Problem Set 4 CS265, Autumn 2023

Due: Friday 10/27 at 11:59pm on Gradescope

Please follow the homework policies on the course website.

---

1. **(4 pt.)** Prove that $(\mathbb{R}^3, \ell_2)$ cannot be embedded into $(\mathbb{R}^2, \ell_2)$ with bounded distortion. In other words, there are no functions $f : \mathbb{R}^3 \to \mathbb{R}^2$ and constants $\alpha, \beta > 0$ such that the following inequality holds for all $x, y \in \mathbb{R}^3$:

$$\beta\|x - y\|_2 \leq \|f(x) - f(y)\|_2 \leq \alpha\beta\|x - y\|_2.$$

[**HINT:** *Try a proof by contradiction. How should the grid* $G_n := \{(i, j, k) : i, j, k \in \{0, 1, \ldots, n\}\}$ *be embedded? Try to pin down the intuition that the embedding of the grid would need to have lots of points fairly close together—within a smallish circle—but each point should not be too close to any other point, and then derive a contradiction from the fact that there just isn't enough area to fit all those points without some being too close....*]

SOLUTION:

2. **(4 pt.)** We showed that Bourgain's embedding allows us to embed an arbitrary metric space (X, d) with $|X| = n$ into $(\mathbb{R}^k, \ell_1)$ with target dimension k being $O((\log n)^2)$ and distortion being $O(\log n)$. Moreover, the embedding can be computed efficiently using a randomized algorithm. Prove that the exact same embedding computed by the randomized algorithm also achieves $O(\log n)$ distortion with high probability when the target metric is ℓ_2. [This actually holds for any ℓ_p metric for any $p \geq 1$, but this problem just asks you to prove it for ℓ_2]. We encourage you to emphasize only the differences from the proof in the lecture notes rather than copying the entire proof.
[**HINT:** *Let* $f : X \to \mathbb{R}^k$ *denote the relevant embedding. For any two points* $x, y \in X$, *we showed that* $\|f(x) - f(y)\|_1 \leq k \cdot d(x, y)$. *Can we say something similar about* $\|f(x) - f(y)\|_2$*?*]
[**HINT:** *For any two points* $a, b \in \mathbb{R}^k$ *it holds that* $\|a - b\|_2 \geq \frac{1}{\sqrt{k}}\|a - b\|_1$. *This is a special case of Hölder's inequality.*]

SOLUTION:

3. **(11 pt.) Johnson-Lindenstrauss with ± 1 entries:** In the lecture notes and videos we showed that a matrix of standard Gaussians can be used to get a dimension reducing map with very little distortion. However, a matrix of arbitrary real numbers can be cumbersome to store and compute with. In this problem you'll show that you can get essentially the same guarantees using random matrices with ± 1 entries. Throughout this problem, let A be an

$m \times d$ matrix who's entries are independently set to $+1$ with probability $1/2$ and otherwise to -1, and $z \in \mathrm{R}^d$ be an arbitrary unit vector.[1]

In this problem, you can use the statements from previous subparts even if you do not successfully prove them.

(a) **(2 pt.)** Show that $\mathbb{E}[\|Az\|_2^2] = m$.

(b) **(2 pt.)** For $Y \sim N(0,1)$, show that for any even $k \geq 0$, $\mathbb{E}[Y^k] \geq 1$, and for odd $k \geq 0$, $\mathbb{E}[Y^k] = 0$.
[**HINT:** *There are many solutions to this. Try to find a short one!*]

(c) **(2 pt.)** Prove that for any independent $X_1, \ldots, X_n$ and independent $Y_1, \ldots, Y_n$, if, for all integers $k \geq 0$ and $i = 1, \ldots, n$,

$$0 \leq \mathbb{E}[(X_i)^k] \leq \mathbb{E}[(Y_i)^k]$$

then for all integers $p \geq 0$,

$$\mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^p\right] \leq \mathbb{E}\left[\left(\sum_{i=1}^n Y_i\right)^p\right]$$

(d) **(4 pt.)** Let B be an $m \times d$ matrix who entries are independently drawn from $N(0,1)$. Prove that, for any $t \geq 0$ and unit vector z, if $\mathbb{E}[e^{t\|Bz\|_2^2}]$ is finite[2], then

$$\mathbb{E}[e^{t\|Az\|_2^2}] \leq \mathbb{E}[e^{t\|Bz\|_2^2}]$$

[**HINT:** *For any random variable X, $\mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbb{E}[X^k]$*]

(e) **(1 pt.)** Show that, for any $\epsilon \in (0,1]$,

$$\Pr[\|Az\|_2^2 \geq m(1+\epsilon)] \leq e^{-\Omega(m\epsilon^2)}.$$

If your proof is similar to that of Theorem 1 in lecture notes 8, we encourage you to emphasize only the differences from the proof in the lecture notes rather than copying the entire proof.

(f) **(0 pt.) [Optional: this won't be graded.]** Show that, for any $\epsilon \in (0,1]$,

$$\Pr[\|Az\|_2^2 \leq m(1-\epsilon)] \leq e^{-\Omega(m\epsilon^2)}.$$

[**HINT:** *We recommend you first show that for any* independent *and* nonnegative *random variables $X_1, \ldots, X_m$, defining $S = \sum_{i=1}^m X_i$, the probability $S \leq \mathbb{E}[S] - \Delta$ is at most* $\exp(-\Omega(\Delta^2 / \sum_{i=1}^m \mathbb{E}[X_i^2]))$. *To do so, use the inequality $e^{-v} \leq 1 - v + v^2/2$ which holds for any $v \geq 0$. Feel free to use the fact that for $Y \sim N(0,1)$, $\mathbb{E}[Y^4] = 3$.*]

---

[1] You may wonder why the proof from the lecture notes doesn't directly apply to ± 1 entries. This is because, when the entries are drawn from a normal distribution, we can use the rotational invariance of Gaussians to rotate z until it is a standard unit vector. That trick no longer applies if the entries are ± 1.

[2] For the purpose of your solutions, feel free to ignore this "is finite."

SOLUTION: