CS519: Special Topics in Computational Geometry, Spring 2015

Homework 3 *

Problem 1. Recall, the C-approximate r-Nearest Neighbor data structure presented in class: a set of points $P \subseteq \mathbb{R}^k$ be stored using $S(n,k)$ space in $T(n,k)$ preprocessing time, to answer queries in $Q(n,k)$ time. For each query point $q \in \mathbb{R}^k$, the data structure guarantees the following properties.

1. If $P \cap B(q,r) \neq \emptyset$ then the data structure returns a point in $P \cap B(q,Cr)$.
2. If $P \cap B(q,Cr) = \emptyset$ then the data structure returns NONE.
3. If $P \cap B(q,r) = \emptyset$, but $P \cap B(q,Cr) \neq \emptyset$ then the data structure either returns a point in $P \cap (q,Cr)$, or it returns NONE.

A C-approximate Nearest Neighbor data structure, on the other hand, guarantees the following property for a query point q. Let p_0 be the closest point of P to q, and let r_0 be the distance between p_0 and q. The data structure returns a point in $B(q,Cr_0)$.

Describe an analyze a $2C$-approximate nearest neighbor data structure using C-approximate r-Nearest Neighbor data structures, for different values of r. (Hint: your time/space requirements can depend on $\Delta = d_{\max}/d_{\min}$, where $d_{\max}$ and $d_{\min}$ are the maximum and minimum distance between points of P, respectively.)

Problem 2. The Johnson-Lindenstrauss lemma tells us that in order to preserve pairwise distances between n vectors up to a $1 \pm \varepsilon$ factor, it is sufficient to project onto $k = O(\log(n)/\varepsilon^2)$ dimensions. What is the required target dimension if we only want to preserve 99% of the pairwise distances?

Problem 3. Let $g_1, \ldots, g_n$ be standard Gaussian random vectors in $\mathbb{R}^n$ (i.e., each coordinate is a standard univariate Gaussian). Show that they are almost orthogonal to each other with high probability and obtain a bound on the maximum inner product between any pair. What happens if you take n^2 vectors?

* Problems 2 and 3 are from notes on advanced algorithms by Moses Charikar. Here is the link https://www.cs.princeton.edu/courses/archive/spring13/cos521/hw3.pdf.