UBC CPSC 536N: Sparse Approximations Winter 2013

Lecture 18 — March 13, 2013

Prof. Nick Harvey *Scribe: Daniel Busto*

1 Preliminaries

First a reminder of what the definition of a spectral expander is.
An infinite family of graphs $\{G_n\}$ are spectral expanders if there is a constant c such that:

1. G_n is connected
2. G_n has $O(n)$ edges
3. $\forall n\ \lambda_{min_{>0}}(L_{G_n}) \geq c$.

Note that the last condition is equivalent to $L_G \succeq c \cdot I_{im_{L_G}}$. We now derive an even simpler condition. Let U be an orthogonal matrix ($U \cdot U^T = I$) with $\frac{\vec{1}}{\sqrt{n}}$ as the last row. Then we have:

$$\begin{aligned} L_G \succeq cI_{im_{L_G}} &\iff UL_GU^{\mathsf{T}} \succeq cUI_{im_{L_G}}U^{\mathsf{T}} \\ &\iff \left(\begin{smallmatrix} M & 0 \\ 0 & 0 \end{smallmatrix}\right) \succeq c\left(\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right) \\ &\iff M \succeq cI \\ &\iff \lambda_{min}(M) \geq c \end{aligned}$$

Thus it makes sense to argue about M in lieu of considering the entirety of L_G.

Recall our definition $L_G = \sum_{uv \in E(G)} y_{uv}y_{uv}^{\mathsf{T}}$ where $y_{uv} = e_u - e_v$. Set v_e to be the first $n-1$ components of $U \cdot y_e$. Note that the last component of Uy_e is zero, since $y_e^T \frac{\vec{1}}{\sqrt{n}} = 0$. Thus $\|v_e\|^2 = \|y_e\|^2 = 2$. Furthermore, since $\sum_{e \in E(K_n)} y_e y_e^{\mathsf{T}} = L_{K_n} = nI_{im_{K_n}}$, we have

$$\begin{aligned} &U\Big(\sum_{e \in E(K_n)} y_e y_e^{\mathsf{T}}\Big)U^{\mathsf{T}} = \left(\begin{smallmatrix} nI & 0 \\ 0 & 0 \end{smallmatrix}\right) \\ \iff &\sum_{e \in E(K_n)} v_e v_e^{\mathsf{T}} = nI \end{aligned}$$

2 How to Construct an Expander

Theorem 2.1. Let $w_1, \ldots, w_m \in \mathbb{R}^p$ satisfy $\sum_i w_i w_i^{\mathsf{T}} = I$ and $\|w_i\|^2 = w_i^{\mathsf{T}} w_i = \frac{p}{m} \forall i$. Choose an integer $k > 4$ arbitrarily.

Then there exists a (multi-)set $F \subseteq [m]$ with $|F| = kp$ such that $\lambda_{min}(\sum_{i \in F} \frac{w_i w_i^{\mathsf{T}}}{w_i^{\mathsf{T}} w_i}) \geq k - 2\sqrt{k}$. Furthermore, there is a deterministic, polynomial-time algorithm to find such a set F.

Proof. The algorithm given below produces an F with the given properties. $\square$

To connect this theorem with expanders, let m be the number of edges in the complete graph, that is $m = \binom{n}{2} = \frac{n(n-1)}{2}$. Then take $p = n - 1$ and let each vector w_i correspond to some $v_e/\sqrt{n}$. Then $\sum_i w_i w_i^\intercal = \sum_e v_e v_e^\intercal/n = I$.

Note that $w_i^\intercal w_i = \frac{v_e^\intercal v_e}{n} = \frac{2}{n}$, so $\frac{w_i w_i^\intercal}{w_i^\intercal w_i} = \frac{v_e v_e^\intercal}{2}$. So the algorithm outputs a (multi-)set F with $|F| = k(n-1)$ such that

$$\sum_{e \in F} \frac{y_e y_e^\intercal}{2} \geq k - 2\sqrt{k}.$$

In other words, letting G be the (multi-)graph with edge set F, we have $\lambda_{min>0}(L_G) \geq 2k - 4\sqrt{k}$, so G is a spectral expander.

2.1 Expander Algorithm

Parameters: $\delta = \frac{1}{p}(1 - \frac{1}{\sqrt{k}}),\ \varepsilon = \frac{p}{\sqrt{k}},\ l_0 = -\frac{p}{\varepsilon} = -\sqrt{k}$
Initially: $F \leftarrow \emptyset,\ A \leftarrow 0,\ l \leftarrow l_0$
Potential: $\Phi_l(A) = trace[(A - lI)^{-1}] = \sum_{i=0}^{p}(\lambda_i - l)^{-1}$ (where λ_i is the i^{th} eigenvalue of A).
For $j = 1, ..., kp$:
  Invariants: (a) $\lambda_{min}(A) > l$, (b) $\Phi_l(A) \leq \varepsilon$
  Find index i such that $\lambda_{min}(A + \frac{w_i w_i^\intercal}{w_i^\intercal w_i}) > l + \delta$, and $\Phi_{l+\delta}(A + \frac{w_i w_i^\intercal}{w_i^\intercal w_i}) \leq \Phi_l(A) \leq \varepsilon$
  $F \leftarrow F \cup i,\ A \leftarrow A + \frac{w_i w_i^\intercal}{w_i^\intercal w_i},\ l \leftarrow l + \delta$
End

Claim 2.2. The invariants hold throughout the running of the algorithm.

Proof. Initially $A = 0$ and $l = l_0$, thus $\lambda_{min}(A) = 0 > l = l_0 = -\sqrt{k}$
Also $\Phi_l(0) = trace[(-lI)^{-1}] = trace[\frac{\varepsilon}{p}I] = \varepsilon$
So the invariants are initially satisfied, and i is chosen specifically as to not violate any invariant, thus the invariants are never violated □

Claim 2.3. The resulting graph is a spectral expander.

Proof. After kp iterations the loop terminates. Thus the number of edges added is $O(n)$.
At this point $l = l_0 + (kp)\delta = -\sqrt{k} + k(1 - \frac{1}{\sqrt{k}}) = k - 2\sqrt{k}$. Thus $\lambda_{min}(A) > k - 2\sqrt{k} = \sum_{i \in F} \frac{w_i w_i^\intercal}{w_i^\intercal w_i}$ (this is positive since $k > 4$).
This implies the graph is connected, since the kernel of its Laplacian is one dimensional. □

Similar to the previous lecture it is not clear that it is possible to find such an index at each iteration. We use a proof very similar to the one used last lecture to show that such an i always exists.
Assume that $\lambda_{min}(A) > l$ and $\Phi_l(A) \leq \frac{1}{\delta}$
Define $M := (A - (l + \delta)I)^{-1},\ N := \frac{M^2}{(\Phi_{l+\delta} - \Phi_l)(A)} - M$.

Observation 2.4. $\lambda_{min}(A) > l + \delta$

Proof. Assume to the contrary that $\lambda_{min}(A) \in (l, l+\delta]$. Then

$$\begin{aligned}\Phi_l(A) &= (\lambda_{min}(A) - l)^{-1} + ... + (\lambda_{max}(A) - l)^{-1} \\ &> (l + \delta - l)^{-1} \\ &= \frac{1}{\delta}\end{aligned}$$

This contradicts our original assumption that $\Phi_l(A) \leq \frac{1}{\delta}$. $\square$

Observation 2.5. $\Phi_{l+\delta}(A) > \Phi_l(A)$.

Proof. $\Phi_{l+\delta}(A) = trace[(A - lI)^{-1}] > trace[(A - (l+\delta)I)^{-1}] = \Phi_l(A)$. $\square$

Lemma 2.6. Set $v \in \mathbb{R}^p$ and $t > 0$ arbitrarily. If $\frac{1}{t} \leq v^\top N v$, then $\Phi_{l+\delta}(A + tvv^\top) \leq \Phi_l(A)$ and $\lambda_{min}(A + tvv^\top) > l + \delta$.

Proof.

$$\begin{aligned}\Phi_{l+\delta}(A + tvv^\top) &= trace[(A + tvv^\top - (l+\delta)I)^{-1}] \\ &= trace[(\underbrace{A - (l+\delta)I}_{M^{-1}} + tvv^\top)^{-1}] \\ &= trace[(M^{-1} + tvv^\top)^{-1}] \\ &= trace[M - \frac{tMvv^\top M}{1 + tv^\top Mv}] \quad \text{By Sherman-Morrison Formula} \\ &= trace[M] - \frac{t \cdot trace(v^\top M^2 v)}{1 + tv^\top Mv} \quad \text{By linearity and cyclic property of trace} \\ &= \Phi_{l+\delta}(A) - \frac{v^\top M^2 v}{\frac{1}{t} + v^\top Mv} \\ &= \Phi_l(A) + (\Phi_{l+\delta}(A) - \Phi_l(A)) - \frac{v^\top M^2 v}{\frac{1}{t} + v^\top Mv}\end{aligned}$$

Thus

$$\begin{aligned}\Phi_{l+\delta}(A + tvv^\top) \leq \Phi_l(A) &\iff (\Phi_{l+\delta}(A) - \Phi_l(A)) - \frac{v^\top M^2 v}{\frac{1}{t} + v^\top Mv} \leq 0 \\ &\iff \frac{1}{t} \leq \frac{v^\top M^2 v}{\Phi_{l+\delta}(A) - \Phi_l(A)} - v^\top Mv = v^\top Nv\end{aligned}$$

And we have $\lambda_{min}(A + tvv^\top) \geq \lambda_{min}(A) > l + \delta$ by observation 2.4. $\square$

Lemma 2.7. $trace(N) \geq \frac{1}{\delta} - \varepsilon$

Proof.

$$\begin{aligned} trace(N) &= \frac{trace(M^2)}{\Phi_{l+\delta}(A) + \Phi_l(A)} - trace(M) \\ &= \frac{\sum_i (\lambda_i - l - \delta)^{-2}}{\sum_i (\lambda_i - l - \delta)^{-1} - \sum_i (\lambda_i - l - \delta)^{-1}} - \sum_i (\lambda_i - l - \delta)^{-1} \\ &= \frac{\sum_i (\lambda_i - l - \delta)^{-2}}{\delta \sum_i (\lambda_i - l - \delta)^{-1} \sum_i (\lambda_i - l - \delta)^{-1}} - \sum_i (\lambda_i - l - \delta)^{-1} \\ &\leq \frac{1}{\delta} - \varepsilon \end{aligned}$$

where the third equality uses the identity $\frac{1}{a} - \frac{1}{a+b} = \frac{b}{a(a+b)}$, and the inequality is a messy use of Cauchy-Schwarz. Details are in the Batson et al. paper. □

Claim 2.8. There is a good index.

Proof.

$$\begin{aligned} \sum_i w_i^\mathsf{T} N w_i &= \sum_i trace(w_i^\mathsf{T} N w_i) \\ &= \sum_i trace(w_i w_i^\mathsf{T} N) \\ &= trace[(\sum_i w_i w_i^\mathsf{T}) N] \\ &= trace[IN] \\ &= trace[N] \\ &\geq \frac{1}{\delta} - \varepsilon \end{aligned}$$

Then the average over the m edges is $\left(\frac{1}{\delta} - \varepsilon\right)/m$, and so $\exists i$ such that:

$$\begin{aligned} w_i^\mathsf{T} N w_i &\geq \frac{\frac{1}{\delta} - \varepsilon}{m} \\ &= \frac{1}{m}\left(p\left(1 - \frac{1}{\sqrt{k}}\right)^{-1} - \frac{p}{k}\right) \\ &\geq \frac{p}{m}\left[\left(1 + \frac{1}{\sqrt{k}}\right) - \frac{1}{\sqrt{k}}\right] \\ &= \frac{p}{m}, \end{aligned}$$

where the second inequality follows from $1/(1-x) \geq 1+x$ for all $x \leq 1$, by Taylor Approximation. So if we set $t = \frac{m}{p} = \frac{\frac{n(n-1)}{2}}{n-1} = \frac{n}{2} = \frac{1}{w_i^\mathsf{T} w_i}$ (from our definitions above) then we can ensure the invariants are satisfied by taking $A \leftarrow A + \frac{w_i w_i^\mathsf{T}}{w_i^\mathsf{T} w_i}$. □

2.2 Comparison with the optimal

The optimal expander graphs, Ramanujan graphs, are d-regular and satisfy

$$\lambda_{min>0}(L_G) \geq d - 2 \cdot \sqrt{d-1}.$$

So their number of edges is exactly $\frac{d}{2}n$.

For the sake of comparison, let us run our algorithm with $k = d/2$, so that $k(n-1) \approx \frac{d}{2}n$. Our algorithm gives us a graph G with $k(n-1)$ edges, and

$$\lambda_{min>0}(L_G) \geq 2k - 4\sqrt{k} = d - 2\sqrt{2} \cdot \sqrt{d}.$$

So the difference between the optimal Ramanujan graphs (which are very difficult to analyze), and the expanders constructed by our algorithm is essentially just a small $\sqrt{2}$ factor in the coefficient of $\sqrt{d}$. Also, Ramanujan graphs are regular whereas G is typically not.