1.022 - Introduction to Network Models

Amir Ajorlou

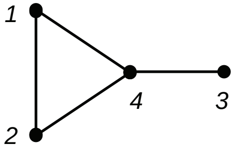
Laboratory for Information and Decision Systems
Institute for Data, Systems, and Society
Massachusetts Institute of Technology

Lecture 6

Graph Laplacian

- Vertex degrees often stored in the diagonal matrix $\mathbf{D}$, where $D_{ii} = d_i$

$$\mathbf{D} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$



- The $|V| \times |V|$ symmetric matrix $\mathbf{L} := \mathbf{D} - \mathbf{A}$ is called graph Laplacian

$$L_{ij} = \begin{cases} d_i, & \text{if } i = j \\ -1, & \text{if } (i,j) \in E \\ 0, & \text{otherwise} \end{cases}, \quad \mathbf{L} = \begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}$$

- Variants of the Laplacian exist, with slightly different interpretations
  - $\Rightarrow$ Normalized Laplacian $\mathbf{L}_n = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$
  - $\Rightarrow$ Random-walk Laplacian $\mathbf{L}_{rw} = \mathbf{D}^{-1}\mathbf{L}$

Laplacian matrix properties

- Smoothness: For any vector $\mathbf{x} \in \mathbb{R}^{|V|}$ of "vertex values", one has

$$\mathbf{x} \;\; \mathbf{L}\mathbf{x} = \sum_{(i,j)\in E} (x_i - x_j)^2$$

  which can be minimized to enforce smoothness of functions on G

- Incidence relation: $\mathbf{L} = \mathbf{B}\mathbf{B}$ where $\mathbf{B}$ has arbitrary orientation
- Positive semi-definiteness: Follows since $\mathbf{x} \;\; \mathbf{L}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^{|V|}$
- Rank deficiency: Since $\mathbf{L1} = \mathbf{0}$, $\mathbf{L}$ is rank deficient

Laplacian matrix properties

Spectrum and connectivity: $\mathbf{L1} = \mathbf{0}$, so 0 is an eigenvalue

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

- The second-smallest eigenvalue λ_2 is called the algebraic connectivity
- If $\lambda_2 = 0$, then G is connected
- If G has k connected components then $0 = \lambda_k < \lambda_{k+1}$

Matrix Tree Theorem: The number of spanning trees of G is

$$t(G) = \lambda_2 \times \ldots \times \lambda_n.$$

- Spanning tree: a subgraph that is a tree which includes all the vertices.

Courant Fischer Theorem

$$\lambda_1 = \min_{x=0} \frac{x^T L x}{x^T x} \qquad v_1 = \operatorname{argmin}_{x=0} \frac{x^T L x}{x^T x}$$

$$\lambda_2 = \min_{\substack{x=0 \\ x \perp v_1}} \frac{x^T L x}{x^T x} \qquad v_2 = \operatorname{argmin}_{\substack{x=0 \\ x \perp v_1}} \frac{x^T L x}{x^T x}$$

Courant Fischer Theorem: M an $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$ and eigenvectors $v_1, \ldots, v_n$.

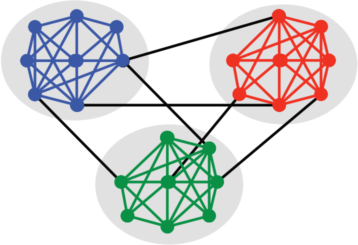
- S_k: the span of $v_1, \ldots, v_k$, $1 \leq k \leq n$ ($S_0 = \{0\}$).
- $S_k^{\perp}$:orthogonal complement of S_k.

Then,

$$\lambda_k = \min_{\substack{x \; =0 \\ x \in S_k^{\perp}}} \frac{x^T M x}{x^T x} \qquad v_k = \operatorname{argmin}_{\substack{x \; =0 \\ x \in S_k^{\perp}}} \frac{x^T M x}{x^T x}.$$

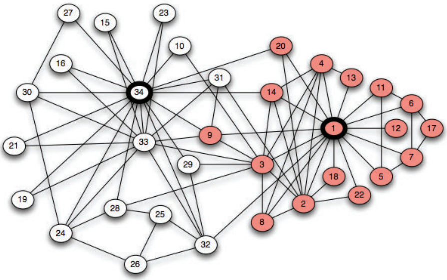
Community detection and spectral clustering

- Nodes in many real-world networks organize into communities
  Ex: families, clubs, political organizations, urban areas, . . .
- Supported by the strength of weak ties theory



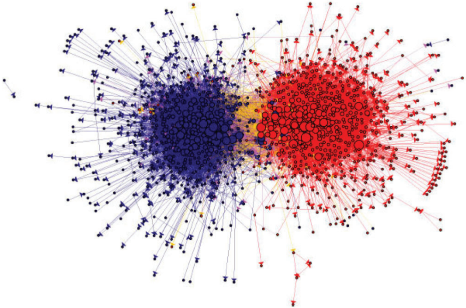
- Community (a.k.a. group, cluster, module) members are:
  ⇒ Well connected among themselves
  ⇒ Relatively well separated from the rest

- Social interactions among members of a karate club in the 70s
  - $\Rightarrow$ Canonical network for community detection methods



- The club split into two during the study (white and red groups)
  - $\Rightarrow$ Offers ground-truth community membership
- Could we have predicted the split only from the network structure?

- The political blogosphere for the US 2004 presidential election



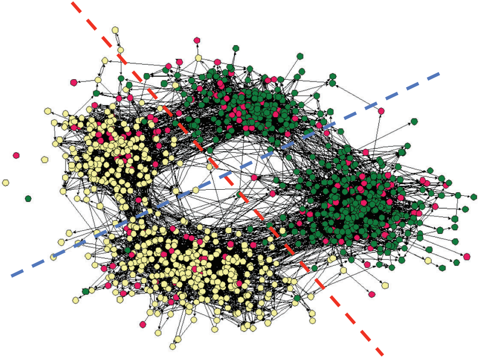
- Community structure of liberal and conservative blogs is apparent

  $\Rightarrow$ Strong evidence of partisan homophily in the network

- Can we detect both parties without looking at the blogs' content?

School students

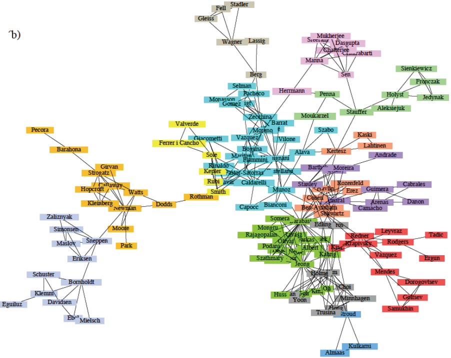
- Social network from a town's middle and high school students



- Two binary divisions are apparent from the structure of the network
  - ⇒ Racial division marked in red
  - ⇒ Age division (middle - high) marked in blue
- Can we estimate race and age of a student from the structure?

Physicists working on Network Science

- Co-authorship network of physicists working on networks
  - ⇒ Edges represent the existence of a collaborative publication



Newman, M. E. J., and M. Girvan. "Finding and evaluating community structure in networks." *Physical Review E* 69 (2004): 026113.

- Tightly-knit subgroups are evident from the network structure
  - ⇒ Some researchers work at the boundary between two groups?
- Can we recover this information without relying on visual inspection?

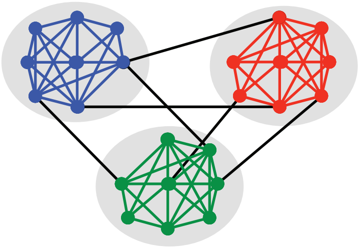
Automatic detection of communities

- Recurring theme in all of the examples provided

  ⇒ How can we automatically detect communities in a network?

- But ... what is a sensible definition of community?

  ⇒ Multiple definitions lead to multiple community detection methods



Graph partitioning and community detection

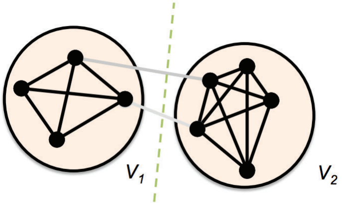
- Community detection is a challenging problem
  - ⇒ No universal definition of community
  - ⇒ No prior knowledge of community number or sizes
  - ⇒ Rare ground-truth data for validation

- We begin with a simpler problem ⇒ Graph partitioning
- Divide V into a given number of non-overlapping groups of a given size
- Graph partitioning is still a hard problem
  - ⇒ Even graph bisection (two groups, equal size) has $\binom{|V|}{|V|/2}$ possibilities
- Exhaustive search intractable beyond small datasets
- Need to rely on tractable relaxations of natural partitioning criteria

Graph partitioning and minimum cuts

- Community members should be well-connected among themselves

  $\Rightarrow$ Loosely connected with members of other communities



- A cut C is the weight of edges between blocks V_1 and $V_2 = V \backslash V_1$

$$C = \mathrm{cut}(V_1, V_2) = \sum_{i \in V_1, j \in V_2} A_{ij}$$

- Find cut that achieves the desired sizes in V_1 and V_2 while minimizing C

Graph partitioning and the Laplacian matrix

- Assign to each node $i \in V$ an identifier $s_i \in \{-1, 1\}$
  - $\Rightarrow$ Form the vector $\mathbf{s} = [s_1, s_2, \dots, s_{|V|}]$
- Notice that $C(\mathbf{s}) = \sum_{ij} A_{ij}$ where $s_i = -1$ and $s_j = +1$
- It can be shown that $C(\mathbf{s}) = \frac{1}{4}\mathbf{s} \;\; \mathbf{L}\mathbf{s}$, where $\mathbf{L}$ is the Laplacian matrix
  - $\Rightarrow$ You will show this in your homework
- We have expressed the cut (relevant graph-related quantity)
  - $\Rightarrow$ In terms of vectors and matrices (amenable algebraic objects)
- Find vector $\mathbf{s} \in \{-1, 1\}^{|V|}$ such that:
  - $\Rightarrow$ $\sum_i s_i = |V_2| - |V_1|$ (desired group sizes), and
  - $\Rightarrow$ Minimizes $C(\mathbf{s}) = \frac{1}{4}\mathbf{s} \;\; \mathbf{L}\mathbf{s}$