Computing Small Hitting Sets for Convex Ranges.

Stefan Langerman* Mudassir Shabbir† William Steiger‡

November 5, 2012

Abstract

Let S be a set of n given points in R^2. If A is a convex subset of R^2 its "size" is defined as $|A \cap S|$, the number of points of S it contains. We describe an $O(n(\log n)^4)$ algorithm to find points $z_1 \neq z_2$, at least one of which must meet any convex set of size greater than $4n/7$; z_1 and z_2 comprise a hitting set of size two for such convex ranges. This algorithm can then be used to construct (i) three points, one of which must meet any convex set of size $> 8n/15$; (ii) four points, one of which must meet any convex set of size $> 16n/31$; (iii) five points, one of which must meet any convex set of size $> 20n/41$.

Let S be a set of n given points in general position in R^2. If A is a convex subset of R^2, its "size" is defined to be $|A \cap S|$, the number of points of S that it contains. The (Tukey) depth of a point $z \in R^2$ is defined as the minimum (over all halfspaces h containing z) of $|S \cap h|$, the size of the smallest halfspace containing z. It is familiar that there always exists a point $z \in R^2$ (z not necessarily in S) with depth $d(z) \geq n/3$. Such a point is called a *centerpoint* for S. The constant $c = 1/3$ is best-possible: for every $c > 1/3$ there are sets S with respect to which NO point has depth cn. The interesting algorithm of Jadhav and Mukhopadhyay [2] computes a centerpoint in linear time.

Alternatively, if z is a centerpoint for S, *every* convex set of size $> 2n/3$ MUST contain z. A centerpoint may thus be said to "hit" all convex subsets of R^2 with more than $2/3$ of the points of S. For this reason, centerpoint z is called a *hitting-set (of size* 1*)* for convex sets of size $> 2n/3$. Mustafa and Ray [4], following related work of Aronov et. al. [1], studied the possibilities for hitting sets with more than one point, a natural extension of the notion of centerpoint. They showed that given $S \subseteq R^2$ there are points $z_1 \neq z_2$ (not necessarily in S) such that every convex set of size $> 4n/7$ must meet at least one of them. In addition they showed the constant $4/7$ to be best possible for hitting sets of size 2: for every $c < 4/7$ there are sets S for which, whatever pair $x \neq y$ be chosen, there is a convex subset containing $> cn$ points of S, but containing *neither* x nor y). In [1] it had been shown that the optimal constant c was in the interval $[5/9, 5/8]$.

Let $c_k \in (0, 1)$ be the smallest constant for which, given any set S of n points in R^2, there are distinct points $z_1, \ldots, z_k$, at least one of which must meet any convex set of size $> c_k n$. We know $c_1 = 2/3$ and $c_2 = 4/7$. Mustafa and Ray were able to show that $c_3 \in (5/11, 8/15]$, that $c_4 \leq 16/31$ and that $c_5 \leq 20/41$.

*Directeur de Recherches du F.R.S.-FNRS. Département d'Informatique, Université Libre de Bruxelles.

†Department of Computer Science, Rutgers University, 110 Frelinghuysen Road, Piscataway, New Jersey 08854-8004. Work supported by grant 0944081, National Science Foundation, USA.

‡Department of Computer Science, Rutgers University, 110 Frelinghuysen Road, Piscataway, New Jersey 08854-8004. Work supported by grant 0944081, National Science Foundation, USA.

Here we address some algorithmic questions about *finding* small hitting sets. The details are contained in the following statement, and its proof.

Theorem 1 *Let S be a set of n given points in general position in R^2 and take $c_2 = 4/7$. Then in $O(n(\log n)^4)$, distinct points z_1, z_2 may be found so that if A is a convex set of size $> c_2 n$, at least one of these points is in A.*

Consider the set $\mathcal{R}$ of all convex subsets of size $> c_2 n$. For each pair $A \neq B$ in $\mathcal{R}$ consider $A \cap B$. Note that $|A \cap B| > n/7$, so there is a point $p_{A,B} = (u, v) \in A \cap B$ of minimal $y-$coordinate. The existence proof in [4] showed that z_1 may be taken as such a point, but one for a pair A', B' where $p_{A',B'} = (u, v)$ has v as large as possible (a point in the intersection of two ranges whose lowest point is highest). They also showed that z_2 may then be taken as the (usual) centerpoint for $S \backslash (A' \cap B')$ and everything works out.

Let $p = (u, v)$ be the lowest point in $A' \cap B'$ - the intersection of two ranges, each of size at least $c_2 n$ - where v is as large as possible (its a highest lowest point). The proof of the theorem relies on understanding what such a point looks like in the line arrangement dual to S. We combine this with tools introduced by Matoušek [3] to compute z_1 in the stated complexity. Once we have z_1, z_2 - the centerpoint of $S \backslash (A' \cap B')$ - can be found in linear time.

It is easy now to show

Corollary 1 *As in Theorem 1, in $O(n(\log n)^4)$ we can find (i) distinct points z_1, z_2, z_3, one of which must meet any convex set of size $> 8n/15$; (ii) points z_1', z_2', z_3', z_4', one of which must meet any convex set of size $> 16n/31$; (iii) points $z_1'', z_2'', z_3'', z_4'', z_5''$, one of which must meet any convex set of size $> 20n/41$.*

References

[1] B. Aronov, F. Aurenhammer, F. Hurtado, S. Langerman, D. Rappaport, S. Smorodinsky, and C. Seara. "Small Weak-epsilon Nets. *Computational Geometry: Theory and Applications, 42(2)* 455-462, (2009).

[2] S. Jadhav and S. Mukhopadhyay. "Computing a Centerpoint of a Finite Planar Set of Points in Linear Time". *Discrete and Computational Geometry 12(1)* (1994), 291-312.

[3] 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.

[4] N. Mustafa and S. Ray. "An Optimal Extension of the Centerpoint Theorem." *Computational Geometry: Theory and Applications 42(7)*, 505-510, (2009).