Lecture 7: Cheeger-type Inequalities for λ_n, cont'd

Last updated 2/25/2016

In which we state an analog of Cheeger's inequalities for the k-th smallest Laplacian eigenvalue, and we discuss the connection between this result and the analysis of spectral partitioning algorithms

1 Review

Let $G = (V, E)$ be a d-regular undirected graph, L its normalized Laplacian, $0 = \lambda_1 \leq \cdots \leq \lambda_n \leq 2$ be the Laplacian eigenvalues, and $\phi_k(G)$ be the order-k expansion of G.

We want to prove

$$\phi_k(G) \leq O(k^{3.5}) \cdot \sqrt{\lambda_k}$$

We will prove the somewhat stronger result that, given k orthonormal vectors $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}$, we can find k disjointly supported vectors $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(k)}$ such that, for every $i = 1, \ldots, k$,

$$R_L(\mathbf{y}^{(i)}) \leq O(k^7) \cdot \max_{j=1,\ldots,k} \quad R_L(\mathbf{x}^{(i)})$$

In order to do that, we define the mapping

$$F(v) := (x_v^{(1)}, \ldots, x_v^{(k)}) \tag{1}$$

of vertices to $\mathbb{R}^k$ and the normalized distance

$$dist(u, v) := \left\| \frac{F(u)}{||F(u)||} - \frac{F(v)}{||F(v)||} \right\| \tag{2}$$

between vertices, and we are going to prove the following two lemmas.

Lemma 1 (Localization) *Given t sets $A_1, \ldots, A_t$ such that, for every $i = 1, \ldots, t$, $\sum_{v \in A_i} ||F(v)||^2 \geq \frac{1}{2}$ and,for every u, v in different sets $dist(u, v) \geq \delta$, we can construct t disjointly supported vectors $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(t)}$ such that for every $i = 1, \ldots, t$, we have*

$$R_L(\mathbf{y}^{(t)}) \leq O(k \cdot \delta^{-2}) \cdot R_L(F)$$

Lemma 2 (Well-Separated Sets) *There are k disjoint subsets of vertices $A_1, \ldots, A_k$ such that*

- *For every $i = 1, \ldots, k$, $\sum_{v \in A_i} ||F(v)||^2 \geq \frac{1}{2}$*
- *For every u and v belonging to different sets, $dist(u, v) \geq \Omega(k^{-3})$*

Recall that, for a function $f : V \to \mathbb{R}^k$ the Rayleigh quotient of f is defined as

$$R_L(f) := \frac{\sum_{\{u,v\}} ||f(u) - f(v)||^2}{d \sum_v ||f(v)||^2}$$

and, by definition of F, we have

$$R_L(F) = \frac{1}{k} \sum_i R_L(\mathbf{x}^{(i)})$$

2 Some Preliminaries

We will prove some simple properties of the embedding $F(\cdot)$ and of the distance function $dist(\cdot, \cdot)$.

First, we observe that

$$\sum_{v \in V} ||F(v)||^2 = \sum_v \sum_i (x_v^{(i)})^2 = \sum_i ||\mathbf{x}^{(i)}||^2 = k$$

Next, we prove the sense in which $F(\cdot)$ "spreads out" vertices across $\mathbb{R}^k$.

Lemma 3 *For every unit vector $\mathbf{w} \in \mathbb{R}^k$,*

$$\sum_{v \in V} \langle F(v), \mathbf{w} \rangle^2 = 1$$

PROOF: Consider the $k \times n$ matrix X whose rows are the vectors $\mathbf{x}^{(i)}$ and whose columns are the points $F(v)$. Then we have

$$\sum_{v \in V} \langle F(v), \mathbf{w} \rangle^2 = ||X^T \mathbf{w}||^2 = \mathbf{w}^T X X^T \mathbf{w} = \mathbf{w}^T \mathbf{w} = 1$$

where we used the fact that the rows of X are orthogonal and so XX^T is the identity. $\square$

This means that, for every direction, the points $F(v)$ correlate with that direction in the same way, regardless of the direction itself.

In the proof of the localization lemma, we will make use of the following inequality: for every two vectors $\mathbf{x}, \mathbf{y}$,

$$|\ ||\mathbf{x}|| - ||\mathbf{y}||\ | \leq ||\mathbf{x} - \mathbf{y}||$$

which is a consequence of Cauchy-Schwarz:

$$\begin{aligned}
(||\mathbf{x}|| - ||\mathbf{y}||)^2 &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2||\mathbf{x}|| \cdot ||\mathbf{y}|| \\
&\leq ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle \\
&= ||\mathbf{x} - \mathbf{y}||^2
\end{aligned}$$

3 Localization

In this section we prove Lemma 1.

3.1 Proof Ideas

The basic idea is that we would like to define the vectors $\mathbf{y}^{(i)}$ as

$$y_v^{(i)} := \mathbf{1}_{A_i}(v) \cdot ||F(v)||$$

The denominator of the Rayleigh quotient of such a vector is, by definition, at least 1/2, and we might hope to upper bound the numerator of the Rayleigh quotient of $\mathbf{y}^{(i)}$ in terms of the numerator of the Rayleigh quotient of F, which is $kR_L(F)$.

Indeed, every edge $\{u, v\}$ with both endpoints outside of A_i contributes zero to the numerator of the Rayleigh quotient of $\mathbf{y}^{(i)}$, and every edge $\{u, v\}$ with both endpoints in A_i contributes

$$(\ ||F(u)|| - ||F(v)|| \ )^2 \leq ||F(u) - F(v)||^2$$

to the numerator of the Rayleigh quotient of $\mathbf{y}^{(i)}$, and the right-hand-side above is the contribution of the edge to the numerator of the Rayleigh quotient of F.

So far, so good, but the problem comes from edges $\{u, v\}$ with one endpoint $u \in A_i$ and one endpoint $v \notin A_i$. Such an edge contributes $||F(u)||^2$ to the Rayleigh quotient of $\mathbf{y}^{(i)}$ and $||F(u) - F(v)||^2$ to the Rayleigh quotient of F, and the former quantity could be much larger than the latter. If $dist(u, v)$ is large, however, $||F(u)||^2$ cannot be much larger than $||F(u) - F(v)||^2$, because of the following fact

Lemma 4

$$||F(v)|| \cdot dist(u, v) \leq 2||F(u) - F(v)|| \tag{3}$$

PROOF:

$$\begin{aligned}
||F(v)|| \cdot dist(u, v) &= ||F(v)|| \cdot \left\| \ \frac{F(u)}{||F(u)||} - \frac{F(v)}{||F(v)||} \ \right\| \\
&= \left\| F(u) \cdot \frac{||F(v)||}{||F(u)||} - F(v) \right\| \\
&\leq \left\| F(u) \cdot \frac{||F(v)||}{||F(u)||} - F(u) \right\| + \|F(u) - F(v)\| \\
&= \left\| F(u) \cdot \left(\frac{||F(v)||}{||F(u)||} - 1 \right) \right\| + \|F(u) - F(v)\| \\
&= ||F(u)|| \cdot \left| \frac{||F(v)||}{||F(u)||} - 1 \right| + \|F(u) - F(v)\| \\
&= | \ ||F(v)|| - ||F(u)|| \ | + \|F(u) - F(v)\| \\
&\leq 2\|F(u) - F(v)\|
\end{aligned}$$

□

We can conclude that the only problem comes from edges $\{u, v\}$ such that $u \in A_i$, $v \notin A_i$, and $dist(u, v)$ is small. To deal with such edges, we will modify the definition of $\mathbf{y}^{(i)}$, and use a "smoothed" version of the indicator function of A_i instead of the actual indicator.

3.2 Proof

If v is a vertex and A is a set of vertices, we define

$$dist(v, A) = \min_{u \in A} dist(v, u)$$

For each i, we define the following *smooth indicator function* of A_i:

$$\tau_i(v) = \begin{cases} 1 & \text{if } v \in A_i \\ 0 & \text{if } dist(v, A_i) \geq \frac{\delta}{2} \\ 1 - \frac{2}{\delta} \cdot dist(v, A_i) & \text{otherwise} \end{cases}$$

Notice that the functions $\tau_i(\cdot)$ are disjointly supported: there cannot be a vertex v such that $\tau_i(v) > 0$ and $\tau_j(v) > 0$ for $i \neq j$, otherwise we would have $dist(v, A_i) < \frac{\delta}{2}$ and $dist(v, A_j) < \frac{\delta}{2}$, contradicting the well-separated condition on the sets A_i.

We define the vectors $\mathbf{y}^{(i)}$ as

$$y_v^{(i)} = \tau_i(v) \cdot ||F(v)||$$

The vectors $\mathbf{y}^{(i)}$ are disjointly supported, and it remains to understand their Rayleigh quotient.

The denominator of the Rayleigh quotient of $\mathbf{y}^{(i)}$ is

$$\sum_{v \in V} \tau_i^2(v) \cdot ||F(v)||^2 \geq \sum_{v \in A_i} ||F(v)||^2 \geq \frac{1}{2}$$

The contribution of an edge $\{u, v\}$ to the numerator is the square of

$$\begin{aligned} |y_v^{(i)} - y_u^{(i)}| &= |\ \tau_i(v) \cdot ||F(v)|| - \tau_i(u) \cdot ||F(u)||\ | \\ &\leq |\ \tau_i(v) \cdot ||F(v)|| - \tau_i(v) \cdot ||F(u)||\ | + |\ \tau_i(v) \cdot ||F(u)|| - \tau_i(u) \cdot ||F(u)||\ | \\ &= \tau_i(v) \cdot ||F(v) - F(u)|| + ||F(u)|| \cdot |\tau_i(v) - \tau_i(u)| \\ &\leq ||F(v) - F(u)|| + ||F(u)|| \cdot \frac{2}{\delta} \cdot dist(v, u) \\ &\leq ||F(v) - F(u)|| \cdot \left(1 + \frac{4}{\delta}\right) \end{aligned}$$

where we used the inequality

$$|\tau_i(v) - \tau_i(u)| \leq \frac{2}{\delta} |dist(v, A_i) - dist(u, A_i)| \leq \frac{2}{\delta} dist(v, u)$$

The numerator of the Rayleigh quotient of $\mathbf{y}^{(i)}$ is thus

$$\sum_{\{u,v\}\in E} |y_v^{(i)} - y_u^{(i)}|^2 \leq O(\delta^{-2}) \sum_{\{u,v\}\in E} ||F(v) - F(u)||^2 = O(\delta^{-2}) \cdot kR_L(F)$$

and this proves Lemma 1.

4 Well-Separated Sets

In this section we prove Lemma 2, which follows easily from the following result.

Lemma 5 *We can find disjoint sets of vertices $T_1, \ldots, T_m$ such that*

- $\sum_{i=1}^m \sum_{v\in T_i} ||F(v)||^2 \geq k - \frac{1}{4}$
- *For every u, v in different sets, $dist(u, v) \geq \Omega(k^{-3})$*
- *For every set T_i, $\sum_{v\in T_i} ||F(v)||^2 \leq 1 + \frac{1}{4k}$*

We can derive Lemma 2 from Lemma 5 as follows. Let us call the quantity $\sum_{v\in A} ||F(v)||^2$ the *mass* of a set A. Starting from the sets $T_1, \ldots, T_m$ promised by Lemma 5 we run the following process: as long as there are two sets both of mass $< \frac{1}{2}$ we merge them. Call $A_1, \ldots, A_t$ the sets of mass $\geq \frac{1}{2}$ obtained at the end of this process; in addition, there may be one more set of mass $< \frac{1}{2}$. Every set has mass $\leq 1 + \frac{1}{4k}$. This means that the total mass of the sets is at most $\frac{1}{2} + t \cdot \left(1 + \frac{1}{4k}\right) \geq k - \frac{1}{4}$, which implies $t \geq k$. Thus we have found k sets of vertices, each of mass at least $1/2$, and such that any two sets have separation $\Omega(k^{-3})$.

Now we turn to the proof of Lemma 5. We are going to use the fact that, for every small cone in $\mathbb{R}^k$, the mass of vertices v such that $F(v)$ is in the cone is also small and, in particular, it can made at most $1 + \frac{1}{4k}$. We will prove the Lemma by covering almost all the points $F(v)$ using a collection of well-separated small cones.

We first formalize the above intuition about cones. If R (for *r*egion) is a subset of the unit sphere in $\mathbb{R}^n$, then the diameter of R is

$$diam(R) := \sup_{\mathbf{w},\mathbf{z}\in R} ||\mathbf{w} - \mathbf{z}||$$

and the *cone* generated by R is the set $\{\alpha\mathbf{w} : \alpha \in \mathbb{R}_{\geq 0}, \mathbf{w} \in R\}$ of non-negative scalar multiples of elements of R. The set of vertices covered by R, denoted $V(R)$ is the set of vertices v such that $F(v)$ is in the cone generated by R or, equivalently,

$$V(R) := \left\{ v \in V : \frac{F(v)}{||F(v)||} \in R \right\}$$

If R has small diameter, then $V(R)$ has small mass.

Lemma 6 *For every subset R of the unit sphere,*

$$\sum_{v \in V(R)} ||F(v)||^2 \leq \left(1 - \frac{1}{2}(diam(R))^2\right)^{-2}$$

Proof: For every two unit vectors $\mathbf{w}$ and $\mathbf{z}$, we have

$$\langle \mathbf{z}, \mathbf{w} \rangle = 1 - \frac{1}{2}||\mathbf{w} - \mathbf{z}||^2$$

For every vertex v, call

$$\bar{F}(v) := \frac{F(v)}{||F(v)||}$$

Let $\mathbf{w}$ be a vector in R. Then we have

$$1 \geq \sum_{v \in V(R)} \langle F(v), \mathbf{w} \rangle^2$$

$$= \sum_{v \in V(R)} ||F(v)||^2 \cdot \langle \bar{F}(v), \mathbf{w} \rangle^2$$

$$= \sum_{v \in V(R)} ||F(v)||^2 \cdot \left(1 - \frac{1}{2}||\bar{F}(v) - \mathbf{w}||^2\right)^2$$

$$\geq \sum_{v \in V(R)} ||F(v)||^2 \cdot \left(1 - \frac{1}{2}(diam(R))^2\right)^2$$

□

In particular, if $diam(R) \leq \frac{1}{\sqrt{5k}}$, then the mass of $V(R)$ is at most

$$\left(1 - \frac{1}{10k}\right)^{-2} \leq \left(1 - \frac{1}{5k}\right)^{-1} = 1 + \frac{1}{5k-1} \leq 1 + \frac{1}{4k}$$

Another observation is that, for every two subsets R_1, R_2 of the unit sphere,

$$\min_{u \in V(R_1), v \in V(R_2)} dist(u, v) \geq \min_{\mathbf{w} \in R_1, \mathbf{z} \in R_2} ||\mathbf{w} - \mathbf{z}||$$

Our approach will be to find disjoint subsets $R_1, \ldots, R_m$ of the unit sphere, each of diameter at most $1/2\sqrt{k}$, such that the total mass of the sets $V(R_1), \ldots, V(R_m)$ is at least $k - \frac{1}{4}$ and such that the separation between any two R_i, R_j is at least $\Omega(k^{-3})$.

To do this, we tile $\mathbb{R}^k$ with axis-parallel cubes of side $L = \frac{1}{2k}$, which clearly have diameter at most $\frac{1}{2\sqrt{k}}$, and, for every cube C, we define its *core* $\tilde{C}$ to be a cube with the same center as C and of side $L \cdot \left(1 - \frac{1}{4k^2}\right)$. Note two points in the core of two different cubes have distance at least $\frac{1}{8k^3}$. Let now $R_1, R_2, \ldots$ be the intersections of the cube cores with the unit sphere. Since each R_i is a subset of a core of a cube, it has diameter at most $\frac{1}{2\sqrt{k}}$, and the distance between any two points in different regions is at least $\frac{1}{8k^3}$. We claim that there is a way to choose the location of the centers of the cubes so that $\sum_i \sum_{v \in V(R_i)} ||F(v)||^2 \geq k - \frac{1}{4}$.

Let us start by a fixed configuration of the cubes and then apply an axis-parallel random shift (by adding to each coordinate, a random real in the range $[0, L]$. Then, for each fixed point in $\mathbb{R}^n$ and, in particular, for each point $\bar{F}(v)$, the probability that it falls in the core of a cube after the shift is at least $1 - \frac{1}{4k}$, so the average mass of the vertices covered by the regions is at least $k - \frac{1}{4}$, and there must exist a shift that is at least as good.