CS 252, Lecture 6: Spectral Graph Theory

Guest Lecture by Pravesh Kothari

1 Introduction

In this lecture, we give an overview of spectral graph theory, where in we use tools from Linear Algebra to study graphs. We demonstrate how we can "read off" combinatorial properties of graphs from their linear algebra properties. In particular, we focus on the problem of connectivity of graphs, and study how we can calculate the number of connected components of a graph from reading eigenvalues of its adjacency matrix.

We denote a graph G by $G = (V, E)$ where V is the set of vertices, and E is the set of edges. In this lecture, we assume all the graphs are undirected, unweighted, and there are no multiple edges or self loops. We also assume that the graphs that we study are d-regular i.e. all the vertices have degree d. Let $n = |V|$ denote the number of vertices in the graph.

We start with defining adjacency matrix and Laplacian of a graph.

Definition 1 (Adjacency matrix)**.** Given a graph G, the adjacency matrix A_G is defined as the follows:

$$A_G(i,j) = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise.} \end{cases}$$

The Laplacian matrix of a graph is defined similar to the adjacency matrix, is easier to use and generalizes well to graphs that are not regular and weighted graphs.

Definition 2 (Laplacian of a graph)**.** Given a graph G, the Laplacian matrix $L_{n\times n}$ is defined as $dI_{n\times n} - A$, where $I_{n\times n}$ is the $n \times n$ identity matrix. In other words, we have

$$L_G(i,j) = \begin{cases} d & \text{if } i = j \\ -1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise.} \end{cases}$$

2 Linear Algebra Background

Consider a real symmetric matrix $M \in \mathbb{R}^{n\times n}$ i.e. $M_{i,j} = M_{j,i}$ for all $i \neq j$.

Definition 3 (Eigenvalue, Eigenvectors)**.** A scalar λ is called *eigenvalue* of M if there exist a vector $x \neq 0$ such that $Ax = \lambda x$. The corresponding vector x is called *eigenvector*.

Fact 4. *The following are standard facts about eigenvalues of a real symmetric matrix M:*

1. *M has n real eigenvalues (including repetitions) $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.*
2. *There exist eigenvectors $v_1, v_2, \ldots, v_n$ such that $Mv_i = \lambda_i v_i$ for all $1 \leq i \leq n$, and the set of vectors $v_1, v_2, \ldots, v_n$ are linearly independent. The eigenvectors corresponding to distinct eigenvectors are orthogonal.*

Henceforth, we let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of M. The corresponding eigenvectors are denoted by $v_1, v_2, \ldots, v_n$. We assume that these n vectors are unit vectors, and are mutually orthogonal i.e. $\langle v_i, v_j \rangle = 0$ for all $i \neq j$.

Definition 5 (Quadratic form)**.** The quadratic form of a matrix M is the function $M(v) = v^T M v$ that outputs a scalar on taking a vector from $\mathbb{R}^n$ as input.

Lemma 6. *Let M be a real symmetric matrix. $v^T M v \geq 0$ for all vectors $v \in \mathbb{R}^n$ if and only if all the eigenvalues of M are non-negative.*

Proof. Suppose that $v^T M v \geq 0$ for all $v \in \mathbb{R}^n$. Then, we claim that all the eigenvalues are non-negative. Suppose for contradiction that there is $\lambda < 0$ and a vector x such that $Mx = \lambda x$. We get $x^T M x = \lambda x^T x = \lambda(\sum_{i=1}^n x_i^2) < 0$, contradicting the fact that $v^T M v \geq 0$ for all vectors v.

Suppose that all the eigenvalues of M are non-negative. Then, for an arbitrary vector v, let $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$. Such an expression is possible for every vector v since the vectors $v_1, v_2, \ldots, v_n$ span the whole of $\mathbb{R}^n$. We have $v^T M v = \sum_{i=1}^n \lambda_i \langle v_i, v_i \rangle \geq 0$. □

A real symmetric matrix with all eigenvalues being non-negative is called as a Positive Semidefinite matrix.

Lemma 7. *(Rayleigh Coefficient) We have*

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

and

$$\lambda_2 = \min_{v \neq 0, \langle v, v_1 \rangle = 0} \frac{v^T M v}{v^T v}$$

Proof. We only prove the first part, the proof of the second part follows very much along the same lines. As $\lambda_1 = \frac{v_1^T M v_1}{v_1^T v_1}$, we get $\lambda_1 \geq \min_{v \neq 0} \frac{v^T M v}{v^T v}$. Furthermore, let $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$. We have $\frac{v^T M v}{v^T v} = \frac{\sum_{i=1}^n \alpha_i^2 \lambda_i}{\sum_{i=1}^n \alpha_i^2} \geq \lambda_1$ for all $v \neq 0$. □

Recall the Laplacian of the graph $G : L_G$. For simplicity, we use L to mean L_G. For a vector $u = (u_1, u_2, \ldots, u_n)$, we have

$$(Lu)_i = du_i - \sum_{j:(i,j) \in E} u_j = \sum_{j:(i,j) \in E} (u_i - u_j)$$

and

$$u^T L u = \sum_{(i,j)\in E} (u_i - u_j)^2 \tag{1}$$

Using this, we can observe that L is Positive Semidefinite matrix.

When we project the vertices of G on to the real line such that the vertex i is mapped to u_i, the quadratic form $u^T L u$ denotes the sum of squares of distances between neighboring vertices of G.

3 Graph Connectivity

Let G be an undirected unweighted d-regular graph and let L be the Laplacian matrix of G. Let $0 \le \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ be the eigenvalues of L.

The smallest eigenvalue of L is always zero:

Lemma 8. $\lambda_1 = 0$.

Proof. As L is a positive semidefinite matrix, $\lambda_1 \ge 0$. Setting $u = (1, 1, \ldots, 1)$, we have $Lu = (0, 0, \ldots, 0)$, thus proving that 0 is an eigenvalue of L. □

We can deduce if the graph G is connected or not by reading the next eigenvalue of L ! We prove the fact below:

Lemma 9. $\lambda_2 > 0$ *if and only if the graph is connected. More generally, t he number of zero eigenvalues of* L *is equal to the number of connected components of* G*.*

Proof. Let k be the number of connected components of G and $S_1, S_2, \ldots, S_k$ be the connected components of G.

First, we will prove that the number of zero eigenvalues is at least the number of connected components. Towards this, we define the following set of k vectors $u^{(1)}, u^{(2)}, \ldots, u^{(k)}$ where $u^{(i)}$ is defined as follows:

$$u^{(i)}(j) = \begin{cases} 1, & \text{if } j \in S_i \\ 0, & \text{otherwise.} \end{cases}$$

Note the following:

1. $\langle u^{(i)}, u^{(j)} \rangle = 0$ for all $i \neq j$.

2. $Lu^{(i)} = 0$ for all $i \in \{1, 2, \ldots, k\}$.

Thus, there are k mutually orthogonal vectors that are all eigenvectors of L corresponding to the eigenvalue 0. In other words, the number of zero eigenvalues of L is at least k.

Next, we will prove that the number of zero eigenvalues is at most k. From Equation (1), we can deduce that $u^T L u = 0$ if and only if $u_i = u_j$ for all $(i, j) \in E$. That is, u should have equal value for all vertices in a connected component. Thus, if $Lu = 0$, then $u^T L u = 0$, which implies that there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_k$ such that $u = \sum_{i=1}^{k} \alpha_i u^{(i)}$. This implies that every eigenvector of L corresponding to the eigenvalue 0 is contained in the subspace spanned by $\{u^{(1)}, u^{(2)}, \ldots, u^{(k)}\}$. Thus, there are at most k linearly independent eigenvectors of L corresponding to the eigenvalue 0. This proves that the eigenvalue 0 of M is repeated at most k times. $\square$

4 Expansion of Graphs

We have seen that $\lambda_2 > 0$ if and only if the graph is connected. Can we say something more general? Does the magnitude of λ_2 have any meaning? It turns out that the absolute value of λ_2 indicates how much "robustly connected" G is. If λ_2 is large, the graph is "very" connected i.e. for every set $S \subseteq V$, a good fraction of edges originating in S cross S.

We first start with a few definitions.

Definition 10 (Cut/Boundary)**.** For a set $S \subseteq V$, the *boundary* of S, denoted by δS is the number of edges that are adjacent to S that go out of S. That is,

$$\delta(S) = \sum_{i \in S} |\{j : (i,j) \in E\}|$$

Definition 11 (Conductance/Isoperimetric ratio of a set)**.** For a set $S \subseteq V$, a scaled version of the fractional number of edges originating in S that cross it is called as the Conductance of S and denoted by $\Theta(S)$. To be precise,

$$\Theta(S) = \frac{\delta(S)}{d\frac{|S|}{n}|V \setminus S|}$$

Note that the above definition is within factor 2 of the fraction of the edges originating in S that go out of S.

Definition 12 (Conductance of a graph)**.** The conductance of a graph G is defined as $\Theta(G) = \min_{S \subseteq V} \Theta(S)$.

The conductance of a graph is directly related to the second eigenvalue λ_2 of the Laplacian of the graph:

Theorem 13 (Cheeger's inequality)**.** $\frac{\lambda_2}{2d} \leq \Theta(G) \leq \sqrt{\frac{2\lambda_2}{d}}$.

Finally, we define Expander graphs, the graphs which are sparse, yet well connected.

Definition 14 (Expander Graphs)**.** A sequence of d-regular graphs $G_1, G_2, \ldots$ is called an Expander if there exist an absolute constant $C > 0$ such that $\frac{\lambda_2(G_i)}{d} \geq C$ for all i.

Expanders have various applications, both in theoretical computer science and beyond, ranging from fault tolerant networks to construction of error correcting codes.