Homework 7

HDP KNU+ Fall 2022

Hints are in the back of this homework set.

As in the previous homework sets, $C, C_1, C_2, \dots$ and $c, c_1, c_2, \dots$ denote positive absolute constants of your choice.

Goemans-Williamson's semidefinite relaxation algorithm (October 10, Lecture 17) yields a 0.878-approximation of the max cut of a graph. In the first problem, you will check that a weaker 0.5-approximation can be achieved by a trivial algorithm: a random cut.

To recall the terminology, consider a graph $G = (V, E)$. For a given partition of the set of vertices V into two subsets V_1 and V_2, the *cut* $|E(V_1, V_2)|$ is defined as the number of edges that connect vertices from V_1 to vertices from V_2.

Problem 1 (A 0.5-approximation of max cut)

Describe a randomized algorithm which, for a given graph $G = (V, E)$, finds a partition $V = V_1 \cup V_2$ that satisfies

$$|E(V_1, V_2)| \geq \frac{1}{2} \cdot \max |E(U_1, U_2)|$$

where the maximum is over all partitions $V = U_1 \cup U_2$.

The *kernel method* in machine learning (October 14, Lecture 19) consists of replacing the inner product $\langle x, y \rangle$ on $\mathbb{R}^d$ with a suitable function $K(x, y)$. Not all functions are allowed, and Mercer's theorem (October 17, Lecture 20) gives a necessary and sufficient condition: K must be a *kernel*. Recall that a kernel is a function $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that, for any number $n \in \mathbb{N}$ and vectors $x_1, \dots, x_n$, the $n \times n$ matrix $\left[K(x_i, x_j)\right]_{i,j=1}^n$ is symmetric and positive semidefinite.

Mercer's condition is not always convenient to check. It might be simpler to build kernels from "building blocks" – simple kernels. Here you will verify some building rules.

Problem 2 (How to build a kernel)

Let $K(x, y)$ and, $M(x, y)$ be kernels. Show that all of the following are kernels, too:

(a) $aK(x, y) + bM(x, y)$, where $a, b > 0$ are constants;

(b) $K(x, y)^p$ where $p \in \mathbb{N}$;

(c) $P(K(x, y))$ where P is a polynomial with nonnegative coefficients;

(d) $K(x, y)\, f(x)\, f(y)$, where $f : \mathbb{R}^d \to \mathbb{R}$ is a function.

Here you will check a version of Grothendieck's identity (October 10, Lecture 17), which we will use in Problem 4.

Problem 3 (A version of Grothendieck's identity)

Let u, v be unit vectors in $\mathbb{R}^d$, and g be a standard normal random vector in $\mathbb{R}^d$, i.e.$g \sim N(0, I_d)$. Prove the following identities.

(a) $\mathbb{E}\langle u, g\rangle\langle v, g\rangle = \langle u, v\rangle$.

(b) $\mathbb{E}\langle u, g\rangle \operatorname{sign}(\langle v, g\rangle) = \sqrt{\frac{2}{\pi}}\langle u, v\rangle$.

(c) Consider the random variable $X_u = \langle u, g\rangle - \sqrt{\frac{\pi}{2}} \operatorname{sign}(\langle u, g\rangle)$, and similarly for X_v. Deduce from (a) and (b) that

$$\frac{\pi}{2}\,\mathbb{E} \operatorname{sign}(\langle u, g\rangle) \operatorname{sign}(\langle v, g\rangle) = \langle u, v\rangle + \mathbb{E}\left[X_u X_v\right]. \tag{1}$$

Here we improve Grothendieck's inequality for positive semidefinite matrices. The general Grothendieck's inequality (October 12, Lecture 18,) holds with constant 1.781. You will improve it to $\pi/2 \approx 1.571$ assuming the matrix (a_{ij}) is positive semidefinite:

$$\max_{u_i \in \mathbb{R}^d \text{ unit}} \sum_{i,j=1}^{n} a_{ij}\langle u_i, u_j\rangle \leq \frac{\pi}{2} \cdot \max_{x_i \in \{\pm 1\}} \sum_{i,j=1}^{n} a_{ij} x_i x_j. \tag{2}$$

More importantly, you will show how to convert a solution (u_i) to a semidefinite program (left hand side of (2); tractable) to an approximate solution (x_i) of the original problem (right-hand side of (2); NP-hard). Previously, we only achieved this for Goemans-Williamson's max-cut relaxation (October 10, Lecture 17) but not for the original problem (2).

The result you will prove was first established in the paper [1] of Alon and Naor from 2004. In just a month and a half of our course, you made it almost to the forefront of contemporary research! Congratulations, and keep doing the great work.

Problem 4 (A semidefinite Grothendieck)

Let $A = [a_{ij}]_{i,j=1}^n$ be a symmetric, positive semidefinite matrix, and let $u_1, \ldots, u_n$ be unit vectors in $\mathbb{R}^d$. Perform the randomized rounding of the vectors u_i, i.e. let $x_i = \operatorname{sign}(\langle g, u_i\rangle)$, where $g \in N(0, I_d)$. Using identity (1), show that

$$\mathbb{E}\left[\frac{\pi}{2} \cdot \sum_{i,j=1}^{n} a_{ij} x_i x_j\right] \geq \sum_{i,j=1}^{n} a_{ij}\langle u_i, u_j\rangle.$$

This immediately implies (2), since if expectation is large, it must be large for some realization of the random labels x_i.

TURN OVER FOR HINTS

Hints

Hint for Problem 1. Consider a *random cut*. For each vertex $v \in V$, flip a coin independently. If it comes up heads, include the vertex in the subset V_1, otherwise include it in the subset V_2. For a given pair of vertices $u, v \in V$, compute the probability of the event $D(u,v) = \{u, v \text{ land in different subsets}\}$. Argue that the cut equals $\sum_{(u,v)\in E} \mathbf{1}_{D(u,v)}$, where $\mathbf{1}_{D(u,v)}$ denotes the indicator random variable (it equals 1 is $D(u,v)$ holds and 0 otherwise). Then take expectation of the sum.

Hint for Problem 2. You may use without proof the standard facts about positive semidefinite matrices. Part (c) should follow from (a) and (b).

Hint for Problem 3. Both identities are proved in the paper [1, Section 5.1] (see the reference below). Don't copy the computations from that paper verbatim; look at them but then write down your argument yourself. You can use without proof the value of the first absolute moment of the standard normal distribution: $\mathbb{E}|g| = \sqrt{\frac{2}{\pi}}$.

Hint for Problem 4. The argument is sketched in the paper [1, Section 5.1] (see the reference below). But I think it would be easier for you to prove it yourself, as follows. Substitute $u = u_i$, $v = u_j$ into equality (1), multiply both sides by a_{ij}, and sum over all i, j. The sum in the right hand side breaks into two sums; check that the second one is nonnegative because the matrix (a_{ij}) is positive semidefinite.

References

[1] N. Alon, A. Naor, *Approximating the cut-norm via Grothendieck's inequality.* In Proceedings of the thirty-sixth annual ACM symposium on Theory of computing, pp. 72–80. 2004. Click to download the paper.