CS 0368-4246: Combinatorial Methods in Algorithms (Spring 2025) June 9th, 2025

Lecture 10: Expanders and Spectral Graph Theory

Instructor: *Or Zamir* Scribes: *Gur Lifshitz*

1 Introduction

Today's lecture is about expander graphs from a spectral graph theory perspective. The main topics are

1) Linear algebra reminder: Spectral decomposition, quadratic form, eigenvalues characterization.

2) Defining the Laplacian of a graph, proving some nice stuff about its eigenvalues and eigenvectors.

3) Cheeger inequality - full proof.

4) Lastly, we will describe an algorithm that finds (fast) a sparse-cut approximation.

2 Linear Algebra Review

Theorem 1 (Spectral Decomposition)**.** *Every real symmetric matrix M (i.e., $M = M^T$) can be diagonalized. That is, there exists an orthonormal basis of eigenvectors $v_1, \ldots, v_n$ such that for all $i \in [n]$,*

$$Mv_i = \lambda_i v_i.$$

In matrix form, this means $M = V^T \Lambda V$, where V is orthogonal and Λ is diagonal.

Definition 2. *Let M be a matrix. The* quadratic form *associated with M is the function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by*

$$x^T M y = \sum_{i,j} M_{i,j} x_i y_j$$

for all vectors x, y.

Claim 3 (Maximum Eigenvalue Characterization)**.** *If the eigenvalues of M are ordered as $\lambda_1 \leq \cdots \leq \lambda_n$, then*

$$\lambda_n = \max_{x \neq 0} \frac{x^T M x}{x^T x}.$$

Proof.

$$\max_{x^T x = 1} x^T M x = \max_{y^T y = 1} y^T \Lambda y = \max \sum \lambda_i y_i^2 \leq \lambda_n,$$

where we used the change of variables $y = V^{-1} x$. □

Claim 4 (Minimum and Successive Eigenvalue Characterization)**.** *With the same ordering of eigenvalues,*

$$\lambda_1 = \min_{x \neq 0} \frac{x^T M x}{x^T x}.$$

Furthermore, if v_1 is an eigenvector corresponding to λ_1, then

$$\lambda_2 = \min_{\substack{x \neq 0 \\ x \perp v_1}} \frac{x^T M x}{x^T x}.$$

We can further extend this equality for every λ_i.

3 The Laplacian of a Graph

Definition 5. *Let A_G be the adjacency matrix of a graph G. The* Laplacian matrix *of G is defined as*

$$L_G = D_G - A_G,$$

where D_G is the diagonal degree matrix, i.e., $D_{i,i}$ equals the degree of vertex $v_i \in V$.

Observation 6. *Note that*

$$L_G = \sum_{(u,v) \in E} L_{(u,v)},$$

where $L_{(u,v)}$ is the Laplacian contribution of the edge (u, v).

Let's compute the quadratic form:

$$x^T L_{(u,v)} x = x_u^2 - 2x_u x_v + x_v^2 = (x_u - x_v)^2.$$

Therefore,

$$x^T L_G x = \sum_{(u,v) \in E} (x_u - x_v)^2.$$

Now consider $S \subseteq V$. Observe that

$$\mathbb{1}_S^T L_G \mathbb{1}_S = e(S, S^c),$$

where $e(S, S^c)$ denotes the number of edges crossing from S to its complement.

Claim 7. *Let G be an undirected graph with Laplacian eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then:*

1. *$\lambda_1 = 0$ with eigenvector $\vec{1}$.*
2. *G is connected if and only if $\lambda_2 > 0$.*

Proof of (1). Since $x^T L_G x \geq 0$ for all x, we have $\lambda_1 = \min_{x^T x = 1} x^T L_G x \geq 0$. Moreover,

$$L_G \vec{1} = (D_G - A_G)\vec{1} = \vec{d} - \vec{d} = 0,$$

so $\lambda_1 = 0$ with eigenvector $\vec{1}$. □

Proof of (2). $\Rightarrow$ If G is disconnected, let $S \subset V$ be the vertex set of a connected component. Then $\mathbb{1}, \mathbb{1}_S$ are in $\ker(L_G)$, so $\dim(\ker(L_G)) \geq 2$ and $\lambda_2 = 0$.

$\Leftarrow$ If $\lambda_2 = 0$, then there exists $x \perp \vec{1}$ with $x^T L_G x = 0$. This implies $(x_u - x_v)^2 = 0$ for all $(u, v) \in E$, so $x_u = x_v$ on each connected component. If G were connected, this would force $x \propto \vec{1}$, contradiction. □

Exercise 8. *Show that $\lambda_k = 0$ if and only if G has at least k connected components.*

Proof. Suppose G has t connected components: $G = G_1 \cup \cdots \cup G_t$. Then

$$L_G = \operatorname{diag}(L_{G_1}, \ldots, L_{G_t}).$$

Thus,

$$\ker(L_G) = \ker(L_{G_1}) \oplus \cdots \oplus \ker(L_{G_t}).$$

Since $\dim \ker(L_{G_i}) = 1$, we have $\dim \ker(L_G) = t$. Therefore, $\lambda_1 = \cdots = \lambda_t = 0$, and $\lambda_{t+1} > 0$ if G has exactly t components. □

Does λ_2 provide stronger information about the connectivity of G?

Definition 9. *The* normalized Laplacian *is defined as*

$$N_G = D_G^{-1/2} L_G D_G^{-1/2} = I - D_G^{-1/2} A_G D_G^{-1/2}.$$

(Note: We could consider $D_G^{-1} L_G$, but it is not symmetric. Observe that $D_G^{-1/2}(D_G^{-1} L_G) D_G^{1/2} = N_G$ and those matrices are similar.)

Observation 10. *The smallest eigenvalue satisfies*

$$\lambda_1(N_G) = 0,$$

with eigenvector $v_1 = D_G^{1/2}\vec{1} = \sqrt{\vec{d}}$.

Claim 11.

$$\lambda_2(N_G) = \min_{\substack{x \neq 0 \\ x \perp \vec{d}}} \frac{x^T L_G x}{x^T D_G x}.$$

Proof.

$$\begin{aligned}
\lambda_2(N_G) = &\min_{\substack{y \neq 0 \\ y \perp \sqrt{\vec{d}}}} \frac{y^T N_G y}{y^T y} \\
&= \min_{\substack{x \neq 0 \\ x \perp \vec{d}}} \frac{x^T D_G^{1/2} N_G D_G^{1/2} x}{x^T D_G x} \\
&= \min_{\substack{x \neq 0 \\ x \perp \vec{d}}} \frac{x^T L_G x}{x^T D_G x},
\end{aligned}$$

where we set $y = D_G^{1/2} x$. □

4 Cheeger Inequality

Definition 12 (Reminder: Graph Conductance)**.** *The conductance of a graph G is defined as*

$$\varphi(G) = \min_{x=\mathbb{1}_S \ 0<\mathrm{Vol}(S)\leq \frac{1}{2}\mathrm{Vol}(V)} \frac{x^T L_G x}{x^T D_G x}.$$

Note that conductance is closely related to λ_2: the expression is the same, but for λ_2, we minimize over all vectors orthogonal to $\vec{d}$, while for conductance we minimize over indicator vectors of subsets.

We will now show that they are indeed closely related. The following inequality was proved by Dodziuk [2] and independently Alon and Milman [1] and states that:

Theorem 13 (Cheeger Inequality)**.**

$$\frac{1}{2}\lambda_2(N_G) \leq \varphi(G) \leq \sqrt{2\lambda_2(N_G)}.$$

Observation 14. *1. G is a good expander (i.e., $\varphi(G)$ is constant) if and only if λ_2 is constant.*

2. If λ_2 is very small (e.g., $O(1/\sqrt{n})$), this inequality is not tight.

Claim 15. *There exists an algorithm that checks whether G is an expander in $O(n^3)$ time (or faster using fast matrix multiplication).*

Proof of Cheeger Inequality. $\Rightarrow$ We aim to prove $\lambda_2(N_G) \leq 2\varphi(G)$.

Take a cut $S \subset V$ with $0 < \mathrm{Vol}(S) \leq \frac{1}{2}\mathrm{Vol}(V)$ where

$$\varphi(G) = \frac{\mathbb{1}_S^T L_G \mathbb{1}_S}{\mathbb{1}_S^T D_G \mathbb{1}_S}.$$

Define $x = \mathbb{1}_S - \sigma\vec{1}$ where σ ensures $x \perp \vec{d}$:

$$\vec{d}^{\,T} x = 0 \Rightarrow \mathrm{Vol}(S) - \sigma\,\mathrm{Vol}(V) = 0 \Rightarrow \sigma = \frac{\mathrm{Vol}(S)}{\mathrm{Vol}(V)}.$$

Since $0 < \mathrm{Vol}(S) \leq \frac{1}{2}\mathrm{Vol}(V)$, we have $0 < \sigma \leq \frac{1}{2}$.

We compute:

$$x^T L_G x = (\mathbb{1}_S - \sigma\,\vec{1})^T L_G\,(\mathbb{1}_S - \sigma\,\vec{1}) = \mathbb{1}_S^T L_G \mathbb{1}_S,$$

where the last equality holds because $L_G\vec{1} = \vec{0}$.

Next, expanding the denominator:

$$\begin{aligned}
x^T D_G x &= (\mathbb{1}_S - \sigma\,\vec{1})^T D_G (\mathbb{1}_S - \sigma\,\vec{1}) \\
&= \mathbb{1}_S^T D_G \mathbb{1}_S - 2\sigma\,\mathbb{1}_S^T D_G \vec{1} + \sigma^2\,\vec{1}^T D_G \vec{1} \\
&= \mathrm{Vol}(S) \ - \ 2\sigma\ \mathrm{Vol}(S) \ + \ \sigma^2\ \mathrm{Vol}(V) \\
&= (1-\sigma)\ \mathrm{Vol}(S) \ \geq \ \tfrac{1}{2}\ \mathrm{Vol}(S).
\end{aligned}$$

Therefore,

$$\frac{x^T L_G x}{x^T D_G x} \leq 2\,\varphi(G).$$

$\Leftarrow$ Start from a "best" vector $x \neq 0$, $x \perp \vec{d}$ that minimizes

$$\frac{x^T L_G x}{x^T D_G x},$$

and convert it into a sparse cut.

First Step – Centralization. Reindex so that $x_1 \leq x_2 \leq \cdots \leq x_n$. and for each j define the prefix set

$$S_j = \{v_1, v_2, \ldots, v_j\}.$$

Take j to be the smallest index satisfying

$$\mathrm{Vol}(S_j) = \sum_{u \in S_j} \deg(u) \ \geq \ \tfrac{1}{2}\ \mathrm{Vol}(V).$$

Then set

$$y = x - x_j\,\vec{1}.$$

By construction, the total degree of the positive entries of y is at most $\frac{1}{2}$ $\mathrm{Vol}(V)$ (and similarly for the negative entries). Moreover,

$$y^T L_G y \ = \ (x - x_j\vec{1})^T L_G (x - x_j\vec{1}) \ = \ x^T L_G x,$$

since $L_G\vec{1} = \vec{0}$.

On the other hand,

$$\begin{aligned} y^T D_G y &= (x - x_j\vec{1})^T D_G (x - x_j\vec{1}) \\ &= x^T D_G x \ - \ 2\,x_j\,(x^T D_G \vec{1}) \ + \ x_j^2\,(\vec{1}^T D_G \vec{1}) \\ &= x^T D_G x \ - \ 2\,x_j\ \mathrm{Vol}(S) \ + \ x_j^2\ \mathrm{Vol}(V) \ \leq \ x^T D_G x. \end{aligned}$$

(Note: although now $y \not\perp \vec{d}$, we no longer require orthogonality—our goal in the next step is to convert y into an indicator vector, so preserving the quotient bound is sufficient.)

Second Step – Split Positive and Negative Parts.

Write

$$y = y^+ - y^-,$$

where $y_u^+ = \max\{y_u, 0\}$ and $y_u^- = \max\{-y_u, 0\}$. Then

$$y^T L_G y = \sum_{(u,v)\in E} (y_u - y_v)^2 = \sum_{(u,v)\in E} \big((y_u^+ - y_v^+) - (y_u^- - y_v^-)\big)^2$$

$$= \sum_{(u,v)\in E} \Big[(y_u^+ - y_v^+)^2 - 2(y_u^+ - y_v^+)(y_u^- - y_v^-) + (y_u^- - y_v^-)^2\Big] \;\geq\; (y^+)^T L_G y^+ \;+\; (y^-)^T L_G y^-,$$

since each term $(y_u^+ - y_v^+)(y_u^- - y_v^-) \geq 0$.

Likewise, for the denominator,

$$y^T D_G y = (y^+ - y^-)^T D_G (y^+ - y^-) = (y^+)^T D_G y^+ \;+\; (y^-)^T D_G y^-,$$

because $y_u^+ \, y_u^- = 0$ for every u.

Therefore,

$$\frac{(y^+)^T L_G y^+ + (y^-)^T L_G y^-}{(y^+)^T D_G y^+ + (y^-)^T D_G y^-} \;\leq\; \frac{y^T L_G y}{y^T D_G y}.$$

Now, using the fact that $\min\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\} \leq \frac{a_1+a_2}{b_1+b_2}$, we get that if z if either y^+ or y^-:

$$\frac{z^T L_G z}{z^T D_G z} \;\leq\; \frac{y^T L_G y}{y^T D_G y}.$$

Claim 16. *Let $\alpha_i \geq 0$ with $\sum_i \alpha_i = 1$, and let $a_i, b_i > 0$. Then*

$$\frac{\sum_i \alpha_i a_i}{\sum_i \alpha_i b_i} \;\geq\; \min_i \frac{a_i}{b_i}.$$

We will use its continuous version soon.

Thus, using z as above, normalize so $\max_i z_i = 1$. For a random threshold τ with $\tau \sim U(0,1)$ define

$$S_\tau = \{\, i : z_i^2 > 1 - \tau \}.$$

Then

$$\mathbb{E}_\tau\big[\mathbb{1}_{S_\tau}^T D_G \mathbb{1}_{S_\tau}\big] = \sum_u d(u)\,\Pr[i \in S_\tau] = \sum_u d(u)\, z_u^2 = z^T D_G z,$$

and

$$\mathbb{E}_\tau\big[\mathbb{1}_{S_\tau}^T L_G \mathbb{1}_{S_\tau}\big] = \sum_{(u,v)\in E} \Pr\big(\tau \in [\,1 - z_u,\, 1 - z_v\,]\big) = \sum_{(u,v)\in E} |z_u^2 - z_v^2| \;\leq\; \sqrt{2\, z^T L_G z \;\cdot\; z^T D_G z},$$

where the last step uses Cauchy–Schwarz and the fact that

$$\sum_{(u,v)\in E} (z_u + z_v)^2 \;\leq\; 2 \sum_u d(u) z_u^2 = 2 z^T D_G z.$$

Therefore, by applying Claim 16 we get

$$\min_\tau \frac{\mathbb{1}_{S_\tau}^T L_G \mathbb{1}_{S_\tau}}{\mathbb{1}_{S_\tau}^T D_G \mathbb{1}_{S_\tau}} \;\leq\; \sqrt{\frac{2\, z^T L_G z}{z^T D_G z}} \;=\; \sqrt{2\,\lambda_2(N_G)},$$

showing there exists τ with conductance at most $\sqrt{2\,\lambda_2(N_G)}$.

□

Claim 17. *Let $x \neq 0$ satisfy $x \perp \vec{d}$ and*

$$\frac{x^T L_G x}{x^T D_G x} \;\leq\; \gamma.$$

Then in $O(m + n\log n)$ time one can find a cut of conductance at most $\sqrt{2\gamma}$.

Proof. We consider only cuts of the form

$$S_\tau = \{\, v : x_v \leq \tau \},\quad S_\tau^c = V \setminus S_\tau.$$

Sort the values x_v in $O(n\log n)$ time. As we sweep through the sorted list, we maintain the current cut and update its boundary size in $O(1)$ per edge (for a total of $O(m)$). Finally, we return the threshold τ that minimizes $\dfrac{e(S_\tau, S_\tau^c)}{\mathrm{Vol}(S_\tau)}$. □

Remark. Computing the exact second eigenvector v_2 can be expensive, but one can use fast approximate solvers to obtain a vector x whose quotient is within a small constant factor of $\lambda_2(N_G)$.

Exercise 18 (Expander Mixing Lemma)**.** *Let G be a d-regular graph on n vertices with λ being the second-largest eigenvalue of A_G. Show that for any $S, T \subseteq V$,*

$$\left| e(S,T) - \tfrac{d}{n}\,|S|\,|T| \right| \;\leq\; \sqrt{\lambda\,|S|\,|T|}.$$

References

[1] Noga Alon and Joel H. Spencer. *The Probabilistic Method.* Wiley Publishing, 4th edition, 2016.

[2] Jozef Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Transactions of the American Mathematical Society*, 284(2):787–794, 1984.