Lecture 9: Spectral Graph Theory

Foundations of Data Science: Algorithms and Mathematical Foundations

Mihai Cucuringu
mihai.cucuringu@stats.ox.ac.uk

CDT in Mathematics of Random System
University of Oxford

22 September, 2023

Spectral methods

Broadly speaking, they define a class of algorithms with

- INPUT: a matrix (most often, a square matrix)
- OUTPUT: insights obtained by leveraging linear algebraic techniques (most often, based on the eigenvectors or singular vectors of the matrix)

Huge success in a variety of domains, including

- network analysis
- graph partitioning
- data analysis
- website ranking
- text classification
- collaborative filtering

Recall the definition of the graph Laplacian

- Graph Laplacian $L = D - A$ (most popular version)
- A is the adjacency matrix of the graph $A_{ij} \geq 0$
- D is a diagonal matrix, D_{ii} denoting the degree of node i

$$L(i,j) \stackrel{\text{def}}{=} \begin{cases} deg(v_i) & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } (i,j) \in E(G) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

- L is symmetric
- eigenvalues $\lambda_0 \leq \lambda_1 \leq \lambda_{n-1}$, eigenvectors $\phi_0, \phi_1, \ldots, \phi_{n-1}$
- every row sum and column sum of L is zero
- thus, $\lambda_0 = 0$, and $\phi_0 = \mathbf{1} \stackrel{\text{def}}{=} [1, 1, \ldots, 1]^T$ since $L\,\mathbf{1} = 0\,\mathbf{1}$
- the second smallest (smallest non-zero) eigenvalue of L is the algebraic connectivity (Fiedler value, spectral gap) of G

Lemma If $G = (V, E)$ is **connected** and $\lambda_0 \leq \lambda_1 \leq \lambda_{n-1}$ are the eigenvalues of its Laplacian L, then it holds true that $\lambda_1 > 0$. (Stronger result: the multiplicity of the zero eigenvalue is equal to the number of connected components).

Properties of the (random-walk) Laplacian matrix P

- Let W denote the adjacency matrix of a weighted graph.
- The random-walk normalized Laplacian matrix is $P = D^{-1}W$;

or $I - D^{-1}W$, but recall

$$(I - P)x = x - Px = x - \lambda x = (1 - \lambda)x \tag{2}$$

Lemma All the eigenvalues of $P = D^{-1}W$ satisfy $|\lambda_i| \leq 1$, $\forall i = 1, \ldots, n$

- Let λ be an eigenvalue of P with associated eigenvector x

$$\lambda\, x = P\, x \tag{3}$$

- Let $i_m = \underset{1 \leq i \leq n}{\operatorname{argmax}} |x_i|$,
- Consider the following

$$\lambda x_{i_m} = \sum_{j=1}^{n} P_{i_m j}\, x_j \tag{4}$$

thus

$$|\lambda| = \left| \sum_{j=1}^{n} P_{i_m j}\, \frac{x_j}{x_{i_m}} \right| \leq \sum_{j=1}^{n} P_{i_m j} \left| \frac{x_j}{x_{i_m}} \right| \leq \sum_{j=1}^{n} P_{i_m j} = 1 \tag{5}$$

Combinatorial Laplacian $L = D - A$

We can further redefine this as follows. Let $G_{1,2}$ be the graph on two vertices u and v and one edge $e_{u,v}$

- Define
$$L_{G_{1,2}} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$
- If $x = [x_1 \ x_2]^T$, note that
$$x^T L_{G_{1,2}} x = (x_1 - x_2)^2$$
- Let $G_{u,v}$ denote a graph on n vertices with only one edge (between u and v)
- Define the Laplacian of $G_{u,v}$ as

$$L_{G_{u,v}}(i,j) = \begin{cases} 1 & \text{if } i = j \quad \text{and} \quad i \in \{u, v\} \\ -1 & \text{if } i = u \quad \text{and} \quad j = v, \text{ or vice versa} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Combinatorial Laplacian

For a general graph $G = (V, E)$ define

$$L(G) \stackrel{\text{def}}{=} \sum_{(u,v)\in E} L_{G_{u,v}}$$

- Note that $L_{G_{1,2}}$ has eigenvalues 0 and 2, and so is positive semidefinite (PSD)
- recall that a symmetric matrix M is PSD if all of its eigenvalues are non-negative
- recall equivalent PSD condition

$$x^T M x \geq 0, \text{ for all } x \in \mathbb{R}^n$$

- using the previous observation that

$$x^T L_{G_{1,2}} x = (x_1 - x_2)^2$$

  we can show that the Laplacian of every graph is PSD

$$x^T L_G\, x = \sum_{(u,v)\in E} (x_u - x_v)^2 \geq 0 \tag{7}$$

Spectral gap of the Combinatorial Laplacian

Lemma If $G = (V, E)$ is a connected graph and $\lambda_0 \leq \lambda_1 \leq \lambda_{n-1}$ are the eigenvalues of its (Combinatorial) Laplacian L, then it holds true that $\lambda_1 > 0$

Proof:

- let x be an eigenvector of L with associated eigenvalue 0

$$Lx = \mathbf{0}$$

- then it must also hold true that

$$x^T L\, x = \sum_{(u,v) \in E} (x_u - x_v)^2 = 0$$

- thus, for each pair of nodes u, v connected by an edge, it holds true that $x_u = x_v$
- since G is connected, $x_u = x_v$ for all pairs of vertices u, v, which implies that x is some constant multiple of the all ones vector $\mathbf{1} = [1, 1, \ldots, 1]$
- eigenspace of $\lambda_0 = 0$ has dimension 1, and thus $\lambda_1 > \lambda_0$.

What about the spectrum of the adjacency matrix A

Let G be a simple graph (no self-loops), w. adjacency matrix A

Because $Tr(A) = 0$, it holds true that the sum of all eigenvalues of a A is always 0.

The complete graph K_n

- the graph where all pairs of nodes are connected
- has an adjacency matrix equal to

$$A = J - I$$

- J is the all-1's matrix
- I is the identity matrix
- For J: $\text{rank}(J) = 1$, only one nonzero eigenvalue equal to n (with an eigenvector $\mathbf{1} = [1, 1, \ldots, 1]$), and all the remaining eigenvalues are 0
- note that subtracting the identity shifts all eigenvalues by -1, because (say x is an eigenvector of J with eigenvalue λ)

$$Ax = (J - I)x = Jx - x = \lambda x - x = (\lambda - 1)x$$

- thus the eigenvalues of K_n are $n - 1$ and -1 (of multiplicity $n - 1$)

d-regular graphs

- Graphs where every node is of degree d
- If G is d-regular, then **1**= [1,1,. . . ,1] is an eigenvector.
- A **1** = d **1**, and hence d is an eigenvalue.
- In general graphs, the largest eigenvalue reveals information about the average degree and largest degree

Lemma: For a d-regular graph $|\lambda_i| \leq d, \forall i = 1, n$

Proof:

- Let (λ, x) be an eig-value, eig-vector pair
- Let x_m be the component of maximum absolute value

$$\lambda x_m = \sum_{j=1}^{n} A_{mj} x_j$$

$$|\lambda x_m| = |\sum_{j=1}^{n} A_{ij} x_j| \leq \sum_{j=1}^{n} A_{ij} |x_j| \leq \sum_{j=1}^{n} A_{ij} |x_m| = d|x_m|$$

$$\text{thus} \quad |\lambda| \leq d$$

d-regular graphs

Complement graph $\overline{G}$: two distinct vertices of $\overline{G}$ are adjacent if and only if they are not adjacent in G.

If $d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are all the eigenvalues of G, then the eigenvalues of the complement graph $\overline{G}$ are given by

$$(n-1-d) \quad \text{and} \quad -1-\lambda_i, \quad \forall i = 2, 3, \ldots, n-1, n$$

- Adjacency matrix of the complement graph $\overline{G}$ is given by

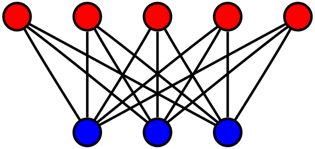
$$\overline{A} = J - I - A \tag{8}$$

- Since $\overline{G}$ is $(n-1-d)-regular$, the largest eigenvalue of $\overline{A}$ is $n-1-d$, and $\mathbf{1}$ is an eigenvector
- Any other eigenvalue must have an eigenvector $x \perp \mathbf{1}$, thus

$$\overline{A}x = (J-I-A)x = (0-1-\lambda)x = (-1-\lambda)x \tag{9}$$

Complete bipartite graph $K_{m,n}$

Exercise: what are its eigenvalues?



- rank(A) = 2
- eigenvalue 0 with multiplicity $n-2$, $\lambda_1 = \lambda_2 = \ldots = \lambda_{n-2} = 0$
- $\lambda_{n-1}, \lambda_n \neq 0$, $\lambda_{n-1} + \lambda_n = 0$
- let $\lambda = |\lambda_{n-1}| = |\lambda_n|$
- Can you find λ by solving $Ax = \lambda x$?
- Hint: due to symmetry, try $x = [\alpha \;\; \beta]^T$