COMPSCI 638: Graph Algorithms	October 30, 2019
Lecture 19	
Lecturer: Debmalya Panigrahi	*Scribe: Kevin Sun*

1 Overview

In the last lecture, we introduced the notion of effective resistances in a network. In this lecture, we give an algorithm that samples according to effective resistances for spectral sparsification, a generalization of graph sparsification that we saw in Lecture 9.

2 Spectral Sparsification via Effective Resistances

Let $G = (V, E)$ be an undirected graph on n vertices and m edges. Recall (from Lecture 9) that the goal of graph sparsification is to find a (edge-weighted) subgraph H on V such that the value of every cut in G is approximately preserved in H. In terms of the Laplacian matrix L_G, this is equivalent to preserving $x^\top L_G x$ for any $x \in \{0, 1\}^n$ (see Lecture 17).

In this lecture, we generalize this notion to spectral sparsification: given $G = (V, E)$ and some $\epsilon \in (0, 1)$, we want to construct a (sparse) subgraph H such that the following is true:

$$(1 - \epsilon)x^\top L_G x \le x^\top L_H x \le (1 + \epsilon)x^\top L_G x \quad \forall x \in \mathbb{R}^n. \tag{1}$$

If H satisfies the condition above with high probability, then H is a spectral sparsifier of G.

The algorithm: The overall strategy is similar to the scheme of Benczúr and Karger [BK15] (see Lecture 10), but instead of using strengths, Spielman and Srivastava [SS11] use effective resistances. Recall (from Lecture 18) that the effective resistance of an edge $e = (a, b) \in E$, denoted $R(e)$, can be thought of as the effective resistance between a and b given by the entire network. Intuitively, edges with higher effective resistance belong to sparser cuts, so p_e should be proportional to $R(e)$.

We now formally state the sampling procedure for constructing a spectral sparsifier. The sampling procedure runs for $q = O(n \log n/\epsilon^2)$ iterations. In each iteration, we sample every edge e with probability p_e, where $p_e = R(e)/\sum_e R(e)$ is proportional to $R(e)$. If e is sampled, we increase the weight of e in H by $1/qp_e$. Thus, if e is sampled x_e times, its final weight in H is x_e/qp_e.

Theorem 1 (Spielman and Srivastava [SS11])**.** *The sampling procedure described above produces a spectral sparsifier H containing $O(n \log n/\epsilon^2)$ edges.*

The outline of our proof is the following: we first reduce the problem to bounding the norm of a matrix. We then show how we construct this matrix. Finally, we apply a matrix concentration theorem due to Rudelson and Vershynin [RV07] to bound the norm of the matrix.

Recall (from Lecture 18) that $L_G = B^\top B$ where B is the signed edge-vertex adjacency matrix of G. To write L_H in a similar form, let $S \in \mathbb{R}^{m \times m}$ be a diagonal matrix defined as follows: $S_{e,e} = x_e/qp_e$, where x_e is the number of times e was sampled. The weight of e in H is $S_{e,e}$, so $L_H = B^\top SB$.

Now notice that proving (1) is equivalent to proving

$$\epsilon \geq \max_{x \in \mathbb{R}^n} \frac{\left|x^\top L_H x - x^\top L_G x\right|}{x^\top L_G x} = \max_{x \in \mathbb{R}^n} \frac{\left|x^\top B^\top S B x - x^\top B^\top B x\right|}{x^\top B^\top B x}.$$

We substitute $y = Bx \in \mathbb{R}^m$ in the above expression and let $im(B)$ denote the image (i.e., column space) of B. By scaling, we can assume $y^\top y = 1$, so our goal is to show

$$\epsilon \geq \max_{y \in im(B)} \frac{\left|y^\top S y - y^T y\right|}{y^\top y} = \max_{y \in im(B)} \left|y^\top (S - I) y\right|, \tag{2}$$

where I denotes the $m \times m$ identity matrix. So we are essentially bounding the *matrix norm* of $S - I$, but the constraint $y \in im(B)$ makes our task more challenging because we are only considering vectors in an n-dimensional subspace of $\mathbb{R}^m$.

The projection matrix: To address this issue, we will define a projection matrix $\Pi : \mathbb{R}^m \to \mathbb{R}^n$ that satisfies the following: if $y \in im(B)$ then $\Pi y = y$, and otherwise, $\Pi y \in im(B)$. The existence of such a matrix allows us to replace y with Πy in (2) and drop the $y \in im(B)$ constraint. Thus, our goal is to find such a Π and prove the following:

$$\left\|\Pi^\top S \Pi - \Pi^\top \Pi\right\|_2 \leq \epsilon. \tag{3}$$

To construct Π in two dimensions, we can project a point v onto a line ℓ by mapping it to its orthogonal projection w on ℓ; this point minimizes $\|v - w\|_2$. In general, we want to map v to a point $w = \Pi v$ such that the following condition is satisfied:

$$w = \Pi v = \arg\min_{x \in im(B)} ||v - x||_2.$$

It can be shown that if $w \in im(B)$ is defined as above, then $B^\top (v - w) = 0$, so $B^\top v = B^\top w$. Since $w = Bx$ for some x, we have $B^\top v = B^\top B x$, and solving for Bx, we get $w = Bx = B(B^\top B)^{-1} B^\top v$. Thus, our projection matrix is $\Pi = B(B^\top B)^{-1} B^\top$; notice Π satisfies $\Pi^\top = \Pi$ and $\Pi^2 = \Pi$.

Matrix concentration: Now that we have defined Π, we can now return our attention to proving (3). By the properties of Π stated above, we can see that

$$\left\|\Pi^\top S \Pi - \Pi^\top \Pi\right\|_2 = \|\Pi S \Pi - \Pi \Pi\|_2.$$

To minimize this quantity, we use the following matrix concentration theorem.

Theorem 2 (Rudelson and Vershynin [RV07]). *Let $y_1, \ldots, y_m \in \mathbb{R}^m$ be vectors that satisfy $\|y_i\|_2 \leq M$ for some $M \in \mathbb{R}$ and every i. Suppose we draw q independent samples, where y_i is drawn with probability p_i, to obtain $\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_q$. If $\mathbb{E}_p[yy^\top] = \sum_{i=1}^m p_i y_i y_i^\top$ satisfies $\left\|\mathbb{E}_p[yy^\top]\right\|_2 \leq 1$, then*

$$\mathbb{E}\left\|\frac{1}{q}\sum_{i=1}^q \tilde{y}_i \tilde{y}_i^\top - \mathbb{E}_p[yy^\top]\right\|_2 = O\left(M\sqrt{\frac{\log q}{q}}\right).$$

Define $y_e = \Pi_e / \sqrt{p_e}$ so that

$$\mathbb{E}_p[yy^\top] = \sum_{e \in E} p_e \frac{\Pi_e}{\sqrt{p_e}} \frac{\Pi_e^\top}{\sqrt{p_e}} = \Pi^2 = \Pi, \tag{4}$$

which implies

$$\left\| \mathbb{E}_p[yy^\top] \right\|_2 = \|\Pi\|_2 = 1.$$

Now suppose we sample the y_e vectors q times according to p_e to obtain $\tilde{y}_1, \ldots, \tilde{y}_q$. Then we have the following:

$$\frac{1}{q} \sum_{i=1}^{q} \tilde{y}_i \tilde{y}_i^\top = \frac{1}{q} \sum_{e \in E} x_e \frac{\Pi_e}{\sqrt{p_e}} \frac{\Pi_e^\top}{\sqrt{p_e}} = \sum_{e \in E} S_{e,e} \Pi_e \Pi_e^\top = \Pi S \Pi. \tag{5}$$

Thus, we can now bound $\mathbb{E}\|\Pi S \Pi - \Pi \Pi\|_2$ by applying Theorem 2 and using (5) and (4):

$$\mathbb{E}\|\Pi S \Pi - \Pi \Pi\|_2 = O\left(M \sqrt{\frac{\log q}{q}} \right) \tag{6}$$

for some M satisfying $\|y_i\|_2 \leq M$ for every i, that we shall now determine. By properties of Π, it can be shown that

$$\|y_e\|_2 = \frac{1}{\sqrt{p_e}} \|\Pi_e\|_2 = \frac{1}{\sqrt{p_e}} \sqrt{\Pi_{e,e}} = \frac{1}{\sqrt{p_e}} \sqrt{R(e)},$$

where $\Pi_{e,e}$ denotes the (e,e)-th entry of Π. In the algorithm, we set $p_e = R(e) / \sum_e R(e)$, so

$$\|y_e\|_2 = \sqrt{\sum_{e \in E} R(e)}.$$

It is known that if we sample a spanning tree of G uniformly at random, then the probability that e is in the tree is exactly $R(e)$. Thus, we can set $M = \sqrt{n-1}$. If we also set $q = cn \log n / \epsilon^2$ for a sufficiently large constant c, then (6) implies

$$\mathbb{E}\|\Pi S \Pi - \Pi \Pi\|_2 = O\left(\sqrt{n-1} \sqrt{\frac{\epsilon^2 \log q}{n \log n}} \right) \leq \frac{\epsilon}{2}.$$

By Markov's inequality, with probability at least one half, $\mathbb{E}\|\Pi S \Pi - \Pi \Pi\|_2 \leq \epsilon$, as desired. Note that this probability can be boosted by the standard trick of repeating the procedure multiple times and taking the median.

3 Summary

In this lecture, we saw how sampling by effective resistances yields a spectral sparsifier, which is a generalized version of cut sparsifiers that we saw in previous lectures. The proof reduces the sparsification condition to a claim about the size of the norm of a matrix and applies a matrix concentration bound to prove the claim.

References

[BK15] András A Benczúr and David R Karger. Randomized approximation schemes for cuts and flows in capacitated graphs. *SIAM Journal on Computing*, 44(2):290–319, 2015.

[RV07] Mark Rudelson and Roman Vershynin. Sampling from large matrices: An approach through geometric functional analysis. *Journal of the ACM (JACM)*, 54(4):21, 2007.

[SS11] Daniel A Spielman and Nikhil Srivastava. Graph sparsification by effective resistances. *SIAM Journal on Computing*, 40(6):1913–1926, 2011.