Deterministic Algorithm for 1-Median 1-Center Two-Objective Optimization Problem

Vahid Roostapour(✉), Iman Kiarazm, and Mansoor Davoodi

Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran
{v.roostapour,i.kiarazm,mdmonfared}@iasbs.ac.ir

Abstract. *k-median* and *k-center* are two well-known problems in facility location which play an important role in operation research, management science, clustering and computational geometry. To the best of our knowledge, although these problems have lots of applications, they have never been studied together simultaneously as a multi objective optimization problem. Multi-objective optimization has been applied in many fields of science where optimal decisions need to be taken in the presence of trade-offs between two or more conflicting objectives. In this paper we consider 1-median and 1-center two-objective optimization problem. We prove that $\Omega(n \log n)$ is a lower bound for proposed problem in one and two dimensions in Manhattan metric. Also, by using the properties of farthest point Voronoi diagram, we present a deterministic algorithm which output the Pareto Front and Pareto Optimal Solutions in $\mathcal{O}(n \log n)$ time.

Keywords: Computational geometry · Pareto optimal solutions · 1-center · 1-median · Multi-objective optimization

1 Introduction

When evaluating different solutions from a design space, it is often the case that more than one criterion comes into play. For example, when choosing a route to drive from one point to another, we may care about the time it takes, the distance traveled and the complexity of the route (e.g. number of turns). When designing a (wired or wireless) network, we may consider its cost, capacity and coverage. Such problems are known as *Multi-Objective Optimization Problems* (MOOP). Multi-objective optimization can be described in mathematical terms as follows:

$$S = \{x \in \mathbb{R}^d : h(x) = 0, g(x) \geq 0\}$$
$$\min\,[f_1(x), f_2(x), \ldots, f_N(x)]$$
$$x \in S,$$

where $N > 1$, f_i is a scalar function for $1 \leq i \leq N$ and S is the set of constraints.

The space in which the objective vector belongs is called *objective space*. The scalar concept of optimality does not apply directly in the multi-objective setting.

M.T. Hajiaghayi and M.R. Mousavi (Eds.): TTCS 2015, LNCS 9541, pp. 164–178, 2016.
DOI: 10.1007/978-3-319-28678-5_12

Here the notion of *Pareto optimality* and *dominance* has to be introduced. In a multi-objective *minimization* problem, a solution $s_1 \in S$ dominates a solution $s_2 \in S$, denoted by $s_1 \prec s_2$, if $f_i(s_1) \leq f_i(s_2)$ for all $i \in \{1, \ldots, N\}$, with at least one strict inequality. A point s^* is said to be a *Pareto optimum* or a *Pareto optimal solution* for the multi-objective problem if and only if there is no $s \in S$ such that $s \prec s^*$. The image of such an efficient set, i.e., the image of all the efficient solutions in the objective space are called *Pareto optimal front* or *Pareto curve*.

One of the common approaches for such problems is evolutionary algorithms [7]. These algorithms are iterative and converge to Pareto front. However they need more time as the complexity of the Pareto front increases. Moreover, all of these approaches have major problems with local optimums. On the other hand there are some classical approaches like *weighted sum* and *ϵ-constraint* which can apply on MOOPs. Although these approaches guarantee finding solutions on the entire Pareto optimal set for problems having a convex Pareto front, they are largely depend on chosen weight and ϵ vectors respectively. Moreover, these approaches require some information from user about the solution space. Furthermore, in most nonlinear MOOPs, a uniformly distributed set of weight vectors wont necessarily find a uniformly distributed set of Pareto optimal solutions. Also there may exist multiple minimum solutions for a specific weight vector [8]. However we find the Pareto front of a MOOP with deterministic algorithm. Here we consider two famous propounded facility location problems [17].

k-median: In this problem the goal is to minimize summation of distances between each demand point and its nearest center. Charikar et al. proposed the first constant time approximation algorithm which its outputs is $6\frac{2}{3}$ times the optimal [5]. This improved upon the best previously known result of $\mathcal{O}(\log p \log\log p)$, which was obtained by refining and derandomizing a randomized $\mathcal{O}(\log n \log\log n)$-approximation algorithm of Bartal [4]. The currently best known approximation ratio is $3 + \epsilon$ achieved by a local search heuristic of Arya et al. [1]. Moreover, Jain et al. proved that the k-median problem cannot be approximated within a factor strictly less than $1 + 2/e$, unless $\mathsf{NP} \subseteq \mathsf{DTIME}[n^{\mathcal{O}(\log\log n)}]$ [12]. This was an improvement over a lower bound of $1 + 1/e$ [16]. Using sampling technique Meyerson, et al. presented an algorithm with running time $\mathcal{O}(p(\frac{p^2}{\epsilon}\log p)^2 \log(\frac{p}{\epsilon}\log p))$. This was the first k-median algorithm with fully polynomial running time that was independent of n, the size of the data set. It presented a solution that is, with high probability, an $\mathcal{O}(1)$-approximation, if each cluster in some optimal solution has $\Omega(\frac{n\cdot\epsilon}{p})$ points [14]. Har-Peled and Kushal presented a (p, ϵ)-coreset of size $O(p^2/\epsilon^d)$ for k-median clustering of n points in $\mathbb{R}^d$, which its size was independent of n [9]. Also, Har-Peled and Mazumdar showed that there exist small coresets of size $\mathcal{O}(p\epsilon^{-d}\log n)$ for the problems of computing k-median clustering for points in low dimension with $(1+\epsilon)$-approximation. Their algorithm has linear running time for a fixed p and ϵ [10]. Moreover, using random sampling for k-median problem Badoiu et al. proposed a $(1+\epsilon)$-approximation algorithm with $2^{(p/\epsilon)^{\mathcal{O}(1)}} d^{\mathcal{O}(1)} n \log^{\mathcal{O}(p)} n$ expected time [3].

k-center: In this problem the goal is to minimize the maximum distance between each demand point from its nearest center. Megiddo and Supowit proved that k-center and k-median are NP-hard even to approximate the k-center problems sufficiently closely [13]. Hochbaum and Shmoys proposed the first constant factor approximation algorithm which its output is 2 times the optimal. It is the best possible algorithm unless $\mathsf{P} \neq \mathsf{NP}$ [11]. It is shown that there is an algorithm with $\mathcal{O}(d^{\mathcal{O}(d)}n)$ time for 1-center problem [6]. In the high dimension, Badoiu and Clarkson presented a $(1+\epsilon)$-approximation algorithm which find a solution in $\lceil 2/\epsilon \rceil$ passes using $\mathcal{O}(nd/\epsilon + (1/\epsilon)^5)$ total time and $\mathcal{O}(d/\epsilon)$ space [2]. Also, for problem of 1-center with outliers, Zarrabi-Zadeh and Mukhopadhyay proposed a 2-approximation one pass streaming algorithm in high dimension which for z, as the number of outliers, needs $\mathcal{O}(zd^2)$ space [19]. Moreover, Zarrabi-Zadeh and Chan presented an streaming one pass 3/2-approximation algorithm for 1-center [18]. Badoiu et al. for 1-center problem, extracted a core-set of size $\mathcal{O}(1/\epsilon^2)$ which its solution is $(1+\epsilon)$-approximation set of points in $\mathbb{R}^d$ [3]. Also, for k-center they presented a $2^{\mathcal{O}((p \log p)/\epsilon^2)}.dn$ time algorithm with $(1+\epsilon)$-approximation solution using previous result.

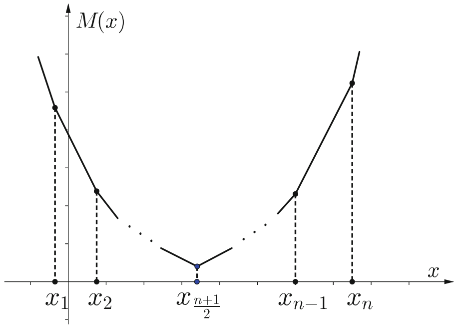
1-median and 1-center are practical problems which have not been considered as a two-objective optimization problem yet. Imagine mayor of a small city wants to build a fire station in a way that minimizes the distance between farthest building to the station, also since the number of fire engines is limited and each fire engine must return to the station after a service, it has to minimizes the total distance of station from all other buildings. As an another example, consider power distribution network. Due to the dependency of energy leakage to wire length, minimizing of the longest wire in the network would be regarded as an essential factor. Also, any decrement in total wire length of network considered as a second objective. The first objective is 1-median, $M(u)$, the summation of distances of demand points from center u and the second objective is 1-center, $C(u)$, the farthest input point from center p. It can be described in mathematical terms as follow:

Definition 1. ***1-Median 1-Center Two-Objective Optimization Problem:*** *Let $P = \{p_1, \dots p_n\}$ be a set of demand points in $\mathbb{R}^d$. Consider functions $M(u) = \sum_{i=1}^{n} D(u, p_i)$ and $C(u) = \max_{1 \leq i \leq n} D(u, p_i)$ are the values of point $u \in \mathbb{R}^d$ as a center for 1-median and 1-center objectives respectively for a certain distance function D. The goal is finding u^* to minimize the objectives.*

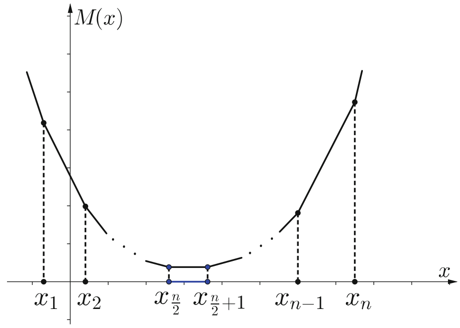
We study this problem in one and two dimensions in Manhattan metric. We assume no input points have the same x or y coordinate.

This is a convex combinatorial multi-objective optimization problem which has been studied with a different approach called ϵ-Pareto. In [15] it is shown that this approximate Pareto curve can be constructed in time polynomial in the size of the instance and $1/\epsilon$, but here we propose a deterministic algorithm for computing the exact Pareto curve because of specifying the problem.

This paper starts with considering 1-median and 1-center as two-objective of MOOP in one dimension. We will find the optimal of objectives and in terms of placement of optimums we will also find the Pareto set in time $\mathcal{O}(n)$ (Lemma 1).



(a) n is odd.



(b) n is even.

Fig. 1. 1-median optimal.

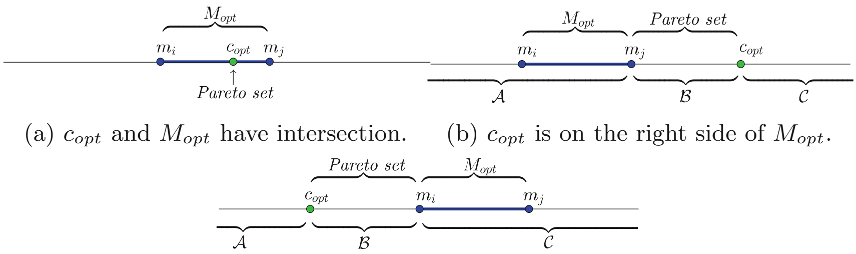
We continue with a proof for convexity of Pareto set. At the end of second section we give an algorithm to compute Pareto optimal front of 1-median 1-center two-objective optimization problem and prove the optimality of the algorithm. In section three the same problem considered in two dimensional space. First we find optimums of 1-median objective. After that by using the properties of farthest point Voronoi diagram we determine the optimum of 1-center. Finally after limiting the solution space to regions which Pareto set lies on, we specifically present Pareto solutions. Convexity of Pareto front is proven in Theorem 2.

2 One Dimensional

Let $P = \{x_1, x_2, \ldots x_n\}$ be a set of input points in one dimension, the goal is to minimize $M(x) = \sum_{i=1}^{n} |x - x_i|$ and $C(x) = \max_{1 \le i \le n} |x - x_i|$. According to the properties of the absolute value function and some simple calculations, it is easy to see that $M(x)$ is a continuous piecewise linear function which its minimum depends on n. The minimum can either be one point or an interval which we denote by M_{opt} in the rest of the paper. Also without loss of generality we assume that input points are sorted increasingly. In one dimensional space, $M_{opt} = [m_i, m_j] \subset \mathbb{R}$ for $1 \le i, j \le n$ such that $m_i = x_i$, $m_j = x_j$. For odd n we have $j = i$ and for even n, $j = i + 1$. Moreover, the function is strictly decreasing before its minimum and is strictly increasing after it (Fig. 1). For $C(x)$ suppose $c_{opt} \in \mathbb{R}$ denote the point which $C(c_{opt})$ is minimum. Obviously $c_{opt} = (x_1 + x_n)/2$. Similarly to $M(x)$, $C(x)$ is strictly decreasing before optimal point and strictly increasing after that.

Lemma 1. *Pareto optimal set in one dimensional* 1*-median* 1*-center two-objective optimization problem is the smallest interval consisting of a solution with* 1*-center optimal and a solution with* 1*-median optimal.*

Proof. Suppose that n is even (the proof is similar for odd n). As shown in Fig. 2, there are three different cases:



(a) c_{opt} and M_{opt} have intersection. (b) c_{opt} is on the right side of M_{opt}.

(c) c_{opt} is on the left side of M_{opt}.

Fig. 2. Pareto set computation in one dimension.

First consider the case that c_{opt} and M_{opt} have an intersection (Fig. 2a). In this case the intersection point is the only member of Pareto optimal solutions. Because not only it is optimal in both objectives, but also it is the only point where $C(x)$ is optimal. So it dominates all the other solutions and no solution dominates it.

As shown in Fig. 2b there are three regions in the second case. In region $\mathcal{C}$ both functions are strictly increasing. Therefore, c_{opt} has the best value in both objectives. It dominates all solutions of this region. In $\mathcal{A}$, $C(x)$ is strictly decreasing, thus $C(m_j)$ is strictly smaller than 1-center objective of all the other solutions. Moreover, $M(m_j)$ is smaller than or equal with 1-median objective of the other solutions. Hence m_j dominates all solutions of $\mathcal{A}$. Finally we claim that $\mathcal{B}$ is Pareto set. By contradiction, suppose it is not true, then there must be a point p which dominates $q \in \mathcal{B}$. It has to be on the left side or right side of q. Let p be on the right side, we know that $M(x)$ is strictly increasing in this side. Hence $M(q) < M(p)$ and it contradicts with dominance of p. Similarly there is a contradiction if p lies on the left side of q, because $C(x)$ is strictly decreasing in this side, i.e. $C(q) < C(p)$. This implies that all the solutions that lie on $\mathcal{B}$ are Pareto set.

The proof is similar for the third case which c_{opt} is on the left side of M_{opt} (Fig. 2c). $\square$

Lemma 2. *Pareto optimal front of one dimensional* 1*-median* 1*-center two-objective optimization problem forms a continuous, convex and piecewise linear function.*

Proof. If there is an intersection between c_{opt} and M_{opt} the lemma is held. Now suppose there is no such intersection and consider c_{opt} is on the right side of M_{opt} (resp. on the left side of M_{opt}). From Lemma 1 for Pareto solutions we have $P_s = [m_j, c_{opt}]$ (resp. $P_s = [c_{opt}, m_i]$). Since $C(x)$ derivation is constant and $M(x)$ is piecewise linear in P_s, the diagram of $M(x)$-$C(x)$ is piecewise linear and break points are $\big(C(x_i), M(x_i)\big)$ such that $m_j \leq x_i \leq c_{opt}$ (resp. $c_{opt} \leq x_i \leq m_i$). The absolute value of slope of $M(x)$ increases on each linear piece in P_s. Thus

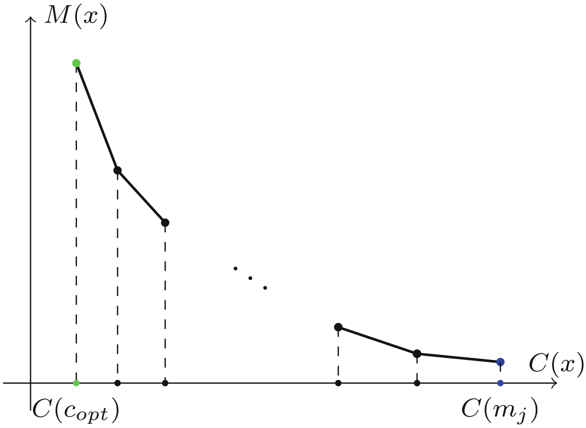


Fig. 3. One dimensional 1-median 1-center two-objective Pareto optimal front.

Pareto optimal front is convex (Fig. 3). Also we can conclude that piecewise linear Pareto front is one-to-one and invertible corresponding to Pareto solutions. □

Lemma 3. *Computing Pareto front of one dimensional* 1*-median* 1*-center two-objective optimization problem requires* $\Omega(n \log n)$ *time.*

Proof. The proof is based on reduction from *sorting problem*. By contradiction assume there is an algorithm which return set $O = \big\{\big(C(\alpha_1), M(\alpha_1)\big) \cdots, \big(C(\alpha_m), M(\alpha_m)\big)\big\}$ –lexicographical ordered break points of the piecewise linear Pareto front function– besides the Pareto solutions interval in $o(n \log n)$ running time. Let $A = \{a_1, \ldots, a_n\}$ is the set of input values of sorting problem, $l = \arg\min_{1 \le i \le n} a_i$ and $h = \arg\max_{1 \le i \le n} a_i$. Suppose $b_1, \ldots, b_{n+1}$ and t are values such that $b_1 < \cdots < b_{n+1} < a_l$ and $t = 2 \cdot a_h - b_1 + 1$, then $B = A \cup \{b_1, \cdots, b_{n+1}, t\}$ is defined in $\mathcal{O}(n)$. For the set B as input points of *one dimensional* 1*-median* 1*-center two-objective optimization*, 1-median optimal interval is $[b_{n+1}, a_l]$ and 1-center optimal point is between a_h and t. Using lemma 2 we conclude that $m = n + 1$ and $\alpha_1 = a_l < \cdots < \alpha_{m-1} = a_h < \alpha_m = \frac{(b_1+t)}{2}$. Therefore, we can sort input points by given algorithm which implies that no algorithms with $o(n \log n)$ running time can compute Pareto front of one dimensional 1-median 1-center two-objective optimization problem. □

Note 1. If the algorithm output the Pareto optimal front as $O = \big\{\big(C(\alpha_1), M(\alpha_1)\big) - \big(C(\alpha_2), M(\alpha_2)\big), \cdots, \big(C(\alpha_{2m-1}), M(\alpha_{2m-1})\big) - \big(C(\alpha_{2m}), M(\alpha_{2m})\big)\big\}$, start points and end points of m segments, since the slope of each segment is an integer of $\mathcal{O}(n)$, the segments can be sorted in $\mathcal{O}(n)$. Therefore, we can have sorted break points of Pareto front function and the above proof holds.

Theorem 1. *Algorithm 1 compute one dimensional* 1*-median* 1*-center two-objective Pareto front and Pareto solutions interval in* $\mathcal{O}(n \cdot \log n)$.

Proof. $C(x)$ can be computed easily in constant time and $M(x)$ can be computed in $\mathcal{O}(\log n)$ using binary search, we obtain that line 13 is $\mathcal{O}(\log n)$ running time. Therefore, we can conclude that Algorithm 1 is $\mathcal{O}(n \cdot \log n)$. □

Corollary 1. *Pareto front of one dimensional* 1*-median* 1*-center two-objective optimization problem can be computed in* $\theta(n \log n)$*.*

Algorithm 1. COMPUTE PARETO OPTIMAL FRONT

Input: Set I s.t. $|I| = n$
Output: P_s(Pareto solutions), P_f(Pareto front)
1: Sort I increasingly to $\{x_1, x_2, \ldots, x_n\}$
2: **if** n is even **then**
3: $\quad b = \frac{n}{2} + 1$
4: **else**
5: $\quad b = \frac{n+1}{2}$
6: **end if**
7: $P_s = [x_b, (x_1 + x_n)/2]$
8: $P_f = \Phi$
9: Add $\big(C(x_b), M(x_b)\big)$ to P_f
10: $i = b$
11: **while** $x_{i+1} < (x_1 + x_n)/2$ **do**
12: $\quad i = i + 1$
13: $\quad$ Add $\big(C(x_i), M(x_i)\big)$ to P_f
14: **end while**
15: Add $\big(C(x_{(x_1+x_n)/2}), M(x_{(x_1+x_n)/2})\big)$ to P_f
16: **return** P_s, P_f

Due to space limitation, Algorithm 1 is just for the case that c_{opt} is on the right side of M_{opt}. The case that c_{opt} is on the left side is similar. If there is an intersection, solution is obviously the intersection point.

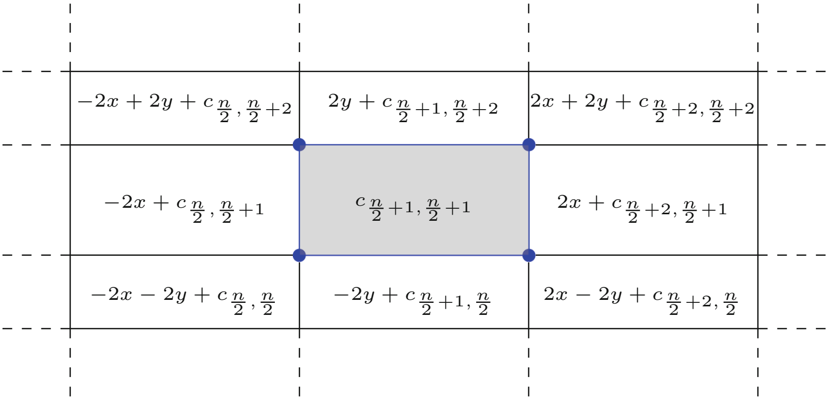
3 Two Dimensional

In this section we consider the problem in $\mathbb{R}^2$. The aim is to find the Pareto front and Pareto solutions in terms of M_{opt} and C_{opt}.

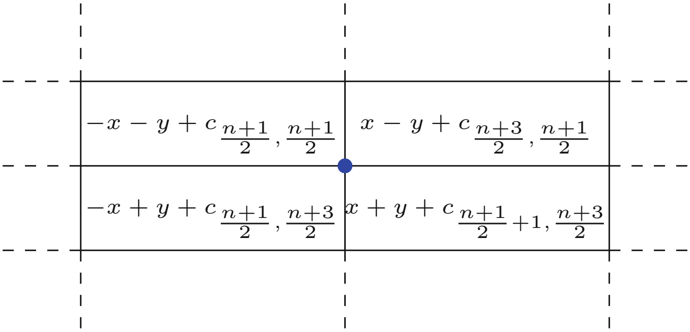
3.1 1-Median Objective

For each point $p \in \mathbb{R}^2$ we have:

$$\begin{aligned} M(u) &= \sum_{i=1}^{n} \|u - p_i\|_1 \\ &= \sum_{i=1}^{n} |u_x - p_{i_x}| + \sum_{i=1}^{n} |u_y - p_{i_y}| \end{aligned} \quad (1)$$



(a) Number of equations is even, 1-median optimal is a rectangular (blue) region.



(b) Number of equations is odd, 1-median is a (blue) point.

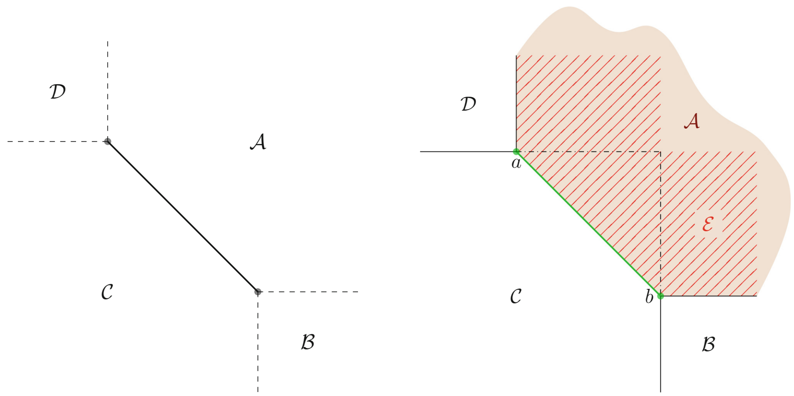
Fig. 4. 1-median optimal and equation of M(p) in middle cells (Color figure online).

We can observe that we need $\mathcal{O}(n)$ time to deterministically minimize Eq. 1. Moreover, because of the assumption that no points have same coordinate the optimal of M_{opt} may be just a point or area of a rectangle.

In the rest of this paper we assume that n is even (all proofs and discussions are similar when n is odd.). Consider lines $y = p_{i_x}$ and $x = p_{i_y}$ such that $1 \leq i \leq n$ which partition the xy-plane into $(n+1)^2$ cells where boundary cells are unbounded. The equation of $M(p)$ for points in each cell is the same because of the absolute value function. Furthermore, for points in a column (resp. row) equation of $\sum_{i=1}^{n} |x - p_{i_x}|$ (resp. $\sum_{i=1}^{n} |y - p_{i_y}|$) do not change but for transformation to upper (resp. right) cell coefficient of y (resp. x) increases by 2 (Fig. 4).

3.2 1-Center Objective

Let $\mathcal{FVD}$ be the *farthest point Voronoi diagram* of input points in Manhattan metric, also let $R_{\mathcal{FVD}}(p)$ denote the region of $\mathcal{FVD}$ which consist of p and



(a) Farthest point Voronoi diagram regions in Manhattan metric. (b) Possible region for site of $\mathcal{C}$ is $\mathcal{A} \setminus \mathcal{E}$.

Fig. 5. Farthest point Voronoi diagram properties.

$S_{\mathcal{FVD}}(\mathcal{R})$ denote the site of region $\mathcal{R}$. According to the definition of 1-center objective, $C(p)$ is $\|p - S_{\mathcal{FVD}}(R_{\mathcal{FVD}}(p))\|_1$. Besides the $\mathcal{FVD}$ partition the plane into at least two and at most four regions (Fig. 5a).

According to the structure of $\mathcal{FVD}$, it is impossible for regions $\mathcal{A}$ and $\mathcal{C}$ to have a common site. However, either $\mathcal{B}$ (resp. $\mathcal{D}$) can merge with $\mathcal{A}$ (resp. $\mathcal{C}$) or $\mathcal{B}$ (resp. $\mathcal{D}$) can merge with $\mathcal{C}$ (resp. $\mathcal{A}$), i.e. $\mathcal{B}$ and $\mathcal{D}$ cannot merge with a common region simultaneously.

Proposition 1. *Site of region $\mathcal{C}$ is in $\mathcal{A} \setminus \mathcal{E}$. Otherwise distances between points on segment ab and $S_{\mathcal{FVD}}(\mathcal{C})$ are not equal and ab is not an edge of $\mathcal{FVD}$ (Fig. 5b).*

From Proposition 1 it can be concluded that $C(p)$ for $p \in \mathcal{C}$ is equal to distance of p from segment ab add up to distance between segment ab and $S_{\mathcal{FVD}}(\mathcal{C})$.

Proposition 2. *As shown in Fig. 6a distances of $m_1, m_2 \in \mathcal{A}$ from line ℓ_1 is equal to their distances from segment ab. For point m_1 both distances are obviously the same and are equal to $\|m_1 - m_1'\|_1$. For point m_2 we have $\mathbf{\Delta} pqm_2'$ and $\mathbf{\Delta} qm_2'a$ as equal isosceles triangles. Therefore, segments qm_2' and qa are equal. Hence $\|m_2 - m_2'\|_1 = \|m_2 - a\|_1$.*

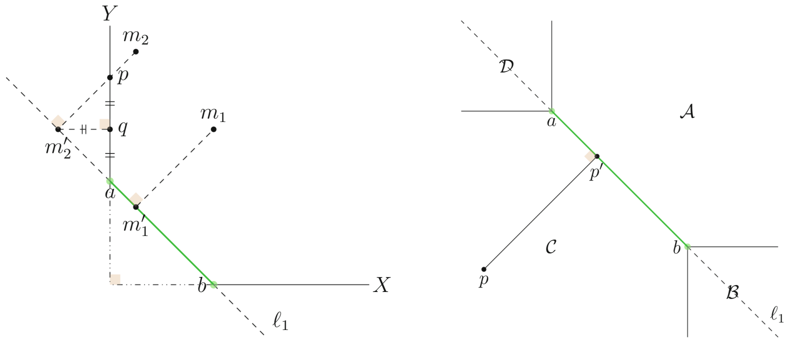
The following two propositions determine the equation of $C(p)$ in the plane and proof that it depends on which region of $\mathcal{FVD}$ includes p.

Proposition 3. *For point $p \in \mathcal{C}$ (resp. $p \in \mathcal{A}$), $C(p) = k_{opt} + c - p_x - p_y$ (resp. $C(p) = k_{opt} - c + p_x + p_y$) where c is y-intercept of ℓ_1 (Fig. 6b).*

Proof. Suppose equation of line ℓ_1 is $y = -x + c$ and distance between site of $\mathcal{C}$ and segment ab is k_{opt}, then projection of point $p = (x, y)$ on ℓ_1 is $p' = (\frac{c-y+x}{2}, \frac{c+y-x}{2})$. Using Propositions 1 and 2 can obtain that:

$$C(p) = k_{opt} + \|p - p'\|_1 = k_{opt} + c - p_x - p_y.$$

□



(a) Property of 1-center optimal segment in Manhattan metric.

(b) Projection of points to line ℓ_1.

Fig. 6. Property of 1-center optimal segment in Manhattan metric.

Proposition 4. *In Fig. 7a since site of $\mathcal{D}$ is in hatched region or on its border, for point $q \in \mathcal{D}$ we have $C(q)$ as the distance of point a from $S_{\mathcal{FVD}}(\mathcal{D})$ add up to distance between point a and point q. Also since point a is an $\mathcal{FVD}$ vertex, we know that distance of point a from $S_{\mathcal{FVD}}(\mathcal{D})$ is equal to its distance from $S_{\mathcal{FVD}}(\mathcal{C})$ and as equal to k_{opt}, hence:*

$$C(q) = k_{opt} + \|a - q\|_1 = k_{opt} - c_1 - q_x + q_y$$
$$c_1 = a_y - a_x$$

Similarly it can be proven that for $q \in \mathcal{B}$:

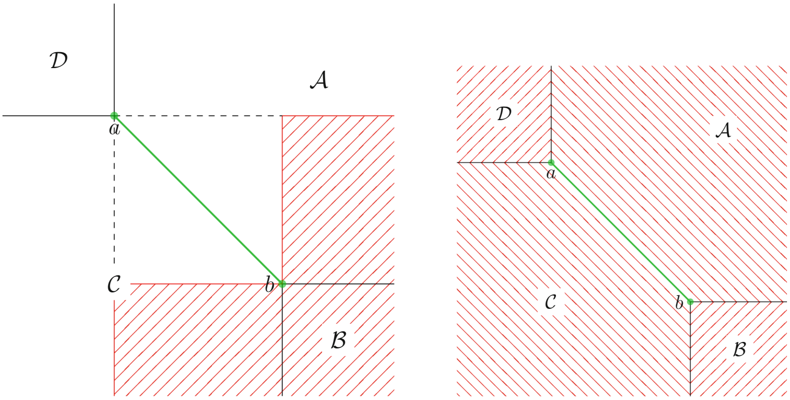
$$C(q) = k_{opt} + \|b - q\|_1 = k_{opt} + c_2 + q_x - q_y$$
$$c_2 = b_y - b_x.$$

Corollary 2. *According to Propositions 3 and 4 we can conclude that points in $\mathcal{A}$ and $\mathcal{C}$ which are on segments parallel to segment ab have the same* 1*-center objective value. Also for $\mathcal{B}$ and $\mathcal{D}$ these points are on segments perpendicular to ab. Moreover, points on ab are optimal of* 1*-center objective (Fig. 7b).*

3.3 Pareto Optimal Solutions

Suppose M_{opt} and C_{opt} are calculated. Obviously if they have intersection, it is the set of Pareto solutions. Hence in the rest of this section we assume that M_{opt} and C_{opt} have no intersection.

Possible Region for Pareto Optimal Set. Here the goal is to find the region $\mathcal{P}$ such that its boundary points dominate all points of the plane, i.e. Pareto set is definitely in $\mathcal{P}$.



(a) Possible region for site of $\mathcal{D}$. (b) Points with the same 1-center value.

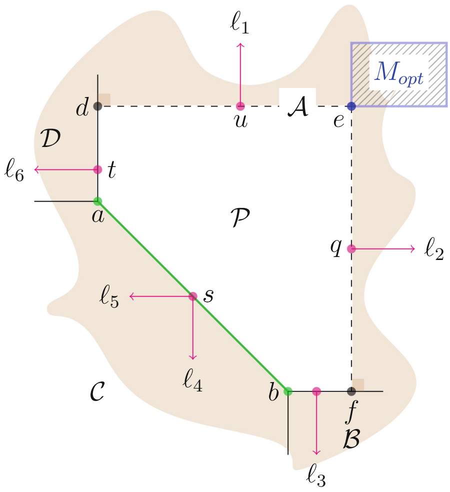
Fig. 7. Possible region for site of $\mathcal{D}$.

According to the optimal of $M(p)$ and $C(p)$, three cases are possible. In the first case M_{opt} is in regions $\mathcal{A}$ or $\mathcal{C}$, in the second case M_{opt} is in regions $\mathcal{B}$ or $\mathcal{D}$ and in the third case M_{opt} intersects with the axis aligned the edges of $\mathcal{FVD}$. For the first case (Fig. 8a) let e be the lower left point of M_{opt} and let ef and ed be the vertical and horizontal segments hitting the edges of $\mathcal{FVD}$. For all points u on line of ed and w on half-line segment ℓ_1 perpendicular to ℓ_{ed}, $C(u) < C(w)$ and $M(u) \leq M(w)$. Thus u dominates all points on ℓ_1. Similarly for point q on ℓ_{ef} and w on half-line segment ℓ_2, $q \prec w$. There are similar results for other edges of $adefb$ which make us able to conclude that polygon $adefb$ is $\mathcal{P}$.

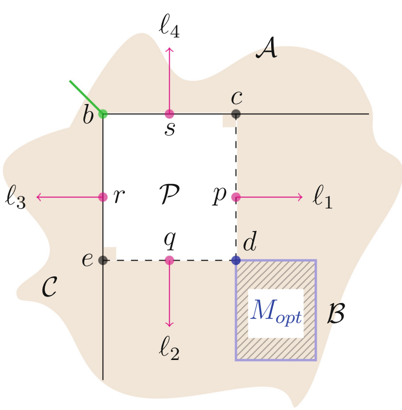
Second and third cases are similar and we consider them simultaneously (Fig. 8b and c). Let $bcde$ be in region $\mathcal{B}$. Obviously above discussion holds for points p, q, r, s and half-line segments ℓ_1, ℓ_2, ℓ_3 and ℓ_4 respectively. Moreover, a dominates all points of $\mathcal{D}$, any point t on ab dominates all points on horizontal (resp. vertical) half-line segment which starts from t and pass through $\mathcal{C}$ (resp. $\mathcal{A}$) and b dominates all points on ab. Therefore, we can conclude that points on the border of $bcde$ dominate all points outside of it and $bcde$ is $\mathcal{P}$. It is the same when $bcde$ is in $\mathcal{D}$.

Pareto Optimal Solutions. We have shown that $M(p)$ partitions the plane to cells in which equation of $M(p)$ is known. According to this partitioning and C_{opt}, seven cases are possible. First three cases happen when M_{opt} is in $\mathcal{A}$ or $\mathcal{C}$ of $\mathcal{FVD}$. Next three cases occur when M_{opt} is in $\mathcal{B}$ or $\mathcal{D}$. Last case occurs when M_{opt} and axis aligned edges of $\mathcal{FVD}$ have intersection.

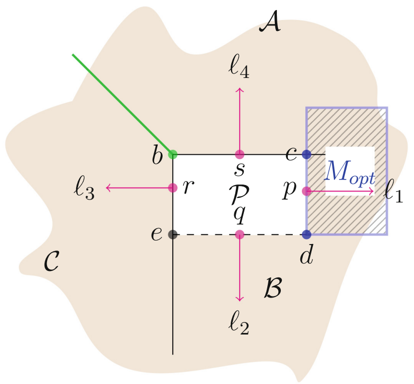
The claim is that cells in $\mathcal{P}$ whose equations are $M(p) = \alpha p_x + \beta p_y + c$ such that $\alpha = \beta$, are part of Pareto set.



(a) M_{opt} is in $\mathcal{A}$.



(b) M_{opt} is in $\mathcal{B}$.



(c) M_{opt} is on the axis aligned edge of $\mathcal{FVD}$.

Fig. 8. M_{opt} is in $\mathcal{A}$.

Proposition 5. *Let $\mathcal{P}$ be in $\mathcal{A}$. For each cell with $M(p) = \alpha p_x + \beta p_y + c$ where $\alpha/\beta > 1$, points on the right and bottom edges dominate other points of the cell.*

Proof. For each point p in the cell, points with the same 1-median values are on a line which is parallel to $y = -\alpha/\beta$. This line will hit the border of the cell in points p' and p'' such that $p'_x > p''_x$, i.e. p' is on bottom or right edge. Since $\alpha/\beta > 1$ we have $C(p') < C(p) < C(p'')$ and p' dominates p and p''. □

Proposition 6. *Suppose $\mathcal{P}$ is in $\mathcal{A}$ (resp. $\mathcal{C}$). In $\mathcal{P}$ let q be a point in a cell with equation $M(q) = \alpha q_x + \alpha q_y + c$ such that $\alpha < 0$ (resp. $\alpha > 0$). Suppose ℓ be the* *line passing through* q *with equation* $y = -x + c'$*. By extending Proposition 5, for all* p *on* ℓ *or bellow (resp. on* ℓ *or above) we have* $M(q) \leq M(p)$*. The same result holds for* $\mathcal{B}$ *and* $\mathcal{D}$ *when* ℓ *is* $y = x + c'$*.*

The following lemma introduces special cells in $\mathcal{P}$ which are part of Pareto set. In the rest of this paper we refer to them as *Pareto cells*.

Lemma 4. *Suppose* $\mathcal{P}$ *is in* $\mathcal{A}$ *(resp.* $\mathcal{C}$*). All points like* p *of cells with* $M(p) = \alpha p_x + \beta p_y + c$ *such that* $\alpha = \beta$ *and* $\alpha < 0$ *(resp.* $\alpha > 0$*), are all or part of Pareto set.*

Proof. Here we assume $\mathcal{P} \subset \mathcal{A}$ but the proof is similar when $\mathcal{P} \subset \mathcal{C}$. Consider $p \in \mathcal{P}$ such that $M(p) = \alpha p_x + \alpha p_y + c$ and $\alpha < 0$. Suppose q dominates p and ℓ be a line passing through p with $y = -x + c'$ equation. if q is above ℓ then $C(q)$ is greater than $C(p)$. Therefore, q is on or bellow ℓ. By Proposition 6, if q is bellow ℓ it means $M(p) < M(q)$ otherwise q is on ℓ; but if both are in the same cell it concludes that $C(p) = C(q)$ and $M(p) = M(q)$, otherwise $M(p) < M(q)$. We can obtain from these contradictions that no point dominates p. □

Similar to Lemma 4, cells with $M(p) = \alpha p_x + \beta p_y + c$ such that $\beta = -\alpha$ and $\alpha > 0$ (resp. $\alpha < 0$) are *Pareto cells* in region $\mathcal{B}$ (resp. $\mathcal{D}$).

For intersection of a *Pareto cell* with edges of $\mathcal{FVD}$ several cases are possible. If the *Pareto cell* intersects with a horizontal (resp. vertical) edge, segment from b (resp. a) to border of the *Pareto cell* will be the rest of Pareto solutions, we refer to this segment as *Pareto segment*. Suppose point q dominates $p \in$ *Pareto segment* and let ℓ be the line passing through p and parallel to $y = -x$, then q must be on or below this line, otherwise $C(q) > C(p)$. But if q is on or below ℓ, since p is in a cell that $\alpha/\beta > 1$, $M(p) < M(q)$. If *Pareto cell* intersects with ab, the part of cell which is in $\mathcal{P}$ is also *Pareto cell* (Fig. 9).

Lemma 5. *Points of a* Pareto cell *in solution space are a segment in objective space.*

Proof. In a *Pareto cell* $M(p) = \alpha p_x + \alpha p_y + c$ ($\alpha < 0$) and $C(p) = p_x + p_y + c'$. Therefore, $M(p) - \alpha C(p) = c''$. This implies that *Pareto cell* in solution space is a segment with $Y - \alpha X = c''$ equation in objective space. It is easy to see that this holds for *Pareto segments*. □

Theorem 2. *Pareto Front of two dimensional* 1*-median* 1*-center two-objective optimization problem is* continuous*,* convex *and* piecewise linear *function.*

Proof. By Lemma 5 we can conclude that Pareto optimal front is piecewise linear. Since in the sequence of *Pareto cells* from M_{opt} to C_{opt} each cell have a common point with the next cell, the sequence of segments of Pareto front is continuous. Moreover, since in each cell the coordinate of x and y in $M(p)$ is smaller than the previous ones, slope of segment of that cell in objective space will be bigger than segments of previous cells which guarantees convexity of Pareto front. □

Corollary 3. *Finding Pareto front and Pareto Solution set of two dimensional* 1*-median* 1*-center two-objective optimization problem is* $\theta(n \log n)$*.*

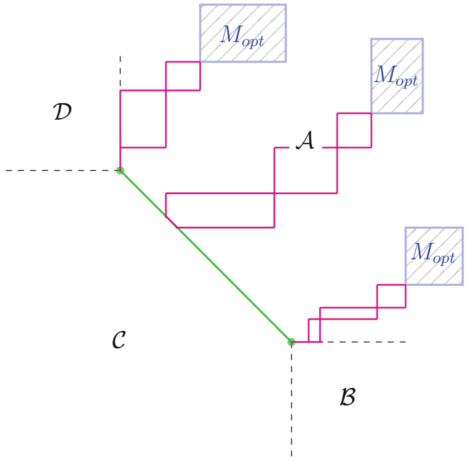


Fig. 9. Intersection of *Pareto cells* with edges of $\mathcal{FVD}$

4 Conclusion and Future Work

In this paper we introduced an important and useful multi-objective optimization problem with 1-median and 1-center in Manhattan metric as its objectives. We considered the problem in one and two dimensional space. We also determined the Pareto optimal front and Pareto set simultaneously. Furthermore we proved finding Pareto front and Pareto solution set of proposed problem is $\theta(n \log n)$.

In higher dimensions, considering Manhattan metric, similar to two dimensional space we can show that optimal of 1-median, i.e. $M(x)$, will be a d dimensional hypercube. Also, it can be computed in $\mathcal{O}(dn)$. For optimal of 1-center, i.e. $C(x)$, the propositions are not straight forward. However finding the smallest circumferential hypercube drives us to the hyperplane which is the locus of cube's center (optimal of 1-center). Moreover, it seems that farthest point Voronoi diagram has the most $2d$ regions. Thus we guess the Pareto optimal set is very similar to two dimensional space; i.e. smallest interval of hypercubes from $M(x)$ to $C(x)$ which are connected by their corners in direction perpendicular to locus of optimal of $C(x)$.

In Euclidean metric, we think this problem will be much harder and the Pareto solutions cannot be computed exactly. In this case we have to approximate Pareto solutions and Pareto front. Moreover, this approximation can be followed for harder objectives such as 2-median and 2-center.

References

1. Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., Pandit, V.: Local search heuristics for k-median and facility location problems. SIAM J. Comput. **33**(3), 544–562 (2004)

2. Badoiu, M., Clarkson, K.L.: Smaller core-sets for balls. In: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics, pp. 801—802 (2003)
3. Bādoiu, M., Har-Peled, S., Indyk, P.: Approximate clustering via core-sets. In: Proceedings of the Thiry-Fourth Annual ACM Symposium on Theory of Computing, pp. 250–257. ACM (2002)
4. Bartal, Y.: Probabilistic approximation of metric spaces and its algorithmic applications. In: Proceedings of 37th Annual Symposium on Foundations of Computer Science, pp. 184–193. IEEE (1996)
5. Charikar, M., Guha, S., Tardos, É., Shmoys, D.B.: A constant-factor approximation algorithm for the k-median problem. In: Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, pp. 1–10. ACM (1999)
6. Chazelle, B., Matoušek, J.: On linear-time deterministic algorithms for optimization problems in fixed dimension. J. Algorithms **21**(3), 579–597 (1996)
7. Coello, C.A.C., Van Veldhuizen, D.A., Lamont, G.B.: Evolutionary Algorithms for Solving Multi-objective Problems, vol. 242. Springer, Verlag (2002)
8. Deb, K.: Multi-objective Optimization Using Evolutionary Algorithms, vol. 16. Wiley, Chichester (2001)
9. Har-Peled, S., Kushal, A.: Smaller coresets for k-median and k-means clustering. In: Proceedings of the Twenty-First Annual Symposium on Computational Geometry, pp. 126–134. ACM (2005)
10. Har-Peled, S., Mazumdar, S.: Coresets for k-means and k-median clustering and their applications, pp. 291–300 (2004)
11. Hochbaum, D., Shmoys, D.: A best possible approximation algorithm for the k-center problem. Math. Oper. **10**, 180–184 (1985)
12. Jain, K., Mahdian, M., Saberi, A.: A new greedy approach for facility location problems. In: Proceedings of the Thiry-Fourth Annual ACM Symposium on Theory of Computing, pp. 731–740. ACM (2002)
13. Megiddo, N., Supowit, K.J.: On the complexity of some common geometric location problems. SIAM J. Comput. **13**(1), 182–196 (1984)
14. Meyerson, A., O'Callaghan, L., Plotkin, S.: A k-median algorithm with running time independent of data size. Mach. Learn. **56**(1–3), 61–87 (2004)
15. Papadimitriou, C.H., Yannakakis, M.: On the approximability of trade-offs and optimal access of web sources. In: Proceedings of 41st Annual Symposium on Foundations of Computer Science, pp. 86–92. IEEE (2000)
16. Shmoys, D.B.: Approximation algorithms for facility location problems. In: Jansen, K., Khuller, S. (eds.) APPROX 2000. LNCS, vol. 1913, pp. 27–32. Springer, Heidelberg (2000)
17. Tansel, B.C., Francis, R.L., Lowe, T.J.: State of the artlocation on net- works: a survey. part i: the p-center and p-median problems. Manage. Sci. **29**(4), 482–497 (1983)
18. Zarrabi-Zadeh, H., Chan, T.M.: A simple streaming algorithm for minimum enclosing balls. In: CCCG. Citeseer (2006)
19. Zarrabi-Zadeh, H., Mukhopadhyay, A.: Streaming 1-center with outliers in high dimensions. In: CCCG, pp. 83–86 (2009)