Median hyperplanes in normed spaces

Horst Martini
Mathematische Fakultät
TU Chemnitz-Zwickau
D–09009 Chemnitz
Germany

Anita Schöbel
Fachbereich Mathematik
Universität Kaiserslautern
D–67653 Kaiserslautern
Germany

Abstract: In this paper we deal with the location of hyperplanes in n–dimensional normed spaces. If d is a distance measure, our objective is to find a hyperplane H which minimizes

$$f(H) = \sum_{m=1}^{M} w_m d(x_m, H),$$

where $w_m \geq 0$ are non-negative weights, $x_m \in R^n, m = 1, \ldots, M$, are M demand points and $d(x_m, H) = \min_{z \in H} d(x_m, z)$ is the distance from x_m to the hyperplane H. In robust statistics and operations research such an optimal hyperplane is called a median hyperplane.

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 (in a certain sense) a halving one with respect to the given point set.

1 Introduction

We consider the problem of approximating a set of points n-dimensional $\mathcal{X} = \{x_1, x_2, \ldots, x_M\} \subset R^n$ by a linear function (the linear fit problem). That problem plays an important role in different mathematical disciplines:

1. In *robust statistics* and *numerical mathematics*, linear fit problems are known as absolute errors regression, 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 a 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 $\mathcal{X} \subseteq R^n$. 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 use a slightly modified definition of halving (which we call pseudo-halving), see Definition 3. Several estimates on the number of halving lines have been developed and will be discussed in Section 2.

3. In *operations research* 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 (the average distance, or the maximum distance) between the existing facilities and the new one. 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]. Extensions to other distances were given by [Sch96b] to block norms and by [Sch96a] to arbitrary distances derived from norms. In the following we will generalize the two latter papers to n-dimensional spaces.

One 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].

We use two different analytical descriptions of hyperplanes, given by the following definition.

Definition 1 *1. Let the real numbers $\tilde{s_1}, \tilde{s_2}, \ldots, \tilde{s_n}$ and b be given, such that at least one of the numbers $\tilde{s_i} \neq 0$. Then we define the hyperplane*

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

2. Let $S := \{s_1, s_2, \ldots, s_{n-1}\} \subset R^n$ be a set of $n-1$ linearly independent vectors and $z \in R^n$. Then we define the hyperplane

$$H_{z,S} := \{x = z + \lambda_1 s_1 + \lambda_2 s_2 + \ldots + \lambda_{n-1} s_{n-1} : \lambda_1, \lambda_2, \ldots, \lambda_{n-1} \in R\}.$$

Now the problem we are dealing with in this paper can be stated mathematically: 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 $x_m \in R^n$ with non-negative 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

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

Some more notations 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 $e_1, \ldots, e_d$, $|\mathcal{A}|$ denotes the number of elements contained in $\mathcal{A}$, and $\text{lin}(\mathcal{A})$, $\text{aff}(\mathcal{A})$ are the sets of all linear and affine combinations of elements of $\mathcal{A}$. For a hyperplane H let H^+,H^-, and H^0 denote the set of demand points x_m lying on either side of H and in H, respectively. In the following we assume that $\mathcal{X}$ contains at least n 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 $f(H) = 0$.
Note that, if $\{x_1, x_2, \ldots, x_n\}$ are n affinely independent vectors in R^n, then $H = \{x : x = \sum_{i=1}^n \lambda_i x_i \text{ with } \sum_{i=1}^n = 1\}$ is the only hyperplane passing through $x_1, x_2, \ldots, x_n$, such that the hyperplane is fixed by those points.

Definition 2 *A hyperplane H is called a halving hyperplane with respect to $\mathcal{X} = \{x_m : m \in \mathcal{M}\}$ if*

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

This definition is the same as the definition of halving given in [NM80],[MN80]. Note that in this definition, it is not required that n of the demand points are on the hyperplane, as it is e.g. in [KM93]. For the Euclidean case, all optimal hyperplanes are halving ones, see [KM90]. That is not necessarily true for more general norms. We therefore introduce the definition of pseudo-halving and 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*

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

We will use the classification scheme of [HN93] which was originally developed for location theory, but is also helpful in this context: In that 5 position scheme our problem can be described as $1H/R^n/\cdot/d/\sum$, 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 distances between the demand points x_m and the hyperplane H ($\sum$). The special case of finding an optimal line in the plane can be described as $1l/P/\cdot/d/\sum$, and if we restrict ourselves to Euclidean distances we would write $1H/R^n/\cdot/l_2/\sum$.

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. Sections 5 and 6 give some algorithmic approaches for the general case and for the case that the distance has been derived from a block norm. The paper is concluded by remarks on possible extensions.

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. Thus, we are concerned with the (weighted) orthogonal L_1 approximation of finite point sets of cardinality M in R^n by hyperplanes.
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 [LC85] sharpened this up 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 for the planar unweighted case $1l/P/w_m = 1/l_2/\sum$ (i.e., all points have equal weight). 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 lines which are halving to the given points $x_1, x_2, \ldots, x_M$ are known:

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

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

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 around 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 around 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 $\Omega(M log M)$, 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] arrived at 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 all whose r-faces, $2 \leq r \leq n-1$, are centrally symmetric) and the weighted orthogonal L_1 approximation reduced to $(n-1)$-subspaces was observed: using necessary conditions for local minima of these support functions, one can prove that *every* optimal $(n-1)$-subspace passes through $n-1$ linearly independent given points. Identifying one given point with the "new" origin, this yields the following necessary condition for the solutions of the weighted orthogonal L_1 approxination problem, due to [KM90] (see also [KM93]): In Euclidean n-space, *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 C_i, such that the intersection of $conv(S_j)$ 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 $1H/R^n/\cdot/d_{hor}/\sum$, that means how to find a best hyperplane minimizing the sum of lengths of the horizontal lines between the given points and the hyperplane. That problem has been well solved in the plane (see e.g. [MT83] and [MN83]). In [Zem84] an n–dimensional formulation is developed, which can be solved efficiently.
At first we give a formula to calculate the distance $d_{hor}(x, H)$ between a point $x \in R^n$ and a hyperplane H.

Let $x = (x_1, x_2, \ldots, x_n) \in R^n$. If $H_{b,\tilde{s_1},\ldots,\tilde{s_n}}$ is a hyperplane according to Definition 1.1, we get:

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

Now we note that for finding a hyperplane H minimizing $f(H)$ we can set $\tilde{s_1} = 1$. Let $H := H_{b,\tilde{s_1},\ldots,\tilde{s_n}}$.

Case 1: If $\tilde{s_1} > 0$, then $H_{b,\tilde{s_1},\ldots,\tilde{s_n}} = H_{\frac{b}{\tilde{s_1}},1,\frac{\tilde{s_2}}{\tilde{s_1}},\ldots,\frac{\tilde{s_n}}{\tilde{s_1}}}$.

Case 2: If $\tilde{s_1} = 0$ and $x_{m1}\tilde{s_1} + \ldots + x_{mn}\tilde{s_n} = b$ for all $m \in \mathcal{M}$, then the problem is trivially solved: $f(H) = 0$ with H a minimizing hyperplane which is a pseudo-halving one passing through all points $x_m \in \mathcal{X}$.

Case 3: If $\tilde{s_1} = 0$ and $x_{k1}\tilde{s_1} + \ldots + x_{kn}\tilde{s_n} \neq b$ for at least one $k \in \mathcal{M}$, then $d(x_k, H) = \infty$, enforcing that $f(H) = \infty$ such that H cannot be optimal.

Summarizing these statements, we get as objective function:

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

Lemma 1 *For a given set $\mathcal{X} = \{x_m : m \in \mathcal{M}\} \subset R^n$ and non-negative 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 $x_m \in \mathcal{X}$.

Proof: Let us assume that $H_{b^*,\tilde{s_2}^*,\ldots,\tilde{s_n}^*}$ is an optimal hyperplane which does not pass through n affinely independent points $x_m \in \mathcal{X}$. Define

$$\mathcal{M}_0 := \{m \in \mathcal{M} : |b^* - x_{m1} - \tilde{s_2}^* x_{m2} - \ldots - \tilde{s_n}^* x_{mn}| = 0\}.$$

Because of continuity of $d(x_m, H) = |b^* - x_{m1} - \tilde{s_2}^* x_{m2} - \ldots - \tilde{s_n}^* x_{mn}|$ there exists an environment $U = U(b^*, \tilde{s_2^*}, \ldots, \tilde{s_n^*})$ around that solution such that for all $(b, \tilde{s_2}, \ldots, \tilde{s_n}) \in U$

$$|b - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| \neq 0 \text{ for all } m \notin \mathcal{M}_0.$$

Now we look at

$$\mathcal{L} := \{(b, \tilde{s_2}, \ldots, \tilde{s_n}) : |b - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| = 0 \text{ for all } m \in \mathcal{M}_0\}.$$

We know that $\mathcal{L} \neq \emptyset$, since $(b^*, \tilde{s_2^*}, \ldots, \tilde{s_n^*}) \in \mathcal{L}$. Since $\{x_m : m \in \mathcal{M}_0\}$ is not affinely independent, the solution of the linear system

$$|b - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| = 0 \text{ for all } m \in \mathcal{M}_0$$

is not unique. That means, there exists another hyperplane $H' = H_{b',\tilde{s_2}',\ldots,\tilde{s_n}'} \neq H^*$ such that $x_m \in H'$ for all $m \in \mathcal{M}_0$. In other words, $(b', \tilde{s_2}', \ldots, \tilde{s_n}') \in \mathcal{L}$. Thus, $\mathcal{L}$ is the solution space of a linear system containing at least two different points, which means that $\dim(\mathcal{L}) \geq 1$. Since the objective function

$$f(H_{b,\tilde{s_2},\ldots,\tilde{s_n}}) = \sum_{m \notin \mathcal{M}_0} w_m |b - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}|$$

is linear for all $(b, \tilde{s_2}, \ldots, \tilde{s_n}) \in U \cap \mathcal{L}$, we can find a solution which is at least as good as $H_{b^*,\tilde{s_2}^*,\ldots,\tilde{s_n}^*}$ and passes through n affinely independent points $x_m \in \mathcal{X}$. q.e.d.

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.

Proof: Suppose that $H = H_{b,\tilde{s_2},\ldots,\tilde{s_n}}$ is optimal, but

$$\sum_{x_m \in H^+} w_m = \sum_{m: x_{m1} + \tilde{s_2} x_{m2} + \ldots \tilde{s_n} x_{mn} > b} w_m > \frac{1}{2} W.$$

We define

$$\mathcal{M}^+ := \{m \in \mathcal{M} : x_{m1} + \tilde{s_2} x_{m2} + \ldots \tilde{s_n} x_{mn} > b\}$$

and $\mathcal{M}^- := \mathcal{M} \setminus \mathcal{M}^+$. Then we choose an $\epsilon \in R$ such that $\epsilon > 0$ and

$$\mathcal{M}^+ = \{m \in \mathcal{M} : x_{m1} + \tilde{s_2} x_{m2} + \ldots \tilde{s_n} x_{mn} > b + \epsilon\}.$$

Evaluating $H_\epsilon = H_{b+\epsilon, \tilde{s_2}, \ldots, \tilde{s_n}}$ leads to

$$\begin{aligned} f(H_\epsilon) &= \sum_{m \in \mathcal{M}^+} w_m |b + \epsilon - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| \\ &\quad + \sum_{m \in \mathcal{M}^-} w_m |b + \epsilon - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| \\ &= \sum_{m \in \mathcal{M}^+} w_m (|b - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| - \epsilon) \\ &\quad + \sum_{m \in \mathcal{M}^-} w_m (|b - x_{m1} - \tilde{s_2} x_{m2} - \ldots - \tilde{s_n} x_{mn}| + \epsilon) \\ &= f(H) + \epsilon \left(\sum_{m \in \mathcal{M}^-} w_m - \sum_{m \in \mathcal{M}^+} w_m \right) < f(H), \end{aligned}$$

contradicting the optimality of H. q.e.d.

In the same way as looking at horizontal distances it is possible to define the distances in the other directions $e_2, e_3, \ldots, e_n$ by

$$d_{e_i}(x, H) = |\frac{b - x_1 \tilde{s_1} - \ldots - x_n \tilde{s_n}}{\tilde{s_i}}|, \; i \in \{2, 3, \ldots, n\},$$

leading to 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

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

one consequence of Lemma 1 and Lemma 2 is that both results also hold for l_1:

Theorem 1 *For rectangular distances $d = l_1$ the following two conditions are satisfied:*

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 [Sch96a] for the two-dimensional case.

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].

We repeat that the distance between a point $x \in R^n$ and a hyperplane H is defined by

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

Let d be any metric derived from a norm, let $x_m \in R^n$ for $m \in \mathcal{M} = \{1, 2, \ldots, M\}$ be a given set of points and let w_m for all $m \in \mathcal{M}$ be non-negative weights. The problem that we want to solve now reads as follows.
Find a hyperplane H minimizing

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

In the classification scheme from [HN93] (see Section 1) we can write this problem as $1H/R^n/\cdot/norm/\sum$.

At first we note that to determine the distance between a point x and a hyperplane H we can increase the unit ball around x until it touches the hyperplane.

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:*

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

Proof:

$$\begin{aligned} d(x, H) &= \min_{z \in H} d(x, z) \\ &= \min_{z \in H} \min\{|\lambda| : z - x \in \lambda B\} \\ &= \min_{z \in H} \min\{|\lambda| : z \in \lambda B + x\} \\ &= \min\{|\lambda| : \exists z \in H \text{ such that } z \in \lambda B + x\} \\ &= \min\{|\lambda| : (x + \lambda B) \cap H \neq \emptyset.\} \qquad \text{q.e.d.} \end{aligned}$$

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 [Sch96a] 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 $t \notin \text{lin}\{s_1, s_2, \ldots, s_{n-1}\}$.
As example we get that the length of the horizontal segment from x_m to H then is $d_{e_1}(x_m, H) = d_{hor}(x_m, H)$ with the first unit vector $e_1 \in R^n$.

Lemma 4 *Let $p, q \in R^n$ and D be a linear transformation with*

1. *$D(p) = q$ and*
2. *$det(D) \neq 0$.*

Then we have

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

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

<u>Proof:</u> The proof has been done in [Sch96a] for two dimensions, but is also valid for more than two dimensions by replacing R^2 by R^n. However, to make the paper self-contained, we shortly give this proof again:
We first show that

$$d_q(D(x), D(y)) = d_p(x, y) \text{ for points } x, y \in R^n.$$

Case 1: $d_p(x, y) = \overline{\alpha} < \infty$: That means, $x - y = \alpha p$ with $|\alpha| = \overline{\alpha}$ and we get

$$\begin{aligned} d_q(D(x), D(y)) &= \gamma_q(D(y) - D(x)) \\ &= \gamma_q(D(y - x)) \\ &= \gamma_q(D(\alpha p)) \\ &= \gamma_q(\alpha D(p)) \\ &= \gamma_q(\alpha q) = |\alpha| = \overline{\alpha}. \end{aligned}$$

Case 2: $d_p(x,y) = \infty$: Then we know that $x-y$ and t are linearly independent, which means that also $D(x-y)$ and $D(t)$ are linearly independent (because $\det(D) \neq 0$) and we get

$$d_q(D(x), D(y)) = \infty.$$

Since

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

we now can conclude that for a hyperplane H and any point x_m

$$\begin{aligned} d_q(D(x_m), D(H)) &= \min_{z \in H} d_q(D(x_m), D(z)) \\ &= \min_{z \in H} d_p(x_m, P) = d_p(x_m, H) \qquad \text{q.e.d.} \end{aligned}$$

With the help of 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 two conditions hold:*

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.*

<u>Proof:</u> To solve $1H/R^n/\cdot/d_t/\sum$ we can proceed as follows:

1. Choose $q = e_1, p = t$ and transform $1H/R^n/\cdot/d_t/\sum$ to $1H/R^n/\cdot/d_{hor}/\sum$ according to Lemma 4. We then know that H is optimal for the problem $1H/R^n/\cdot/d_t/\sum$ if and only if $D(H)$ is optimal for $1H/R^n/\cdot/d_{hor}/\sum$.
2. Solve $1H/R^n/\cdot/d_{hor}/\sum$ and get an optimal solution H^*_{hor} which passes through n affinely independent points $D(x_m)$, $x_m \in \mathcal{X}$ (Lemma 1). We also know from Lemma 2 that all optimal hyperplanes are pseudo-halving ones for the transformed point set $D(\mathcal{X}) = \{D(x_m) : x_m \in \mathcal{X}\}$
3. Determine $D_t = D^{-1}(H^*)$. Since

$$D(x_m) \in H^*_{hor} \Longleftrightarrow x_m \in D^{-1}(H^*_{hor})$$

and affine independency is invariant with respect to the transformation D, we know that H_t passes through n affinely independent points $x_m \in \mathcal{X}$. For the second statement we use that under the transformation D either no x_m or all x_m change the side of H such that all optimal hyperplanes have to be pseudo-halving ones for the distances d_t. q.e.d.

In the following we will show that for any distance d derived from a norm and any hyperplane with fixed slopes $s_1, s_2, \ldots, s_{n-1} \in R^n$ there exists a $t \in R^n$ such that

$$d(x_m, H) = d_t(x_m, H) \text{ for all } m \in \mathcal{M}.$$

Thus, when evaluating the objective function $f(H)$ we can replace d by d_t.

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 set of slopes $S = \{s_1, s_2, \ldots, s_{n-1}\} \subset R^n$ be given and let $t \notin \mathrm{lin(S)}$. Then there exists a constant $C := C(S, d, l_2)$ such that for all $z \in R^n$ and all $x \in R^n$*

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

Proof: Consider at first $x = 0$. For a fixed point $z_0 \notin \mathrm{lin}(S)$ we know that $l_2(0, H_{z_0,S}) \neq 0$ and $0 \neq d(0, H_{z_0,S}) < \infty$, and therefore we find a real number $C \neq 0$ such that $d(0, H_{z_0,S}) = Cl_2(0, H_{z_0,S})$. Now we look at a hyperplane $H = H_{z,S} \neq H_{z_0,S}$. Since $z_0 \notin \mathrm{lin}(S)$, H can be written as $H_{z,S}$ with $z = \beta z_0$ for a real number β. Hence we get

$$\begin{aligned}
\beta d(0, H_{z_0,S}) &= \beta \min_{y \in H_{z_0,S}} d(0, y) \\
&= \min_{\alpha_1, \ldots, \alpha_{n-1} \in R} \beta d(0, z_0 + \alpha_1 s_1 + \ldots + \alpha_{n-1} s_{n-1}) \\
&= \min_{\alpha_1, \ldots, \alpha_{n-1} \in R} \beta \gamma(z_0 + \alpha_1 s_1 + \ldots + \alpha_{n-1} s_{n-1}) \\
&= \min_{\alpha_1, \ldots, \alpha_{n-1} \in R} \gamma(\beta(z_0 + \alpha_1 s_1 + \ldots + \alpha_{n-1} s_{n-1})) \\
&= \min_{\alpha'_1, \ldots, \alpha'_{n-1} \in R} \gamma(z + \alpha'_1 s_1 + \ldots + \alpha'_{n-1} s_{n-1}) \\
&= d(0, H_{z,s})
\end{aligned}$$

$$\Longrightarrow d(0, H_{z,S}) = \beta d(0, H_{z_0,S}) = \beta C l_2(0, H_{z_0,S}) = C l_2(0, H_{z,S}),$$

using the above equation for both d and in the special case l_2. For any point $x \in R^n$ with $x \neq 0$ we finally get:

$$\begin{aligned}
d(x, H_{z,S}) &= \min_{y \in H_{z,S}} d(x, y) \\
&= \min_{\alpha_1, \ldots, \alpha_{n-1} \in R} d(x, z + \alpha_1 s_1 + \ldots + \alpha_{n-1} s_{n-1}) \\
&= \min_{\alpha_1, \ldots, \alpha_{n-1} \in R} d(0, z - x + \alpha_1 s_1 + \ldots + \alpha_{n-1} s_{n-1}) \\
&= d(0, H') \\
&\quad \text{with } H' = \{z - x + \alpha_1 s_1 + \ldots + \alpha_{n-1} s_{n-1} : \alpha_1, \ldots, \alpha_{n-1} \in R\} \\
&= Cl_2(0, H') \text{ because the set of slopes of } H' \text{ remains } S \\
&= Cl_2(x, H_{z,S}). \qquad \text{q.e.d.}
\end{aligned}$$

Note that in Lemma 5 the properties of the Euclidean distance l_2 have not been used, such that l_2 can be replaced by any other distance derived from a norm or by distances derived from γ_t with $t \notin \text{lin}(S)$. If d_1, d_2, and d_3 are such distances and S is a set of slopes, we get

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

In particular, if $H = \{(x_1, x_2) : x_2 = \tilde{s}x_1 + b\}$ is a line in R^2 with $\tilde{s}, b \in R$, $\tilde{s} \neq 0$, we obtain

$$C(\begin{pmatrix} 1 \\ \tilde{s} \end{pmatrix}, d_{hor}, d_{ver}) = |\tilde{s}|.$$

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

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

As an immediate consequence of Lemma 5 we get: For a given set of slopes $S \subset R^n$ the optimal hyperplane H with slopes $S = \{s_1, s_2, \ldots, s_{n-1}\}$, that means the hyperplane $H_{z^*,S}$ minimizing

$$\min_{z \in R^n} f(H_{z,S})$$

is the same for all norms d and distances d_t.
There is another reason for introducing the distances d_t. Namely, the following relation between any distance $d(x, y) = \gamma(y - x)$ derived from a norm γ and the distances d_t holds.

Lemma 6

$$d(x, H) = \min_{t \in R^n, \gamma(t)=1} d_t(x, H).$$

Proof: To make the paper self-contained we repeat here the proof originally given in [Sch96a]: Because of Lemma 3 we know that

$$d(x_m, H) = \min\{|\lambda| : x_m + \lambda B \cap H \neq \emptyset\} =: \lambda^0.$$

That means there exists $t^0 \in R^n$ with $\gamma(t^0) = 1$ such that $x_m + \lambda^0 t^0 \in H$. (Note that $\gamma(t) = 1$ if and only if $t \in \text{boundary}(B)$.) Using the definition of d_{t^0}, that means:

$$d(x_m, H) = d_{t^0}(x_m, H).$$

For all t' with $\gamma(t') = 1$ we can calculate that

$$\begin{aligned} d_{t'}(x_m, H) &= \min\{|\lambda| : x_m + \lambda t' \in H\} \\ &\geq \min\{|\lambda| : x_m + \lambda B \cap H \neq \emptyset\} \\ &= d(x_m, H) \text{ using Lemma 3 again} \qquad \text{q.e.d.} \end{aligned}$$

Lemma 7 *Let H be a hyperplane with S as set of slopes, 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*

$$d(x,H) = d_t(x,H) \text{ for all } x \in R^n.$$

Proof: Let $x \in R^n$. According to Lemma 6 we can find a direction $u \in R^n$ such that $\gamma(u) = 1$ and

$$d(x,H) = d_u(x,H) \leq d_t(x,H) \text{ for all t} \in R^n.$$

Suppose that there exist points $y \in R^n$ and $v \in R^n$ with $\gamma(v) = 1$ and

$$d(y,H) = d_v(y,H) < d_u(y,H).$$

Note that $u \notin \text{lin}(S)$ and $v \notin \text{lin}(S)$ because $d(x,H) \neq \infty$ and $d(y,H) \neq \infty$. With Lemma 5 we know that there exists a constant $C := C(S, d_u, d_v)$ such that

$$d_u(x,H) = Cd_v(y,H) \text{ and } d_u(x,H) = Cd_v(y,H),$$

yielding $C > 1$ in the first case and $C \leq 1$ in the second case, which is impossible. q.e.d.

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

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.*

Proof:

ad 1: Suppose H^* is an optimal hyperplane, but does not pass through n affinely independent demand points. Choose t^* such that $d(x_m, H^*) = d_{t^*}(x_m, H^*)$ for all $m \in \mathcal{M}$ according to Lemma 7.

By Theorem 2 we know that the first result holds for $1H/R^n/\cdot/d_{t^*}/\sum$, and therefore we can choose a hyperplane H^0 minimizing the sum of all distances d_{t^*} and passing through n affinely independent demand points.

Now let t^0 be given such that $d(x_m, H^0) = d_{t^0}(x_m, H^0)$ for all $m \in \mathcal{M}$ according to Lemma 7 again. Then we get:

$$\begin{aligned} f(H^*) &= \sum_{m \in \mathcal{M}} w_m d(x_m, H^*) \\ &= \sum_{m \in \mathcal{M}} w_m d_{t^*}(x_m, H^*) \end{aligned}$$

$$\begin{aligned}
&\geq \sum_{m \in \mathcal{M}} w_m d_{t^*}(x_m, H^0) \\
&\geq \sum_{m \in \mathcal{M}} w_m d_{t^0}(x_m, H^0) \quad \text{because of Lemma 6} \\
&= \sum_{m \in \mathcal{M}} w_m d(x_m, H^0) \\
&= f(H^0) \geq f(H^*) \text{ by the optimality of } H^*.
\end{aligned}$$

Thus, also H^0 is an optimal hyperplane. That hyperplane passes through n affinely independent demand points which completes the proof.

ad 2: We assume that there is an optimal hyperplane H^* with

$$\sum_{x_m \in (H^*)^-} w_m > \frac{W}{2}$$

. With the same notation as in the first part of the proof we know from Theorem 2 that any hyperplane H^0 minimizing d_{t^*} satisfies

$$\sum_{x_m \in (H^0)^- w_m} \leq \frac{W}{2} \text{ and } \sum_{x_m \in (H^0)^+} w_m \leq \frac{W}{2}.$$

Therefore we get

$$\begin{aligned}
f(H^*) &= \sum_{m \in \mathcal{M}} w_m d_{t^*}(x_m, H^*) \\
&> \sum_{m \in \mathcal{M}} w_m d_{t^*}(x_m, H^0) \\
&\geq f(H^0),
\end{aligned}$$

contradicting the optimality of H^*. q.e.d.

5 Algorithmical approaches for general norms

By Lemma 3 and Lemma 6 the distance $d(x, H)$ strictly depends on the shape of the unit ball B which can be an arbitrary convex body centred about 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 widely ignore the difficulties mentioned above, and so it turns out that Theorem 3 yields approaches analogous to the Euclidean case discussed in [KM90], [HII+93], and [KM93]. But once more we underline a basic difference occuring here. Namely, [KM90] developed an $O(M^n)$ time and $O(M)$ space algorithm for the weighted Euclidean case on the basis that *each* optimal hyperplane has to pass through n affinely independent given points. In principle the same algorithm was established by [HII+93], but only on the basis of the weaker statement that there *exists* an optimal hyperplane satisfying that incidence criterion. In general, part 1 of Theorem 3 above cannot be sharpened to the state as in the Euclidean situation (a simple example for the l_1-case is given in the final section), and so all our algorithmic procedures refer to only one of possibly many optimal hyperplanes.

With these remarks as starting point, we only have to give a brief outline of the computational approaches from [HII+93] and [KM93], 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 line 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 leads, together with some further considerations, 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 possibilities for further improvements are perhaps obtainable with the help of pseudo-halving hyperplanes.

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 solve the problem more efficiently. In the plane that was done in [Sch96b]. The same idea can be extended to n–dimensional space.

If B is 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

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

In the classification scheme the problem of locating a hyperplane with block norms can now be written as $1H/R^n/\cdot/d_B/\sum$.

Lemma 8 *Let d_B be derived from a block norm γ_B. Then*

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

Proof: For determining $d(x, H)$ we know from Lemma 3 that we can increase the unit ball around x until it touches H. In our case the unit balls are polytopes. Obviously, a hyperplane touches an n–dimensional polytope in at least one vertex of that polytope, see e.g. [Sha78]. Thus the formula

$$d(x, H) = \min_{t \in R^n, \gamma(t)=1} d_t(x, H)$$

from Lemma 6 simplifies for block norms to

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

q.e.d.

With the help of the above lemma we can decompose our problem in 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 that 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, w_m 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

$$z^* := f(H_g^*)$$
$$l^* := D^{-1}(H_g^*)$$

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

That 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.

A common and simple block norm is l_1. The problem $1/P/\cdot/l_1/\sum$, to locate a line in the plane with rectangular distances, has been well solved by many authors. Most algorithms for this problem separate it into two subproblems; namely to find the best lines minimizing the sum of vertical and the sum of horizontal distances, respectively (and then choosing the better one). Most of the l_1–algorithms in the plane also give a transformation from the horizontal line problem to the vertical line problem or vice versa. We remark that in this sense the usual planar algorithms for l_1 are special cases of the above algorithm.

7 Concluding remarks

We have shown that for all distances in R^n derived from norms, and all point sets $\mathcal{X}$ containing n affinely independent points, there exists a hyperplane, minimizing the sum of distances to all points in $\mathcal{X}$ and passing through n affinely independent points $x_1, \ldots x_n \in \mathcal{X}$. This is a generalisation of [Sch96a] where the above statement was proved in the two-dimensional case.

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.

Consider the following instance of our problem with rectangular distances in the plane $(1l/P/\cdot/l_1/\sum)$: The unit ball B then 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 e.g. the lines $x_1=0$ and $x_2=0$ have the same 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 possibe 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 (that *each* median hyperplane is necessarily the affine hull of n affinely independent points).

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 with respect to approximation problems of the type discussed here.

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.

[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.

[HN93] H.W. Hamacher and S. Nickel. Multicriteria planar location problems. Technical Report 243, Universität Kaiserslautern, 1993. accepted by EJOR.

[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 related topics. In J. Pach, editor, *New Trends in Discrete and Computational Geometry*, chapter 6, pages 135–162. Springer-Verlag, 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, 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 problems. *Algorithmica*, 1:193–211, 1986.

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

[Mes95] J.A. Mesa. Continuous location of dimensional structures. Workshop on Locational Analysis, 1995.

[Min67] H. Minkowski. *Gesammelte Abhandlungen, Band 2*. Chelsea Publishing 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 extensions of the basic problem. *EJOR*, 12:90–94, 1983.

[MT82] N. Megiddo and A. Tamir. On the complexity of locating linear facilities 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 configurations 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. *Numerical 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, 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*, volume 1. Addison-Wesley, 1964.

[RL87] P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection.* 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.

[Sch96a] A. Schöbel. Locating least-distance lines in the plane. Technical Report 3, Universität Kaiserslautern, 1996. submitted to Mathematics of Operation Research.

[Sch96b] A. Schöbel. Locating least-distant lines with block norms. In S. Nickel H.W.Hamacher, K. Klamroth and F. Plastria, editors, *Studies in Locational Analysis*, 1996. to appear.

[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.

[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.

[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.