An Optimal Algorithm for the Hyperplane Depth in R^2

Stefan Langerman
Computer Science
Rutgers University

William Steiger
Computer Science
Rutgers University

Abstract

Given an arrangement of n hyperplanes $h_1, \ldots, h_n \in R^d$ the *hyperplane depth* of a point $P \in R^d$ is the minimum number of hyperplanes that a ray from P can meet. The *hyperplane depth* of the arrangement is the maximal depth of points P not in any h_i. We give an optimal $O(n \log n)$ deterministic algorithm to compute the hyperplane depth of an arrangement in dimension $d = 2$.

Summary

Given a set $S = \{h_1, \ldots, h_n\}$ of n hyperplanes in R^d, Rousseeuw and Hubert [7] defined the hyperplane depth of a point $x \in R^d$ to be

$$\delta(x) = \min_{u:\|u\|=1} (\rho(x, u)), \tag{1}$$

where $\rho(x, u)$ is the number of $h_i \in S$ that meet the ray $\{x + tu, t \geq 0\}$ through x in direction u. A median is a point of maximal depth. When $d = 1$, S is a set of n given reals and (1) measures the smaller of $|\{h_i \in S : h_i \leq x\}|$ and $|\{h_i \in S : h_i \geq x\}|$. A deepest point μ has depth $\lceil n/2 \rceil$ and if we insist that $\mu \in S$, it is the usual median. The depth in (1) describes a multivariate generalization of ranks and order statistics.

There are many situations - for example in Statistics - where a notion of multivariate rank or depth is useful, and several definitions for depth in R^d have already been suggested [2], [4]. In one of the most familiar ones, we are given a set $S = \{P_1, \ldots, P_n\}$ of n points in R^d. The depth, or Tukey depth of $x \in R^d$ is defined as the minimal number of points of S that must be in a closed halfspace containing x. A median is a point of maximal depth. Helly's theorem implies that a median always has Tukey depth at least $n/(d+1)$. For the case $d = 2$, Matoušek [6], described an algorithm that finds a planar Tukey median in time $O(n(\log n)^5)$.

The hyperplane depth in (1) was only recently proposed by Rousseeuw and Hubert, and was motivated by problems in robust regression. It has attracted much recent interest. On the combinatorial side it was just shown [1] that like the Tukey median, the hyperplane median must have depth at least $n/(d+1)$, and the lower bound is sharp. Other interesting combinatorial questions remain open.

On the computational side, a main question is to determine the complexity of selection; that is, computing points of given depth, or of maximal depth. We use the unit cost RAM model. The hyperplanes in S partition R^d into a complex of $O(n^d)$ convex cells, called the *arrangement* of S, and written $A(S)$. The hyperplane depth is the same for every point in a cell. The depth of the arrangement, $\delta(A(S))$ is the depth of the deepest cell.

In the original paper, Rousseeuw and Hubert point out that for any x, $\delta(x)$ can be computed in $O(n^{d-1} \log n)$, and since there are $O(n^d)$ vertices in $A(S)$, selection has cost $O(n^{2d-1} \log n)$. For $d = 2$ they mention an $O(n^3)$ algorithm [8]. In Amenta et. al. [1] it was observed that $A(S)$ can be constructed in $O(n^d)$ and then, using breadth-first-search on the graph of adjacent cells, the depth of every cell is obtained in the same $O(n^d)$ time. This is the analogue of using sorting to find the ordinary median.

For the case $d = 2$, vanKreveld et.al. [5] recently described a beautiful $O(n(\log n)^2)$ algorithm for hyperplane depth in R^2. This shows that it is not necessary to find the depth of every cell in order to find the deepest one, and is analogous to the fact that sorting is not optimal for the usual median.

In the present paper we prove the following:

Theorem 1 *Given a set S of n lines in the plane, the hyperplane depth of $A(S)$, and a witness point of that depth can be obtained in $\Theta(n \log n)$ time and $O(n)$ space.*

The proof is via an algorithm, shown to have the asserted complexity. The lower bound is via a connected components argument showing that $\Omega(n \log n)$ steps are needed by an algebraic decision tree that can decide if the depth of a given point is greater than k, a given integer. An easy consequence is that in the same time bounds, given $k \leq \delta(A(S))$, a point of depth k can be found, as long as the lines are in general position.

It is interesting that our algorithm combines parametric search and *prune-and-search* to improve the result in [5], in much the same way that the linear-time algorithm for the usual median improves on sorting. We are searching for a point μ in a cell C of maximal depth. In fact, we will determine for each line in S its above/below relationship with μ. Each step of the algorithm solves the parametric problem of finding that above/below relationship for a constant fraction $a < 1$ of the m (initially n) remaining lines. These lines are then pruned out and the search continues with the remaining lines. We are able to do the pruning step in $O(m \log n + n)$. Since there are at most $n(1-a)^j$ unpruned candidates after the j^{th} step, and at most $\log n$ steps in all, the complexity is at most

$$bn \log n + \sum_{j=0}^{\infty} b(1-a)^j n \log n = O(n \log n).$$

References

[1] N. Amenta, M. Bern, D. Eppstein, and S.-H. Teng. Regression depth and Center Points. *Discrete and Comp. Geom.* (to appear).

[2] B. Chazelle. Geometric Order statistics. *Proc ACM Geometry*, 1988.

[3] R. Cole, J. Salowe, W. Steiger, and E. Szemerédi. An Optimal Time Algorithm for Slope Selection, *SIAM J. Comp. 18*, 792-810 (1989).

[4] J. Gill, W. Steiger, and A. Wigderson. Geometric Medians, *Discrete Math. 108*, 37-51. (1992),

[5] M. van Kreveld, J. Mitchell, P. Rousseeuw, M. Sharir, J. Snoeyink, and B. Speckmann. Efficient Algorithms for Maximum Regression Depth. *Discrete and Comp. Geom.* (to appear).

[6] J. Matoušek. Computing the Center of a Planar Point Set. *Discrete and Computational Geometry: Papers from the DIMACS Special Year* Amer. Math. Soc., J.E. Goodman, R. Pollack, W. Steiger, Eds, (1992), 221-230.

[7] P. Rousseeuw and M. Hubert. Depth in an Arangement of Hyperplanes. *Discrete and Comp. Geom.* (1999)

[8] P. Rousseeuw and M. Hubert. Regression Depth. *J. Amer. Statist. Assoc.*(to appear in June 1999).