Discrete and Polyhedral Geometry, SS 2011 — exercise sheet 1

due date: Tuesday, April 19th, 2011, 14:00
(with CORRECTED version of exercise 2a)

---

1. (10 credits)

   Show that for any n and m, $5n \leq m \leq \binom{n}{2}$, there is a graph with n vertices and m edges that can be drawn in the plane with $O(m^3/n^2)$ crossings.

2. (10 credits)

   (a) Prove the following version of the Lovász lemma in the planar case: For a set $X \subset \mathbb{R}^2$ in general position, every vertical line ℓ intersects the interiors of at most $2(k+1)$ of the k-edges. More precisely, the number of *directed* k-edges crossing ℓ from left to right is at most $k+1$.

   (b) Using (a), prove that the number of k-sets is $O(n\sqrt{k+1})$.

3. (0 credits)

   (Exact planar Lovász lemma) Let $X \subset \mathbb{R}^2$ be a $2n$-point set in general position, and let ℓ be a vertical line having k points of X on the left and $2n-k$ points on the right. Prove that ℓ crosses exactly $\min(k, 2n-k)$ halving edges of X.

4. (10 credits)

   This exercise shows limits for what can be proved about k-sets using the Lovász lemma alone.

   Construct an n-point set $X \subset \mathbb{R}^2$ and a collection of $\Omega(n^{3/2})$ segments with endpoints in X such that no line intersects more than $O(n)$ of these segments.

5. (0 credits)

   (Duality). A nonvertical hyperplane h can be uniquely written in the form

   $$h = \{\, x \in \mathbb{R}^d : x_d = a_1 x_1 + \cdots + a_{d-1} x_{d-1} - a_d \,\}.$$

   We set $\mathcal{D}(h) = (a_1, \ldots, a_{d-1}, a_d)$. Conversely, the point $a = (a_1, \ldots, a_{d-1}, a_d)$ maps back to the hyperplane $\mathcal{D}(a) := h$.

   Show the following statements:

   (a) $p \in h \Leftrightarrow \mathcal{D}(h) \in \mathcal{D}(p)$.

   (b) A point p lies above a hyperplane h if and anly if the point $\mathcal{D}(h)$ lies above the hyperplane $\mathcal{D}(p)$.

6. Closest pairs (0 credits)

   The closest pair graph for a set of points is the graph in which the edges represent all pairs that have the minimum pairwise distance

   (a) Show that the closest pair graph for n points in the plane is non-crossing, when drawn with straight segments.

   (b) What is the maximum possible vertex degree in this graph?

   (c) Show that there are at most $3n$ closest pairs.

7. Largest distances (0 credits)

   The farthest pair graph is defined analogously.

   (a) Show that no two edges in the farthest pair graph are disjoint: two edges must either share a vertex or cross.

   (b) Show that there are at most $O(n)$ farthest pairs.

8. Least-squares clustering (0 credits)

   We want to partition a set of $2n$ points P into two clusters $C_1 \cup C_2$ of n points, minimizing the sum of the squared intra-cluster distances

$$\sum_{p,q\in C_1} \|p-q\|^2 + \sum_{p,q\in C_2} \|p-q\|^2.$$

   The following exercise helps to show that the two clusters C_1 and C_2 can be separated by a hyperplane. (Can you find a different or simpler proof?)

   (a) Show that the sum of all pairwise squared Euclidean distances can be written as the sum of squared Euclidean distances from the centroid

$$\bar{p} = \frac{1}{|C|}\sum_{q\in C} q,$$

   up to a constant factor:

$$\sum_{p,q\in C} \|p-q\|^2 = \sum_{p\in C} \|p-\bar{p}\|^2 \cdot \text{const}$$

   What is the "constant" factor?

   (b) Show that the point x that minimizes $\sum_{p\in C} \|p-x\|^2$ is the centroid $x = \bar{p}$.

   (c) Show that the following problem $(*)$ is equivalent to the original problem. Find a partition $P = C_1 \cup C_2$ into two equal-size sets together with two "cluster centers" x_1, x_2 such that

$$\sum_{p\in C_1} \|p-x_1\|^2 + \sum_{p\in C_2} \|p-x_2\|^2$$

   is minimized.

   (d) Show that a partition $P = C_1 \cup C_2$ in which C_1 and C_2 are not separated by a hyperplane perpendicular to the line connecting the two cluster centroids $\bar{p}_1$ and $\bar{p}_2$ cannot be optimal for $(*)$, and hence for the original problem.

   (e) * Does the separability statement remain valid if we minimize the sum of distances (instead of squared distances)?

   (f) * Does the separability statement remain valid if we partition n points into two clusters of different given sizes n_1 and n_2, with $n_1 + n_2 = n$?