CAS CS 599 B: Mathematical Methods for TCS

Lecturer: Mark Bun | Spring 2022

Lecture Notes 14:

Conductance, Cheeger's inequality

Reading.

- Trevisan Chapter 4
- Spielman Chapters 20 and 21

Let's recap a few definitions. For an undirected graph $G = (V, E)$, the Laplacian quadratic form is

$$Q[f] = \sum_{(u,v)\in E} (f(u) - f(v))^2.$$

It's so named because it is the quadratic form associated to the Laplacian operator L, which you can either define to be the operator for which $\langle f, Lf\rangle = Q[f]$ or $L = D - A$ where D is the diagonal matrix of degrees of G and A is the adjacency matrix.

Last time we started studying the eigenvalues of L and began associating them to combinatorial properties of G. In particular, we showed that all the eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \dots \lambda_n$ and that $\lambda_2 = 0$ iff G is disconnected.

1 Courant-Fischer Theorem

Theorem 1 (Spectral Theorem)**.** *Let M be a real symmetric matrix. Then there exists a diagonal matrix Λ and an orthogonal matrix V such that $M = V\Lambda V^T = \sum_{i=1}^n \lambda_i v_i v_i^T$. The eigenvalues of M are the entries λ_i of Λ and the columns v_i of V are the eigenvectors.*

Definition 2. Given an operator M and a function f, define the Rayleigh quotient

$$R_M[f] := \frac{\langle f, Mf\rangle}{\langle f, f\rangle}.$$

Theorem 3. *If M is a symmetric matrix with eigenvalues $\lambda_1 \leq \dots \leq \lambda_n$ and corresponding eigenvectors $f_1, \dots, f_n$. Then*

$$\lambda_1 = \min_{f\neq 0} R_M[f],$$

$$\lambda_2 = \min_{f\neq 0, f\perp f_1} R_M[f], \dots$$

$$\lambda_k = \min_{f\neq 0, f\in S_{k-1}^{\perp}} R_M[f]$$

where $S_{k-1} = \mathrm{span}\{f_1, \dots, f_{k-1}\}$

Proof sketch. By the spectral theorem for real symmetric matrices, we know that we can take the eigenvectors $f_1, \ldots, f_n$ to be orthogonal. Thus, the condition $f \in S_{k-1}^{\perp}$ implies $f \perp f_i$ for all $i = 1, \ldots, k-1$. So for $f \in S_{k-1}^{\perp}$, we have

$$\langle \sum_{i=k}^{n} c_i f_i, M \sum_{i=k}^{n} c_i f_i \rangle = \langle \sum_{i=k}^{n} c_i f_i, \sum_{i=k}^{n} \lambda_i c_i f_i \rangle = \sum_{i=k}^{n} \lambda_i c_i^2 \geq \lambda_k \sum_{i=k}^{n} c_i^2.$$

The denominator of $R_M[f]$, on the other hand, is $\sum_{i=k}^{n} c_i^2$. So the Rayleigh quotient itself is always at least λ_k. Meanwhile, plugging in $f = f_k$ itself yields $R_M[f_k] = \lambda_k$. $\square$

Recall that one motivation for studying this "variational" characterization of eigenvalues of the Laplacian was from graph cuts. For a set of vertices S, we introduced a quantity called the "isoperimetric ratio":

$$\theta(S) := \frac{|\partial S|}{|S|}.$$

This quantity doesn't treat S and $V \setminus S$ symmetrically, so we'll consider it only when $|S| \leq n/2$. Sometimes you'll see this asymmetry dealt with by normalizing by $\min\{|S|, |V \setminus S|\}$.

Claim 4. *If $S \subseteq V$ with $|S| \leq n/2$, then*

$$\frac{|\partial S|}{|S|} \geq \lambda_2/2.$$

Proof. We want to use the variational characterization of λ_2 of the Laplacian

$$\lambda_2 = \min_{f \neq 0, f \perp \mathbf{1}} \frac{\langle f, Lf \rangle}{\langle f, f \rangle}.$$

To do this, we need to associate to every cut S a "test function" f that makes the Rayleigh quotient small.

Based on our previous discussion, your first instinct might be to set $f = \mathbf{1}_S$, but this is not orthogonal to $\mathbf{1}$. So we'll fix this by instead taking

$$f = \mathbf{1}_S - \sigma \mathbf{1}$$

where $\sigma = |S|/n$.

Now we can check:

$$\langle f, Lf \rangle = \sum_{(u,v) \in E} (f(u) - f(v))^2 = \sum_{(u,v) \in E} (\mathbf{1}_S(u) - \mathbf{1}(v))^2 = |\partial S|,$$

and

$$\langle f, f \rangle = \sum_{v \in S} (1 - \sigma)^2 + \sum_{v \notin S} \sigma^2 = |S|(1 - \sigma).$$

Thus we get

$$\frac{|\partial S|}{|S|(1-\sigma)} \geq \min_{f \neq 0, f \perp \mathbf{1}} \frac{\langle f, Lf \rangle}{\langle f, f \rangle} = \lambda_2$$

and rearranging,

$$\theta(S) \geq \lambda_2(1 - s) \geq \frac{\lambda_2}{2}.$$

$\square$

If you don't like the inductive statement of Theorem 3, you may find comfort in the following closely related statement called the Courant-Fischer Theorem:

Theorem 5 (Courant-Fischer)**.** *If M is a symmetric matrix with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$,*

$$\lambda_k = \max_{\dim(S)=n-k+1} \min_{f \neq 0, f \in S} R_M[f] = \min_{\dim(T)=k} \max_{f \neq 0, f \in T} R_M[f].$$

2 Cheeger's Inequality

Claim 4 gave a lower bound on the isoperimetric ratio of a graph in terms of λ_2. It turns out that λ_2 actually *characterizes* this quantity. The result is actually cleaner and tighter to state if we instead study a related quantity called the graph conductance, and relate it to a normalized version of the Laplacian.

Definition 6. Let G be an undirected graph and let $S \subseteq V$ be a set of vertices. The *conductance* of S is defined to be

$$\phi(S) = \frac{|\partial S|}{\sum_{v \in S} \deg(v)}.$$

The conductance of G itself is defined as

$$\phi(G) = \min_{1 \leq |S| \leq |V|/2} \phi(S).$$

The new normalization makes it natural to study the following *generalized* Rayleigh quotient

$$\frac{\langle f, Lf \rangle}{\langle f, Df \rangle}$$

where D is the diagonal degree matrix of G. It'll be convenient to express this in terms of the ordinary Rayleigh quotient, which we can do be performing the change of variables $g = D^{1/2} f$. This turns the ratio into

$$\frac{\langle D^{-1/2}g, LD^{-1/2}g \rangle}{\langle D^{-1/2}g, D^{1/2}g \rangle} = \frac{\langle g, D^{-1/2}LD^{-1/2}g \rangle}{\langle g, g \rangle},$$

which is the ordinary Rayleigh quotient for the operator $N := D^{-1/2}LD^{-1/2}$. This operator is called the *normalized Laplacian*. Observe that:

- Since $L = D - A$, we have $N = I - D^{-1/2}AD^{-1/2}$.
- If G is d-regular, then $D = dI$ and $D^{-1/2} = -\frac{1}{\sqrt{d}}I$, so $N = \frac{1}{d}L$.

Theorem 7 (Cheeger's Inequality)**.** *For every undirected graph G with normalized Laplacian N with eigenvalues $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n$, we have*

$$\frac{\nu_2}{2} \leq \phi(G) \leq 2\sqrt{\nu_2}.$$

3 Proof of Cheeger's Inequality

Cheeger's Inequality consists of an "easy" direction and a "hard" direction. We've already essentially seen the easy direction in our proof of Claim 4, so I won't repeat it. The idea is to show that a set with small conductance can be used to construct a test function that makes that Rayleigh quotient of N small.

The "hard" direction is also constructive. The idea is to show that if f is an eigenvector of N for eigenvalue λ_2, then f can be used to construct a set with small conductance. We will follow a proof due to Trevisan, specialized to the case where G is d-regular.

Proposition 8. *Let G be a d-regular graph, and let $f \perp 1$. Then there exists a number t such that $S_t = \{u : f(u) \geq t\}$ satisfies*

$$\phi(S_t) \leq \sqrt{2R_N[f]}$$

Note for d-regular graphs,

$$\phi(S_t) = \frac{|\partial S_t|}{d|S_t|}$$

while

$$R_N[f] = \frac{\langle f, Nf \rangle}{\langle f, f \rangle} = \frac{Q[f]}{d\langle f, f \rangle}.$$

Proof. We first do some preprocessing to f to make it into a nicer function g with a smaller Rayleigh quotient. First, assume $V = [n]$ and that the vertices are sorted so that

$$f(1) \leq f(2) \leq \cdots \leq f(n).$$

Second, by setting $g = f - f(1)\mathbf{1}$, we can get a function for which $g(1) = 0$ and hence $g(u) \geq 0$ for all u. Note that this doesn't increase the Rayleigh quotient: $Q[g] = Q[f]$, while since $f \perp 1$, we have $\langle g, g \rangle = \langle f, f \rangle + f(1)^2 \langle \mathbf{1}, \mathbf{1} \rangle \geq \langle f, f \rangle$. Hence $R_N[g] \leq R_N[f]$.

Finally, by multiplying by a scalar, we can assume that $g(n)^2 = 1$ without changing its Rayleigh quotient.

Our goal is to define a distribution on t such that

$$\mathbb{E}_t[|\partial(S_t)|] \leq d \cdot \mathbb{E}_t[|S_t|] \cdot \sqrt{2R_N[g]},$$

(where now $S_t = \{u : g(u) \geq t\}$)) which will imply the existence of a t satisfying the claim.

A distribution that works is the following: Sample $t \in [0, 1]$ such that t^2 is uniformly distributed in $[0, 1]$.

We now compute an upper bound on $\mathbb{E}_t[|\partial S_t|]$. First, we write this expectation as

$$\sum_{(u,v) \in E} \Pr_t[(u, v) \in \partial S_t].$$

Now an edge (u, v) where $g(u) \leq g(v)$ is in ∂S_t iff $g(u)^2 \leq t^2 < g(v)^2$. The probability that this happens is $|g(u)^2 - g(v)^2| = |g(u) - g(v)| \cdot (g(u) + g(v))$. Applying Cauchy-Schwarz twice:

$$\begin{aligned}
\mathbb{E}_t[|\partial S_t|] &= \sum_{(u,v)\in E} |g(u) - g(v)| \cdot (g(u) + g(v)) \\
&\leq \sqrt{\sum_{(u,v)\in E} (g(u) - g(v))^2} \cdot \sqrt{\sum_{(u,v)\in E} (g(u) + g(v))^2} \\
&\leq \sqrt{Q[g]} \cdot \sqrt{\sum_{(u,v)\in E} 2g(u)^2 + 2g(v)^2} \\
&\leq \sqrt{Q[g]} \cdot \sqrt{2d \sum_{u\in V} g(u)^2} \\
&= \sqrt{2R_N[g]} \cdot d \cdot \sum_{u\in V} g(u)^2
\end{aligned}$$

Now we observe that $\mathbb{E}_t[|S_t|] = \sum_u \Pr_t[g(u) \geq t] = \sum_u \Pr_t[g(u)^2 \geq t^2] = \sum_u g(u)^2$. Putting everything together gives us

$$\mathbb{E}_t[|\partial(S_t)|] \leq d \cdot \mathbb{E}_t[|S_t|] \cdot \sqrt{2R_N[g]}$$

as we wanted.

□