College of Computer & Information Science — Spring 2016
Northeastern University — 26 February 2016
CS 7880: Network Algorithms — Scribe: Tanay Mehta

- Finish Cut Sparsifiers
- Spectral Sparsifiers

1 Cut Sparsifiers

Last time, we gave a method to probabilistically create a cut sparsifier H of a given graph $G = (V, E)$. More specifically, H has the same vertex set V and each edge $e \in E$ is sampled with probability p_e. We take $p_e \propto \frac{1}{\lambda_e}$ where λ_e is the connectivity of e (i.e. the size of the smallest cut containing e). Specifically,

$$p_e = \frac{c \ln^2 n}{\epsilon^2 \lambda_e}$$

If e is sampled, it is given weight $1/p_e$ in H. Last time we showed that the expected number of edges in H is $O\left(\frac{n \ln^2 n}{\epsilon^2}\right)$ where n is the number of vertices. Today, we will prove the following lemma.

Lemma 1. *For all cuts $S \subseteq V$, we have that $w(H(S, \bar{S})) = (1 \pm \epsilon)G(S, \bar{S})$ with high probability.*

w is the weight of the cut, i.e. the sum of weights of edges across the cut.

Proof. Let Δ be the size of the cut $(S, \bar{S})$.

Briefly recall the proof from last lecture for the graph generated for constant $p_e = p$. In that case, we had that

$$\mathbb{E}(H(S, \bar{S})) = p \cdot \Delta$$

We then applied Chernoff bounds to obtain that

$$\begin{aligned} \Pr[|H(S, \bar{S})| \le (1 \pm \epsilon) p\Delta] &= e^{\frac{-\epsilon^2 p \Delta}{3} \frac{c \ln^2 n}{\epsilon^2 \Delta^*}} \\ &= e^{-\frac{\Delta}{\Delta^*} \ln^2 n} \end{aligned}$$

In this case, we cannot follow the same proof since the probabilities are different for each edge. We will get around this issue by looking at families of edges with probabilities bounded in some range. Define

$$R_i := \{e \in (S, \bar{S}) \ : \ 2^i \le \lambda_e \le 2^{i+1}\}$$

Note that

$$|R_i| \cdot \frac{c \ln^2 n}{\epsilon^2 2^{i+1}} \le \mathbb{E}[\# \text{ of edges in } R_i \text{ selected}] \le |R_i| \cdot \frac{c \ln^2 n}{\epsilon^2 2^i}$$

We can apply Chernoff bounds on each family R_i, and take a union bound over all families. For $e \in R_i$ define $X_e = 1/p_e$, the weight of e. Define $X = \sum_{e \in R_i} X_e$. Then, we have that

$$\mathbb{E}[X] = |R_i|$$

We can see from this that in expectation, the weight of the cut in H is equal to the weight of the cut in G. To observe the behavior of the tail distribution of the cut, we proceed with a Chernoff bound on R_i. However, we need a modified Chernoff bound since the X_e are not Bernoulli random variables.

Chernoff bound:

$$\begin{aligned}
&\text{For } 0 \le X_i \le R, \text{ and } \mu = \mathbb{E}[X] \quad \text{where } X = \textstyle\sum X_i \\
&\Pr[X \le (1+\zeta)\mu] \quad \le \exp\left(\tfrac{-\zeta^2\mu}{3R}\right), \qquad \zeta < 1 \\
&\qquad\qquad\qquad\qquad\quad \le \exp\left(\tfrac{-\zeta\mu}{3R}\right), \qquad \zeta \ge 1 \\
&\Pr[X \ge (1-\zeta)\mu] \quad \le \exp\left(\tfrac{-\zeta\mu}{2R}\right), \qquad 0 \le \zeta \le 1
\end{aligned}$$

Applying this to our situation, we obtain

$$\Pr[X \ge (1+\epsilon)|R_i|] \le \exp\left(\frac{-\epsilon^2|R_i|}{3 \cdot 2^{i+1}\epsilon^2} \cdot c\ln^2 n\right)$$

where the value of R in the Chernoff formula is

$$R = \frac{2^{i+1}\epsilon^2}{c\ln^2 n}$$

Choose ζ such that $(1+\zeta)|R_i| = |R_i| + \frac{\epsilon\Delta}{\log n}$. Then, we have that

$$\begin{aligned}
\Pr[X \ge |R_i| + \frac{\epsilon\Delta}{\log n}] &\le \exp\left(\frac{-|R_i|c\ln^2 n}{3 \cdot 2^{i+1} \cdot \epsilon^2}\frac{\epsilon\Delta}{\log n|R_i|}\right) \\
&= \exp\left(-\Omega\left(\frac{\ln n \cdot \Delta}{2^i}\right)\right) \\
&= n^{-\Omega(\Delta/2^i)}
\end{aligned}$$

Let us take stock of the situation at this point. We have a fixed cut $(S, \bar{S})$ and a tail bound for the weight of edges selected from each family R_i. We want to generalize the tail bound for all edges selected from $(S, \bar{S})$. Take the union bound over all the R_i families to get the tail bound for all edges. To do this, we need to show that there aren't too many R_i families that comprise of the cut $(S, \bar{S})$. The following lemma (due to Fung et al) will give us that the number of possibilities for (S, R_i) where $|G(S)| \le \alpha \cdot 2^j$ is $\le n^{2\alpha}$.

Cut Projection Lemma: $\forall \alpha \ge 1$ we have that

$$|\{\delta(U) \cap B \ : \ |\delta(U)| \le \alpha \cdot \min_{e \in B} \lambda_e \text{ where } U \subseteq V, B \subseteq E\}| \le n^{2\alpha}$$

where $\delta(U) = \text{cut}(U, \bar{U})$.

In application of this lemma, B corresponds to the family of edges R_i and U corresponds to the cut S. Since $|G(S)| \le \alpha \cdot 2^j$, the union bound gives us

$$n^{-\Omega(2^j/2^i)} \cdot n^{2 \cdot \frac{2^i}{2^j}} \le \frac{1}{n^4},$$

$\square$

where we set the constant c hidden in the Ω-notation to be sufficiently large.

2 Spectral Sparsifiers

First, let us state an equivalent definition for the cut sparsifier requirement. For all cuts S,

$$\sum_{e \in (S, \bar{S})} w_e \in (1 \pm \epsilon)|\delta(S)|$$

where w_e is the weight of e in H. Building on this, here is another equivalent definition. For all cuts S, define the vector x such that for all vertices $u \in V$

$$x_u = \begin{cases} 1 & \text{if } u \in S \\ -1 & \text{if } u \in \bar{S} \end{cases}$$

Then, the cut sparsifier requirement is that

$$\sum_{e \in E} w_e (x_u - x_v)^2 \in (1 \pm \epsilon) \sum_{e \in E} (x_u - x_v)^2$$

where $e = (u, v)$. Note that the sum $\sum_{e \in E} (x_u - x_v)^2$ can be represented (after suppressing a 1/4 factor) as

$$\sum_{e \in E} (x_u - x_v)^2 = x^T L_G x$$

where $L_G = D - A$ is the Laplacian of G. We can represent the Laplacian through the following notation. Let e_i denote the unit vector with 1 in the i-th coordinate. Let $\chi_e = e_u - e_v$ where $e = (u, v)$. Then, we have that

$$L_G = \sum_{e \in G} \chi_e \chi_e^T$$

Then, the cut sparsification requirement is equivalent to the following. For all $x \in \{-1, 1\}^n$

$$(1 - \epsilon) x^T L_H x \ \leq \ x^T L_G x \ \leq \ (1 + \epsilon) x^T L_H x$$

For **spectral sparsifiers**, the above condition is required for all $x \in \mathbb{R}^n$.

Notation: For matrices M, N, we write

$$M \preceq N \iff \forall x \ : \ x^T M x \leq x^T N x$$

Note that $M \preceq N$ is equivalent to saying that $N - M \succeq 0$, or that $N - M$ is positive semi-definite.

In order to build spectral sparsifiers, we will take $p_e \propto R_e$, the effective resistance of e. We will review resistance and other relevant concepts in the next lecture.