Problem Set 5

Electronic submission via Gradescope due **11:59pm Tuesday 10/29**. You are strongly encouraged to submit a homework with a partner—that is, submit one homework with both of your names.

[You may discuss these problems with classmates. Feel free to look at wikipedia, course notes, etc. for reference material, but do not try to specifically search online for solutions to the problems. ***Your submission must be the original work of you and your partner, and you must understand everything that is written on your submission.*** *We strongly suggest that you write solutions using LaTex—see the course website for a latex solution template.]*

Theorem 1. *(Johnson and Lindenstrauss, 84') Given any $\epsilon \in (0,1)$, and a set $X \subset \mathbb{R}^k$ with $|X| = n$, there exists a randomized linear map $f : \mathbb{R}^k \to \mathbb{R}^d$ with $d = O(\frac{\log n}{\epsilon^2})$ that embeds (X, ℓ_2) into $(\mathbb{R}^d, \ell_2)$ so that with high probability*

$$(1-\epsilon)\|x-y\|_2^2 \leq \|f(x)-f(y)\|_2^2 \leq (1+\epsilon)\|x-y\|_2^2$$

for any $x, y \in X$.

1. (6 points) Prove that the logarithmic dependence on the number of points, n, in the above theorem cannot be improved. Specifically, define a set of n points $X \subset \mathbb{R}^k$ and some constant c such that it is impossible to embed them into $(\mathbb{R}^{c \log n}, \ell_2)$ without incurring distortion at least 2. Note that you are free to choose the dimension k to be anything you want! [Hint: Feel free to assume that a sphere in d dimensions of radius r has volume in the interval $[r^d/(d/2)^{d/2}, 2^{d/2} r^d/(d/2)^{d/2}]$. More specifically, come up with a set of pairwise distances that can be instantiated in $k = n$ dimensional space, but has the property that if an embedding into d dimensional space exists, then the image of all the points must lie within a ball of radius 1 in order to avoid having distortion more than 2. Now, leverage the fact that a sphere in d dimensions (for even d) of radius r has volume $\frac{\pi^{d/2}}{(d/2)!} r^d \in [r^d/(d/2)^{d/2}, 2^{d/2} r^d/(d/2)^{d/2}]$, to argue that there simply is not enough room in $d = 0.001 \log n$ dimensions to accommodate the images of all the points without distorting some of their distances too much.]

2. In class we saw a very simple probabilistic argument that in any graph (V, E), the max-cut cuts at least $|E|/2$ edges. Here we consider two improvements/variants:

   (a) (4 points) Consider the greedy algorithm that, given an ordering $v_1, \ldots, v_n$ of the vertices, assigns v_1 to set A, then greedily partitions the other vertices (by sequentially assigning each unassigned vertex v to either A or B according to whether v has more neighbors already assigned to B or more neighbors already assigned to A.) Assume that ties are broken by assigning the point to set A. Prove that the cut found by this greedy algorithm cuts at least $\frac{|E|+|B|}{2}$ edges, where $|B|$ is the size of set B at the end of the algorithm. [Hint: think about what happens when a vertex actually gets put in set B.]

   (b) (4 points) Prove that if the graph has an even number of nodes, n, then there exists a cut with at least $\frac{|E|}{2}\frac{n}{n-1}$ edges crossing the cut. [Hint: the fact that that n is even is just meant to suggest that it can be divided into two equal sets....]

(c) **Double Bonus:** Can you turn the above probabilistic argument into an efficient deterministic algorithm? [This problem will not be graded...it is simply food for thought.]

3. (6 points) An *independent set* in a graph is a subset of vertices with the property that no pair of them are connected via an edge. Given a graph $G = (V, E)$ with $|V| = n$, prove that G has an independent set of size at least

$$\sum_{i=1}^{n} \frac{1}{d_i + 1},$$

where d_i is the degree of vertex v_i. [Hint: Given a permutation $\pi = v_{\pi(1)}, \ldots, v_{\pi(n)}$ of the vertices, define the corresponding (independent) set $S(\pi)$ by including vertex $v_{\pi(i)}$ if none of $v_{\pi(1)}, \ldots, v_{\pi(i-1)}$ is a neighbor of $v_{\pi(i)}$. Note that $S(\pi)$ is always an independent set. Now, define a randomized scheme for generating a permutation π, such that $\mathbf{E}[|S(\pi)|]$ is the claimed quantity above.]

4. Thursday night is halloween, and there is a student-faculty costume party mixer. Assume n people show up to the party: some of them are students, and the rest are faculty. Assume that each person has at least $1 + \log_2 n$ different costumes available to choose from (my set of costumes might or might not overlaps with your set). Note: its kindof funny if you show up with the same costume as your friends, but its horribly embarrassing if you have the same costume as one of the professors. Is it possible for all n people to coordinate their costumes in such a way that no student has the same costume as a professor? You will show, via the probabilistic method, that such a coordination is possible. hooray!

(a) (6 points) Recall that the probabilistic method is strictly BYOP (bring your own pumpkin, I mean probability distribution). Define a probabilistic scheme for choosing costumes, such that with positive probability, the chosen costumes satisfy the desired property (i.e. no student has the same costume as a professor). [Hint: try using a probability distribution over costumes that has some coordination between people. If each person selects one of their costumes uniformly at random from their set of $\geq 1 + \log_2 n$ costumes, then the resulting scheme does NOT seem to work.]