Assignment 4 - Geometric Optimization (0368-4144)

Due: January 14, 2014

Problem 1

Let $P = \{p_1, \ldots, p_n\}$ be a set of n points in the plane.

(a) Solve the 2-mean problem for P: Find two points c_1, c_2, such that

$$\sum_{i=1}^{n} \min\{\|p_i - c_1\|^2,\ \|p_i - c_2\|^2\}$$

is minimized.

(b) Solve (a simpler variant of) the 2-line mean problem for P: Find two lines ℓ_1, ℓ_2, such that

$$\sum_{i=1}^{n} \min\{d^2(p_i, \ell_1),\ d^2(p_i, \ell_2)\}$$

is minimized, where $d(p, \ell)$ is the *vertical* distance from point p to line ℓ (that is, if $p = (\xi, \eta)$ and $\ell :\ y = ax + b$ then $d(p, \ell) = |\eta - a\xi - b|$).

The first algorithm should run in nearly quadratic time, and the second one in nearly cubic time. (**Hint:** First recall / show how to solve the 1-mean and (this version of) the 1-line mean problem. Then try to partition P into two subsets and to apply the 1-mean or the 1-line mean solution to each subset.)

Problem 2

Clustering by partitioning. Let P be a set of n points in the plane. We want to partition P into two sets P_1, P_2, so as to minimize some objective function. In this exercise, the function is the sum of the areas of the smallest axis-parallel squares that enclose P_1 and P_2.

Show that there is always an optimal solution in which P_1 and P_2 are separated from each other by a line. Using this, and duality, derive an efficient, near-quadratic algorithm for this problem.

Problem 3

Complete the details of the extension of the algorithm for the exact k-center problem that we studied in class, to the L_∞-distance in three dimensions, so that it runs in time $O\left(n^{ck^{2/3}}\right)$, for some constant c.

Problem 4

The discrete 2-center problem. Let P be a set of n points in the plane. We want to find two congruent disks of smallest radius, each *centered at some point of* P, whose union covers P. Give a near-quadratic algorithm for this problem. (**Hints:** (a) Show that the discrete 1-center problem can be solved in near-linear time. (b) Solve the decision problem in near-quadratic time. (c) What are the critical radii? Show that there is no need for real parametric searching.)

Problem 5

(1) Let P be a set of n points in d dimensions, and let μ be the center of mass of P. Let c be any point in $\mathbb{R}^d$. Show that

$$\sum_{p \in P} d(p, \mu) \leq 2 \sum_{p \in P} d(p, c).$$

(2) Use (1) to give a factor-2 approximation for the 1-median problem for a set P of n points in any dimension. Can you suggest a way of using this idea to approximate the 2-median problem in any dimension?