Lecture 14: Sparsest cut and Leighton-Rao algorithm*

Rajat Mittal

IIT Kanpur

We will look at the problem of sparsest cut and an approximation algorithm (Leighton-Rao) based on the relaxation-rounding technique.

1 Sparsest cut

We looked at the edge-expansion problem in the last lecture and gave an algorithm for a regular graph. For a d-regular graph, remember that for a subset S of vertices V, its edge expansion is defined as,

$$\phi(S) = \frac{E(S, V-S)}{d|S|}.$$

To generalize it to all graphs, the correct notion of expansion would be,

$$\phi(S) = \frac{E(S, V-S)}{\min(E(S), E(V-S))}.$$

Here $E(S)$ denotes the total number of edges from subset S (which will be $d|S|$ in the regular case). Then, expansion for a graph is the minimum expansion for any non-empty subset S with at most half the vertices.

A related but slightly different quantity of interest is called sparsest cut. For a subset S,

$$\theta(S) = \frac{E(S, V-S)}{|S||V-S|}.$$

Similar to the edge expansion case, the sparsest cut is the minimum cut over all nonempty subsets with at most half the vertices.

$$\theta(G) = \min_{S \subset V : 1 \leq |S| \leq \frac{|V|}{2}} \theta(S).$$

The quantities, $\phi(G)$ and $\theta(G)$ are closely related in the case of d-regular graphs. Notice that $|V-S|$ is between $|V|$ and $|V|/2$. You can easily prove the following exercise.

Exercise 1. Show that for a d-regular graph,

$$d\phi(G) \leq |V|\theta(G) \leq 2d\phi(G).$$

So, for a d-regular graph, finding an approximation for edge-expansion or sparsest cut is similar. In these notes, we will see how to solve sparsest cut problem for a general graph approximately. Actually, we will solve a more general problem than the sparsest cut problem, known as the *non-uniform sparsest cut* problem.

1.1 Non-uniform sparsest cut

The non-uniform sparsest cut problem generalizes both, sparsest cut as well as min $s-t$ cut problem we looked at before.

Suppose we are given two graphs G and H instead of just one, the corresponding quantity of interest for a subset S is,

$$\Theta(S) = \frac{E_G(S, V-S)}{E_H(S, V-S)}.$$

* The content of these notes is largely taken from Luca Trevisan's course notes

Here, E_G denotes the edges in G and E_H in H.

The non-uniform sparsest cut is defined as,

$$\Theta(G) = \min_{S \subset V: 1 \leq |S| \leq \frac{|V|}{2}} \Theta(S).$$

You can check that uniform sparsest cut corresponds to the case when H is chosen to be a clique. You will prove in the assignment that min $s-t$ cut problem is also a special case of non-uniform sparsest cut.

Can we find an integer program to represent $\Theta(G)$? Again, we are interested in the subset S that minimizes $\Theta(S)$. Like before, we will keep variable x_i to denote if the i-th vertex is in S. Then, for any edge (i,j), $|x_i - x_j|$ is the indicator that (i,j) is in the set $E(S, v-S)$.

Hence,

$$E(S, V-S) = \sum_{(i,j) \in E} |x_i - x_j|.$$

This implies that $\Theta(G)$ can be written as the integer program,

$$\Theta(G) = \min_{x \in \{0,1\}^{|V|}, 1 \leq \sum_i x_i \leq |V|/2} \frac{\sum_{(i,j) \in E_G} |x_i - x_j|}{\sum_{(i,j) \in E_H} |x_i - x_j|}.$$

We already know that integer programs are hard to solve. So, we will approximate $\Theta(G)$ by relaxing the integer program. The relaxation follows by noting that a feasible vector x for $\Theta(G)$ defines a pseudo-metric over the vertices V.

2 Metric and Leighton-Rao relaxation

Before we look at the relaxation, we need to define metric/pseudo-metric over a set. A metric is an abstraction of the notion of distance between points.

Given a set S, a *metric* d is a function $d : S \times S \to \mathbb{R}_+$ satisfying the following conditions.

1. $d(x,y) \geq 0$ with equality iff $x = y$,
2. $d(x,y) = d(y,x)$,
3. $d(x,z) \leq d(x,y) + d(y,z)$.

The last condition is called *triangle inequality.*

Exercise 2. Does a feasible solution x of the integer program for $\Theta(G)$ define a metric over V with function $d(x_i, x_j) = |x_i - x_j|$?

You will realize that x is not a metric over V because *iff* in the first condition does not hold. Such functions are called *pseudo-metric.* A function from $S \times S$ to $\mathbb{R}_+$ is a pseudo-metric if it satisfies the following.

1. $d(x,y) \geq 0$ with equality if $x = y$,
2. $d(x,y) = d(y,x)$,
3. $d(x,z) \leq d(x,y) + d(y,z)$.

Notice the small difference in the first condition for a pseudo-metric.

Exercise 3. Prove, a feasible solution x of the integer program for $\Theta(G)$ defines a pseudo-metric over V with function $d(x_i, x_j) = |x_i - x_j|$?

Leighton-Rao relaxation for $\Theta(G)$ follows by relaxing $|x_i - x_j|$ to any arbitrary pseudo-metric.

$$LR(G) = \min_{d:\text{ d is a metric}} \frac{\sum_{(i,j) \in E_G} d(x_i, x_j)}{\sum_{(i,j) \in E_H} d(x_i, x_j)}.$$

Why is this relaxation easy to compute? It turns out that it can be written as a linear program. The relaxation is equivalent to,

$$\begin{aligned} \min \ & \frac{\sum_{(i,j)\in E_G} d(x_i,x_j)}{\sum_{(i,j)\in E_H} d(x_i,x_j)} \\ \text{s.t. } & d(x_i,x_j) \le d(x_i,x_k) + d(x_k,x_j) \quad \forall i,j,k \in V \\ & d(x_i,x_j) \ge 0. \end{aligned}$$

It is almost a linear program except the objective function. That can be taken care of by noting that scaling d does not change our program.

$$\begin{aligned} LR(G) = \min \ & \sum_{(i,j)\in E_G} d(x_i,x_j) \\ \text{s.t.} \ & \sum_{(i,j)\in E_H} d(x_i,x_j) = 1 \\ & d(x_i,x_j) \le d(x_i,x_k) + d(x_k,x_j) \quad \forall i,j,k \in V \\ & d(x_i,x_j) \ge 0. \end{aligned}$$

Our next task will be to show that $LR(G)$ is closely related to $\Theta(G)$.

Exercise 4. Show that $\Theta(G) \ge LR(G)$.

We will also give a $O(\log|V|)\cdot LR(G)$ upper bound on $\Theta(G)$ by rounding. This proof utilizes an intermediate relaxation between $LR(G)$ and $\theta(G)$ using something called an *L1 metric*.

First, we need to define an L1 metric.

An L1 metric d of dimension m on S is defined by a function $f : S \to \mathbb{R}^m$,

$$d(x,y) := \sum_{i\in[m]} |f(x)_i - f(y)_i|.$$

Here, $f(x)_i$ denotes the i-th co-ordinate of $f(x)$. The distance, $d(x,y)$, in case of L1 metric is also denoted by $\|f(x)-f(y)\|_1$ due to its connection with L1 norm.

Exercise 5. Show that an L1 metric is a pseudo-metric.

Call $LR1$ to be a relaxation of $\Theta(G)$ where only L1 metric are allowed.

$$LR1(G) = \min_{d\text{: d is an L1 metric}} \frac{\sum_{(i,j)\in E_G} d(x_i,x_j)}{\sum_{(i,j)\in E_H} d(x_i,x_j)}.$$

Since space of L1 metrices is a subset of all metrices, we know that $LR(G) \le LR1(G)$.

You can also check: a feasible solution x of the integer program for $\Theta(G)$ defines an L1 metric over V with dimension 1 by the function $f(i) = x_i$. This shows that the feasible space of $\Theta(G)$ is contained in the feasible space of $LR1(G)$, which is contained in the feasible space of $LR(G)$. So, we can extend the previous inequality,

$$LR(G) \le LR1(G) \le \Theta(G).$$

The rounding of $LR(G)$ solution to a solution of $\Theta(G)$ involves two steps.

- Convert any general metric to an *L1 metric* over V with a loss of $\log|V|$ factor (Bourgain's theorem).
- Show that the best L1 metric is the one used in $\Theta(G)$.

These two facts and the observation that $LR(G)$ is a linear program gives a $O(\log |V|)$ approximation algorithm for sparsest cut.

For the first fact, Bourgain's theorem provides a generic conversion of any pseudo-metric to an L1 metric without much loss. We will not cover the proof of Bourgain's theorem.

Theorem 1 (Bourgain). *Given a pseudo-metric d over a set S, we can find an $f : S \to \mathbb{R}^m$ for some m, such that,*

$$\|f(x) - f(y)\|_1 \le d(x, y) \le \|f(x) - f(y)\|_1 \cdot O(\log |V|).$$

Moreover, this mapping f can be found efficiently.

Bourgain's theorem implies that we can restrict our attention to L1 metric with a loss of $\log |V|$ factor.

Exercise 6. Show that,

$$LR1(G) \le LR(G) \cdot O(\log |V|).$$

We just need to show $LR1(G) = \Theta(G)$ to complete our proof.

3 L1 distance metric

We need to prove that we don't loose anything by relaxing to L1 metric. The proof will follow in two steps.

1. We can restrict L1 metric to be one dimensional.
2. For any one dimensional L1 metric, there always exist a feasible solution of $\Theta(G)$ which does better (or not worse).

3.1 L1 metric to one dimensional L1 metric

The first part follows from the observation that any m-dimensional L1 metric can be thought of as a convex combination of one dimensional L1 metric and vice versa. We focus on all possible convex combinations of one dimensional L1 metric as a feasible set. Since $LR(G)$ is a linear program, for any feasible set, optimum on the feasible set will be obtained at the vertices.

Exercise 7. Why is the optimum always at the vertices (hint: remember simplex method)?

The conversion from m-dimensional L1 metric to a convex combination of 1-dimensional L1 metric is natural. If $(f(x)_1, f(x)_2, \cdots, f(x_{|V|}))$ is the m-dimensional metric, consider m one dimensional metrices with $f_i(x) = f(x)_i$. Then the distance with respect to f is the sum of the distance of f_i's.

Since scaling the metric does not change the objective value, we can divide distance from f by m. Hence, distance from f/m is a convex combination of distance from $f_1, f_2, \cdots, f_m$.

For the reverse direction, if $f = \sum_i \alpha_i f_i$, we can take the m dimensional f to be $(\alpha_1 f_1, \cdots, \alpha_m f_m)$.

3.2 One dimensional L1 metric

We have a one-dimensional L1 metric. We need to show that the objective value is smaller than some feasible solution for $\Theta(G)$.

Again, since scaling and shifting of the function f of our metric does not change the feasibility and objective of our optimization program. We can assume that f is a function from V to $\{0, 1\}$.

Arrange the function values in the ascending order, $0 \le f(x_{i_1}), f(x_{i_2}), \cdots, f(x_{i_{|V|}}) \le 1$.

Let S_l, where l ranges from 1 to $|V|$, to be the set $\{i_1, i_2, \cdots, i_l\}$. Let x_l be the feasible solution corresponding to set S_l in the optimization program for $\Theta(G)$.

Then distance function of f is a convex combination of x_l's with coefficients $f(x_{i_l}) - f(x_{i_{l-1}})$. Again, by the properties of linear programming, the objective value at f is at least the minimum of objective value at x_l's.

This proves that generalizing the pseudo-metric in $\Theta(G)$ to be any one dimensional L1 metric does not change the objective value. From the previous section, $LR1(G) = \Theta(G)$.

4 Assignment

Exercise 8. Show that min $s - t$ cut is a special case of non-uniform sparsest cut problem.

Exercise 9. Prove that the Euclidean distance over $\mathbb{R}^n$ is a metric.

Exercise 10. Suppose we are given m L1 metrices of dimension 1; $f_1, \cdots, f_m$. Show that the distance function is same for $f = \sum_i \alpha_i f_i$ or the m-dimensional f with i-th coordinate being $\alpha_i f_i$.