CAS CS 599 B: Mathematical Methods for TCS

Lecturer: Mark Bun Spring 2022

Lecture Notes 18:

Resistor Networks, Spectral Sparsification

Reading.

- Spielman 11.7-11.9, 12.1-12.4, 32

1 Resistor Networks

Let G be an undirected *weighted* graph. Think of G as a resistor network, where each edge e is a resistor with resistance r_e. Given a *voltage* function $v : V \to \mathbb{R}$ on the vertices, Ohm's Law ($V = IR$) implies that the current from a to b is

$$i(a,b) = \frac{v(a) - v(b)}{r_{(a,b)}} = w_{(a,b)} \cdot (v(a) - v(b)).$$

Note that even though the graph is undirected, this is a directed quantity: $i(a, b) = -i(b, a)$. The net current from a to the rest of the network is

$$i_{ext}(a) = \sum_{b \in V} w_{(a,b)} \cdot (v(a) - v(b)) = (Lv)(a)$$

where L is the usual Laplacian matrix, but taking into account edge weights. That is, L is the linear operator such that

$$\langle v, Lv \rangle = \sum_{(a,b) \in E} w_{(a,b)} (f(a) - f(b))^2.$$

The principle of conservation of flow implies the net current flowing in an out of a vertex must be 0. So if $i_{ext}(a) \neq 0$, that means some external current is being applied at a. One can think of the vertices for which $i_{ext}(a) \neq 0$ as "boundary vertices" through which current is entering or exiting the network, and the other vertices as "internal vertices."

Let's think about what it means for $i_{ext} \equiv 0$. This corresponds to a voltage function for which $Lv = 0$, which for a connected graph G, occurs if and only if v is a constant function.

Now let's consider a more general version of this problem. Given a vector of external currents i_{ext}, what are the induced voltages on the vertices in the graph? Algebraically, this is a solution v to the linear system

$$Lv = i_{ext}.$$

This system has problems owing to the fact that L is singular. Nevertheless, we have:

1. $Lv = i_{ext}$ has a solution iff $i_{ext} \perp \mathbf{1}$, i.e., the net current flowing into the network is zero.

2. Solutions to the system are unique up to constants. That is, v is a solution if and only if $v + c\mathbf{1}$ is a solution for every $c \in \mathbb{R}$.

A canonical solution to this system is given by the Moore-Penrose pseudo-inverse of L. The general definition is kind of complicated, but for symmetric matrices, it's characterized as follows:

Definition 1. The pseudo-inverse of a symmetric matrix L, written L^+, is the matrix such that

1. $\mathrm{Im}(L^+) = \mathrm{Im}(L)$, and
2. $LL^+ = \Pi$ where Π is the symmetric matrix that projects onto $\mathrm{Im}(L)$.

As usual, it's easiest to think about L^+ in terms of the eigendecomposition of L. Let $0 = \lambda_1 \leq \lambda_2 \leq \dots \lambda_n$ be the eigenvalues of L with corresponding orthonormal eigenvectors $\psi_1 = \mathbf{1}, \psi_2, \dots, \psi_n$. Then we can write

$$L = \sum_{i=1}^{n} \lambda_i \psi_i \psi_i^T = \sum_{i=2}^{n} \lambda_i \psi_i \psi_i^T.$$

The pseudoinverse is given by

$$L^+ = \sum_{i=2}^{n} \frac{1}{\lambda_i} \psi_i \psi_i^T.$$

So given a linear system $Lv = i_{ext}$ where $i_{ext} \perp \mathbf{1}$, we can take the canonical solution $v = L^+ i_{ext}$. This is the "balanced" solution, i.e., the one such that $v \perp \mathbf{1}$.

2 Effective Resistance

Fix two vertices a, b in a resistor network. We can think of the network as a one large resistor connecting a and b. What is the resistance of this equivalent resistor?

Recalling Ohm's Law, $i(a,b) = (v(a) - v(b))/r_{(a,b)}$, we can define the effective resistance $R_{\text{eff}}(a,b) = v(a) - v(b)$ where v is the vector of induced potentials when one unit of current enters the network at a and exits at b. That is, v is the solution to

$$Lv = \mathbf{1}_a - \mathbf{1}_b,$$

i.e., $v = L^+(\mathbf{1}_a - \mathbf{1}_b)$ and we are interested in $v(a) - v(b)$. Thus,

$$R_{\text{eff}}(a,b) = \langle \mathbf{1}_a - \mathbf{1}_b, L^+(\mathbf{1}_a - \mathbf{1}_b) \rangle = \| L^{+/2}\mathbf{1}_a - L^{+/2}\mathbf{1}_b \|^2$$

where $L^{+/2}$ is the matrix square root of positive semidefinite L^+, explicitly given by

$$L^{+/2} = \sum_{i=2}^{n} \frac{1}{\sqrt{\lambda_i}} \psi_i \psi_i^T$$

using the decomposition above.

Example 2. Consider the path graph between vertices 1 and n with resistors of resistance $r_{1,2}, \dots, r_{n-1,n}$ connecting them. To compute the effective resistance between 1 and n, we can repeatedly use Ohm's Law to solve for the induced potentials when one unit of current enters at 0 and exits at n. WLOG we can take $v(0) = 1$. Conservation of flow through vertex 1 tells us that one unit of current must flow across edge $(1,2)$. By Ohm's Law, this means $v(1) = -r_{1,2}$. Continuing, we get that $v(n) = -r_{1,2} - \dots - r_{n-1,n}$. And so the effective resistance $R_{\text{eff}}(1,n) = r_{1,2} + \dots + r_{n-1,n}$.

Example 3. Now consider two vertices a and b with parallel edges of resistance $r_1, \ldots, r_k$ connecting them. The weight of an edge is the inverse of the resistance, so this is a weighted graph with k parallel edges of weights $1/r_1, \ldots, 1/r_k$. By linearity, this is equivalent to a graph with a single edge of weight $1/r_1 + \cdots + 1/r_k$ between a and b, which corresponds to effective resistance $1/(1/r_1 + \ldots, 1/r_k)$.

You can interpret effective resistance as a measure of distance between vertices in a graph. (Actually, it literally is a metric.) When two vertices are connected only by a few long paths, the effective resistance is large. When they are connected by many short paths, the effective resistance is small.

3 Spectral Sparsification

Given a dense graph G, can we find another graph H that approximates G in the sense that

1. H preserves the cuts of G: $(1-\varepsilon)\phi(G) \leq \phi(H) \leq (1+\varepsilon)\phi(G)$,
2. Effective resistances in H approximate those in G,
3. Solutions v to linear equations $L_H v = b$ approximate those to $L_G v = b$,
4. The eigenvalues of L_G and L_H are similar?

The following definition of approximation implies all of the above:

Definition 4. A weighted graph H is an ε-approximation to G if

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

Here, $A \preceq B$ means that A precedes B in the Loewner order on PSD matrices: $B - A$ is a positive semidefinite matrix.

One interpretation of expanders is that they are are sparse approximations to the complete graph. We can see that this is true in the above sense.

Proposition 5. *If H is a d-regular γ-spectral expander on n vertices, then H is a $(1-\gamma)$-approximation to $G = \frac{d}{n}K_n$.*

Proof. I'll only prove that $\gamma L_G \preceq L_H$. The other direction is similar. To do this, it suffices to show that for every f, we have $\langle f, (L_H - \gamma L_G) f\rangle \geq 0$. Since $L_G f = L_H f = 0$ whenever $f \in \operatorname{span}\{\mathbf{1}\}$, it suffices to show this for $f \perp \mathbf{1}$. Since for any $f \perp \mathbf{1}$ we have $L_G f = df$, and by spectral expansion, $\langle f, L_H f\rangle \geq \gamma d$, it follows that

$$\langle f, (L_H - \gamma L_G) f\rangle = \langle f, L_H f\rangle - \gamma d\|f\|^2 \geq 0.$$

$\square$

Theorem 6. *Every weighted undirected graph G has an ε-approximation H with $O(n \log n/\varepsilon^2)$ edges.*

Proof. We give a simple randomized algorithm for sampling H. For every $e \in E$, set

$$w_e^H = \begin{cases} \frac{w_e^G}{p_e} & \text{with probability } p_e \\ 0 & \text{with probability } 1 - p_e \end{cases}$$

where $p_e = C \cdot w_e^G \cdot R_{\text{eff}}(e)$ for $C = O(\log n/\varepsilon^2)$ to be chosen later. That is, we include each candidate edge with probability proportional to its effective resistance, and when we do include it, we give it weight w_e/p_e. We can assume that each $p_e \leq 1$ by splitting it into multiple candidate edges otherwise.

If you believe me that it suffices to take $C = O(\log n/\varepsilon^2)$, then in expectation, the total number of edges in the graph is

$$\begin{aligned}
\sum_{(a,b)\in E} p_e &= C \sum_{(a,b)\in E} w_{(a,b)} \langle \mathbf{1}_a - \mathbf{1}_b, L^+(\mathbf{1}_a - \mathbf{1}_b)\rangle \\
&= C \sum_{(a,b)\in E} w_{(a,b)} \operatorname{Tr}(L^+(\mathbf{1}_a - \mathbf{1}_b)(\mathbf{1}_a - \mathbf{1}_b)^T) \\
&= C \sum_{(a,b)\in E} w_{(a,b)} \operatorname{Tr}(L^+ L_{(a,b)}) \\
&= C \operatorname{Tr}\left(L^+ \sum_{(a,b)\in E} w_{(a,b)} L_{(a,b)}\right) \\
&= C \operatorname{Tr}(L^+ L) \\
&= C \operatorname{Tr}(\Pi_{\mathrm{Im}(L)}) \\
&= C(n-1) = O(n \log n/\varepsilon^2).
\end{aligned}$$

Here, $L_{(a,b)}$ is the combinatorial Laplacian of the single edge (a, b), i.e., the operator for which $\langle f, L_{(a,b)} f\rangle = (f(a) - f(b))^2$. A Chernoff bound shows that the number of edges is within a constant factor of this bound with all but exponentially small probability.

Remark 7. What's going on here? A consequence of Kirchoff's "matrix tree theorem" is that $w_e \cdot R_{\text{eff}}(e)$ is the probability that an edge e appears in a random spanning tree of G when a tree is sampled with probability proportional to its edge weights. Since every spanning tree has $n - 1$ edges, the sum of these probabilities is $n - 1$. The connection between spanners and effective resistance sampling has led to subsequent fast algorithms for constructing sparsifiers, e.g., Kapralov and Panigraphy, "Spectral sparsification via random spanners."

It now remains to show that H is indeed a good spectral sparsifier. By construction, it has the right expectation:

$$\mathbb{E}[L_H] = \sum_{(a,b)\in E} \mathbb{E}[w_{(a,b)}(\mathbf{1}_a - \mathbf{1}_b)(\mathbf{1}_a - \mathbf{1}_b)^T] = L_G.$$

The basic idea is to use a *matrix* Chernoff bound to show that L_H is highly concentrated around its expectation.

Theorem 8. *Let $X_1, \ldots, X_m$ be independent $(n \times n)$ symmetric PSD matrices where $\|X_i\| \leq R$ almost surely for every i. Let $X = \sum_{i=1}^m X_i$ and let $\mu_{\min}$ and $\mu_{\max}$ be the minimum and maximum eigenvalues of $\mathbb{E}[X]$. Then for $\varepsilon \in (0, 1)$,*

$$\begin{aligned}
\Pr[\lambda_{\min}(X) \leq (1-\varepsilon)\mu_{\min}] &\leq n e^{-\varepsilon^2 \mu_{\min}/2R} \\
\Pr[\lambda_{\max}(X) \geq (1+\varepsilon)\mu_{\max}] &\leq n e^{-\varepsilon^2 \mu_{\max}/3R}.
\end{aligned}$$

I'll only prove that $L_H \preceq (1+\varepsilon)L_G$ with high probability. The lower bound is similar, but requires a bit of care as we want to carry out the argument in the subspace orthogonal to $\mathbf{1}$. First, we'll convert the statement we want to show into an easier form to work with. By pre- and post-multiplying by the PSD matrix $L_G^{+/2}$, we see that it's equivalent to

$$L_G^{+/2} L_H L_G^{+/2} \preceq (1+\varepsilon) L_G^{+/2} L_G L_G^{+/2} = (1+\varepsilon)\Pi$$

where Π is the projection onto $\mathrm{Im}(L_G)$. So it's enough to show that the maximum eigenvalue of $L_G^{+/2} L_H L_G^{+/2}$ is at most $(1+\varepsilon)$ with high probability.

We will now prove this using the matrix Chernoff bound. Let

$$X_e = \begin{cases} \frac{w_e}{p_e} L_G^{+/2} L_e L_G^{+/2} & \text{with probability } p_e \\ 0 & \text{otherwise.} \end{cases}$$

Let $X = \sum_{e \in E} X_e$ so that $X = L_G^{+/2} L_H L_G^{+/2}$.

We first observe that by linearity,

$$\mathbb{E}[X] = \mathbb{E}[L_G^{+/2} L_H L_G^{+/2}] = L_G^{+/2} \mathbb{E}[L_H] L_G^{+/2} = L_G^{+/2} L_G L_G^{+/2} = \Pi$$

which has maximum eigenvalue 1.

Next, we check that

$$\begin{aligned} \|X_e\| &\leq \frac{w_e}{p_e} \|L_G^{+/2} L_e L_G^{+/2}\| \\ &\leq \frac{w_e}{p_e} \operatorname{Tr}(L_G^{+/2} (\mathbf{1}_a - \mathbf{1}_b)(\mathbf{1}_a - \mathbf{1}_b)^T L_G^{+/2}) \\ &= \frac{w_e}{p_e} \operatorname{Tr}(L_G^{+/2} L_G^{+/2} (\mathbf{1}_a - \mathbf{1}_b)(\mathbf{1}_a - \mathbf{1}_b)^T) \\ &= \frac{w_e}{p_e} \operatorname{Tr}(L_G^{+} (\mathbf{1}_a - \mathbf{1}_b)(\mathbf{1}_a - \mathbf{1}_b)^T) \\ &= \frac{w_e}{p_e} R_{\text{eff}}(e) = \frac{1}{C}. \end{aligned}$$

Thus, the matrix Chernoff bound tells us that

$$\Pr[\lambda_{\max}(X) \geq 1+\varepsilon] \leq n e^{-3C\varepsilon^2}.$$

To get this below a constant, it suffices to take $C = O(\log n/\varepsilon^2)$.

$\square$