Seminar on Algorithms and Geometry 2014B – Problem Set 1

Robert Krauthgamer

April 6, 2014

We discussed in class an algorithm for $(1+\varepsilon)$-approximate Nearest Neighbor Search, by relying on the doubling dimension of the dataset.

1. Consider the metric space $M = \mathbb{R}^k$ with distances according to the ℓ_∞-norm, i.e., $d(x, y) = \|x - y\|_\infty$. Prove that $\mathrm{ddim}(M) \le O(k)$.

2. Let $k = \mathrm{ddim}(M)$ and define k' similarly using diameter instead of radius (covering by sets of half the diameter). Prove that $k' \le O(k)$.

3. Design a variant of the query algorithm, where instead of keeping a set of points Z_i, we keep only one point z_i, which is computed (iteratively) as the point closest to q inside the list $L_{z_{i+1},i+1}$. Show that your variant finds an $O(1)$-approximate NNS.

   For simplicity (avoiding handling special cases) you may assume that $OPT = d(q, S) \in [10, \mathrm{diam}(S)/10]$. If needed, change the constant in the definition of $L_{y,i}$.