Median hyperplanes in normed spaces — a survey

Horst Martini

Mathematische Fakultät

TU Chemnitz

D–09009 Chemnitz

Germany

Anita Schöbel

Fachbereich Mathematik

Universität Kaiserslautern

D–67653 Kaiserslautern

Germany

Abstract: In this survey we deal with the location of hyperplanes in n–dimensional normed spaces, i.e., we present all known results and a unifying approach to the so-called median hyperplane problem in Minkowski spaces. We describe how to find a hyperplane H minimizing the weighted sum $f(H)$ of distances to a given, finite set of demand points. In robust statistics and operations research such an optimal hyperplane is called a median hyperplane.

After summarizing the known results for the Euclidean and rectangular situation, we show that for all distance measures d derived from norms one of the hyperplanes minimizing $f(H)$ is the affine hull of n of the demand points and, moreover, that each median hyperplane is a halving one (in a sense defined below) with respect to the given point set. Also an independence of norm result for finding optimal hyperplanes with fixed slope will be given. Furthermore we discuss how these geometric criteria can be used for algorithmical approaches to median hyperplanes, with an extra discussion for the case of polyhedral norms. And finally a characterization of all smooth norms by a sharpened incidence criterion for median hyperplanes is mentioned.

1   Introduction

We consider the problem of approximating a set of arbitrarily given points $\{x_1, x_2, \ldots, x_M\}$ with weights $w_1, w_2, \ldots, w_M$ in n-dimensional normed spaces (Minkowski spaces) by a linear function (*the linear fit problem*). Especially, but not only, the Euclidean subcase of this location problem plays an important role in different mathematical disciplines.

1. In *robust statistics* and *numerical mathematics*, linear fit problems are mainly studied with respect to the Euclidean, the Manhattan and the Chebyshev distance, and they are known as absolute errors regression,

1

median problems, L_1 regression and orthogonal/vertical L_1–fit problems, respectively. Related investigations are going back to the 18th century, see [Bos57], [Edg87], and [Edg88]. It should be noticed that the basic geometric criteria for *orthogonal* and *vertical* L_1-fit procedures are strongly related to each other, see also Section 3 below. The importance of L_1 regression (e.g., instead of the known least squares regression) for robust statistics is based on the fact that exactly for $p = 1$ the corresponding L_p estimates are technically robust in the sense that they provide protection against arbitrary outliers, cf. the survey [NW82] and [RL87]. On the other hand, certain approximation problems in numerical mathematics (e.g., the approximation of given functions by linear functions) lead in a natural way to the same type of problems, see [Ric64] and [PFTV86]. In particular, [SW87] present a numerical algorithm for linear approximation of finite point sets (regarding orthogonal distances) which corresponds to a concave quadratic programming algorithm.

2. The strong development of *computational geometry* has provided new insights into various (classical) research areas. In this sense, also a large variety of location problems was enriched by new methods and algorithmical motivations, see the surveys [Lee86], [LW86], [Kor89], and [KM93]. In particular, the time complexity of linear fit problems (in computational geometry also called *linear L_1 approximation problems*) was investigated by several authors, cf. [MT83], [YKII88], [HIIR89], [KM90], [KM93], and [HII$^+$93].

And as a second point of view, a special case of one of the most interesting problems in discrete and computational geometry (namely the k–set problem) turns out to be related to our considerations below. This subcase is the problem of counting the number of halving hyperplanes (i.e., the number of $\frac{M}{2}$–sets) with respect to an M-element set $\mathcal{X} \subseteq R^n$, see [PSS92], [DE94], and [ZV92]. Namely, here a hyperplane is said to be halving with respect to $\mathcal{X}$ if it is spanned by a subset of $\mathcal{X}$ and the number of points on each side differ at most by one. In this paper we also use a slightly modified definition of halving (which we call pseudo-halving), see Definition 3. Several estimates on the number of halving lines and hyperplanes have been developed and will be discussed in Section 2.

3. In *operations research* and *location science* the two-dimensional version of the linear fit problem is known as the *line facility location problem*, which belongs to the area of *path location*.

Path location is an extension of classical facility location. The set $\mathcal{X}$ of demand points can be seen as a set of existing facilities or demand points (in the plane) where the weights represent the importance of the existing

facilities. In classical facility location the objective is to find a good point-shaped facility (see, e.g., the books or surveys of [LMW88], [Pla95], and [Ham95]), whereas the problem of path location is to locate a dimensional facility such as a line or a curve in the plane. The objective function is the same as in classical facility location, namely to minimize the sum of distances (or the maximum distance) between the existing facilities and the new one, where, so far, mainly the Euclidean and the Manhattan distance have been considered. An application in that area is the planning of new railways or motorways, where the existing facilities can be cities and the weights the number of their inhabitants. Path location can also be used to determine the location of pipelines, drainage or irrigation ditches, or in the field of plant layout, see for example [MN80]. A recent survey about the location of dimensional structures in the plane is [Mes95]. Using Euclidean and rectangular distances, line location problems in the plane were discussed by [Wes75], [MN80], [MN83], [MT82], [MT83], [LW86], and [LC85], for higher dimensions see again [KM90], [KM93], and [HII+93]. Extensions to other distances in two dimensions were given by [Sch96] to block norms and by [Sch98] to arbitrary distances derived from norms. In the following we will show the way to generalize these results to n-dimesnional normed spaces, cf. also [MS97].

We use the following standard description of hyperplanes.

Definition 1 *Let the real numbers $s_1, s_2, \ldots, s_n$ and b be given, such that $\sqrt{s_1^2 + s_2^2 + \ldots + s_n^2} = 1$. Then we define the hyperplane $H_{s_1, s_2, \ldots, s_n, b}$ by*

$$H_{s_1, s_2, \ldots, s_n, b} := \{(x_1, x_2, \ldots, x_n) : s_1 x_1 + s_2 x_2 + \ldots + s_n x_n + b = 0\}$$

and $\mathfrak{n} = (s_1, s_2, \ldots, s_n) \in R^n$ as its normal vector. Thus we can also write $H_{\mathfrak{n}, b}$ instead of $H_{s_1, s_2, \ldots, s_n, b}$.

Now the problem we are dealing with in this paper can be stated as follows: Given a distance measure d, an index set $\mathcal{M} := \{1, 2, \ldots, M\}$ and a set

$$\mathcal{X} = \{x_m : m \in \mathcal{M}\} \subset R^n$$

of demand points with positive weights w_m for all $m \in \mathcal{M}$, we are looking for a hyperplane H such that

$$f(H) = \sum_{m \in \mathcal{M}} w_m d(x_m, H)$$

is minimized, where the distance between a point x and the hyperplane H is given by

$$d(x, H) = \min_{z \in H} d(x, z).$$

Any optimal hyperplane is called a *median hyperplane*. Some more notation should be introduced. In particular, $W = \sum_{m \in \mathcal{M}} w_m$ denotes the sum of weights of all demand points x_m, the usual unit vectors in R^n are given by $e_1, \ldots, e_n$, and for $\mathcal{A}$ an arbitrary subset of R^n, aff($\mathcal{A}$) is the set of all affine combinations of elements of $\mathcal{A}$, i.e., its *affine hull*.

For a hyperplane H let H^+ and H^- denote the two open halfspaces separated by the hyperplane H. In the following we assume that $\mathcal{X}$ contains at least $n+1$ affinely independent points, since all other cases are trivial. Namely, in these cases the optimal hyperplane H will pass through all demand points and satisfy $f(H) = 0$.

Definition 2 *A hyperplane H is called a halving hyperplane with respect to $\mathcal{X} = \{w_m x_m : m \in \mathcal{M}\}$ if it is the affine hull of n points in $\mathcal{X}$ and*

$$\sum_{x_m \in H^+} w_m < \frac{W}{2} \ and \ \sum_{x_m \in H^-} w_m < \frac{W}{2}.$$

This definition has been given by [KM93] and the second part of it is equivalent to the definition of halving given in [NM80],[MN80]. For the Euclidean case, all optimal hyperplanes are halving ones, see [KM90]. Since this is not necessarily true for more general norms, we have to introduce the term *pseudo-halving*, and we will show in Section 4 that all optimal hyperplanes (for any norm) are pseudo-halving.

Definition 3 *A hyperplane H is called a pseudo-halving hyperplane with respect to $\mathcal{X} = \{x_m : m \in \mathcal{M}\}$ if*

$$\sum_{x_m \in H^+} w_m \leq \frac{W}{2} \ and \ \sum_{x_m \in H^-} w_m \leq \frac{W}{2}.$$

Note that in this definition it is not required that n of the demand points are on the hyperplane, as it is in the definition before.

We will use the classification scheme of [HN96] which was originally developed for location theory, but is also helpful in this context: in that 5 position scheme our problem can be described by $1H/R^n/ \cdot /d/\sum$, meaning in short that we want to locate one hyperplane ($1H$) in n-dimensional space R^n with no special assumptions ($\cdot$), for example about the weights; this should be done by using the distance measure d, and we want to minimize the sum of weighted distances between the demand points x_m and the hyperplane H ($\sum$).

In the next two sections some results for Euclidean and rectangular distances are given. Section 4 extends these results to distance measures derived from arbitrary norms in R^n. The sequence of lemmas and theorems In Sections 3 and 4 below should be understood as a unifying approach to the median hyperplane problem

in Minkowski spaces coming from the vertical L_1 approximation in $\mathbb{R}^n$. Having such a unified representation as one aim of this survey, we slightly modified related approaches from [Sch98] and [MS97], and for proofs of particular statements the reader should consult these two papers. Sections 5 and 6 give some algorithmic approaches for the general case and for the case that the distance has been derived from a polyhedral norm (block norm). The paper is concluded by remarks on possible extensions and on a characterization of smooth norms by a strong incidence criterion for median hyperplanes.

2   Results for Euclidean distances in R^n

Now we shall give a survey on the results for the Euclidean version $(1H/R^n/\cdot/l_2/\sum)$ of our problem. Our starting point is the planar weighted case. In [MN80] it was shown that each optimal line has to pass through two of the given points, and this was used to get an $O(M^3)$ time and linear space algorithm. This result was improved by [MT83] to an $O(M^2 log M)$ time and $O(M^2)$ space algorithm, and further on [LC85] improved this to $O(M^2)$ time and space. Finally, [YKII88] and, independently, [KM90] derived an $O(M^2)$ time and $O(M)$ space approach, see also [KM93].
Much more is known about the special case $n = 2$ with identical weights for all points, i.e., about the planar unweighted case $1l/R^2/w_m = 1/l_2/\sum$. Namely, [MN80] observed also a second criterion: optimal lines have to be halving ones in the sense of Definition 2. Hence, for the unweighted situation purely combinatorial properties of the given point set become interesting, since the following subquadratic bounds on the number $h(M)$ of halving lines (which are the affine hull of n of the given points) are known:

$$h(M) \quad < \quad M^{\frac{3}{2}} \text{ (cf. [Lov71] and [ELSS73])}, \tag{1}$$

$$h(M) \quad < \quad M^{\frac{3}{2}} log^{-\frac{1}{100}} M \text{ (see [PSS92])}. \tag{2}$$

For using these bounds to improve the time complexities given above, it is necessary to implement the halving line procedure due to [Lov71] and explained in the following for a given point set in general position. Starting with an arbitrary halving line $H_1 = \text{aff}(x_1, x_2)$ with initial orientation from x_1 to x_2, one rotates it clockwise about x_2 (while preserving the orientation as intrinsic) until it hits some further point x_3 to obtain $H_2 = \text{aff}(x_2, x_3)$. Then H_2 is rotated clockwise about x_3 to get H_3, and so on, until one returns to the starting position. For odd M, all lines H_i are halving ones, and for even M the line H_i is halving if and only if it is oriented from x_i to x_{i+1} (otherwise it is an $(\frac{M}{2}, \frac{M}{2} - 2)$–divider). Using a certain data structure of [OvL81], the rotation procedure of [Lov71] may be implemented in $O(h(M) log^2 M)$ time, bearing in mind that the number of $(\frac{M}{2}, \frac{M}{2} - 2)$–dividers in the even case has asymptotically the same upper bound as the number of halving lines. This led [YKII88] to an $O(M^{\frac{3}{2}} log^2 M)$ time and linear space algorithm

by using (1), and by (2) this can be improved to an $O(M^{\frac{3}{2}}log^{2-\frac{1}{100}}M)$ time and $O(M)$ space approach, see [KM93]. The question for the time optimal algorithm remains to be answered, yet. The known lower bound is $\Omega(MlogM)$, proved in [YKII88] by reduction from the so-called uniform gap on a circle problem.

Regarding the weighted orthogonal L_1 approximation for $n \geq 3$, already the paper [NM80] contains the statement that there exists an optimal hyperplane spanned by n affinely independent given points, and a direct generalization of the halving criterion is also mentioned (at least for the unweighted case). Using that incidence criterion and basic techniques from computational geometry (such as point-hyperplane dual transforms and sweep techniques applied to hyperplane arrangements), [HII$^+$93] obtained an $O(M^n)$ worst-case time and $O(M)$ space algorithm for getting *one* optimal hyperplane. (It should be noticed that this approach was obtained already in 1988 by the same authors.) Independently, [KM90] derived an equivalent algorithmical approach (i.e., $O(M^n)$ time and linear space), but on a much stronger geometrical basis, see also [KM93]. Namely, in [Mar87] a relation between support functions of zonotopes (i.e., vector sums of line segments or, equivalently, convex n-polytopes whose r-faces, $2 \leq r \leq n - 1$, are all centrally symmetric) and the weighted orthogonal L_1 approximation was observed: using necessary conditions for local minima of these support functions, one can prove that *every* optimal hyperplane has to pass through n affinely independent points of the given set.

Unfortunately, until now there exists no spatial analogue to the computational evaluation of the line rotating procedure in the plane. However, one can hope to improve the $O(M^n)$ time complexity in the unweighted case by recent results on the number $h(M)$ of halving hyperplanes to sets of M given points. For $n = 3$, the first non-trivial upper bound was given by [BFL90], namely

$$h(M) \leq O(M^{3-c}), \ c = \frac{1}{343},$$

and [ACE$^+$91] presented

$$h(M) \leq O(M^{\frac{8}{3}}log^{\frac{5}{3}}M).$$

Finally, [DE94] improved this by the polylogarithmic factor to the best known bound
$$h(M) \leq O(M^{\frac{8}{3}}).$$

For $n \geq 4$ dimensions, the following bound was obtained by [ZV92]:

$$h(M) \leq O(M^{(n-\epsilon_n)})$$

with $\epsilon_n = t^{-(n+1)}$, where t is the smallest integer with the property that for every system $C_1, \ldots, C_{n+1}$ of finite point sets in R^n, each of size at least t, there exist $n + 1$ pairwise disjoint sets S_j, each containing at least one member from each

6

C_i, such that the intersection of the sets $conv(S_j)$ is nonempty. The authors say that $4n + 3$ is a good estimate for t, and they actually prove that $4n + 3$ is an upper bound for t. For related considerations, we also refer to [VZ94].

For the weighted case, it even remains to be answered whether cM^n is the worst case number of halving hyperplanes.

3 Results for horizontal distances in R^n

In this section we describe how to solve $1H/R^n/\cdot/d_{hor}/\sum$, i.e., how to find a hyperplane minimizing the sum of horizontal distances between the given points and the hyperplane.

This problem is particularly interesting from the viewpoint of statistical linear regression, see, e.g., [Wag59], [Fis61], [RS72], and [Sch73] for different methods to solve its planar version. However, all these approaches were improved by [MN83] and [MT83]. In the latter paper an $O(Mlog^2 M)$ time algorithm was presented, and [Zem84] even gave a linear time algorithm for any fixed dimension. An analysis of the planar version of this problem with the help of a dual interpretation is given in [Sch97]. We shortly give two lemmas about the geometric side of that problem, since these lemmas form an important building block for our main results in Section 4. Furthermore, they can easily be extended to the rectangular distance using the fact that the horizontal direction can be replaced by any of the unit vectors $e_2, \ldots, e_n$. Thus, before introducing the distances d_{e_i}, we give the definition of the *horizontal distance* $d_{hor}(x, H) = d_{e_1}$. For $x = (x_1, x_2, \ldots, x_n) \in R^n$ and $H_{s_1, \ldots, s_n, b}$ a hyperplane according to Definition 1, we have

$$d_{hor}(x, H_{s_1, \ldots, s_n, b}) = \begin{cases} \left| \frac{b + s_1 x_1 + \ldots + s_n x_n}{s_1} \right| & \text{if } s_1 \neq 0 \\ 0 & \text{if } s_1 = 0 \text{ and } s_1 x_1 + \ldots + s_n x_n + b = 0 \\ \infty & \text{if } s_1 = 0 \text{ and } s_1 x_1 + \ldots + s_n x_n + b \neq 0. \end{cases}$$

Now we note that for finding a hyperplane H minimizing $f(H)$ we can assume $s_1 \neq 0$. Thus we get as objective function

$$\begin{aligned} f(H_{s_1, \ldots, s_n, b}) &= \sum_{m \in \mathcal{M}} w_m d_{hor}(x_m, H_{s_1, \ldots, s_n, b}) \\ &= \sum_{m \in \mathcal{M}} w_m \left| \frac{b + s_1 x_{m1} + s_2 x_{m2} + \ldots + s_n x_{mn}}{s_1} \right| . \end{aligned}$$

Lemma 1 *For a given set $\mathcal{X} = \{x_m : m \in \mathcal{M}\} \subset R^n$ and positive weights w_m for all $m \in \mathcal{M}$ there always exists a hyperplane minimizing*

$$f(H) = \sum_{m \in \mathcal{M}} w_m d_{hor}(x_m, H)$$

and passing through n affinely independent points $x_m \in \mathcal{X}$.

The proof is based on the fact, that the above problem is a convex, piecewise linear optimization problem. The same holds for the following

Lemma 2 *For d_{hor} every hyperplane H^* minimizing*

$$f(H) = \sum_{m \in \mathcal{M}} w_m d_{hor}(x_m, H)$$

is a pseudo-halving one.

Defining the distances in the other directions $e_2, e_3, \ldots, e_n$ by

$$d_{e_i}(x, H) = \begin{cases} \left| \frac{b + s_1 x_1 + \ldots + s_n x_n}{s_i} \right| & \text{if } s_i \neq 0 \\ 0 & \text{if } s_i = 0 \text{ and } s_1 x_1 + \ldots + s_n x_n + b = 0 \\ \infty & \text{if } s_i = 0 \text{ and } s_1 x_1 + \ldots + s_n x_n + b \neq 0, \end{cases}$$

one gets the same results as in the horizontal case $d_{hor}(x, H) = d_{e_1}(x, H)$. Since the *rectangular distance* between a point $x_m \in R^n$ and a hyperplane H is given by

$$\begin{aligned} l_1(x, H) &= \min_{i=1,2,\ldots,n} \frac{1}{s_i} |b + s_1 x_1 + \ldots + s_n x_n| \\ &= \min_{i=1,2,\ldots,n} d_{e_i}(x, H), \end{aligned}$$

one consequence of Lemma 1 and Lemma 2 is that both results also hold for l_1. Therefore we can formulate

Theorem 1 *For the rectangular distance $d = l_1$ the following holds:*

1. *There exists a median hyperplane which passes through n affinely independent points $x_m \in \mathcal{X}$.*

2. *All median hyperplanes are pseudo-halving ones.*

4 Locating hyperplanes in normed spaces

In this section we extend the results of Section 3 to *all distances d derived from norms*. The method we use has been developed in [Sch98] for the two-dimensional case, and it was extended to higher dimensions by [MS97].

Let B be a compact, convex set containing the origin in its interior. Moreover, let B be symmetric with respect to the origin and let $x \in R^n$. The gauge

$$\gamma_B(x) := \min\{|\lambda| : x \in \lambda B\}$$

then defines a norm with the unit ball B. On the other hand, all norms can be characterized by their unit balls, see [Min67] and, for a modern representation, [Tho96], Section 1.1. At first we note that to determine the distance between a point x and a hyperplane H we can dilate the unit ball with respect to x until it is supported by the hyperplane. This yields immediately

Lemma 3 *For any norm γ with unit ball B and the derived distance d, any hyperplane H, and any point $x \in R^n$ the following equality holds:*

$$d(x, H) = \min\{|\lambda| : (x + \lambda B) \cap H \neq \emptyset\}.$$

Definition 4 *Let $t \in R^n$ be a given direction. For $x \in R^n$ and any hyperplane $H \subset R^n$ we define*
$$d_t(x, H) := \min\{|\lambda| : x + \lambda t \in H\},$$
where $\min \emptyset := \infty$.

In [Sch98] it has been shown that this distance between any point and a hyperplane can be derived from the following distance between two points $x, y \in R^n$:

$$d_t(x, y) := \gamma_t(y - x),$$

where

$$\gamma_t(x) := \begin{cases} |\alpha| & \text{if } x = \alpha t \\ \infty & \text{else.} \end{cases}$$

Thus we get $d_t(x, H) = \min_{z \in H} d_t(x, z)$. Note that $0 < d_t(x_m, H) < \infty$ if and only if t is not orthogonal to the normal vector $\mathfrak{n}$ of H. For example, the length of the horizontal segment from x_m to H then is $d_{e_1}(x_m, H) = d_{hor}(x_m, H)$. This yields the following lemma.

Lemma 4 *Let $p, q \in R^n$ and D be a linear transformation with $D(p) = q$ and $\det(D) \neq 0$. Then we have*

$$d_q(D(x), D(H)) = d_p(x, H),$$

where $D(H) := \{D(y) : y \in H\}$.

With Lemma 4 we can easily extend the results of Section 3 to the distances d_t.

Theorem 2 *For all distances d_t the following holds:*

1. *There exists a median hyperplane which passes through n affinely independent points $x_m \in \mathcal{X}$.*

2. *All median hyperplanes are pseudo-halving ones.*

The following observations say that for any distance d derived from a norm and any hyperplane with fixed normal vector $\mathfrak{n} \in R^n$ there exists a $t \in R^n$ such that $d(x_m, H) = d_t(x_m, H)$ for all $m \in \mathcal{M}$. Thus, when evaluating the objective function $f(H)$ we can replace d by d_t. Writing (as usual) l_2 for the Euclidean norm, we can formulate

Lemma 5 *Let γ be a norm or $\gamma = \gamma_t$ for some vector $t \in R^n$ and let $d(x,y) = \gamma(y-x)$ be the corresponding distance. Let a vector $\mathfrak{n} \in R^n$ be given and let t be not orthogonal to $\mathfrak{n}$. Then there exists a constant $C := C(\mathfrak{n}, d, l_2)$ such that for all $z \in R^n$ and all $x \in R^n$*

$$d(x, H_{\mathfrak{n},z}) = C \ l_2(x, H_{\mathfrak{n},z}).$$

Note that instead of l_2 we can use any other distance derived from a norm or distances derived from γ_t with t and $\mathfrak{n}$ not orthogonal. If $d_1, d_2,$ and d_3 are such distances and $\mathfrak{n}$ is a normal vector, we get

$$C(\mathfrak{n}, d_1, d_2) = \frac{C(\mathfrak{n}, d_1, d_3)}{C(\mathfrak{n}, d_2, d_3)}.$$

In particular, if $H_{s,-1,b} = \{(x_1, x_2) : x_2 = sx_1 + b\}$ is a line in the plane we obtain

$$C((s,-1), d_{hor}, d_{ver}) = |s|.$$

Another example is given by the following relation, holding for a hyperplane $H := H_{s_1,\ldots,s_n,b} = H_{\mathfrak{n},b}$ (see Section 3):

$$C(\mathfrak{n}, l_1, d_{hor}) = \min_{i=1,2,\ldots,n} \frac{s_1}{s_i}.$$

Now we are ready to formulate the announced independence of norm result for finding optimal hyperplanes with fixed slope.

Corollary 1 *For a given $\mathfrak{n} \in R^n$ the optimal hyperplanes H with normal $\mathfrak{n}$, i.e., the hyperplanes $H_{\mathfrak{n},z^*}$ minimizing $f(H_{\mathfrak{n},z})$, are the same for all norms d and distances d_t.*

There is another reason for introducing the distances d_t. Namely, based on Lemma 5 one can show the following relation between any distance d derived from a norm and distances d_t.

Lemma 6 *Let H be a hyperplane, and $d(x,y) = \gamma(y-x)$ be a distance derived from a norm γ. Then there exists a direction $t \in R^n$ such that for all $x \in R^n$*

$$d(x, H) = d_t(x, H)$$

Furthermore, for all $x \in R^n$ this direction t satisfies

$$d_t(x, H) \le d_{t'}(x, H) \text{ for all } t' \in R^n.$$

With the help of Lemma 6 and Theorem 2 one can prove the following theorem.

Theorem 3 *For all distances d derived from norms the following holds:*

1. *There exists a median hyperplane which passes through n affinely independent points $x_m \in \mathcal{X}$.*

2. *All median hyperplanes are pseudo-halving ones.*

5 Algorithmical approaches for general norms

By Lemma 3 the distance $d(x, H)$ strictly depends on the shape of the unit ball B which can be an arbitrary convex body centered at the origin. Thus, for certain unit balls (e.g., having smooth boundary which might be sufficiently complicated describable) the calculation of $d(x, H)$ is impossible by discrete methods in the spirit of computational geometry. On the other hand, there are norms (like the Euclidean one) giving a direct motivation and basis for computational approaches, and in Section 6 we will show that for polyhedral norms the time complexity is even more reducable.

In the following we will ignore such calculation difficulties, and from this point of view Theorem 3 yields approaches analogous to the Euclidean case discussed in [KM90], [HII$^+$93], and [KM93], such that we only have to give a brief outline of these computational approaches since for (one of) the best hyperplanes the basic incidence criteria coincide (Lemma 6.3 in [HII$^+$93], Theorem 2 in [KM93] and part 1 of Theorem 3 above). It is trivial to see that one can get an optimal hyperplane in $O(M^{n+1})$ time and $O(M^n)$ space, namely by enumerating all $C_M^n = O(M^n)$ candidate hyperplanes and computing the corresponding sums of weighted distances. (Enumeration algorithms spending constant time per candidate-k-subset can be taken from [RND77], Section 5.2.2.) The further reduction of the time complexity to $O(M^n)$ and of the high space cost to $O(M)$ can be obtained by constructing a certain homogeneous hyperplane arrangement in $(n+1)$-space and by using the topological hyperplane sweep technique, which is due to [EG89]. The first step is based on an incremental algorithm due to [EOS86] (and yields $O(M^n)$ time and space), and the second one, together with some further considerations, leads to the linear space requirements. (The details of these approaches can be taken from [HII$^+$93], pp 227-230, and [KM93], pp. 138-142.) Thus, one gets finally $O(M^n)$ time and $O(M)$ space requirements, and further improvements are perhaps obtainable with the help of the pseudo-halving criterion.

6 Algorithm for block norms

In the special case that the distance measure d is derived from a block norm (i.e., the unit ball B is a polytope) it is possible to find a median hyperplane more efficiently. For the plane that was done in [Sch96], and for $n \geq 3$ see [MS97].

Let B be a compact, convex polytope with nonempty interior and extreme points

$$ext(B) = \{b_1, b_2, \ldots, b_G, -b_1, -b_2, \ldots, -b_G\}, \quad b_i \in R^n, i = 1, \ldots, G.$$

We see that $\gamma_B(x) := \min\{|\lambda| : x \in \lambda B\}$ is a block norm and can be expressed by

$$\gamma_B(x) = \min\{\sum_{g=1}^{G} |\lambda_g| : x = \sum_{g=1}^{G} \lambda_g b_g\}.$$

11

Lemma 7 *Let d_B be derived from a block norm γ_B. Then*

$$d_B(x_m, H) = \min_{g=1,\ldots,G} d_{b_g}(x_m, H)$$

and the minimizer is the same for all $x \in R^n$.

To see this, we use the argument that the unit ball can be dilated with respect to x until it is supported by H (cf. Lemma 3). Obviously, a hyperplane touches an n–dimensional polytope in at least one vertex of that polytope, see, e.g., [Sha78]. Combining this fact with Lemma 6, one sees that there exists an index $g \in \{1, 2, \ldots, G\}$ such that for all $x \in R^n$ we have $d_B(x_m, H) = d_{b_g}(x_m, H)$, i.e., the direction t (such that d can be replaced by d_t) can always be chosen from the set $\{b_1, \ldots, b_G\}$.

Hence we can decompose our problem into G independent subproblems. Thus, for solving $1H/R^n/\cdot/d_B/\sum$ it is sufficient to find the best hyperplane H_g^* minimizing $\sum_{m \in \mathcal{M}} w_m d_{b_g}(x_m, H)$ for $g = 1, 2, \ldots, G$, and then to choose the hyperplane H_g^* with the smallest objective value. How to find the best hyperplanes H_g^* is described in Lemma 4. Therefore we get the following algorithm.

Algorithm

Input: block norm distance d_B, x_m, and $w_m > 0$ for all $m \in \mathcal{M}$

Output: hyperplane H^* which solves $1H/R^n/\cdot/d_B/\sum$

1. $z^* := \infty$.

2. For $g = 1$ to G do

    1. Determine a transformation D such that $D(b_g) = e_1$ and $det(D) \neq 0$.

    2. For $m \in \mathcal{M}$ do: $x_m^D = D(x_m)$.

    3. Find a hyperplane H_g^* minimizing $f(H) = \sum_{m \in \mathcal{M}} w_m d_{hor}(x_m^D, H)$.

    4. If $f(H_g^*) < z^*$ then set $z^* := f(H_g^*)$ and $H^* := D^{-1}(H_g^*)$.

3. Output: H^* with objective value z^*.

The algorithm runs in O(GR), where R is the complexity to solve the corresponding problem with horizontal distances ($1H/R^n/\cdot/d_{hor}/\sum$). In [Zem84] it is shown that this can be done in linear time for all dimensions n, such that our algorithm runs in O(GM) time.

7 Concluding remarks

We have clarified that for all distances in R^n derived from norms, and all weighted point sets $\mathcal{X}$ containing $n+1$ affinely independent points, there exists a hyperplane minimizing the sum of weighted distances to all points in $\mathcal{X}$ and passing through n affinely independent points.

As already mentioned, it was shown in [KM90] that *each* median hyperplane in Euclidean n-space is spanned by n affinely independent points of the given (weighted) set. Our Theorem 3 (part 1), referring to all finite-dimensional normed spaces, says that there *exists* a median hyperplane passing through n such given points. In this general setting, the latter statement cannot be sharpened (in the direction of the Euclidean incidence criterion), as the following simple example will demonstrate. We consider rectangular distances in the plane, i.e., our problem is described by $1l/R^2/\cdot/l_1/\sum$. The unit ball B is given by the convex hull of the four points $\{(1,0), (-1,0), (0,1), (0,-1)\}$. Furthermore, let the non-weighted point set $\mathcal{X}$ be given by the four points $\mathcal{X} = \{(1,1), (1,-1), (-1,1), (-1,-1)\}$. It is easy to see that each line passing through two of the four given points has the (minimal) distance sum 4 with respect to $\mathcal{X}$; but also the lines $x_1 = 0$ and $x_2 = 0$ have this distance sum with respect to $\mathcal{X}$. Hence there exist normed spaces with median hyperplanes containing *no* point of a suitably given set (a situation which is not possible in Euclidean spaces). Thus, one is motivated to ask for geometric characterisations of those normed spaces (or unit balls) which enforce the stronger incidence criterion. This problem was recently solved by the authors. Namely, a Minkowski space has the stronger incidence criterion if and only if its unit ball B is a *smooth* convex body centered at the origin, i.e., each boundary point of B belongs to a unique supporting hyperplane of B, see [MS98].

In addition, one might extend the investigations to k–dimensional affine flats approximating finite point sets in normed spaces regarding the distance sum, where $k \in \{0, \ldots, n-2\}$. For $k = 0$, one obtains an immediate generalisation of the well-known Weber-Problem (or Fermat-Torricelli problem or minisum problem) of location theory. And also further non-Euclidean spaces, like those of constant curvature etc., might be taken into consideration.

References

[ACE+91] B. Aronov, B. Chazelle, H. Edelsbrunner, L.Guibas, M. Sharir, and R. Wenger. Points and triangles in the plane and halving planes in space. *Discrete and Computational Geometry*, 6:435–442, 1991.

[BFL90] I. Bárány, Z. Füredi, and L. Lovász. On the number of halving planes. *Combinatorica*, 10:175–183, 1990.

[Bos57]   R.J. Boscovich. De litteraria expeditione per pontificiam ditionem, et synopsis amplioris operis, ac habentur plura ejus ex exemplaria etiam sensorum impressa. *Bononiensi Scientiarum et Artum Instituto atque Academia Commentarii*, 4:353–396, 1757.

[DE94]   T.K. Dey and H. Edelsbrunner. Counting triangle crossings and halving planes. *Discrete Computational Geometry*, 12:281–289, 1994.

[Edg87]   F.Y. Edgeworth. On observations relating to several quantities. *Hermathena*, 6(13):279–285, 1887.

[Edg88]   F.Y. Edgeworth. On a new method of reducing observations relating to several quantities. *Phil. Magazine (Series 5)*, 25:184–191, 1888.

[EG89]   H. Edelsbrunner and L.J. Guibas. Topologically sweeping an arrangement. *Journal Comput. Systems Sci.*, 38:165–194, 1989.

[ELSS73]   P. Erdös, L. Lovász, A. Simmons, and E. Strauss. Dissection graphs of planar point sets. In *A Survey of Combinatorial Theory*, pages 139–149. J. Srivastara et al., 1973.

[EOS86]   H. Edelsbrunner, J. O'Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM J. Comput.*, 15:341–363, 1986.

[Fis61]   W.D. Fisher. A note on curve fitting with minimum deviations by linear programming. *Journal Amer. Statist. Association*, 56:359–362, 1961.

[Ham95]   H.W. Hamacher. *Mathematische Lösungsverfahren für planare Standortprobleme*. Vieweg, Braunschweig, 1995.

[HII$^+$93]   M.E. Houle, H. Imai, K. Imai, J.-M. Robert, and P. Yamamoto. Orthogonal weighted linear L_1 and L_∞–approximation and applications. *Discrete Applied Mathematics*, 43:217–232, 1993.

[HIIR89]   M.E. Houle, H. Imai, K. Imai, and J.-M. Robert. Weighted orthogonal linear L_∞-approximation and applications. *Lecture Notes Computer Science*, 382:183–191, 1989.

[HN96]   H.W. Hamacher and S. Nickel. Multicriteria planar location problems. *European Journal of Operational Research*, 94:66–86, 1996.

[KM90]   N.M. Korneenko and H. Martini. Approximating finite weighted point sets by hyperplanes. *Lecture Notes Computer Science*, 447:276–286, 1990.

[KM93]   N.M. Korneenko and H. Martini. Hyperplane approximation and re-
          lated topics. In J. Pach, editor, *New Trends in Discrete and Com-
          putational Geometry*, chapter 6, pages 135–162. Springer-Verlag, New
          York, Inc., 1993.

[Kor89]   N.M. Korneenko. Optimal lines in the plane. *Math. Research*, 51:43–
          51, 1989.

[LC85]   D.T. Lee and Y.T. Ching. The power of geometric duality revisited.
          Inform. Process. Letters, 21:117–122, 1985.

[Lee86]   D.T. Lee. Geometric location problems and their complexity. *Lecture
          Notes Computer Science*, 233:154–167, 1986.

[LMW88]   R.F. Love, J.G. Morris, and G.O. Wesolowsky. *Facilities Location*,
          chapter 3.3, pages 51–60. North-Holland, Amsterdam, 1988.

[Lov71]   L. Lovász. On the number of halving lines. In *Ann. Univ. Eötvös
          Loránd*, number 14 in Sekt. Math., pages 107–108. Budapest, 1971.

[LW86]   D.T. Lee and Y.F. Wu. Geometric complexity of some location prob-
          lems. *Algorithmica*, 1:193–211, 1986.

[Mar87]   H. Martini. Some results and problems around zonotopes. In
          K. Böröczky and G. Fejes Tóth, editors, *Intuitive Geometry*, num-
          ber 48 in Coll.Math.Soc. J. Bolyai, pages 383–418. North-Holland,
          1987.

[Mes95]   J.A. Mesa. Continuous location of dimensional structures. Working
          paper, presented at the Workshop on Locational Analysis, 1995.

[Min67]   H. Minkowski. *Gesammelte Abhandlungen, Band 2*. Chelsea Publish-
          ing Company, New York, 1967.

[MN80]   J.G. Morris and J.P. Norback. A simple approach to linear facility
          location. *Transportation Science*, 14(1):1–8, 1980.

[MN83]   J.G. Morris and J.P. Norback. Linear facility location - solving exten-
          sions of the basic problem. *European Journal of Operational Research*,
          12:90–94, 1983.

[MS97]   H. Martini and A. Schöbel. The minisum hyperplane problem in
          Minkowski spaces. Technical Report 18, University of Kaiserslautern,
          Wirtschaftsmathematik, 1997. submitted.

[MS98]   H. Martini and A. Schöbel. A characterization of smooth norms. Tech-
          nical report, 1998. in preparation.

[MT82]   N. Megiddo and A. Tamir. On the complexity of locating linear facil-
          ities in the plane. *Operations Research Letters*, 1(5):194–197, 1982.

[MT83]   N. Megiddo and A. Tamir. Finding least-distance lines. *SIAM J. on
          Algebraic and Discrete Methods*, 4(2):207–211, 1983.

[NM80]   J.P. Norback and J.G. Morris. Fitting hyperplanes by minimizing
          orthogonal deviations. *Math. Programming*, 19:102–105, 1980.

[NW82]   S.C. Narula and J.F. Wellington. The minimum sum of absolute errors
          regression: a state of the art survey. *International Statistical Review*,
          50:317–326, 1982.

[OvL81]   M.H. Overmars and J. van Leeuwen. Dynamically maintaining con-
          figurations in the plane. *Journal Comput. Syst. Sci*, 23:166–204, 1981.

[PFTV86] W.H. Press, B.P. Flannery, S.A. Teulosky, and W.T. Vetterling. *Nu-
          merical Recipes*. Cambridge University Press, Cambridge, New York,
          1986.

[Pla95]   F. Plastria. Continuous location problems. In Z. Drezner, editor,
          Facility Location: A survey of applications and methods, chapter 11,
          pages 225–262. Springer, New York, Inc., 1995.

[PSS92]   J. Pach, W. Steiger, and E. Szemeredi. An upper bound on the number
          of planar *k*–sets. *Discrete Comput. Geometry*, 7:109–123, 1992.

[Ric64]   J. Rice. *The Approximation of Functions: The Linear Theory*.
          Addison-Wesley, 1964. Volume 1.

[RL87]   P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier De-
          tection*. Wiley & Sons, New York, 1987.

[RND77]   E.M. Reingold, J. Nievergelt, and N. Deo. *Combinatorial Algorithms:
          Theory and Practice*. Prentice Hall, Englewood Cliffs, New Jersey,
          1977.

[RS72]   M.R. Rao and V. Srinivasan. A note on Sharpe's algorithm for min-
          imizing the sum of absolute deviations in a simple regression problem.
          Management Science, 19:222–225, 1972.

[Sch73]   E.J. Schlossmacher. An iterative technique for absolute deviations
          curve fitting. *Journal Amer. Statist. Association*, 68:857–859, 1973.

[Sch96]   A. Schöbel. Locating least-distant lines with block norms. *Studies in
          Locational Analysis*, 10:139–150, 1996.

[Sch97]   A. Schöbel. Locating line segments with vertical distances. *Studies in Locational Analysis*, 11:143–158, 1997.

[Sch98]   A. Schöbel. Locating least distant lines in the plane. *European Journal of Operational Research*, 106(1):152–139, 1998.

[Sha78]   M.I. Shamos. *Computational Geometry*. PhD thesis, Department of Computer Science, Yale University, New Haven, 1978.

[SW87]   H. Späth and G.A. Watson. On orthogonal linear L_1 approximation. *Numer. Math.*, 51:531–543, 1987.

[Tho96]   A.C. Thompson. *Minkowski Geometry*. Cambridge University Press, 1996.

[VZ94]   Vrećica and R. Zivaljević. New cases of the colored Tverberg's theorem. In *Jerusalem Combinatorics '93*, volume 178 of *Contemp. Math.*, pages 325–334, 1994.

[Wag59]   H.M. Wagner. Linear programming techniques for regression analysis. *Journal Amer. Statist. Association*, 54:206–212, 1959.

[Wes75]   G.O. Wesolowsky. Location of the median line for weighted points. *Environment and Planning A*, 7:163–170, 1975.

[YKII88]   P. Yamamoto, K. Kato, K. Imai, and H. Imai. Algorithms for vertical and orthogonal L_1 linear approximation of points. Proc. 4th Ann. Sympos. Comput. Geom., pages 352–361, 1988.

[Zem84]   E. Zemel. An O(n) algorithm for the linear multiple choice knapsack problem and related problems. *Inform. Process. Letters*, 18:123–128, 1984.

[ZV92]   R. Zivaljević and S. Vrećica. The colored Tverberg's problem and complexes of injective functions. *Journal Combin. Theory Series A*, 61:309–318, 1992.