Seminar on Algorithms and Geometry – Handout 7

Robert Krauthgamer

June 4, 2009

Today's topics

We study the distance estimation problem, particularly in ℓ_1, from a communication complexity perspective. We will see the following upper and lower bounds.

Theorem 1 (Kushilevitz, Ostrovsky, and Rabani, 2000)**.** *For every $\varepsilon > 0$ there is a randomized simultaneous protocol for estimating the ℓ_1-distance within factor $1+\varepsilon$ (in the decision version) using $O(1/\varepsilon^2)$ bits of communication.*

Theorem 2 (Woodruff, 2004)**.** *For every $\varepsilon > 0$, one-way distance estimation within factor $1+\varepsilon$ (in the decision version) requires communication $\Omega(1/\varepsilon^2)$.*

We will also discuss the connection between these bounds to:

- dimension reduction in ℓ_1 (weak analogue to Johnson-Lindenstrauss)
- Near Neighbor Search algorithms for ℓ_1

Research Directions. It follows that a metric M that admits a low-distortion embedding into ℓ_1 also admits a distance estimation protocol (with low approximation and communication). However, the general relation between these two is not clear.

It is also interesting to design NNS algorithms for metrics M that do not admit a "good" distance estimation protocol. One known example is ℓ_∞^d.

Reading material. For more details, see these papers. We will see in class a more elementary proof of Theorem 2 due to Jayram, Kumar, and Sivakumar, 2008. The course webpage will contain exact details and links for these references.

References

[JKS08] T.S. Jayram, R. Kumar, and D. Sivakumar. The One-Way Communication Complexity of Hamming Distance *Theory of Computing*, 4(1):129-135, 2008.

[KOR00] E. Kushilevitz, R. Ostrovsky, and Y. Rabani. Efficient search for approximate nearest neighbor in high dimensional spaces. *SIAM J. Comput.*, 30(2):457–474, 2000.

[Woo04] D. Woodruff. Optimal space lower bounds for all frequency moments. In *15th annual ACM-SIAM symposium on Discrete algorithms*, pages 167–175. SIAM, 2004.