a_3$. In this case (17) holds as an equation. We distinguish again two cases:
  - *Case* 2a: $b_1 + |a_4| - a_2 > 0$. We determine $\delta_2 := \min(a_3, |a_4|, b_1 + |a_4| - a_2)$ and set $a_3 := a_3 - \delta_2, a_4 := a_4 + \delta_2$. This transformation retains the feasibility of the solution and does not increase the cost. If afterwards $a_4 = 0$, we are done. So we assume $a_4 < 0$. The case $a_3 = 0$ cannot occur, since $\delta_1 = b_2 + |a_4| - a_3$ implies $|a_4| = a_3 - b_2 < 0$, which is a contradiction to $\min(a_3, |a_4|) = a_3$. If, however, $a_3 > 0$ and $|a_4| > 0$, both inequalities (16) and (17) now hold as equations and we come to Case 2b.
  - *Case* 2b: Eq. (16) *also holds as an equation*. We show that this cannot occur. By adding Eqs. (16) and (17) we get

    $$a_2 + a_3 = b_1 + b_2 + 2|a_4| = D + WY_{=} + 2|a_4|.$$

    Introducing Eq. (15) we get

    $$a_2 + a_3 = a_1 + a_2 + a_3 + WY_{=} + |a_4|$$

    and therefore

    $$0 = a_1 + WY_{=} + |a_4|.$$

    Since $a_1 \geqslant 0$ and $|a_4| > 0$ according to the assumptions, and $WY_{=} \geqslant 0$ according to the assumption of nonnegative weights, we arrive at a contradiction. Therefore Case 2b cannot occur, provided that we start with a feasible solution.

Thus we can reach a feasible solution with $a_1, a_2, a_3, a_4 \geqslant 0$.

Once these coefficients are fixed to nonnegative values, the positive coefficients r_0, r_1, r_2 can be decreased and the negative coefficients r_3, r_4, r_5 can be increased until all values r_0, r_1, r_2 are nonnegative and r_3, r_4, r_5 are nonpositive. This reduction is feasible with respect to the weight bounds, does not increase the cost and keeps the values $a_1, a_2, a_3, a_4 \geqslant 0$ unchanged. Therefore the feasibility conditions (12)–(14) are also met. □

The following striking lemma yields the last argument for developing a greedy algorithm.

Lemma 3.5. *If the problem is feasible, then increasing the weights of some points in* $X_{\geqslant}$ *and decreasing the weights of some points in* $X_{<}$ in a greedy fashion *yields an optimal solution.*

Proof. Because of Lemma 3.4 it just remains to show that changing the cheapest possible weights which increase $WX_{\geqslant}$ or decrease $WX_{<}$ leads to an optimal solution of the inverse 1-median problem, i.e., a solution which fulfills (12)–(14). We define

$$X_{\sim}Y_{\approx} := \{i \mid x_i \sim x_0, y_i \approx y_0\}, \qquad WX_{\sim}Y_{\approx} := \sum_{i \in X_{\sim}Y_{\approx}} w_i,$$

where $\sim$ and $\approx$ are any of the relations $=, <, \leqslant, >, \geqslant$. Using this notation, D, b_1 and b_2 are composed in the following way:

$$\begin{aligned} WX_<Y_< + WX_<Y_> &+ WX_<Y_= - WX_\geqslant Y_< - WX_\geqslant Y_> - WX_\geqslant Y_= = D,\\ WX_<Y_> &+ WX_<Y_= - WX_\geqslant Y_< = b_1,\\ WX_<Y_< &+ WX_<Y_= - WX_\geqslant Y_> = b_2. \end{aligned}$$

By combining these equations we get

$$WX_<Y_= + WX_\geqslant Y_= = b_1 + b_2 - D.$$

Since the initial weights are all positive, this shows

$$WX_<Y_= \leqslant b_1 + b_2 - D. \tag{18}$$

Then the positivity of the lower bounds for the weights implies that $WX_<Y_=$ can be reduced by at most b_1+b_2-D units, or in our notation,

$$a_4 \leqslant b_1 + b_2 - D. \tag{19}$$

Therefore, we can add constraint (19) to constraints (12)–(14) without affecting the optimal solution:

$$\begin{aligned} &\min \sum_{i=0}^{n} c_i p_i + d_i q_i\\ &\text{s.t. } a_1 + a_2 + a_3 + a_4 = D,\\ &\qquad a_2 + a_4 \leqslant b_1,\\ &\qquad a_3 + a_4 \leqslant b_2,\\ &\qquad a_4 \leqslant b_1 + b_2 - D,\\ &\quad w_i + p_i \leqslant \bar{w}_i, \quad i = 0, 1, \ldots, n,\\ &\quad w_i - q_i \geqslant \underline{w}_i, \quad i = 0, 1, \ldots, n,\\ &\quad p_i \geqslant 0, q_i \geqslant 0, \quad i = 0, 1, \ldots, n. \end{aligned} \tag{20}$$

If the greedy algorithm finds a solution with $a_4 = 0$ it is clearly optimal, since the values of a_1, a_2 and a_3 do not affect each other. So, assume that in the current step a_4 should be increased. This is possible up to $\delta := \min\{b_1, b_2, b_1+b_2-D\}$ at most. Afterwards,

$$\begin{aligned} a_1 + a_2 + a_3 &= D - \delta,\\ a_2 &\leqslant b_1 - \delta,\\ a_3 &\leqslant b_2 - \delta. \end{aligned}$$

The inequalities

$$b_1 - \delta + b_2 - \delta = b_1 + b_2 - 2\delta \geqslant WX_<Y_= + D - 2\delta \geqslant D - \delta,$$

using (18) and $WX_<Y_= \geqslant \delta$ show that the change of a_4 also does not affect the values of a_1, a_2, a_3 and vice versa. That is, the greedy algorithm yields an optimal solution of problem (20). □

Remark. In general, linear programs of the form

$$\begin{aligned} \min \quad & \sum_{i=1}^{4} c_i a_i\\ \text{s.t.} \quad & a_1 + a_2 + a_3 + a_4 = D,\\ & a_2 + a_4 \leqslant b_1,\\ & a_3 + a_4 \leqslant b_2,\\ & a_i \geqslant 0, \quad i = 1, \ldots, 4 \end{aligned}$$

are not solvable by a greedy method. Consider the problem

$$\begin{aligned} \min \quad & 99a_1 + 9a_2 + 9a_3 + a_4 \\ \text{s.t.} \quad & a_1 + a_2 + a_3 + a_4 = 8, \\ & a_2 + \qquad a_4 \leqslant 2, \\ & \qquad a_3 + a_4 \leqslant 5, \\ & a_i \geqslant 0, i = 1, \dots, 4. \end{aligned}$$

Here, the greedy algorithm obtains the solution $a_4 = 2, a_3 = 3, a_1 = 3, a_2 = 0$ with cost 326. The optimal solution, however, is $a_1 = 1, a_2 = 2, a_3 = 5, a_4 = 0$ with cost 162.

Lemma 3.5 implies that retaining feasibility we can find an optimal solution by performing the cheapest weight changes, until the optimality condition $D = 0$ is reached. To do so we arrange the given points P_i, $i = 0, 1, \dots, n$, in one list such that the coefficients c_i of points $P_i \in X_{\geqslant}$ and the coefficients d_i of $P_i \in X_{<}$ are sorted increasingly. Then the weights of the points are changed in this order. This ensures feasibility for X, since each such change reduces D. Additional restrictions for the weight changes are needed in order to obtain a solution which is feasible for Y, i.e. meets

$$WY_{<} - WY_{\geqslant} \leqslant 0$$

and

$$0 \leqslant WY_{\leqslant} - WY_{>}.$$

The actual value of D indicates the remaining value to be changed. After a new change by δ units there are $D - \delta$ units left to be changed. Then the solution must meet the feasibility conditions for Y, too. We consider three cases when deriving the additional restrictions:

- If $WY_{<}$ is increased by δ units, we get the conditions

  $$WY_{<} - WY_{\geqslant} \leqslant -\delta + (D - \delta) \tag{21}$$

  and

  $$-\delta + (D - \delta) \leqslant WY_{\leqslant} - WY_{>}. \tag{22}$$

  Inequality (21) implies that

  $$\delta \leqslant \tfrac{1}{2}(WY_{\geqslant} - WY_{<} + D) = b_1.$$

  If (21) is fulfilled, (22) is met anyway. An analogous restriction is needed when decreasing $WY_{>}$.
- If $WY_{<}$ is decreased it has to be assured that

  $$WY_{<} - WY_{\geqslant} \leqslant \delta - (D - \delta) \tag{23}$$

  and

  $$\delta - (D - \delta) \leqslant WY_{\leqslant} - WY_{>}. \tag{24}$$

  Inequality (24) implies

  $$\delta \leqslant \tfrac{1}{2}(WY_{\leqslant} - WY_{>} + D) = b_2.$$

  If (24) is met, then (23) is met, too. An analogous result is obtained for increasing $WY_{>}$.
- The inequalities which have to be met after decreasing $WY_{=}$ are

  $$WY_{<} - WY_{\geqslant} \leqslant \delta + (D - \delta) \tag{25}$$

  and

  $$\delta - (D - \delta) \leqslant WY_{\leqslant} - WY_{>}. \tag{26}$$

  Therefore, in this case

  $$\delta \leqslant \tfrac{1}{2}(WY_{\geqslant} - WY_{<} + D) = b_1$$

  and

  $$\delta \leqslant \tfrac{1}{2}(WY_{\leqslant} - WY_{>} + D) = b_2.$$

These considerations imply that we have to distinguish 6 cases for the change of the weights. Let us recall that initially $D > 0$, $b_1 = 1/2 \cdot (D - D_1)$ and $b_2 = 1/2 \cdot (D - D_2)$ holds.

- *Case* 0: $P_i \in R_0$. The weight can be increased by

$$\delta := \min(\bar{w}_i - w_i, D).$$

  After this change the value D has to be adjusted to $D - \delta$.
- *Case* 1: $P_i \in R_1$. The weight can be increased by

$$\delta := \min(\bar{w}_i - w_i, D, b_1).$$

  After this change D and b_1 have to be adjusted to $D - \delta$ and $b_1 - \delta$, respectively.
- *Case* 2: $P_i \in R_2$. The weight can be increased by

$$\delta := \min(\bar{w}_i - w_i, D, b_2).$$

  After this change D and b_2 have to be adjusted to $D - \delta$ and $b_2 - \delta$, respectively.
- *Case* 3: $P_i \in R_3$. The weight can be decreased by

$$\delta := \min(w_i - \underline{w}_i, D, b_1, b_2).$$

  After this change D, b_1 and b_2 are decreased by δ.
- *Case* 4: $P_i \in R_4$. The weight can be decreased by

$$\delta := \min(w_i - \underline{w}_i, D, b_2).$$

  After this change D and b_2 are decreased by δ.
- *Case* 5: $P_i \in R_5$. The weight can be decreased by

$$\delta := \min(w_i - \underline{w}_i, D, b_1).$$

  After this change D and b_1 are decreased by δ.

Before we state the algorithm we get rid of the assumption that $D = WX_< - WX_\geqslant$. In order to describe the coming arguments in a compact form, we introduce the following notation. Let Z be either X or Y. If $Z = X$, then $\bar{Z} = Y$, and vice versa. Moreover, let the relations A and B denote either $<$ or $>$. If A denotes $<$, then B denotes $>$ and $\bar{B}$ denotes $\geqslant$. Vice versa, if A denotes $>$, then B denotes $<$ and $\bar{B}$ denotes $\leq$. With this notation the preceding considerations can be summarized in Algorithm 2.

Altogether, we get

Proposition 3.6. *The inverse* 1-*median problem in the plane can be solved in* $O(n \log n)$ *time, provided the distances between the points are measured in the Manhattan metric.*

Algorithm 2. Solves the inverse 1-median problem under l_1-norm

1. Determine

$$D := \max\{\max_{Z=X,Y} |WZ_< - WZ_>| - WZ_=, 0\}.$$

2. **If** $D = 0$, **then** stop: an optimal solution has been found.
   Else let $Z \in \{X, Y\}$ and $A \in \{<, >\}$ be such that

$$D = WZ_A - WZ_{\bar{B}}.$$

   Define $\bar{Z} := Y$, **if** $Z = X$, and $\bar{Z} := X$, **if** $Z = Y$.
   Define $B := <$, **if** $A = >$, and $B := >$, **if** $A = <$.
   Define

$$b_1 := 1/2 \cdot (D + W\bar{Z}_\geqslant - W\bar{Z}_<),$$
$$b_2 := 1/2 \cdot (D + W\bar{Z}_\leqslant - W\bar{Z}_>).$$

3. Sort the cost coefficients c_i for points $P_i \in Z_{\bar{B}}$ and the cost coefficients d_i for points $P_i \in Z_A$ increasingly in one list.

4. Consider the point P_i with the smallest not yet considered cost. **If** there does not exist such a coefficient, **then** stop: no feasible solution exists.

   - **If** $P_i \in Z_{\bar{B}} \cap \bar{Z}_{=}$, **then**
$$\delta := \min(\bar{w}_i - w_i, D),$$
$$w_i := w_i + \delta, \quad D := D - \delta.$$
   - **If** $P_i \in Z_{\bar{B}} \cap \bar{Z}_{<}$, **then**
$$\delta := \min(\bar{w}_i - w_i, D, b_1),$$
$$w_i := w_i + \delta, \quad D := D - \delta, \quad b_1 := b_1 - \delta.$$
   - **If** $P_i \in Z_{\bar{B}} \cap \bar{Z}_{>}$, **then**
$$\delta := \min(\bar{w}_i - w_i, D, b_2),$$
$$w_i := w_i + \delta, \quad D := D - \delta, \quad b_2 := b_2 - \delta.$$
   - **If** $P_i \in Z_A \cap \bar{Z}_{=}$, **then**
$$\delta := \min(w_i - \underline{w}_i, D, b_1, b_2),$$
$$w_i := w_i - \delta, \quad D := D - \delta, \quad b_1 := b_1 - \delta, \quad b_2 := b_2 - \delta.$$
   - **If** $P_i \in Z_A \cap \bar{Z}_{<}$, **then**
$$\delta := \min(w_i - \underline{w}_i, D, b_2),$$
$$w_i := w_i - \delta, \quad D := D - \delta, \quad b_2 := b_2 - \delta.$$
   - **If** $P_i \in Z_A \cap \bar{Z}_{>}$, **then**
$$\delta := \min(w_i - \underline{w}_i, D, b_1),$$
$$w_i := w_i - \delta, \quad D := D - \delta, \quad b_1 := b_1 - \delta.$$

5. **If** $D = 0$, **then** stop: an optimal solution has been found. **Else** return to Step 4.

3.1. The inverse 1-median problem under l_∞-norm in the plane

In this section, denote the Manhattan distance between the points $P_i = (x_i, y_i)$ and $P_j = (x_j, y_j)$ by

$$l_1((x_i, y_i), (x_j, y_j)) := |x_i - x_j| + |y_i - y_j|$$

and the Tschebychev distance by

$$l_\infty((x_i, y_i), (x_j, y_j)) := \max\{|x_i - x_j|, |y_i - y_j|\}.$$

The Tschebychev distance can be led back to the Manhattan distance:

Lemma 3.7 (Hamacher [8]). *For $P = (x, y) \in \mathbb{R}^2$ let $T(P) = (\frac{1}{2}(x + y), \frac{1}{2}(-x + y))$. Then for all $P_i = (x_i, y_i)$ and $P_j = (x_j, y_j) \in \mathbb{R}^2$:*

$$l_\infty(P_i, P_j) = l_1(T(P_i), T(P_j)).$$

When applying transformation T of Lemma 3.7, the feasibility conditions in the l_∞-norm case,

$$\sum_{i=0}^{n} (w_i + p_i - q_i)[l_\infty(P_i, P_0) - l_\infty(P_i, P)] \leqslant 0, \quad \text{for all } P \in \mathbb{R}^2,$$

become

$$\sum_{i=0}^{n} (w_i + p_i - q_i)[l_1(T(P_i), T(P_0)) - l_1(T(P_i), T(P))] \leqslant 0, \quad \text{for all } P \in \mathbb{R}^2.$$

This shows that after having transformed all given points $P_i = (x_i, y_i)$, $i = 0, 1, \dots, n$, to $P_i = (\frac{1}{2}(x_i + y_i), \frac{1}{2}(-x_i + y_i))$, $i = 0, 1, \dots, n$, Algorithm 2 yields an optimal solution of the problem:

Proposition 3.8. *The inverse* 1*-median problem under l_∞-norm in the plane can be solved in* $O(n \log n)$ *time with Algorithm* 2 *after having applied transformation T of Lemma* 3.7 *to the points.*

3.2. Outlook

In the discrete case only the inverse 1-median problem on trees is easy to solve. For the problem on general graphs no strongly polynomial algorithm is known. The inverse 2-median problem on trees is also unsolved. In the planar case, the inverse 1-median problem under l_1- or l_∞-norm can be solved with a greedy-type algorithm. For spaces with a dimension higher than two, polynomial algorithms are not yet known. Note that the above algorithm cannot be adopted for spaces $\mathbb{R}^n (n \geqslant 3)$, since there a greedy algorithm will not always yield an optimal solution. In particular, the weight changes in an optimal solution do not necessarily sum up to the optimality gap.

Example. Consider an inverse 1-median problem in $\mathbb{R}^3$ endowed with the Manhattan metric. Let the following data be given. The points are $\{P_0=(5,4,6), P_1=(1,1.5,1), P_2=(8.5,5.5,3.3), P_3=(7.5,3.5,6.5), P_4=(7,5,7.8), P_5=(4.3,4.5,6)\}$. The initial weights are $w=(4,7,10,5,2,5)$, the costs are $c=(80,50,1,1,10,1)$, $d=(90,10,50,1,20,50)$. All weights have a lower bound 0 and an upper bound 10. The analogous greedy algorithm in $\mathbb{R}^3$ would yield the solution $q_2=1$ with cost 50. The optimal solution, however, is $p_3=1, q_4=1, p_5=1$ with cost 22.

Further open problems concern continuous inverse p-median problems, $p \geqslant 1$, where the distances are measured by l_2- or general l_p-norms as well as obnoxious inverse p-median problems.

Acknowledgements

We thank Clemens Heuberger and Bettina Klinz for their constructive remarks during the preparation of this paper. Further, we thank two anonymous referees for their comments on the first version of this paper.

Appendix

In order to prove Lemma 3.3 we formulate the inverse 1-median problem as the following linear program:

$$
\begin{aligned}
\min \quad & \sum_{i=0}^{n} (c_i p_i + d_i q_i) \\
\text{s.t.} \quad & \sum_{i\in R_0}(p_i-q_i) + \sum_{i\in R_1}(p_i-q_i) + \sum_{i\in R_2}(p_i-q_i) - \sum_{i\in R_3}(p_i-q_i) - \sum_{i\in R_4}(p_i-q_i) - \sum_{i\in R_5}(p_i-q_i) \geqslant D, \\
& \sum_{i\in R_0}(p_i-q_i) - \sum_{i\in R_1}(p_i-q_i) + \sum_{i\in R_2}(p_i-q_i) + \sum_{i\in R_3}(p_i-q_i) - \sum_{i\in R_4}(p_i-q_i) + \sum_{i\in R_5}(p_i-q_i) \geqslant D_1, \\
& \sum_{i\in R_0}(p_i-q_i) + \sum_{i\in R_1}(p_i-q_i) - \sum_{i\in R_2}(p_i-q_i) + \sum_{i\in R_3}(p_i-q_i) + \sum_{i\in R_4}(p_i-q_i) - \sum_{i\in R_5}(p_i-q_i) \geqslant D_2, \\
& -p_i \geqslant w_i - \bar{w}_i, \quad i=0,1,\ldots,n, \\
& -q_i \geqslant \underline{w}_i - w_i, \quad i=0,1,\ldots,n, \\
& p_i \geqslant 0, q_i \geqslant 0, \quad i=0,1,\ldots,n.
\end{aligned}
\tag{A.1}
$$

The corresponding dual program is

$$
\begin{aligned}
\max \quad & D y_1 + D_1 y_2 + D_2 y_3 + \sum_{i=0}^{n}(w_i - \bar{w}_i)\bar{z}_i + \sum_{i=0}^{n}(\underline{w}_i - w_i)\underline{z}_i \\
\text{s.t.} \quad & y_1 + y_2 + y_3 - \bar{z}_i \leqslant c_i, \quad \forall i \in R_0, \\
& y_1 - y_2 + y_3 - \bar{z}_i \leqslant c_i, \quad \forall i \in R_1, \\
& y_1 + y_2 - y_3 - \bar{z}_i \leqslant c_i, \quad \forall i \in R_2, \\
& -y_1 + y_2 + y_3 - \bar{z}_i \leqslant c_i, \quad \forall i \in R_3, \\
& -y_1 - y_2 + y_3 - \bar{z}_i \leqslant c_i, \quad \forall i \in R_4, \\
& -y_1 + y_2 - y_3 - \bar{z}_i \leqslant c_i, \quad \forall i \in R_5, \\
& -y_1 - y_2 - y_3 - \underline{z}_i \leqslant d_i, \quad \forall i \in R_0,
\end{aligned}
$$

$$
\begin{aligned}
-y_1 + y_2 - y_3 - \underline{z}_i &\leqslant d_i, \quad \forall i \in R_1,\\
-y_1 - y_2 + y_3 - \underline{z}_i &\leqslant d_i, \quad \forall i \in R_2,\\
y_1 - y_2 - y_3 - \underline{z}_i &\leqslant d_i, \quad \forall i \in R_3,\\
y_1 + y_2 - y_3 - \underline{z}_i &\leqslant d_i, \quad \forall i \in R_4,\\
y_1 - y_2 + y_3 - \underline{z}_i &\leqslant d_i, \quad \forall i \in R_5,\\
y_1 \geqslant 0, y_2 \geqslant 0, y_3 \geqslant 0, \bar{z}_i \geqslant 0, \underline{z}_i &\geqslant 0, \quad i = 0, 1, \ldots, n.
\end{aligned}
$$

Let a feasible solution $(y_1, y_2, y_3, \bar{z}_1, \ldots, \bar{z}_n, \underline{z}_1, \ldots, \underline{z}_n)$ of the dual problem be given. We shall prove that it is possible to obtain a new solution (by changing y_1, y_2 and y_3 to $\bar{y}_1, \bar{y}_2$ and $\bar{y}_3$) where $\bar{y}_1 > 0$ and the new objective value are not smaller than the old one. Then it follows from the complementary slackness conditions that there is always an optimal solution where the first inequality in the primal problem holds with equality. For our purpose it suffices to consider the values of $\bar{z}_1, \ldots, \bar{z}_n, \underline{z}_1, \ldots, \underline{z}_n$ as fixed. So, setting

$$\hat{c}_i := c_i + \bar{z}_i, \quad i = 0, 1, \ldots, n,$$

$$\hat{d}_i := d_i + \underline{z}_i, \quad i = 0, 1, \ldots, n,$$

we just have to show that $\bar{y}_1, \bar{y}_2$ and $\bar{y}_3$ with $\bar{y}_1 > 0$ is a better or equally good feasible solution (than that one initially given) of:

$$
\begin{aligned}
\max \quad & D y_1 + D_1 y_2 + D_2 y_3\\
\text{s.t.} \quad & y_1 + y_2 + y_3 \leqslant \min\{\hat{c}_i \mid i \in R_0\} =: \tilde{c}_0,\\
& y_1 - y_2 + y_3 \leqslant \min\{\hat{c}_i | i \in R_1\} =: \tilde{c}_1,\\
& y_1 + y_2 - y_3 \leqslant \min\{\hat{c}_i | i \in R_2\} =: \tilde{c}_2,\\
& -y_1 + y_2 + y_3 \leqslant \min\{\hat{c}_i | i \in R_3\} =: \tilde{c}_3,\\
& -y_1 - y_2 + y_3 \leqslant \min\{\hat{c}_i | i \in R_4\} =: \tilde{c}_4,\\
& -y_1 + y_2 - y_3 \leqslant \min\{\hat{c}_i | i \in R_5\} =: \tilde{c}_5,\\
& -y_1 - y_2 - y_3 \leqslant \min\{\hat{d}_i | i \in R_0\} =: \tilde{d}_0,\\
& -y_1 + y_2 - y_3 \leqslant \min\{\hat{d}_i | i \in R_1\} =: \tilde{d}_1,\\
& -y_1 - y_2 + y_3 \leqslant \min\{\hat{d}_i | i \in R_2\} =: \tilde{d}_2,\\
& y_1 - y_2 - y_3 \leqslant \min\{\hat{d}_i | i \in R_3\} =: \tilde{d}_3,\\
& y_1 + y_2 - y_3 \leqslant \min\{\hat{d}_i | i \in R_4\} =: \tilde{d}_4,\\
& y_1 - y_2 + y_3 \leqslant \min\{\hat{d}_i | i \in R_5\} =: \tilde{d}_5,\\
& y_1 \geqslant 0, y_2 \geqslant 0, y_3 \geqslant 0.
\end{aligned}
$$

Assume that $y_1 = 0$ (otherwise we are finished) and set

$$\tilde{c} := \min\{\tilde{c}_0, \ldots, \tilde{c}_5, \tilde{d}_0, \ldots, \tilde{d}_5\}.$$

We first consider the case $\tilde{c} > 0$. In order to prove that there is a feasible solution with $\bar{y}_1 > 0$ and objective value (obj) of $(\bar{y}_1, \bar{y}_2, \bar{y}_3) \geqslant$ objective value of (y_1, y_2, y_3) we distinguish 4 cases. Notice that $D \geqslant D_1$ and $D \geqslant D_2$.

1. $y_2 = 0, y_3 = 0$.
   - Set $\bar{y}_1 := \tilde{c}, \bar{y}_2 := y_2 = 0, \bar{y}_3 := y_3 = 0$.
   - This solution is feasible.
   - $obj(\bar{y}_1, \bar{y}_2, \bar{y}_3) = \tilde{c}D > 0 = obj(y_1, y_2, y_3)$.
2. $y_2 > 0, y_3 = 0$.
   - Set $\bar{y}_1 := \varepsilon, \bar{y}_2 := y_2 - \varepsilon, \bar{y}_3 := y_3 = 0$,
     where $\varepsilon > 0, \varepsilon \leqslant \frac{y_2}{2}$.

- This solution is feasible, since:

$$\bar{y}_1 + \bar{y}_2 = \varepsilon + y_2 - \varepsilon = y_1 + y_2 = y_2,$$
$$-\bar{y}_1 + \bar{y}_2 = -\varepsilon + y_2 - \varepsilon < y_2,$$
$$-\bar{y}_1 - \bar{y}_2 = -\varepsilon - y_2 + \varepsilon = -y_2,$$
$$\bar{y}_1 - \bar{y}_2 = \varepsilon - y_2 + \varepsilon \leqslant y_2 - y_2 = 0 \leqslant \tilde{c}.$$

- $obj(\bar{y}_1, \bar{y}_2, \bar{y}_3) = \varepsilon D + D_1(y_2 - \varepsilon) = \varepsilon(D - D_1) + D_1 y_2 \geqslant D_1 y_2 = obj(y_1, y_2, y_3).$

3. $y_2 = 0, y_3 > 0$
   can be dealt with in a manner analogous to the previous case.
4. $y_2 > 0, y_3 > 0.$
   (a) $-y_2 + y_3 < \min\{\tilde{c}_1, \tilde{d}_5\}.$
   - Set $\bar{y}_1 := \varepsilon, \bar{y}_2 := y_2 - \varepsilon, \bar{y}_3 := y_3,$
     where $\varepsilon > 0, \varepsilon \leqslant y_2, \varepsilon \leqslant \frac{1}{2}(\min\{\tilde{c}_1, \tilde{d}_5\} + y_2 - y_3), \varepsilon \leqslant (y_2 + y_3 + \tilde{d}_3)/2.$
   - This solution is feasible, because:

$$\bar{y}_1 + \bar{y}_2 + \bar{y}_3 = \varepsilon + y_2 - \varepsilon + y_3 = y_2 + y_3 \leqslant \tilde{c}_0,$$
$$\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = \varepsilon - y_2 + \varepsilon + y_3 = -y_2 + y_3 + 2\varepsilon \leqslant -y_2 + y_3 + y_2 - y_3 + \min\{\tilde{c}_1, \tilde{d}_5\} \leqslant \tilde{c}_1,$$
$$\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = \varepsilon + y_2 - \varepsilon - y_3 = y_2 - y_3 \leqslant \tilde{c}_2,$$
$$-\bar{y}_1 + \bar{y}_2 + \bar{y}_3 = -\varepsilon + y_2 - \varepsilon + y_3 = y_2 + y_3 - 2\varepsilon \leqslant y_2 + y_3 \leqslant \tilde{c}_3,$$
$$-\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = -\varepsilon - y_2 + \varepsilon + y_3 = -y_2 + y_3 \leqslant \tilde{c}_4,$$
$$-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = -\varepsilon + y_2 - \varepsilon - y_3 = y_2 - y_3 - 2\varepsilon \leqslant y_2 - y_3 \leqslant \tilde{c}_5,$$
$$-\bar{y}_1 - \bar{y}_2 - \bar{y}_3 = -\varepsilon - y_2 + \varepsilon - y_3 = -y_2 - y_3 \leqslant \tilde{d}_0,$$
$$-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = -\varepsilon + y_2 - \varepsilon - y_3 = y_2 - y_3 - 2\varepsilon \leqslant y_2 - y_3 \leqslant \tilde{d}_1,$$
$$-\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = -\varepsilon - y_2 + \varepsilon + y_3 = -y_2 + y_3 \leqslant \tilde{d}_2,$$
$$\bar{y}_1 - \bar{y}_2 - \bar{y}_3 = \varepsilon - y_2 + \varepsilon - y_3 = -y_2 - y_3 + 2\varepsilon \leqslant -y_2 - y_3 + y_2 + y_3 + \tilde{d}_3 \leqslant \tilde{d}_3,$$
$$\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = \varepsilon + y_2 - \varepsilon - y_3 = y_2 - y_3 \leqslant \tilde{d}_4,$$
$$\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = \varepsilon - y_2 + \varepsilon + y_3 = -y_2 + y_3 + 2\varepsilon \leqslant -y_2 + y_3 + y_2 - y_3 + \min\{\tilde{c}_1, \tilde{d}_5\} \leqslant \tilde{d}_5.$$

   - $obj(\bar{y}_1, \bar{y}_2, \bar{y}_3) = \varepsilon D + D_1(y_2 - \varepsilon) + D_2 y_3 = \varepsilon(D - D_1) + D_1 y_2 + D_2 y_3 \geqslant D_1 y_2 + D_2 y_3 = obj(y_1, y_2, y_3).$

   (b) $-y_2 + y_3 = \min\{\tilde{c}_1, \tilde{d}_5\}.$
   - Set $\bar{y}_1 := \varepsilon, \bar{y}_2 := y_2, \bar{y}_3 := y_3 - \varepsilon,$
     where $\varepsilon > 0, \varepsilon \leqslant y_3, \varepsilon \leqslant (-y_2 + y_3 + \tilde{c}_2)/2, \varepsilon \leqslant (-y_2 + y_3 + \tilde{d}_4)/2, \varepsilon \leqslant (y_2 + y_3 + \tilde{d}_3)/2.$
   - This solution is feasible because:

$$\bar{y}_1 + \bar{y}_2 + \bar{y}_3 = \varepsilon + y_2 + y_3 - \varepsilon = y_2 + y_3 \leqslant \tilde{c}_0,$$
$$\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = \varepsilon - y_2 + y_3 - \varepsilon = -y_2 + y_3 \leqslant \tilde{c}_1,$$
$$\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = \varepsilon + y_2 - y_3 + \varepsilon = y_2 - y_3 + 2\varepsilon \leqslant y_2 - y_3 + y_3 - y_2 + \tilde{c}_2 \leqslant \tilde{c}_2,$$
$$-\bar{y}_1 + \bar{y}_2 + \bar{y}_3 = -\varepsilon + y_2 + y_3 - \varepsilon = y_2 + y_3 - 2\varepsilon \leqslant y_2 + y_3 \leqslant \tilde{c}_3,$$
$$-\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = -\varepsilon - y_2 + y_3 - \varepsilon = -y_2 + y_3 - 2\varepsilon \leqslant -y_2 + y_3 \leqslant \tilde{c}_4,$$
$$-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = -\varepsilon + y_2 - y_3 + \varepsilon = y_2 - y_3 \leqslant \tilde{c}_5,$$
$$-\bar{y}_1 - \bar{y}_2 - \bar{y}_3 = -\varepsilon - y_2 - y_3 + \varepsilon = -y_2 - y_3 \leqslant \tilde{d}_0,$$
$$-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = -\varepsilon + y_2 - y_3 + \varepsilon = y_2 - y_3 \leqslant \tilde{d}_1,$$
$$-\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = -\varepsilon - y_2 + y_3 - \varepsilon = -y_2 + y_3 - 2\varepsilon \leqslant -y_2 + y_3 \leqslant \tilde{d}_2,$$
$$\bar{y}_1 - \bar{y}_2 - \bar{y}_3 = \varepsilon - y_2 - y_3 + \varepsilon = -y_2 - y_3 + 2\varepsilon \leqslant -y_2 - y_3 + y_2 + y_3 + \tilde{d}_3 \leqslant \tilde{d}_3,$$
$$\bar{y}_1 + \bar{y}_2 - \bar{y}_3 = \varepsilon + y_2 - y_3 + \varepsilon = y_2 - y_3 + 2\varepsilon \leqslant y_2 - y_3 - y_2 + y_3 + \tilde{d}_4 \leqslant \tilde{d}_4,$$
$$\bar{y}_1 - \bar{y}_2 + \bar{y}_3 = \varepsilon - y_2 + y_3 - \varepsilon = -y_2 + y_3 \leqslant \tilde{d}_5.$$

   - $obj(\bar{y}_1, \bar{y}_2, \bar{y}_3) = \varepsilon D + D_1 y_2 + D_2(y_3 - \varepsilon) = \varepsilon(D - D_2) + D_1 y_2 + D_2 y_3 \geqslant D_1 y_2 + D_2 y_3 = obj(y_1, y_2, y_3).$

This shows that—for $\tilde{c} > 0$—given a feasible solution of the dual problem, a better or equal good solution with $\bar{y}_1 > 0$ can be achieved. By the complementary slackness conditions, there is always an optimal solution where the first inequality in the linear program holds with equality, i.e. $a_1 + a_2 + a_3 + a_4 = D$.

The case $\tilde{c} = 0$ can be reduced to the positive one in the following way. The inverse 1-median problem consists in solving a linear program of the form:

$$\begin{aligned} \min \quad & c^{\mathrm{T}} x \\ \text{s.t.} \quad & Ax \leqslant b \\ & x \geqslant 0. \end{aligned} \tag{A.2}$$

Consider a perturbed problem of (A.2):

$$\begin{aligned} &\min\,(c+\varepsilon)^{\mathrm{T}}x \\ &\text{s.t.}\;\; Ax \leqslant b \\ &\qquad x \geqslant 0, \end{aligned} \tag{A.3}$$

with $\varepsilon > 0$. In [6] it is shown that the optimal solution of the dual problem of (A.3) converges to an optimal solution of the dual problem of (A.2) as $\varepsilon \to 0$. Then it follows from strong duality that the optimal solution of (A.3) converges to an optimal solution of (A.2), too. Thus, there is an optimal solution with $a_1 + a_2 + a_3 + a_4 = D$ also in the case $\tilde{c} = 0$. □

References

[1] O. Berman, D.I. Ingco, A.R. Odoni, Improving the location of minisum facilities through network modification, Ann. Oper. Res. 40 (1992) 1–16.

[2] O. Berman, D.I. Ingco, A.R. Odoni, Improving the location of minimax facilities through network modification, Networks 24 (1994) 31–41.

[3] D. Burton, Ph.L. Toint, On an instance of the inverse shortest path problem, Math. Programming 53 (1992) 45–61.

[4] M.C. Cai, X.G. Yang, J.Z. Zhang, The complexity analysis of the inverse center location problem, J. Global Optim. 15 (1999) 213–218.

[5] P. Cappanera, A survey on obnoxious facility location problems, Report TR-99-11, Dip. di Informatica, Univ. di Pisa, 1999.

[6] G.B. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, NJ, 1998.

[7] A.J. Goldman, Optimal center location in simple networks, Transportation Sci. 2 (1962) 77–91.

[8] H.W. Hamacher, Mathematische Lösungsverfahren für planare Standortprobleme, Vieweg, Braunschweig/Wiesbaden, 1995.

[9] C. Heuberger, Inverse optimization: a survey on problems, methods, and results, SFB Report No. 219, Institute of Mathematics B, Graz University of Technology, Graz (Austria), May 2001; J. Combin. Optim. (2004), to appear.

[10] Hua, et al., Applications of mathematical models to wheat harvesting, Acta Mathematica Sinica 11 (1961) 63–75 (in Chinese) (English translation in Chinese Math. 2, 1962, 77–91).

[11] F. Plastria, Optimal location of undesirable facilities: a selective overview, Belg. J. Oper. Res. Statist. Comput. Sci. 36 (1996) 109–127.

[12] J. Zhang, Z. Liu, Z. Ma, Some reverse location problems, European J. Oper. Res. 124 (2000) 77–88.