CS 496 Lecture 5: Spectral Algorithms for Planted Problems

Sidhanth Mohanty

September 30, 2025

Last lecture, we proved the Cheeger inequalities, which established a connection between edge expansion and spectral expansion. Today we will discuss how eigenvalues control various combinatorial quantities in graphs such as the independence number, chromatic number and cuts, and then we will see how this can be used to algorithmically find planted cuts, colorings, and independent sets in random graphs.

1 Connections between eigenvalues and cuts, chromatic number, and independent sets

Let G be a graph, and let $L_G = D_G - A_G$ be the *unnormalized* Laplacian of G.

Lemma 1.1. *The size of the max cut in G is bounded by* $\lambda_{\max}(L_G) \cdot n$.

Proof. For any cut described by a set S, let $x_S \in \{\pm 1\}^n$ denote the vector with $+1$ for entries in S, and -1 for entries outside S. The number of edges that participate in the cut is exactly equal to $x_S^\top L_G x_S$, which is at most $\lambda_{\max}(L_G) \cdot \|x_S\|^2 \leqslant \lambda_{\max}(L_G) \cdot n$. □

Corollary 1.2. *The max cut in a d-regular λ-two-sided spectral expander cuts at most $\frac{1+\lambda}{2}$-fraction of the edges.*

Next, we describe a connection between the eigenvalues of L_G and its chromatic number.

Lemma 1.3. *Suppose G has m edges and is q-colorable. Then,* $\lambda_{\max}(L_G) \geqslant \frac{2m}{n} \cdot \frac{q}{q-1}$. *Specialized to a d-regular graph, this is saying that for the adjacency matrix A_G:*

$$-\lambda_{\min}(A_G) \geqslant d \cdot \left(1 - \frac{1}{q-1}\right).$$

Proof. For a coloring $\pi : V \to [q]$, let X_π be the matrix that has blocks of 1s on the diagonal blocks corresponding to the color classes, and $-\frac{1}{q-1}$ on the off-diagonal blocks. It can be verified that X_π is a positive semidefinite matrix since it is block diagonally dominant. On one hand, we have:

$$\langle L_G, X_\pi \rangle = 2m \cdot \frac{q}{q-1}$$

since for any edge e we have $\langle L_e, X_\pi \rangle = \frac{2q}{q-1}$. On the other hand, $\langle L_G, X_\pi \rangle \leqslant \lambda_{\max}(L_G) \cdot \mathrm{tr}(X_\pi) = \lambda_{\max}(L_G) \cdot n$. Combining these gives the desired inequality. □

Remark 1.4. A similar connection can be drawn between the independence number and the top eigenvalue of the Laplacian.

2 Planted problems

We now describe a couple of algorithmic problems on random graphs.

Problem 2.1 (Hidden clique in random graph)**.** Let $\boldsymbol{G}$ be an n-vertex random graph generated by choosing k vertices at random, placing a clique on them, and for every other pair in the graph, placing an edge independently with probability $\frac{1}{2}$. The algorithmic task is to recover the *planted* clique given $\boldsymbol{G}$ as input.

Problem 2.2 (Stochastic block model)**.** Let $\boldsymbol{G}$ be an n-vertex random graph generated by coloring each vertex `red` or `blue` independently, and then placing an edge between vertices of the same color with probability p, and between vertices of opposite colors with probability q.

We focus on the hidden clique model from above with edge probability $1/2$. Let $S \subseteq [n]$ be the planted k-clique and write $\mathbf{1}_S \in \{0,1\}^n$ for its indicator. Define the *centered adjacency* matrix

$$\underline{A}_{\boldsymbol{G}} := A_{\boldsymbol{G}} - \mathbf{E} A_{\boldsymbol{G}},$$

The spectral algorithm to recover the hidden clique, due to Alon, Krivelevich, and Sudakov [AKS98] is as follows:

- Compute the top eigenvector v_1 of $\underline{A}_{\boldsymbol{G}}$.
- Let S' be the k vertices with largest coordinates in v_1.
- Return the set of all vertices with $\geqslant k/2$ neighbors in S'.

It turns out that the returned set is equal to S once $k \gtrsim \sqrt{n}$. Roughly, one may model $\underline{A}_{\boldsymbol{G}}$ as $\boldsymbol{W} + \mathbf{1}_S \mathbf{1}_S^\top$ where $\boldsymbol{W}$ is the "noise" part (it is a random matrix where all the clique entries are 0, and all the nonclique entries are independent uniform ± 1 random variables.

Lemma 2.3 (Spectral norm of the noise)**.** *With probability $1 - o(1)$, $\|\boldsymbol{W}\| \leqslant (2 + o(1))\sqrt{n}$ for an absolute constant C.*

We will not prove this in the current lecture, but will see some of the proof ideas in the future when studying expansion of random graphs.

Lemma 2.4 (Outlier in the clique direction)**.** *We have with high probability*

$$\frac{1}{k}\mathbf{1}_S^\top \underline{A}_{\boldsymbol{G}} \mathbf{1}_S = k \pm O(\sqrt{n})\,.$$

Proof. By expanding $\underline{A}_{\boldsymbol{G}}$ as $\boldsymbol{W} + \mathbf{1}_S \mathbf{1}_S^\top$, we see that $\frac{1}{k}\mathbf{1}_S^\top \underline{A}_{\boldsymbol{G}} \mathbf{1}_S = k + \mathbf{1}_S^\top \boldsymbol{W} \mathbf{1}_S^\top$, where the second term is $\pm O(\sqrt{n})$ by Lemma 2.3. $\square$

In particular, the above proves that $\lambda_{\max}(\underline{A}_{\boldsymbol{G}}) \geqslant k - O(\sqrt{n})$.

Observation 2.5. By Cauchy interlacing, there is only one eigenvalue larger than $k - O(\sqrt{n})$ and it must have magnitude at most $k + O(\sqrt{n})$.

Write v_1 for a unit top eigenvector. We now show that v_1 must correlate strongly with $\mathbf{1}_S$. We may write $\underline{A}_{\boldsymbol{G}} = \widetilde{\boldsymbol{W}} + \lambda v_1 v_1^\top$ where λ is the corresponding eigenvalue, and $\left\|\widetilde{\boldsymbol{W}}\right\| \leqslant (2+o(1))\sqrt{n}$. We have:

$$\begin{aligned} k^2 - O\left(k\sqrt{n}\right) \leqslant \mathbf{1}_S^\top \underline{A}_{\boldsymbol{G}} \mathbf{1}_S &\leqslant \left(k + O(\sqrt{n})\right)\langle \mathbf{1}_S, v_1\rangle^2 + \mathbf{1}_S^\top \widetilde{\boldsymbol{W}} \mathbf{1}_S \\ &\leqslant \left(k + O(\sqrt{n})\right)\langle \mathbf{1}_S, v_1\rangle^2 + O\left(k\sqrt{n}\right). \end{aligned}$$

After some rearrangement, we see for $k = C\sqrt{n}$:

$$\langle \mathbf{1}_S, v_1\rangle^2 \geqslant 1 - O(1/C).$$

It turns out the stochastic block model, which is essentially a planted model for cuts, and also a corresponding model for planted colorings can be algorithmically solved with a spectral algorithm, where the planted solution sticks out as an outlier eigenvector.

References

[AKS98] Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3-4):457–466, 1998. 2