Sublinear Time and Space Algorithms 2020B – Problem Set 3

Robert Krauthgamer

Due: June 22, 2020

General instructions: Please keep your answers short and easy to read. You can use results, calculations or notation seen in class without repeating them, unless asked explicitly to redo them.

1. Let $h', h'' \in \mathbb{R}^n$ be supported on disjoint sets $T', T'' \subset [n]$ respectively, and let the matrix $A \in \mathbb{R}^{m \times n}$ be $(|T'| + |T''|, \varepsilon_0)$-RIP. Show that

   $$|\langle Ah', Ah'' \rangle| \leq \varepsilon_0 \|h'\|_2 \|h''\|_2.$$

   Hint: Consider first h', h'' that have unit length, and apply the formula $\|u+v\|_2^2 - \|u-v\|_2^2 = 4\langle u, v\rangle$ to $u = Ah'$ and $v = Ah''$.

2. Let $y \in \mathbb{R}^n$ be the frequency vector of an input stream in the turnstile model (i.e., allowing insertions and deletions), and suppose its coordinates are integers in the range $[-n^2, n^2]$.

   Design a linear sketch that detects whether $|\operatorname{supp}(y)| = 1$ using storage requirement of $O(1)$ words (i.e., $O(\log n)$ bits), not counting storage of the algorithm's random coins. Its success probability should be at least $1 - 1/n$.

   Hint: Use a variant of the AMS sketch with large random coefficients.

3. Recall that in our ℓ_0-sampling algorithm, $h : [n] \to [\log n]$ is a hash function such that each $h(i)$ is distributed like

   $$\Pr[h(i) = l] = 2^{-l}, \qquad \forall l \in [\log n],$$

   and since these probabilities do not add to 1, in the remaining probability $h(i) = \text{nil}$.

   Assume now that $h(1), \ldots, h(n)$ are pairwise independent and show that for every $x \neq 0$, there is a level $l \in [\log n]$ for which

   $$\Pr\left[|\operatorname{supp}(y^{(l)})| = 1\right] = \Omega(1).$$

   (We proved this in class but assuming full independence.) Recall that $y^{(l)}$ is obtained from x by zeroing coordinates $i \notin h^{-1}(l)$. If needed, slightly modify the distribution, e.g., extend it to one more level by picking $l \in [1 + \log n]$.

   Remark: Unfortunately, pairwise independence is not enough to guarantee that the "unique surviving" coordinate is uniform.

Extra credit:

4. Show that for an arbitrary $x \in \mathbb{R}^n$, if some $\tilde{x}$ satisfies the ℓ_1/ℓ_2 guarantee

$$\|x - \tilde{x}\|_2 \leq \frac{O(1)}{\sqrt{k}} \|x_{tail(k)}\|_1$$

(e.g., as seen in class using an RIP measurement matrix), then $x^* = \tilde{x}_{top(k)}$ satisfies the ℓ_1/ℓ_1 guarantee

$$\|x - x^*\|_1 \leq O(1)\|x_{tail(k)}\|_1.$$

Hint: Let T be the top k coordinates of x, and $\tilde{T}$ the top k coordinates of $\tilde{x}$. Split the coordinates into $\tilde{T}$, $T \setminus \tilde{T}$, and the rest.