# Randomized Algorithms 2023A – Problem Set 2

Robert Krauthgamer and Moni Naor

Due: Jan. 16, 2023

1. Given parameters $d, \varepsilon, n$, design a randomized linear map $L : \mathbb{R}^d \to \mathbb{R}^t$ into dimension $t = O(\varepsilon^{-2} \log n)$, such that for every $n$ points $x_1, \ldots, x_n \in \mathbb{R}^d$, with high probability (at least 0.9), $L$ preserves the area of every triangle $\{x_i, x_j, x_k\}$ within factor $1 \pm \varepsilon$.

   (a) Prove that every set $S \subset \mathbb{R}^d$ of $n$ points has a set $S' \subset \mathbb{R}^d$ of $n' = O(n^3)$ points, such that if a linear map $L$ preserves the area of every *right-angle* triangle in $S \cup S'$ within factor $1 \pm \varepsilon$, then $L$ also preserves the area of every triangle in $S$ within factor $1 \pm \varepsilon$.

   Hint: For every triangle, find a point that "breaks" that triangle into two right-angle triangles.

   (b) Show that when $L$ is randomized as above, for every set of $n + n'$ points in $\mathbb{R}^d$, with high probability, the area of every right-angle triangle $\{y, y', y''\}$ among these points is preserved within factor $1 \pm \varepsilon$.

   Hint: Denote the triangle's sidelengths by $v = y' - y$ and $w = y'' - y$, and write the area of the image triangle using their image sidelengths $Lv$ and $Lw$, e.g., $\frac{1}{2}\sqrt{\|Lv\|^2 \, \|Lw\|^2 - \langle Lv, Lw \rangle^2}$.

2. Given parameters $d, \varepsilon, \delta$, 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

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

   Hint: This $L$ is exactly the JL construction seen in class but for general $\delta$ (instead of $\delta = 1/n^3$). First, prove a weaker version of (1), without the absolute value, by modifying the claim seen in class to show that if $Y$ has a chi-squared distribution with parameter $k$, then

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

   Then bound the other case (i.e., where the absolute-value operator is replaced by negation). Finally, fully prove (1) by combining the two cases, e.g., using

   $$\forall z \in \mathbb{R}, \qquad \max\{0, |z| - \varepsilon\} \leq \max\{0, z - \varepsilon\} + \max\{0, -z - \varepsilon\}.$$

   Remark: You can use the following bound for deviation below the expectation (it is similar to what was seen in class, no need to prove it):

   $$\forall \varepsilon \in (0, 1), \qquad \Pr[Y \geq (1 - \varepsilon)^2 k] \leq e^{-\varepsilon^2 k/2}.$$

3. Given parameters $d$ and $\varepsilon$, 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 is independent of $n$.

   Hint: Use the preceding question.

**Extra credit:**

4. Let $\|A\|$ be the spectral norm (i.e., largest singular value), which is also the operator norm, i.e., $\|A\| = \sup\{\|Ax\| : \ \|x\| = 1\} = \sup\{y^\top Ax : \ \|x\| = \|y\| = 1\}$.

   Let $S \in \mathbb{R}^{s\times n}$ be an $(\varepsilon, \delta, d)$-OSE matrix. Prove that for every $A, B \in \mathbb{R}^{n\times m}$ where $\text{rank}(A) + \text{rank}(B) \leq d$, with probability at least $1 - \delta$,

   $$\|(SA)^\top (SB) - A^\top B\| \leq O(\varepsilon) \cdot \|A\|\|B\|.$$

   Hint: Assume WLOG that $\|A\| = \|B\| = 1$, consider an orthonormal basis $U$ for the column space of $\{A, B\}$, and use the bound seen in class for $\langle Sx, Sy\rangle - \langle x, y\rangle$.