Randomized Algorithms 2025A – Problem Set 3

Robert Krauthgamer and Moni Naor

Due: Dec. 23, 2024

1. (a) Is there a deterministic version of the JL Lemma? This mean that for every d, ε, n there is a *single* matrix $L \in \mathbb{R}^{k\times d}$ for $k = \mathrm{poly}(\varepsilon^{-1}\log n)$, such that for all $x_1, \ldots, x_n \in \mathbb{R}^d$,

$$\forall i, j \in [n], \qquad \|Lx_i - Lx_j\|_2 \in (1 \pm \varepsilon)\|x_i - x_j\|_2.$$

(b) Pick your favorite LLM, ask it this question, and evaluate its solution in one paragraph, for instance which LLM is it, whether it got the question correctly (and/or answered other questions), and if its proof/argument is valid.

2. Given parameters d, ε, δ, consider a randomized linear mapping $L = G/\sqrt{k}$ where $G \in \mathbb{R}^{k\times d}$ is a random matrix of standard Gaussians for a suitable $k = O(\varepsilon^{-2}\log\frac{1}{\delta})$. Prove that

$$\forall 0 \neq v \in \mathbb{R}^d, \qquad \mathbb{E}\Big[\max\Big\{0, \tfrac{\|Lv\|}{\|v\|} - 1 - \varepsilon\Big\}\Big] \le \delta. \tag{1}$$

Hint: L is the JL construction seen in class but for general δ (instead of $\delta = 1/n^3$). Modify the claim seen in class to show that if Y has a chi-squared distribution with parameter k, then

$$\forall s \ge 2, \qquad \Pr[Y \ge sk] \le e^{-sk/100}.$$

Then split the possible values of $\frac{\|Lv\|}{\|v\|}$ into the ranges $[1+\varepsilon, 2]$ and $(2, \infty)$.

3. Given parameters d and ε, consider a randomized linear mapping $L = G/\sqrt{k}$ where $G \in \mathbb{R}^{k\times d}$ is a random matrix of standard Gaussians for a suitable $k = O(\varepsilon^{-2}\log\frac{1}{\varepsilon})$. Prove that for every n points $x_1, \ldots, x_n \in \mathbb{R}^d$,

$$\mathbb{E}\Big[\sum_{i,j\in[n]} \|Lx_i - Lx_j\|\Big] \in (1 \pm \varepsilon) \sum_{i,j\in[n]} \|x_i - x_j\|.$$

Notice that the target dimension d is independent of n.

Hint: You can use a slightly stronger version of question 2 above (it has a similar proof), where (1) is replaced by

$$\forall 0 \neq v \in \mathbb{R}^d, \qquad \mathbb{E}\Big[\max\Big\{0, \Big|\tfrac{\|Lv\|}{\|v\|} - 1\Big| - \varepsilon\Big\}\Big] \le \delta. \tag{2}$$

4. Given parameters d and ε, consider a suitable $k = O(\varepsilon^{-2} \log \frac{1}{\varepsilon})$ and a randomized linear mapping $L = G/\sqrt{k}$ where $G \in \mathbb{R}^{k \times d}$ is a random matrix of standard Gaussians.

(a) Let $x_1, \ldots, x_n \in \mathbb{R}^d$ and let $W := \sum_{i<j} \|x_i - x_j\|$ (viewed as total "edge weight"). Prove that with high probability (say at least 0.9),

$$\forall S \subseteq V, \qquad \sum_{i \in S, j \in \bar{S}} \|Lx_i - Lx_j\| \in (1 \pm \varepsilon) \sum_{i \in S, j \in \bar{S}} \|x_i - x_j\| \pm \varepsilon W. \tag{3}$$

(Informally, this shows that L preserves the value of every cut of these n points up to small multiplicative and additive errors.)

Hint: To bound the LHS from above, consider for each $\{i, j\}$ whether it belongs to

$$P^+ := \big\{ \{i, j\} : \ \|Lx_i - Lx_j\| > (1 + \varepsilon) \|x_i - x_j\| \big\}.$$

Again, you can use a slightly stronger version of question 2 above, where (1) is replaced by (2).

(b) Improve the guarantee (3) to have only additive error, i.e.,

$$\forall S \subseteq V, \qquad \sum_{i \in S, j \in \bar{S}} \|Lx_i - Lx_j\| \in \sum_{i \in S, j \in \bar{S}} \|x_i - x_j\| \pm \varepsilon W. \tag{4}$$

(c) Prove that (for every of $x_1, \ldots, x_n$) with high probability L preserves the Max-Cut value up to $(1 + \varepsilon)$-approximation, i.e.,

$$\max_{T \subset [n]} \sum_{i \in T, j \in \bar{T}} \|Lx_i - Lx_j\| \in (1 \pm \varepsilon) \max_{S \subset [n]} \sum_{i \in S, j \in \bar{S}} \|x_i - x_j\|.$$