Computational Geometry: Lecture 17

Don Sheehy

March 16, 2010

1 Geometric Medians

This begins our third unit of the course. We have moved past Sorting and Planar Graphs and are now on to selection. The classic selection problem is the Median problem. Given a set of n numbers, find the one in the middle. Like many of the problems in this course, we are interested in interpreting this problem for geometric point sets rather than just points on the real line.

Recall that a closed *halfspace* is a set $\{x \in \mathbb{R}^d : a \cdot x \leq b, a \in \mathbb{R}^d, b \in \mathbb{R}\}$. In 1D, this is just all the points to one side of (and inlcuding) a point b/a on the real line.

Definition 1.1. *The* Tukey depth *of a point x with respect to a set P is the minimum number of points of P in any halfspace containing x.*

Tukey depth is also known as *halfspace depth* or *location depth*. We will just use the term *depth* rather than Tukey depth for brevity.

In R, the median is a point of depth $n/2$. In this case, depth is just the order or rank of the point from the beginning or end of the list. The reason, we rephrase rank in terms of depth is that it allows us to take our definition with us when we go up in dimension.

It is important to note here that the point x need not be an element of P. We can measure the depth of any point in the space with respect to P.

A natural notion of a median in higher dimensions is just to pick a point of maximum depth. It is not hard to see that we will not be able to get points of depth $n/2$ in $\mathbb{R}^2$. Consider just three points in convex position. Every point in the plane is contained in a halfspace that contains at most one point of P. Thus, the maximum depth of all points in the plane is only $n/3$.

The main theorem for today will imply that this is the worst case. That is, for every input set $P \subset \mathbb{R}^d$, there is a point (not necessarily in P) of depth at least $n/3$. More generally, we will show that for $P \subset \mathbb{R}^d$, there is a point of depth at least $\frac{n}{d+1}$.

Definition 1.2. *A point x is a* centerpoint *with respect to a set $P \subset R^d$ if the depth of p is at least $\frac{n}{d+1}$.*

So, another way to state the main theorem for today is: **Centerpoints always exist!**

2 Classic Convexity Theorems

Theorem 2.1 (Radon's Theorem)**.** *Given $n \geq d+2$ points $P \subset \mathbb{R}^d$, there exists a partition (U_+, U_-) of P such that $CC(U_+) \cap CC(U_-) \neq \emptyset$.*

Proof. Since $n \geq d+2$, the points $p_i \in P$ are affinely dependent. This means there exist coefficients $\alpha_1, \ldots, \alpha_n$ such that

$$\sum_{i=1}^{d+2} c_i s_i = 0 \text{ and } \sum_{i=1}^{d+2} c_i = 0.$$

Let $U_+ = \{p_i : \alpha_i \leq 0\}$ and $U_- = \{p_i : \alpha_i \leq 0\}$.

Let $k = \sum_{i:p_i \in U_+} \alpha_i = \sum_{i:s_i \in U_-} \alpha_i$. The point $p = \sum_{p_i \in U_+} \frac{\alpha_i}{k} p_i = \sum_{p_j \in U_-} \frac{-\alpha_j}{k} p_j$ is in $\operatorname{pos} U_+ \cap \operatorname{pos} U_-$. Moreover, the coefficients c_i/k have the property that $\sum_{p_i \in U_+} \frac{p_i}{k} = \sum_{p_j \in U_-} \frac{-c_j}{k} = 1$ so $p \in \operatorname{aff} U_+ \cap \operatorname{aff} U_-$. For any set U, $CC(U) = \operatorname{aff} U \cap \operatorname{pos} U$ so we conclude that $p \in CC(U_+) \cap CC(U_-)$. □

Theorem 2.2 (Helly's Theorem)**.** *Given a collection of convex sets $S_1, \ldots, S_n \subset \mathbb{R}^d$. If every $d+1$ of these sets have a common intersection, then the whole collection has a common intersection.*

Proof. We will prove the Theorem only for the special case of $n = d+2$ and provide only a sketch for how to extend the proof to higher numbers of sets.

Choose $x_i \in \bigcap_{j \neq i} S_j$ for $i = 1 \ldots d+2$. The hypothesis of the theorem ensures that such points exist. Let (U_+, U_-) be a Radon partition of $\{x_1, \ldots, x_{d+2}\}$ and let p be the corresponding Radon point. We will show that $p \in \bigcap_{i=1}^{d+2} X_i$. It will suffice to show $p \in X_i$ for an arbitrary X_i. Without loss of generality, say $x_i \in U_+$. Because of the way we chose the other x_j's we know that $x_j \in X_i$ for all $j \neq i$. In particular, every $x_j \in U_-$ is in X_i and thus $U_- \subseteq X_i$. Since X_i is convex, $U_- \subset X_i$ implies that $CC(U_-) \subset X_i$. Since p is a Radon point, $p \in CC(U_-) \subset X_i$.

When $n > d+2$, we can simply use induction on n and replace a pair of sets X_{n-1}, X_n with their intersection. You are encouraged to try to figure the details for yourself. □

3 The Centerpoint Theorem

We are now ready to put these together to prove the centerpoint theorem. First, we give a more formal statement of the result.

Theorem 3.1 (The Centerpoint Theorem)**.** *Let P be a set of n points in R^d. There exists a point $c \in \mathbb{R}^d$ such that every closed halfspace containing c contains at least $\frac{n}{d+1}$ points of P.*

Proof. Say that an open halfspace is *heavy* if it contains more than $\frac{dn}{d+1}$ points of P. Say that a closed halfspace is *light* if it contains fewer than $\frac{n}{d+1}$ points of P. Clearly, the complement of a heavy halfspace is light and vice versa. A point lies in a light halfspace if and only if it has depth less than $\frac{n}{d+1}$. This follows directly from the definition of depth. Thus a point c is a centerpoint if and only if it is contained in every heavy halfspace.

It will suffice to prove that the intersection of all heavy halfspaces is nonempty. We need to be a little careful here an observe that we don't really have to consider *all* heavy halfspaces. We can instead focus on only a finite set of them. This is important so that we can apply Helly's Theorem in the form stated above. There are several ways to reduce the number of halfspaces to consider only a finite set of heavy convex sets. You are invited to think about how to do this. For now, we'll just assume it.

Let $H_1, \ldots, H_{d+1}$ be heavy halfspaces. Each heavy halfspace omits fewer than $\frac{n}{d+1}$ points of P. So, by the pigeonhole principle, there must be some point $p \in P$ that is not omitted by any H_i. Thus $p \in \bigcap i = 1^{d+1} H_i$ and so we have shown that every set of $d + 1$ heavy halfspaces has a common intersection. We can therefore apply Helly's Theorem to the finite collection of heavy spaces to complete the proof.

□

An Algorithm? An important thing to gather about this proof is that it implies an algorithm. The desired point is one guaranteed by Helly's Theorem, which in turn is constructed by Radon points. The proof of Radon's theorem also implied an algorithm. Recall that all we did, was to compute an affine dependence and then partition by the sign of the coefficients. Computing the affine dependence is just solving a linear system, which can be done in time polynomial in d. Unfortunately, the algorithm as a whole is not efficient. This is because there could be many heavy halfspaces that we need to consider. We will talk about some more efficient algorithms for computing centerpoints soon.

The shape of centerpoints Another critical observation about the preceding proof is that the set of all centerpoints is the intersection of a finite set of halfspaces. This implies that the set of centerpoints is convex, and moreover, it forms a polytope.