

Solution Sketch for Problem Set 3

1. (15 points) Local improvement algorithm for Max-Cut

Recall that in the Max-Cut problem, we are given an undirected graph $G = (V, E)$ and we would like to partition the vertices into two sets V_1 and V_2 so as to maximize the number of edges with one end-point in V_1 and the other end-point in V_2 . In this exercise, we consider a simple local improvement algorithm for the problem discussed in class.

Initially, partition the vertices arbitrarily. Consider the following local improvement step. Pick a node u , say in V_1 , that has fewer edges to vertices in V_2 than to other vertices in V_1 ; if such a node u exists, then move u from V_1 to V_2 .

The local improvement algorithm executes the above step as long as there exists a node that has fewer edges in the cut than edges to other vertices in its set.

Prove that the above algorithm is a polynomial-time 2-approximation algorithm for Max-Cut.

Answer: The above greedy algorithm continues to move vertices until for each vertex u the number of edges between u and vertices in its side of the partition is at most the number of edges between u and the vertices in the opposite side of the partition. In other words, at termination, for each vertex u , the number of edges adjacent to u in the cut (V_1, V_2) is at least $\delta(u)/2$, where $\delta(u)$ is the degree of u . Thus, if we add over all the vertices u , we obtain that the number of edges in the cut is at least $(\sum_u \delta(u)/2)/2 = |E|/2$. (The division by 2 is because each edge is counted twice in the summation, once for each endpoint.) Since the maximum cut is of size at most $|E|$, the greedy algorithm returns a 2-approximate cut on termination.

But does the algorithm always terminate? Yes. Since every iteration increases the cut size by 1, the algorithm terminates within at most $|E|$ iterations.

2. (15 points) Capacity of a binary erasure channel

Consider an erasure channel in which an input bit (0 or 1) is erased (output as '?') by the channel independently with probability p . What is the capacity of this channel?

Answer: Intuitively, the answer is $1 - p$. Let us derive this formally.

Let us consider a given input distribution with $q = \Pr[X = 0]$. Then, we compute the mutual information

$$I(X, Y) = H(X) - H(X|Y).$$

After some elementary algebra, we can calculate this to be

$$I(X, Y) = (1 - p)(q \log(1/q) + (1 - q) \log(1/(1 - q))).$$

This is maximized when $q = 1/2$, yielding a capacity of $1 - p$.

3. (15 points) Supermultiplicative and submultiplicative functions

Let f be a function mapping positive integers to positive reals. We say that f is supermultiplicative if $f(m+n) \geq f(m) \cdot f(n)$, and submultiplicative if $f(m+n) \leq f(m) \cdot f(n)$. Prove that if f is supermultiplicative, then

$$\sup_n f(n)^{1/n} = \lim_{n \rightarrow \infty} f(n)^{1/n}.$$

Prove that if f is submultiplicative, then

$$\inf_n f(n)^{1/n} = \lim_{n \rightarrow \infty} f(n)^{1/n}.$$

You may assume that $f(n)^{1/n}$ is bounded; that is, there exists a real M such that for any positive integer n , $f(n)^{1/n} \leq M$.

We used the above two facts for establishing that the fractional chromatic number of a graph is equal to the fractional clique number.

Answer: The above problem as stated implicitly assumes the existence of the limit. One can, in fact, prove that a limit exists, but we do not do this here.

Assuming the limit exists, we can establish the desired claim for the supermultiplicative case as follows. Let L denote the limit of the sequence $f(n)^{1/n}$. Then, for any $\varepsilon > 0$, there exists n_0 such that $|f(n)^{1/n} - L| < \varepsilon$ for all $n \geq n_0$. Clearly,

$$\sup_n f(n)^{1/n} \geq L,$$

since otherwise, we can find an $n \geq n_0$ such that $f(n)^{1/n} > \sup_n f(n)^{1/n}$. We now argue the other direction. Suppose

$$K = \sup_n f(n)^{1/n} > L.$$

Fix an $\varepsilon < (K - L)$. There exists an $n = n_0$ such that $f(n)^{1/n} \geq L + \varepsilon$, since otherwise K is not the supremum. By the supermultiplicativity condition, we have that

$$f(nk)^{1/(nk)} \geq f(n)^{1/n}, \text{ for all } k$$

Therefore, for infinitely many n , $f(n)^{1/n} \geq L + \varepsilon$, violating the assumption that L is the limit of $f(n)^{1/n}$.

The argument for submultiplicativity is similar.

4. (15 points) Embedding a hypercube in a sphere

Consider an embedding of the 2^k -point hypercube in the k -dimensional sphere. Show that any angle formed by three points of the hypercube is at most 90 degrees.

Answer: Take any three nodes u, v, w . Let us consider the angle θ between uv and vw . Without loss of generality set v to be the origin. Then, it is easy to check that $\cos \theta \geq 0$ since all points can be placed in the positive “quadrant” (all coordinates are positive).

5. (15 points) Random vector in high dimensions

Let r be a unit vector, chosen uniformly at random from an m -dimensional unit sphere. Prove the following two inequalities.

$$\Pr \left[|r \cdot e_z| < \frac{t}{\sqrt{m}} \right] \leq \Theta(t)$$

$$\Pr \left[|r \cdot e_z| > \frac{t}{\sqrt{m}} \right] \leq e^{-\Theta(t^2)}.$$

Answer: It can be shown using first principles that any coordinate of a point chosen at random from the surface of the unit sphere is distributed normally with mean 0 and standard deviation $1/\sqrt{m}$. (One of Knuth's units on the Art of Computer Programming has a proof, and I am sure you will find this in math texts or on the web.) So the random variable $r \cdot e_z$ is normally distributed with mean 0 and standard deviation $1/\sqrt{m}$. The two inequalities now follow easily.

$$\begin{aligned} \Pr \left[|r \cdot e_z| < \frac{t}{\sqrt{m}} \right] &= \int_{-t}^t \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^t e^{-x^2/2} dx \\ &= \Theta(t) \end{aligned}$$

since $e^{-x^2}/2$ is decreasing in $[0, 1]$ and has value 1 at $x = 0$.

$$\begin{aligned} \Pr \left[|r \cdot e_z| > \frac{t}{\sqrt{m}} \right] &= 2 \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &\leq 2 \int_t^\infty \frac{x}{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \frac{2}{\sqrt{2\pi}t} e^{-t^2/2} \\ &= e^{-\Theta(t^2)}. \end{aligned}$$

6. (25 points) Conditions in the Master Structure Theorem

The Master Structure Theorem (in Arora-Rao-Vazirani's sparsest cut result) says that given n points v_1, \dots, v_n , in \mathbf{R}^m such that (a) the average ℓ_2^2 distance among them is $\Omega(1)$, (b) the ℓ_2^2 distances form a metric, and (c) $\|v_i\| \leq 1$, there exist subset S and T of the points such that $|S| = \Omega(n)$, $|T| = \Omega(n)$ and $d(S, T) = \Omega(1/\sqrt{\log n})$.

Show that the desired claim does not hold if any of the three conditions is dropped.

Answer: If condition (a) is dropped, then we can put all points in the same position. If (c) is dropped, then place all but one points at the origin and the other point far away so that the average ℓ_2^2 distance is $\Omega(1)$; clearly, the desired claim does not hold.

Condition (b) is the tricky one. I was hoping one of you would solve it :-). One possible counterexample seems to be the following one. Consider n points chosen uniformly at random from the surface of a unit sphere of dimension $\log n$ – let us call this set X . In this case, we want to argue that (a) and (c) are true, yet (with high probability – even probability > 0 would suffice) there do not exist subsets S and T with the desired property.

Condition (c) is trivially true from our choice. For condition (a), consider any two points u and v . By part (a) of Problem 3 above, with constant probability each z -coordinate is at least $1/\sqrt{\log n}$ from 0. This yields that with constant probability the difference of their z -coordinates is $\Omega(1/\sqrt{\log n})$. So the expected ℓ_2^2 distance between them is $\Omega(1)$ (by summing over all coordinates). So the expected average distance is $\Omega(1)$. And this can be converted into a high-probability bound as well.

Now, we would like to argue that with some non-trivial probability, there do not exist subset S and T of X such that $|S| = \Omega(n)$, $|T| = \Omega(n)$ and $d(S, T) = \Omega(1/\sqrt{\log n})$. I do not have a proof for this; I am trying out one approach using a measure concentration lemma – Lemma 9 of ARV’s paper (full version available from Sanjeev Arora’s webpage). I will add more details if I am able to complete it, but don’t hold your breath on this :-).