Lecture Feb 03, 2020: Graph Sparsifiers

Lecturer: Gary L Miller   *Scribes: Zichao Dong*

1.1 Motivation

Suppose we come across the following problem:

   Input. $G = (V, E, W)$, with $|E| >> |V|$.

   Output. $H = (V, E', W')$, with $|E'| \approx |V|$.

In words, our problem is:

   *Given a **dense** weighted graph $G = (V, E, W)$, how to find a weighted graph $H = (V, E', W')$ on the same vertex set such that H is <u>similar</u> to G?*

By now, the problem is not specific at all, as we haven't figured out the meaning of *similar*. Here are two possible answers:

   Answer 1. We want the *similarity* to be *cut-preserving*, which is for every $S \subseteq V$,

$$\sum_{i \in S, j \notin S} w_{ij} \approx \sum_{i \in S, j \notin S} w'_{ij}.$$

   Answer 2. We want the *similarity* to be *spectral-preserving,*, which is

$$(1 - \epsilon)L_G \preceq L_H \preceq (1 + \epsilon)L_G.$$

Here L is for the Laplacian of a graph, and the notation $\preceq$ is from the following definition:

Definition 1.1 *(Loewner Order)*

For symmetric $n \times n$ matrices A and B, we say $A \preceq B$, if $B - A$ is positive semidefinite, i.e.

$$\forall x \in \mathbb{R}^n, \qquad x^T A x \le x^T B x.$$

Exercise. Show that spectral-preserving implies cut-preserving.

The exercise above implies we can focus on the study of **Answer 2** and forget about **Answer 1**.

1.2 Random Edge Sampling

The very basic idea is to take a random sampling of the edges of E with respect to G. If we can sample it well such that (E', W') is attained as the mean (expectation), and if we can also prove some kind of tight concentration, then we may claim that our random algorithm is successful.

Let's say we take each edge e_i from E with probability p_i (picking edges independently seems reasonable). Definitely, $\sum_{i=1}^{m} p_i = 1$ and $p_i \geq 0$, where $|E| = m$. Since $\mathbb{E}(\text{sample}) = (E, W)$ is wanted, we return $w_i' = \frac{w_i}{e_i}$ when e_i gets picked.

Suppose we have a bunch, say k, of samples. Then a natural thing to do is to take the average of the k Laplacians as the Laplacian of G'. Now our concern is how well G' can approximate G. Definitely, we need some concentration results, and the following Ahlswede-Winter theorem is the right thing:

Theorem 1.2 *(Ahlswede-Winter)*

Let $Y_1, Y_2, \cdots, Y_k$ be a sequence of i.i.d random variables with outcomes being a set Ω of symmetric semidefinite $n \times n$ matrices. If $\mathbb{E}(Y_i) = Z$ and $Y_i \preceq \mu Z$ for every i and every possible Y_i from Ω, then

$$\mathbb{P}\left((1 - \epsilon)Z \preceq \sum_{i=1}^{k} \frac{Y_i}{k} \preceq (1 + \epsilon)Z \right) \geq 1 - 2n e^{\frac{\epsilon^2 k}{4\mu}}.$$

We shall use E_{ij} for the matrix with a 1 at (i, j) entry and 0's everywhere else. Let's see some examples:

Example 1. Set $Z = I_n$ ($n \times n$ identity matrix). Pick $i \in \{1, 2, \cdots, n\}$ uniformly at random and return $Y = nE_i$, then

$$\mathbb{E}(Y) = E_1 + \cdots + E_n = I_n = Z.$$

It is easily seen that the best dominating constant $\mu = n$, as nI_n exactly dominates nE_i. Thus, by Ahlswede-Winter theorem,

$$\mathbb{P} := \mathbb{P}\left((1 - \epsilon)Z \preceq \sum_{i=1}^{k} \frac{Y_i}{k} \preceq (1 + \epsilon)Z \right) \geq 1 - 2n e^{\frac{\epsilon^2 k}{4n}}.$$

If we ask for $\mathbb{P} \geq 1 - \frac{2}{n}$, then $k = c'n \log n$ works, where $c' = \frac{8}{\epsilon^2}$. Thus, taking a series of $\sim \frac{n \log n}{\epsilon^2}$ samples is plausible.

Example 2. Set $Z = L_G$ for some graph G. Take $H_{ij} = H_e := E_{ii} + E_{jj} - E_{ij} - E_{ji}$, which is the unit Laplacian for a single edge. Then $L = \sum_{e \in E(G)} w_e H_e$ and we assume $w_e = 1$ hence the graph is unit-weighted.

Now let's try picking each edge uniformly at random and return mH_e as Y, where m is the number of edges in G. Then $\mathbb{E}(Y) = L_G$, by linearity.

Exercise. Show that $H_e \leq \mu L_G$ if and only if $\mu \geq ER_e$, here ER is for the effective resistance.

According to the exercise above, we see that $\mu = e$ suffices as we assumed G to be unit-weighted. Thus, by the same analysis as in **Example 1**, we conclude that $k = O_\epsilon(m \log m)$ works. Note that we generate our graph G' with $\Theta(m \log m)$ edges, which is reasonably sparse.

In fact, we see that **Example 2** is basically the general case, modulo the actual weights of our graph. For a weighted graph, a first observation is that picking edges uniformly at random is probably not good.

Suppose we choose edge e with probability p_e, then we return $Y = \frac{w_e}{p_e} H_e$ such that $\mathbb{E}(Y) = L_G$. Now we need μ such that $\frac{w_e}{p_e} H_e \preceq \mu L_G$, which is equivalent to

$$\mu \geq \frac{w_e ER_e}{p_e} \tag{*}$$

for every $e \in E$, by the **Exercise**.

Now the following result of Foster gives us the correct p_e to optimize μ:

Theorem 1.3 *(Foster)*

In a simple weighted graph $G = (V, E, W)$,

$$\sum_{e \in E} w_e ER_e = n - 1,$$

where w_e is the weight of edge e and ER_e is the effective resistance corresponding to edge e.

If G is a unit-weighted tree, then Foster's theorem is obviously true. The proof will be given later.

Due to Foster's theorem, we choose $p_e = \frac{w_e ER_e}{n-1}$ to optimize μ. Then $(*)$ shows $\mu \geq n - 1$, hence $\mu = n$ works for our purposes. Ahlswede-Winter theorem implies that an $O_\epsilon(n \log n)$-sampling gives us a good concentration, hence a likely graph sparsifier. We sum it up as below:

Theorem 1.4 *(Spectral Sparsifier)*

If H is obtained from G by sampling proportional to ER_e $O_\epsilon(n \log n)$ times, then with high probability,

$$(1 - \epsilon)L_G \preceq L_H \preceq (1 + \epsilon)L_G.$$

1.3   Questions

So far we have the following questions:

Question 1. Can we find spectral graph sparsifiers with linear many edges?

Question 2. Can we find $O(n \log n)$ spectral graph sparsifiers with a quick algorithm?

Question 3. Do we need the exact effective resistances for each edge?

Question 4. How to prove Ahlswede-Winter theorem?

Question 5. How to prove Foster's theorem?

In fact, the answers for **Question 1** and **Question 2** are both YES. But we will not go into the discussion for these topics in the class.

For **Question 4** and **Question 5**, we shall prove both of the theorems later.

Finally, we conclude this lecture by answering **Question 3**. The answer is NO.

Suppose $R_e \geq ER_e$ for every $e \in E(G)$, and set $t = \sum_{e \in E} w_e R_e$. Then our probabilities become $p'_e = \frac{w_e R_e}{t}$, and Foster's theorem implies $t \geq n - 1$. We pick e with probability p'_e, and generate $Y = \frac{w_e}{p'_e} H_e = \frac{t}{H_e} H_e$ if e get picked. Clearly, $\mathbb{E}(Y) = G$, and we need μ such that

$$\frac{t}{R_e} H_e \preceq \mu L_G$$

for every e. This is equivalent to

$$ER_e \leq \mu \frac{R_e}{t}$$

for every e, by **Exercise**, hence $\mu = t$ works for our purposes.

By Ahlswede-Winter theorem, we need $O_\epsilon(t \log t)$ samples for the concentration, which is reasonable as we replaced n by t. As long as R_e's are chosen closed to ER_e enough, we only lose a little bit.