Introduction to Spectral Graph Theory

Ekene Ezeunala

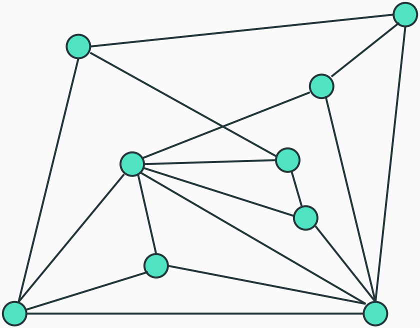
September 30th, 2023

Graphs & Matrices

Graphs

Definition

A graph is a pair $G = (V, E)$, where V is a set whose elements are called vertices and E is a set of paired vertices, whose elements are called edges.



Graphs

We will (mostly) consider cases where G:

- is finite
- has no $v \in V$ s.t. $\deg(v) = 0$
- could have self-loops and parallel edges
- is not necessarily connected
- unweighted

Adjacency matrix

Definition (Adjacency matrix)

Let $\mathcal{G}$ be a graph with vertices $v_1, v_2, \ldots, v_n$. Then the *adjacency matrix* of $\mathcal{G}$ is the matrix $A \in \mathrm{Mat}(n \times n; \{0, 1\})$, whose (i, j) entry, denoted by $[A]_{i,j}$, is defined by

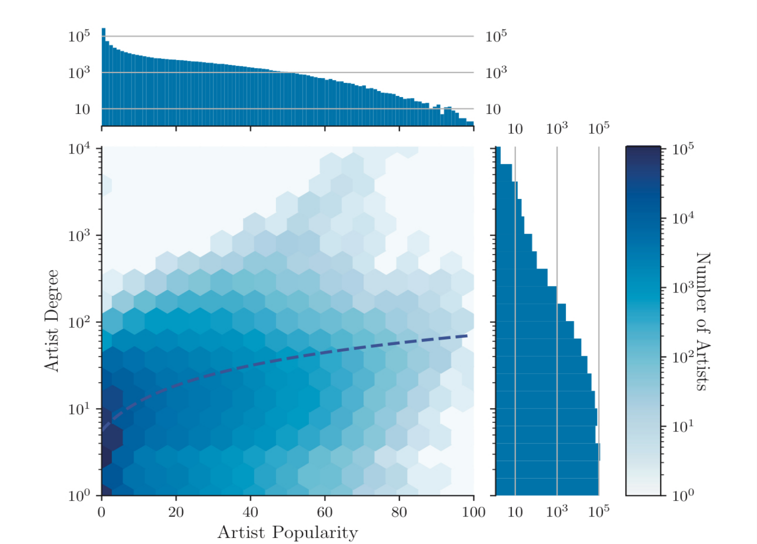
$$[A]_{i,j} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

A network analysis of Spotify [SRM'21]

- A network of all the artists on Spotify connected by who they worked with
- 1,250,065 artists (vertices on undirected graph)
- 3,766,631 collaborations (edges on undirected graph)
- Snowball sampling starting with Kanye West
- Get metadata on these artists (eg popularity, etc)
- Represent using adjacency matrix:

$$A = \begin{array}{c} \\ \text{Kanye} \\ \text{Drake} \\ \text{Taylor} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Kanye} & \text{Drake} & \text{Taylor} \end{array} \\ \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

Relative popularity



Eigenvector centrality

- Calculate the *eigenvector centrality*: $Av = \lambda v$

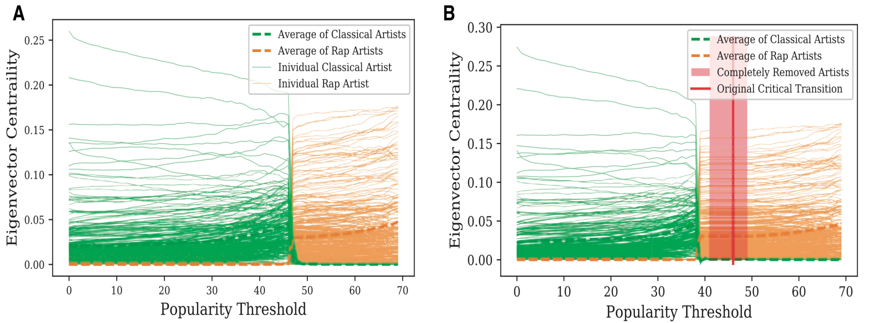
Filter popularity + eigenvector centrality

- Take the most popular contemporary artists

Figure 1: Dominance shifts from classical music to rappers.

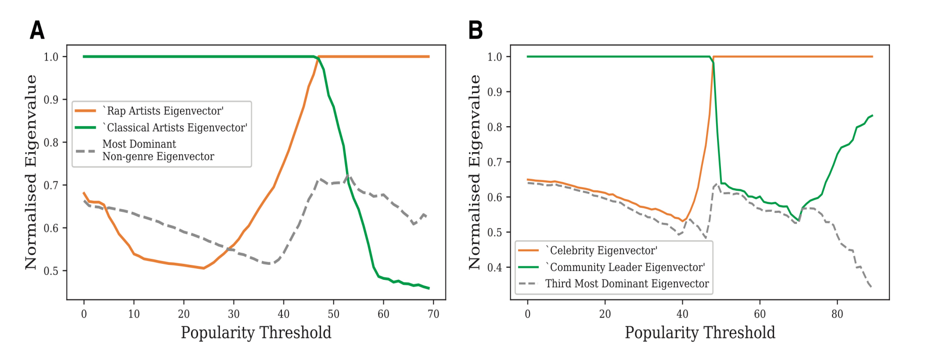
When does it all change?

- Delete nodes with a popularity of 10 or less
- Nothing changes in critical region, but location of the transition is shifted
- Even unpopular artists can change the structure of graph!



Why? It's all in the eigenvalues

- The vectors representing classical dominance and rap dominance exist the entire time, but they change ranking:



Graph Laplacian

Definition (Laplacian matrix)

Define the degree matrix $D \in \mathrm{Mat}(n \times n; \mathbb{R})$ as

$$D_{i,j} = \begin{cases} \deg(v_i), & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases}$$

The Laplacian matrix is then defined as $L = D - A$.

- Analogous to analytic definition of the Laplacian!

Functions on Graphs

A function on graphs

- Suppose a function $f : V \to \mathbb{R}$ on G:
  - could be voltage, temperature, 0/1 indicator for a subset $S \subseteq V$, etc...

A function on graphs

- Suppose a function $f : V \to \mathbb{R}$ on G:
  - could be voltage, temperature, 0/1 indicator for a subset $S \subseteq V$, etc...
- ...then would have

$$f : V \to \mathbb{R} \equiv \begin{bmatrix} f(v_1) \\ f(v_2) \\ \vdots \\ f(v_n) \end{bmatrix}$$

- f acting on V behaves like $\mathbb{R}^n$ and forms a vector space.

Local variance

Definition (Local variance)

The *local variance* of f is defined as:

$$\mathcal{E}(f) := \mathbb{E}_{u \sim v}\left[(f(u) - f(v))^2\right],$$

where $u \sim v$ is a probability distribution over the edges.

Local variance

Definition (Local variance)

The *local variance* of f is defined as:

$$\mathcal{E}(f) := \mathbb{E}_{u\sim v}\left[(f(u) - f(v))^2\right],$$

where $u \sim v$ is a probability distribution over the edges.

$\implies \mathcal{E}(f)$ is "small" iff f is "smooth" over the edges.

Local variance

Definition (Local variance)

The *local variance* of f is defined as:

$$\mathcal{E}(f) \coloneqq \mathbb{E}_{u\sim v}\left[(f(u)-f(v))^2\right],$$

where $u \sim v$ is a probability distribution over the edges.

$\implies \mathcal{E}(f)$ is "small" iff f is "smooth" over the edges.

Can immediately observe that:

- $\mathcal{E}(f) \geq 0$
- $\mathcal{E}(c \cdot f) = c^2 \cdot \mathcal{E}(f)$
- $\mathcal{E}(f + c) = \mathcal{E}(f)$.

Local variance: example

Let $S \subseteq V$ and $f = \mathbb{1}\{v \in S\}$.

Local variance: example

Let $S \subseteq V$ and $f = \mathbb{1}\{v \in S\}$.

The local variance is then

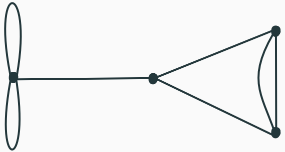
$$\begin{aligned}
\mathcal{E}(f) &= \frac{1}{2} \cdot \mathbb{E}_{u \sim v}\left[\left(\mathbb{1}\{u \in S\} - \mathbb{1}\{v \in S\}\right)^2\right] \\
&= \frac{1}{2} \cdot \mathbb{E}_{u \sim v}\left[\mathbb{1}\{(u, v) \text{ ``crosses'' } S\}\right] \\
&= \frac{1}{2} \cdot \{\text{fraction of edges on } \partial S\} \\
&= \mathbb{P}_{u \sim v}\{u \to v \text{ is ``stepping'' out of } S\}.
\end{aligned}$$

Random Vertices

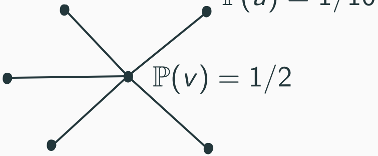
- Procedure for choosing a random vertex:
  - Choose $\mathbf{u} \sim \mathbf{v}$ a random edge.
  - Output $\mathbf{u}$.

Random Vertices

- Procedure for choosing a random vertex:
  - Choose $\mathbf{u} \sim \mathbf{v}$ a random edge.
  - Output $\mathbf{u}$.
- This *stationary distribution* on the vertices, π, prioritises vertices based on the number of adjacent edges (evenly spread probability only in the case of regular graph).



$\mathbb{P}(n_i) = const.$



Facts for the walk

- The probability of picking u, $\pi(u)$, is proportional to $\deg(u)$. (In fact it is $\pi(u) = \deg(u)/2|E|$.)

- The process of picking $\mathbf{u}$ from π and then picking $\mathbf{v}$ as a uniformly random neighbour of $\mathbf{u}$ is the same as drawing an edge uniformly at random $\mathbf{u} \sim \mathbf{v}$.

- Let $t \in \mathbb{N}$. Pick $\mathbf{u} \sim \pi$. Now do a random walk starting at $\mathbf{u}$ taking t steps. Then the distribution of $\mathbf{v}$, the endpoint of the walk, is π. (π is invariant.)

Some Linear Algebra

The Spectral Theorem

We say that ν is an eigenvector of $M \in \mathrm{Mat}(n \times n; \mathbb{R})$ with eigenvalue λ if $M\nu = \lambda\nu$.

Then λ is an eigenvalue iff $\lambda\mathbb{I} - M$ is singular.

Theorem (Spectral Theorem)

If M is an n-by-n, real, symmetric matrix, then there exist real numbers $\lambda_1, \ldots, \lambda_n$ and n mutually orthogonal unit vectors $\nu_1, \ldots, \nu_n$ and such that ν_i is an eigenvector of M of eigenvalue λ_i, for each i.

Proof Sketch: Induction

If $n = 1$, then $M = \lambda$ and can pick any $v \neq 0$ as a basis for $\mathbb{R}$.

Hypothesis. Every k-by-k matrix for $k = 1, \ldots, n-1$ satisfies the spectral theorem.

Step.

- Obtain orthonormal basis $B = v_1, \ldots, v_n$ for $\mathbb{R}^n$ by choosing $\lambda_1 \in \mathbb{R}$
- Use $P = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$ to obtain a block form for $A = P^\top MP$
- Then show that $\exists$ orthogonal R s.t. $R^\top MR$ is diagonal via induction hyp. $\implies \exists$ orthonormal basis for $\mathbb{R}^n$ consisting of eigenvalues of M.

Weird Applications

Cover time of graph

Given, for graph G:

- Laplacian matrix $L = D - A$
- Eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$
- Eigenvectors $\nu_1, \nu_2, \ldots, \nu_n$

Problem. Want to find how long it takes to cover G.

Cover time of graph

Theorem (DLP'11)

If $g_1, \ldots, g_n \sim_{\text{i.i.d.}} \mathcal{N}(0,1)$, *then*

$$n \cdot \left\| \sum_{i=2}^{n} \frac{g_i}{\sqrt{\lambda_i}} \cdot \nu_i \right\|_\infty^2 \asymp \textit{ cover time of } G,$$

up to $o(1)$ *error.*

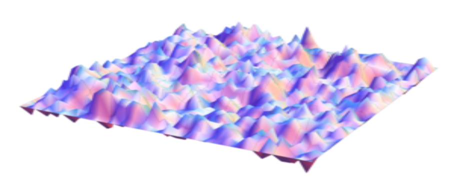


Figure 2: Gaussian free field induced by graph

Unique games problems

Problem: Suppose you are given p prime and a family of 2-variate linear equations over the variables $x_1, \dots, x_n$:

$$\begin{aligned} x_{13} - x_7 &\equiv 4 \pmod p \\ x_4 - x_7 &\equiv 9 \pmod p \\ x_7 - x_{12} &\equiv 1 \pmod p \\ x_{11} - x_4 &\equiv 0 \pmod p \\ &\vdots \end{aligned}$$

If 99% of these equations are satisfiable, then can we find a solution that works for 1% of these equations?

Unique games problems

Conjecture (Khot'02). This problem is computationally intractable.

The problem *can* be reduced to finding small clusters corresponding to specific λ.

Unique games problems

Conjecture (Khot'02)**.** This problem is computationally intractable.

Theorem (LOGT'12)

For $S \subseteq V$, let $\phi_G(S) = |E(S, S^c)|/d|S|$, and let the k-way expansion constant be

$$\rho_G(k) = \min_{S_1,\ldots,S_k} \max\{\phi_G(S_i) : i = 1, \ldots, k\}.$$

Then, for every graph G and every $k \in \mathbb{N}$,

$$\lambda_k/2 \leq \rho_G(k) \leq \mathcal{O}(k^2)\sqrt{\lambda_k}$$

$\implies$ Can always find small clusters!

Nearly linear-time Laplacian solver

Problem: Solve the system $Lx = b$.

- Will be culmination of seminar!
- We can quickly compute the spectral object x [ST'04, KMP'10, KOSZ'13]
- This in turn is useful for computing max flows and min cuts on a graph [CKMST'11, LSR'13, Madry'13, etc]

Fin

Sign up for talks!!!

1. Courant-Fischer, graph Laplacian
2. Random walk
3. Electronic network
4. Sampling spanning tree
5. Graph sparsifier
6. Cheeger inequality
7. Laplacian solver
8. (Nearly) Linear-time Laplacian solver