College of Computer & Information Science
Northeastern University
CS 7880: Network Algorithms

Spring 2016
01 March 2016
Scribe: Tanay Mehta

- Spectral Sparsifiers

1 Spectral Sparsifiers

We are given a graph $G = (V, E)$ and want to create a graph $H = (V, E')$ with weights w_e such that

- The number of edges with non-zero weight are small (near linear)
- $\forall x \in \mathbb{R}^n$ where $|V| = n$ we have

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

  where $L = D - A$ is the Laplacian and $L_H = \sum_{e=(i,j)} w_e (x_i - x_j)^2$.

We will use the same recipe for sampling edges from G to create H except take p_e proportional to the effective resistance. If we view the graph as an electrical network with edges having resistance equal to their weight (1 in an unweighted graph), then the effective resistance of an edge (u, v) is the resistance of that edge if 1 A of current enters u and exits v. Before we formally define effective resistance, let us give some notation and state some basic results.

Let e_u for the unit vector for $u \in V$. That is, e_u has a 1 in the u-th coordinate and 0 elsewhere. Then, define

$$\chi_e := e_u - e_v$$

where $e = (u, v)$. Let the incidence matrix χ be the $n \times m$ matrix with columns χ_e for all $e \in E$. Note that the Laplacian $L_G = \chi\chi^T = \sum_{e \in E} \chi_e \chi_e^T$. Furthermore, $L_H = \sum_{e \in E} w_e \chi_e \chi_e^T$.

Kirchoff's law: Let c_{ext} be the external current coming into the graph. This can be found by adding all currents in and subtracting all currents going out. Then, we have

$$\chi \cdot i = c_{\text{ext}}$$

where i is the $m \times 1$ current vector.

Ohm's law: Let v be the $n \times 1$ voltage vector. We then have the following relation: $v_a - v_b = i_{a \to b}$, the current from a to b. More generally,

$$\chi^T v = i$$

Note that since the all edges are 1 Ω resistors, the effective resistance is the potential difference between the two nodes of the edge.

We can combine Ohm and Kirchoff's laws to get the following

$$L_G v = \chi \chi^T v = c_{\text{ext}}$$

To calculate the effective resistance of the edge (x, y), set $c_{\text{ext}} = e_x - e_y$. This implies that $L_G v = e_x - e_y$. Solve for v to get

$$v = L_G^{-1}(e_x - e_y)$$

Since the effective resistance $R_{\text{eff}} = v_x - v_y$, we can sample with each edge with probability $p_{(x,y)} = R_{\text{eff}}$ to get

$$p_{(x,y)} = v_x - v_y = (e_x - e_y)^T \cdot v$$

Expanding out the v, we arrive at the following form.

$$p_{(x,y)} = (e_x - e_y)^T L_G^{-1}(e_x - e_y)$$

However, there is a serious problem with what we have just done. L_G is singular and therefore not invertible. This can be shown by verifying that 0 is an eigenvalue of L_G.

We can fix this issue by using the *pseudoinverse* of L_G. Let M be a real, symmetric matrix. By the spectral theorem,

$$M = \sum_{i=1}^{n} \lambda_i v_i v_i^T$$

where the v_i are normalized eigenvectors. Define the *pseudoinverse* M^+ to be

$$M^+ = \sum_{\lambda_i \neq 0} \frac{1}{\lambda_i} v_i v_i^T$$

Therefore, we have that for an edge e

$$p_e = \chi_e^T L_G^{-1} \chi_e$$

where we take L_G^{-1} to be L_G^+. This also gives us a well defined matrix $L_G^{-1/2}$ implying

$$\begin{aligned} p_e &= \chi_e^T L_G^{-1/2} L_G^{-1/2} \chi_e \\ &= (L_G^{-1/2}\chi_e)^T \cdot (L_G^{-1/2}\chi_e) \end{aligned}$$

Define $x_e = L_G^{-1/2}\chi_e$. Finally, we have the effective resistance of e as

$$p_e = x_e^T x_e$$

Suppose each edge e is selected with probability p_e. This implies that

$$\mathbb{E}[\# \text{ edges selected}] = \sum_e p_e = \sum_e x_e^T x_e$$

We can analyze $x_e^T x_e$ by looking at the following matrix of outer products, and taking its trace.

$$\begin{aligned} \sum_e x_e x_e^T &= \sum L_G^{-1/2} \cdot \chi_e \cdot \chi_e^T \cdot L_G^{-1/2} \\ &= L_G^{-1/2} (\sum_e \chi_e \chi_e^T) L_G^{-1/2} \\ &= L_G^{-1/2} \cdot L_G \cdot L_G^{-1/2} \\ &= I \end{aligned}$$

This is not entirely accurate, however. Remember that we are using the pseudoinverse of L_G. Let's continue and see where it doesn't work. Let us expand out the sum $\sum_e x_e x_e^T =$

$$\begin{array}{cccc} \begin{matrix} x_{11} & x_{11} & \dots & x_{1n} \\ \vdots & & & \\ x_{1n} & & & \end{matrix} & = & x_1 x_1^T = & \begin{pmatrix} x_{11}^2 & & \\ & \ddots & \\ & & x_{1n}^2 \end{pmatrix} \\ & + & + & \\ & \vdots & \vdots & \\ & + & + & \\ \begin{matrix} x_{n1} & x_{n1} & \dots & x_{nn} \\ \vdots & & & \\ x_{nn} & & & \end{matrix} & = & x_n x_n^T = & \begin{pmatrix} x_{n1}^2 & & \\ & \ddots & \\ & & x_{nn}^2 \end{pmatrix} \end{array}$$

This implies that

$$\begin{aligned}\sum_e p_e &= \mathrm{Tr}\left(\sum_e x_e x_e^T\right)\\ &= n-1\end{aligned}$$

where the -1 appears because we used the pseudoinverse. The number of non-zero eigenvalues is $n-1$ assuming a connected graph.

Random Sampling Algorithm:

$\rho \leftarrow \frac{c \cdot \ln n}{\epsilon^2}$
$\forall e:\ w_e \leftarrow 0$
$\forall e:\ p_e \leftarrow x_e^T x_e; x_e = L_G^{-1/2}\chi_e$
for $j \leftarrow 1, \ldots, \rho$ **do**
  for e **do**
    $Z_{j,e} := \begin{cases} 1 \text{ w.p. } p_e \\ 0 \text{ otherwise}\end{cases}$
  end for
  for e **do**
    $w_e \leftarrow w_e + \frac{Z_{j,e}}{\rho p_e}$
  end for
end for

The largest for-loop runs ρ times in order to boost the probability that the graph is connected. Note that the expected weight of e

$$\mathbb{E}[w_e] = \sum_j \frac{\mathbb{E}[Z_{j,e}]}{\rho \cdot p_e} = \sum_j \frac{p_e}{\rho \cdot p_e} = 1$$

since j ranges from 1 to ρ.

Theorem 1. *For the graph H generated from G by the above sampling algorithm,*

(a) The expected number of nonzero-weight edges is $O\left(\frac{n \ln n}{\epsilon^2}\right)$.

(b) For all $x \in \mathbb{R}^n$,

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

Proof. (a) The expected number of non-zero weight edges $\leq \sum_e \sum_j \mathbb{E}[Z_{j,e}]$ because of double counting. Then, we see that that

$$\sum_e \sum_j \mathbb{E}[Z_{j,e}] = \rho \cdot \sum_e p_e = \rho(n-1)$$

(b) Note that $L_H = \sum_e w_e \chi_e \chi_e^T$. Then, we can restate (b) as

$$\begin{array}{rcccl} (1-\epsilon)L_G & \preceq & L_H & \preceq & (1+\epsilon)L_G \\ (1-\epsilon)L_G & \preceq & \sum_e w_e \chi_e \chi_e^T & \preceq & (1+\epsilon)L_G \\ (1-\epsilon)\underbrace{L_G^{-1/2} L_G L_G^{-1/2}}_{I} & \preceq & \sum_e w_e \underbrace{(L_G^{-1/2}\chi_e)}_{x_e}\underbrace{(\chi_e^T L_G^{-1/2})}_{x_e^T} & \preceq & (1+\epsilon)\underbrace{L_G^{-1/2} L_G L_G^{-1/2}}_{I} \end{array}$$

where the last line follows from the next claim. Note that statements of the form $(1-\epsilon)I \preceq A$ mean that for all x we have $x^T A x \geq 1-\epsilon$.

Claim 1.

$$M \preceq N \Leftrightarrow C^T M C \preceq C^T N C$$

where C is a nonsingular symmetric matrix.

The claim is not true for general C but it works in the case $C = L_G^{-1/2}$. So, we now have that (b) is equivalent to

$$(1-\epsilon)I \preceq \sum_e w_e x_e x_e^T \preceq (1+\epsilon)I$$

This is identical to the statement that all eigenvalues of $L_G^{-1/2} L_H L_G^{-1/2}$ are between $(1-\epsilon)$ and $(1+\epsilon)$. Thus, it is sufficient to prove the following.

Claim 2. *The eigenvalues of $M := \sum_e w_e x_e x_e^T$ are in $[1-\epsilon, 1+\epsilon]$ where $x_e = L_G^{1/2}\chi_e$.*

Proof. Recall that $w_e = \sum_j \frac{Z_{j,e}}{\rho \cdot p_e}$. Then, we have that

$$\begin{aligned}
M &= \sum_e \sum_j \frac{Z_{j,e}}{\rho \cdot p_e} x_e x_e^T \\
&= \sum_e \sum_j \frac{Z_{j,e}}{\rho} \left(\frac{x_e x_e^T}{x_e^T x_e} \right) \\
\mathbb{E}[M] &= \sum_e \sum_j \frac{\mathbb{E}[Z_{j,e}]}{\rho} \left(\frac{x_e x_e^T}{x_e^T x_e} \right) \\
&= \sum_e x_e x_e^T = I
\end{aligned}$$

Now, we need a matrix version of the Chernoff bounds to show that the eigenvalues are outside our goal with low probability. The matrix Chernoff bounds are due to Tropp. Consider m $n \times n$ matrices $X_1, \ldots, X_m$ such that $0 \preceq X_i \preceq R \cdot I$. Let $X = \sum_i X_i$. For

$$\mu_{\min} I \preceq \mathbb{E}[X] \preceq \mu_{\max} I$$

we have the following bounds.

$$\begin{aligned}
\Pr[\lambda_{\max}(X) \geq \mu_{\max}(1+\epsilon)] \leq &\quad \exp\left(\tfrac{-\epsilon^2 \mu_{\max}}{3R}\right), \quad 0 \leq \epsilon < 1 \\
\Pr[\lambda_{\min}(X) \leq \mu_{\min}(1-\epsilon)] \leq &\quad \exp\left(\tfrac{-\epsilon^2 \mu_{\min}}{2R}\right), \quad 0 < \epsilon < 1
\end{aligned}$$

We can bound the eigenvalues of M by taking $X = \sum_j \frac{Z_{j,e}}{\rho} \left(\frac{x_e x_e^T}{x_e^T x_e} \right)$ for $R = 1/\rho$. This gives us a bound of $\exp(\frac{-\epsilon^2 \rho}{3}) = \frac{1}{n^{c/3}}$ since $\rho = \frac{c \ln n}{\epsilon^2}$. $\square$

$\square$