

Gauge Distances and Median Hyperplanes^{1,2}

F. PLASTRIA³ AND E. CARRIZOSA⁴

Communicated by J. P. Crouzeix

Abstract. A median hyperplane in d -dimensional space minimizes the weighted sum of the distances from a finite set of points to it. When the distances from these points are measured by possibly different gauges, we prove the existence of a median hyperplane passing through at least one of the points. When all the gauges are equal, some median hyperplane will pass through at least $d - 1$ points, this number being increased to d when the gauge is symmetric, i.e. the gauge is a norm.

Whereas some of these results have been obtained previously by different methods, we show that they all derive from a simple formula for the distance of a point to a hyperplane as measured by an arbitrary gauge.

Key Words. Gauges, distance to a hyperplane, hyperplane fitting.

1. Gauge Distance to a Hyperplane

Let γ be a gauge on \mathbb{R}^d with unit ball B ; i.e., B is a compact convex set containing the origin in its interior such that

$$\gamma(x) = \min\{t \geq 0 \mid x \in tB\};$$

see e.g. Refs. 1–2. Given a hyperplane H in \mathbb{R}^d , the γ -distance of a point $a \in \mathbb{R}^d$ to H is defined as

$$d_\gamma(a, H) \stackrel{\text{def}}{=} \min\{\gamma(x - a) \mid x \in H\}.$$

¹The research of the second author was partially supported by a DGES Grant, Madrid, Spain.

²The authors thank two anonymous referees for many suggestions which helped streamline this paper.

³Professor, Department of Management Informatics, Vrije Universiteit, Brussels, Belgium.

⁴Professor, Facultad de Matemáticas, Universidad de Sevilla, Sevilla, Spain.

Let γ° be the dual (or polar) gauge of γ , given by

$$\gamma^\circ(v) \stackrel{\text{def}}{=} \max\{\langle v; y \rangle \mid \gamma(y) \leq 1\},$$

which is well-defined and also a gauge on (the dual space of) \mathbb{R}^d ; see e.g. Ref. 3. This definition implies directly the following well-known generalized Cauchy–Schwartz inequality (see e.g. Ref. 3, p. 129):

$$\langle v; y \rangle \leq \gamma^\circ(v)\gamma(y), \quad \forall v, y \in \mathbb{R}^d, \quad (1)$$

in which for any fixed $v \neq 0$ equality holds iff $y = \lambda z$, for some $\lambda \geq 0$ and some $z \in \partial\gamma^\circ(v)$, where $\partial\gamma^\circ(v)$ denotes the (nonempty) subdifferential of the dual gauge at v ; see e.g. Ref. 2. Note that equality in (1) for $v, y \neq 0$ also implies $\langle v; y \rangle > 0$.

We will denote the hyperplane of equation $\langle u; x \rangle = \beta$, $u \neq 0$, by $H(u, \beta)$, and the set of all hyperplanes in \mathbb{R}^d by \mathcal{H} .

The following theorem gives a simple expression for the gauge distance to a hyperplane. The use of (1) enables us to simplify the proof given in Ref. 4 for a similar problem.

Theorem 1.1. For any gauge γ and any hyperplane $H(u, \beta)$, we have

$$d_\gamma(a, H(u, \beta)) = \begin{cases} [\beta - \langle u; a \rangle] / \gamma^\circ(u), & \text{when } \langle u; a \rangle \leq \beta, \\ [\langle u; a \rangle - \beta] / \gamma^\circ(-u), & \text{when } \langle u; a \rangle > \beta. \end{cases}$$

Any γ -closest point of $H(u, \beta)$ to a is found as the unique intersection point of $H(u, \beta)$ with the line through a having as direction any subgradient of γ° at u when $\langle u; a \rangle \leq \beta$, and at $-u$ when $\langle u; a \rangle > \beta$.

Proof. Let $u \neq 0$, and assume first that

$$\langle u; a \rangle \leq \beta.$$

For any $x \in H(u, \beta)$, after substituting v by $u (\neq 0)$ and y by $x - a$ in the generalized Cauchy–Schwartz inequality (1), we have that

$$\gamma(x - a) \geq \langle u; x - a \rangle / \gamma^\circ(u) = [\beta - \langle u; a \rangle] / \gamma^\circ(u),$$

where equality happens at $x \in H(u, \beta)$ iff

$$x - a \text{ is of the form } \lambda z, \quad \text{for some } z \in \partial\gamma^\circ(u). \quad (2)$$

Moreover, such an x exists. Indeed, since $u \neq 0$, for any given $z \in \partial\gamma^\circ(u)$ we have

$$\gamma^\circ(z) = 1,$$

so that by (1)

$$\langle u; z \rangle = \gamma(z)\gamma^\circ(u) = \gamma^\circ(u) > 0;$$

thus, the function

$$\lambda \geq 0 \mapsto \langle u; a + \lambda z \rangle - \beta = \langle u; a \rangle - \beta + \lambda \langle u; z \rangle$$

has a unique root in $[0, +\infty[$. In other words, there exist some $\lambda \geq 0$ and $x \in H(u, \beta)$ satisfying (2).

The same reasoning can be used for the case $\langle u; a \rangle > \beta$ and will not be repeated here. □

When γ is symmetric [$\gamma(-x) = \gamma(x)$, for all $x \in \mathbb{R}^d$], i.e. γ is a norm, then its dual enjoys the same property, and the following simplified formula arises directly (compare with Ref. 5, which uses a proof based on the Kuhn–Tucker conditions).

Corollary 1.1. For any norm v , we have

$$d_v(a, H(u, \beta)) = |\beta - \langle u; a \rangle| / v^\circ(u).$$

Note also that, for the particular case of the l_p -distances, $1 \leq p \leq +\infty$, this also proves directly the formula (painstakingly derived by Ref. 6)

$$d_p(a, H(u, \beta)) = |\beta - \langle u; a \rangle| / l_q(u), \quad 1/p + 1/q = 1,$$

by the well-known duality

$$l_p^\circ = l_q, \quad \text{with } 1/p + 1/q = 1,$$

including their limits

$$p = 1, q = +\infty \quad \text{or} \quad p = +\infty, q = 1.$$

The proof above also shows that in fact this is a direct consequence of the Hölder inequality

$$\langle v; y \rangle \leq l_q(v)l_p(y).$$

2. Median Hyperplanes

Given a finite set $A \subset \mathbb{R}^d$, together with corresponding positive weights w_a ($a \in A$) and gauges γ_a on \mathbb{R}^d , any hyperplane H^* minimizing the weighted sum of the gauge distances from A is called a median hyperplane; i.e.,

$$H^* \in \arg \min \{ f(H) \mid H \in \mathcal{H} \},$$

where

$$f(H) = \sum_{a \in A} w_a d_{\gamma_a}(a, H).$$

Since for any $\lambda \neq 0$, we have

$$H(\lambda u, \lambda \beta) = H(u, \beta),$$

the function

$$H: \mathbb{R}^d \setminus \{0\} \times \mathbb{R} \rightarrow \mathcal{H}: (u, \beta) \mapsto H(u, \beta)$$

is surjective, with the property that, for each one-dimensional linear space X in $\mathbb{R}^d \times \mathbb{R}$, $X \neq \{0\} \times \mathbb{R}$, the set $H(X \setminus \{0\})$ is reduced to a singleton; i.e., all the nonzero points in X are mapped to the same hyperplane.

It is a well-established fact from topology that no continuous bijective mapping may exist between a subset of $\mathbb{R}^d \times \mathbb{R}$ and the projective space \mathcal{H} . Several continuous and bijective restrictions of H may however be considered.

For example, consider the restriction of H on the cylinder in $\mathbb{R}^d \times \mathbb{R}$ with base some unit sphere of \mathbb{R}^d (i.e., $S_{d-1} \times]0; +\infty[$), where

$$S_{d-1} \stackrel{\text{def}}{=} \{u \in \mathbb{R}^d \mid \|u\| = 1\},$$

and where $\|\cdot\|$ denotes the standard Euclidean norm (or any other norm). This is an injection, the image of which contains all \mathcal{H} except the hyperplanes through the origin, i.e. of type $H(u, 0)$. Note that allowing also $\beta = 0$ leads to loss of injectivity, since

$$H(u, 0) = H(-u, 0).$$

Restriction of H to some nonvertical hyperplane in $\mathbb{R}^d \times \mathbb{R}$ (which we will rather call a superplane, to distinguish it from hyperplanes in \mathbb{R}^d) not passing through the origin, and excepting its single point on the β -axis, yields another continuous injection to \mathcal{H} : consider the superplane $S(c, \lambda, \mu)$, $\mu \neq 0, \lambda \neq 0$, with equation

$$\langle c; u \rangle + \lambda \beta = \mu,$$

from which the point $(0, \mu/\lambda)$ is deleted; then, the only hyperplanes not represented will be those of form $H(v, \mu/\lambda)$, where $v \neq 0$ is any vector orthogonal to c .

The set of all hyperplanes in \mathbb{R}^d normal to some fixed $u \neq 0$ is denoted by

$$\mathcal{H}_u = \{H(u, \beta) \mid \beta \in \mathbb{R}\}.$$

Lemma 2.1. For any fixed $u \in \mathbb{R}^d \setminus \{0\}$, there exists a hyperplane H_u^* minimizing f on \mathcal{H}_u which passes through some point $a \in A$.

Proof. For fixed $u \neq 0$, Theorem 1.1 shows that, for any $a \in A$, the distance $\gamma_a(a, H(u, \beta))$ is a convex piecewise linear function of β : it consists of two unbounded pieces with breakpoint $\beta_u^a = \langle u, a \rangle$, linearly decreasing with slope $-1/\gamma^\circ(-u)$ on $]-\infty; \beta_u^a]$ and linearly increasing with slope $1/\gamma^\circ(u)$ on $[\beta_u^a; +\infty[$. Note also that it is coercive, i.e. asymptotically equal to $+\infty$ in any direction.

It follows that, as the positively weighted sum over all $a \in A$ of such distance functions, $f(H(u, \beta))$ is convex, coercive, and piecewise linear, with breakpoints $\beta_u^a, a \in A$, and therefore reaches its minimum at least at one of these breakpoints, say $\beta_u^{a^0}$ with $a^0 \in A$. But

$$\langle u, a^0 \rangle = \beta_u^{a^0}$$

means that the corresponding minimizing hyperplane,

$$H_u^* = H(u, \beta_u^{a^0}),$$

passes through a^0 . □

Let us introduce the notations $A^\#(u, \beta)$ for any $\# \in \{<, \leq, >, \geq\}$ by

$$A^\#(u, \beta) = \{a \in A \mid \langle u, a \rangle \# \beta\}.$$

In fact, the problem of minimizing f on \mathcal{H}_u may be seen as a one-dimensional asymmetric distance Weber problem (Ref. 7),

$$\min_{\beta \in \mathbb{R}} \sum_{a \in A^>(u, \beta)} [w_a/\gamma_a^\circ(-u)] |\langle u, a \rangle - \beta| + \sum_{a \in A^<(u, \beta)} [w_a/\gamma_a^\circ(u)] |\langle u, a \rangle - \beta|, \tag{3}$$

for which a fixed-point optimality property was derived. This interpretation of the problem also enables us to state an important property of median hyperplanes.

Definition 2.1. The hyperplane $H(u, \beta)$ halves A if

$$\sum_{a \in A^{\geq}(u, \beta)} w_a/\gamma_a^\circ(-u) \geq \sum_{a \in A^{<}(u, \beta)} w_a/\gamma_a^\circ(u), \tag{4}$$

$$\sum_{a \in A^{>}(u, \beta)} w_a/\gamma_a^\circ(-u) \leq \sum_{a \in A^{\leq}(u, \beta)} w_a/\gamma_a^\circ(u). \tag{5}$$

The following result generalizes analogous results in Ref. 8, and the first part proves a conjecture made there on p. 182.

Theorem 2.1. There exists a median hyperplane which passes through some point $a \in A$. Moreover, any median hyperplane halves A .

Proof. Lemma 2.1 shows that, for every fixed $u \neq 0$, the minimum of $f(H)$ is reached on \mathcal{H}_u at some hyperplane $H(u, \beta)$ with $\beta \in [\beta'_u, \beta^h_u]$, where

$$\beta'_u = \min\{\beta^a_u | a \in A\}, \quad \beta^h_u = \max\{\beta^a_u | a \in A\}.$$

The values β'_u and β^h_u are clearly continuous in u , so they reach their extreme values

$$\beta^l = \min\{\beta'_u | u \in S_{d-1}\}, \quad \beta^h = \max\{\beta^h_u | u \in S_{d-1}\}$$

on the (compact) unit sphere S_{d-1} .

It follows that finding the minimum of f on \mathcal{H} is equivalent to finding the minimum of $f(H(u, \beta))$ on $S_{d-1} \times [\beta^l, \beta^h]$, which is compact. Since f is continuous, this minimum will be reached, establishing the existence of a median hyperplane. Then, the existence of a median hyperplane which meets A follows immediately from Lemma 2.1.

Finally, if $H(u^0, \beta^0)$ defines a median hyperplane, then β^0 must solve (3) for $u = u^0$. The left and right directional derivatives of this convex function of β are respectively

$$\begin{aligned} & \sum_{a \in A^{<}(u^0, \beta^0)} w_a / \gamma^{\circ}_a(u^0) - \sum_{a \in A^{\cong}(u^0, \beta^0)} w_a / \gamma^{\circ}_a(-u^0), \\ & \sum_{a \in A^{\cong}(u^0, \beta^0)} w_a / \gamma^{\circ}_a(u^0) - \sum_{a \in A^{>}(u^0, \beta^0)} w_a / \gamma^{\circ}_a(-u^0). \end{aligned}$$

A necessary and sufficient condition for a minimum is that the first should be nonpositive and the second nonnegative; in other words, (u^0, β^0) should satisfy (4)–(5). □

Observe that, when there exists a hyperplane in \mathbb{R}^d containing A , then this is clearly a median hyperplane with objective value 0. Therefore, we further assume that this is not the case; i.e.

$$\dim(A) = d,$$

meaning that A contains at least $d + 1$ affinely independent points.

When all gauges are symmetric and equal (i.e., when all distances are measured by a same norm), we may then derive a much stronger result which was obtained already by other means in Refs. 8–9.

In the sequel, we will write f' for $f \circ H$, i.e.,

$$f'(u, \beta) \stackrel{\text{def}}{=} f(H(u, \beta)),$$

and consider the median hyperplane determination as the problem of minimizing f' on $\mathbb{R}^d \times \mathbb{R}$.

Theorem 2.2. For distances measured by a fixed norm, i.e. $\gamma_a = v$, $a \in A$, with $v(-x) = v(x)$ for all $x \in \mathbb{R}^d$, when $\dim(A) = d$, some median hyperplane passes through d affinely independent points of A .

Proof. Let $H(u^0, \beta^0)$ be a median hyperplane, the existence of which follows from Theorem 2.1. Without loss of generality, we may assume that $v^\circ(u^0) = 1$. Define $\mathcal{T} \subset \mathbb{R}^d \times \mathbb{R}$ by

$$(u, \beta) \in \mathcal{T} \quad \text{iff} \quad \begin{cases} \langle u; a \rangle \geq \beta, & \forall a \in A^{\geq}(u^0, \beta^0), \\ \langle u; a \rangle \leq \beta, & \forall a \in A^{<}(u^0, \beta^0), \end{cases}$$

and the following linear function:

$$g: \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R} \\ : (u, \beta) \mapsto \sum_{a \in A^{\geq}(u^0, \beta^0)} w_a(\langle u; a \rangle - \beta) + \sum_{a \in A^{<}(u^0, \beta^0)} w_a(\beta - \langle u; a \rangle).$$

Consider the set P in $\mathbb{R}^d \times \mathbb{R}$ defined as

$$P = \{(u, \beta) \in \mathcal{T} \mid g(u, \beta) = g(u^0, \beta^0)\}.$$

P is a closed polyhedral set in $\mathbb{R}^d \times \mathbb{R}$: it is defined by linear inequalities and one linear equality in $d + 1$ variables.

P is also bounded. Indeed, if unbounded, there would exist some $(u, \beta) \neq (0, 0)$ such that

$$(\lambda u + u^0, \lambda \beta + \beta^0) \in P, \quad \text{for all } \lambda \geq 0.$$

This means that we would have

$$\begin{aligned} \langle u; a \rangle - \beta &\geq 0, & \forall a \in A^{\geq}(u^0, \beta^0), \\ \langle u; a \rangle - \beta &\leq 0, & \forall a \in A^{<}(u^0, \beta^0), \\ g(u, \beta) &= 0, \end{aligned}$$

which by the definition of g implies that

$$\langle u; a \rangle - \beta = 0, \quad \text{for all } a \in A;$$

in other words, $A \subset H(u, \beta)$, which contradicts the assumption $\dim(A) = d$. Therefore, P is a polytope of (at most) dimension d . It contains the optimal solution (u^0, β^0) , so any solution optimizing f' on P is also a global optimal

solution. Moreover,

$$f'(u, \beta) = g(u, \beta) / v^\circ(u) = g(u^0, \beta^0) / v^\circ(u), \quad \forall (u, \beta) \in P.$$

Hence, minimizing f' on P turns out to be equivalent to maximizing the function $(u, \beta) \mapsto v^\circ(u)$ on P . Since this function is convex, it attains its maximum on the polytope P at some extreme point (u^1, β^1) . Since (u^1, β^1) is obtained as the point common to d hyperplanes bounding linearly independent halfspaces defining P , it corresponds to the hyperplane $H(u^1, \beta^1)$ passing through d affinely independent points of A . \square

That the previous theorem does not hold for an asymmetric gauge is shown by the counterexample in Ref. 8. However, we may show the following only slightly weaker result.

Theorem 2.3. For distances measured by a fixed gauge (i.e. $\gamma_a = \gamma$, $a \in A$), when $\dim(A) = d$, some median hyperplane passes through $d - 1$ affinely independent points of A .

Proof. Let $H(u^0, \beta^0)$ be a median hyperplane, and define the functions f^0 and g on $\mathbb{R}^d \times \mathbb{R}$,

$$\begin{aligned} f^0(u, \beta) &= \frac{1}{\gamma^\circ(u)} \left[\sum_{a \in A^{\geq}(u^0, \beta^0)} w_a(\langle u; a \rangle - \beta) \right] \\ &\quad + \frac{1}{\gamma^\circ(-u)} \left[\sum_{a \in A^<(u^0, \beta^0)} w_a(\beta - \langle u; a \rangle) \right] \\ &\stackrel{\text{def}}{=} f^{\geq}(u, \beta) + f^{<}(u, \beta), \\ g(u, \beta) &= \frac{1}{\gamma^\circ(u^0)} \left[\sum_{a \in A^{\geq}(u^0, \beta^0)} w_a(\langle u; a \rangle - \beta) \right] \\ &\quad + \frac{1}{\gamma^\circ(-u^0)} \left[\sum_{a \in A^<(u^0, \beta^0)} w_a(\beta - \langle u; a \rangle) \right]. \end{aligned}$$

Observe that g is linear, f^0 is nonlinear, while

$$g(u^0, \beta^0) = f^0(u^0, \beta^0) = f'(u^0, \beta^0).$$

Let the subset $P \subset \mathbb{R}^d \times \mathbb{R}$ be defined by the constraints

$$\begin{aligned} \langle u; a \rangle &\geq \beta, & \forall a \in A^{\geq}(u^0, \beta^0), \\ \langle u; a \rangle &\leq \beta, & \forall a \in A^<(u^0, \beta^0), \\ g(u, \beta) &= f'(u^0, \beta^0). \end{aligned}$$

It is easy to see by similar arguments as those used in Theorem 2.2 that P is a polytope of (at most) dimension d . Since for all $(u, \beta) \in P$, we have

$$A^\#(u, \beta) = A^\#(u^0, \beta^0), \# \in \{ \geq, < \},$$

it follows that on P we have $f^0 = f'$. But P contains the optimal solution (u^0, β^0) of f' on $\mathbb{R}^d \times \mathbb{R}$, so any minimum of f^0 on P will also be a global optimum of f' and yield a median hyperplane.

Since

$$f^0 = f^\geq + f^<,$$

any solution minimizing f^0 on P is an efficient solution to the biobjective problem of minimizing both f^\geq and $f^<$ on P .

Each of the functions f^\geq and $f^<$ is quasiconcave on P (see e.g. Ref. 10), since their upper level sets at level α are given respectively by inequalities of the form

$$\alpha \gamma^\circ(u) - \sum_{a \in A^\geq(u^0, \beta^0)} w_a (\langle u; a \rangle - \beta) \leq 0,$$

$$\alpha \gamma^\circ(-u) - \sum_{a \in A^<(u^0, \beta^0)} w_a (\beta - \langle u; a \rangle) \leq 0,$$

which, by the convexity of γ° , define convex sets in $\mathbb{R}^d \times \mathbb{R}$.

It was shown in Ref. 11 that, in this case, the set of edges (one-dimensional faces) of P constitutes a dominator of P ; i.e., for any $(u, \beta) \in P$, there exists some (u', β') on some edge of P with

$$f^\geq(u', \beta') \leq f^\geq(u, \beta),$$

$$f^<(u', \beta') \leq f^<(u, \beta),$$

and hence

$$f^0(u', \beta') \leq f^0(u, \beta).$$

And any edge of P is the intersection of $d - 1$ hyperplanes bounding linearly independent halfspaces defining P . Therefore, there exists some minimum of f^0 on P (and hence a global minimum of f') satisfying as equality $d - 1$ linearly independent inequalities among those defining P . But this corresponds to a hyperplane H passing through $d - 1$ affinely independent points of A . □

References

1. DURIER, R., and MICHELOT, C., *Geometrical Properties of the Fermat-Weber Problem*, European Journal of Operational Research, Vol. 20, pp. 332-343, 1985.

2. MICHELOT, C., *The Mathematics of Continuous Location*, Studies in Locational Analysis, Vol. 5, pp. 59–83, 1993.
3. ROCKAFELLAR, T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
4. THACH, P. T., *The Design Centering Problem as a DC Programming Problem*, Mathematical Programming, Vol. 41, pp. 229–248, 1988.
5. MANGASARIAN, O. L., *Arbitrary Norm Separating Plane*, Operations Research Letters, Vol. 24, pp. 15–23, 1999.
6. MELACHRINOUDIS, E., *An Analytical Solution to the Minimum L_p -Norm of a Hyperplane*, Journal of Mathematical Analysis and Applications, Vol. 211, pp. 172–189, 1997.
7. DREZNER, Z., and WESOLOWSKY, G. O., *The Asymmetric Distance Location Problem*, Transportation Science, Vol. 23, pp. 201–207, 1989.
8. SCHÖBEL, A., *Locating Lines and Hyperplanes*, Kluwer Academic Publishers, Dordrecht, Holland, 1998.
9. NORBACK, J. P., and MORRIS, J. G., *Fitting Hyperplanes by Minimizing Orthogonal Deviations*, Mathematical Programming, Vol. 19, pp. 102–105, 1980.
10. AVRIEL, M., DIEWERT, W. E., SCHAIBLE, S., and ZHANG, I., *Generalized Concavity*, Plenum Press, New York, NY, 1988.
11. CARRIZOSA, E., and PLASTRIA, F., *Dominators for Multiple-Objective Quasi-convex Maximization Problems*, Journal of Global Optimization, Vol. 18, pp. 35–58, 2000.