COS 521: Advanced Algorithm Design
Homework 1
Due: Thu, April 25

Collaboration Policy: You may collaborate with other students on these problems. Collaboration is limited to discussion of ideas only, and you should write up the solutions entirely *on your own* and *list your collaborators* as well as *cite any references* (book, paper, etc.) you may have used. Limit your answers for each problem to two pages or less (one page for most problems) – you need to give enough detail to convince the grader.

1. Consider a set of n elements with distances $d(i,j)$ defined for every pair of elements such that $d(i,j) = d(j,i)$. An embedding of these distances into ℓ_2 with distortion α is a mapping $f : \{1,\ldots,n\} \to R^d$ (d can be chosen to be as large as needed) such that

$$\forall i,j \in \{1,\ldots,n\}, \quad d(i,j) \le ||f(i) - f(j)||_2 \le \alpha \cdot d(i,j)$$

   Express the problem of finding the minimum distortion embedding into ℓ_2 as a semidefinite program of polynomial size.

2. A graph is k-colorable if every vertex can be assigned a label (color) in $\{1,\ldots,k\}$ such that adjacent vertices have distinct colors. Consider the following SDP relaxation for graph coloring: We have a unit vector v_i for every vertex i, with the following constraints:

$$\forall (i,j) \in E \quad v_i \cdot v_j = \frac{-1}{k-1} \tag{1}$$

   Note that there is no objective function for this SDP. We will be interested in whether there is a feasible solution that satisfies the SDP constraints.

   (a) Show that if a graph is k-colorable, then there is a feasible solution to this SDP.

   (b) Show that the SDP is not feasible for a clique of size k if the RHS of constraint (1) is replaced by a smaller quantity.

3. As with linear programs, semidefinite programs have duals. The dual of the MAX CUT SDP we looked at is

$$\frac{1}{2}\sum_{i<j} w_{ij} + \frac{1}{4}\min \sum_i \gamma_i$$

$$\text{subject to:} \qquad W + diag(\gamma) \text{ symmetric, positive semidefinite,}$$

where the matrix W is the symmetric matrix of the edge weights and the matrix $diag(\gamma)$ is the matrix with zeroes on the off-diagonal entries and γ_i as the ith entry on the diagonal. Show that the value of any feasible solution for this dual is an upper bound on the cost of any cut.

4. Given a graph $G(V, E)$, the *sparsest cut* problem is the problem of computing

$$\min_{S \subset V} \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

Consider the following vector program with a vector v_i for every $i \in V$.

$$\min \sum_{(i,j) \in E} (v_i - v_j)^2$$
$$\text{such that} \quad \sum_{i<j} (v_i - v_j)^2 = 1$$

(a) Show that the vector program is a relaxation for sparsest cut.

(b) Suppose that graph G is a d-regular graph. Show that the optimum value of the vector program is the second eigenvalue of the Laplacian of G (suitably scaled).

5. The Johnson-Lindenstrauss lemma tells us that in order to preserve pairwise distances between n vectors up to a $1 \pm \epsilon$ factor, it is sufficient to project onto $k = O(\log(n)/\epsilon^2)$ dimensions. What is the required target dimension if we only want to preserve 99% of the pairwise distances?

6. Let $g_1, \ldots, g_n$ be standard Gaussian random vectors in R^n (i.e., each coordinate is a standard univariate Gaussian). Show that they are almost orthogonal to each other with high probability and obtain a bound on the maximum inner product between any pair. What happens if you take n^2 vectors?