1.022 - Introduction to Network Models

Amir Ajorlou

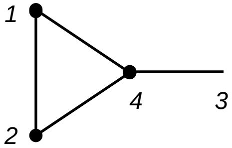
Laboratory for Information and Decision Systems
Institute for Data, Systems, and Society
Massachusetts Institute of Technology

Lecture 7

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 $\mathbf{L} := \mathbf{D} \quad \mathbf{A}$ is called graph Laplacian

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

Courant Fischer Theorem

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

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

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

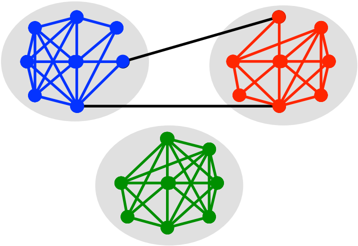
- S_k: the span of $v_1, \ldots, v_k$, $1 \leq k \leq n$ ($S_0 = \{0\}$).
- $S_k^\perp$:orthogonal complement of S_k.

Then,

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

Community detection and spectral clustering

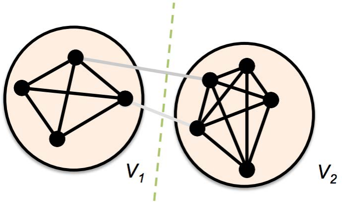
- Nodes in many real-world networks organize into communities
  Ex: families, clubs, political organizations, urban areas, . . .
- Supported by the strength of weak ties theory



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

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 $V_2 = V \setminus V_1$

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

- 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 $\mathbf{s} = [s_1, s_2, \ldots, s_{|V|}]$
- Notice that $C(\mathbf{s}) = \sum_{ij} A_{ij}$ where $s_i = \ \ 1$ and $s_j = +1$
- It can be shown that $C(\mathbf{s}) = \frac{1}{4}\mathbf{s}^\top \mathbf{L}\mathbf{s}$, 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 $C(\mathbf{s}) = \frac{1}{4}\mathbf{s}^\top \mathbf{L}\mathbf{s}$

Spectral graph partitioning

- Finding such optimal $\mathbf{s}$ is still challenging
  - $\Rightarrow$ Due to the integer restriction $\mathbf{s} \in \{\ \ 1, 1\}^{|V|}$
- Relax this requirement into $\sum_i s_i^2 = |V|$
  - $\Rightarrow$ Note that $\mathbf{s} \in \{\ \ 1, 1\}^{|V|}$ implies $\sum_i s_i^2 = |V|$ but not vice versa
- New relaxed problem: Find $\mathbf{s}$ that minimizes $C(\mathbf{s}) = \frac{1}{4}\mathbf{s}^\top \mathbf{L}\mathbf{s}$
  - $\Rightarrow$ Subject to $\sum_i s_i = |V_2| \quad |V_1|$ and $\sum_i s_i^2 = |V|$
- The optimal $\mathbf{s}^*$ is given by

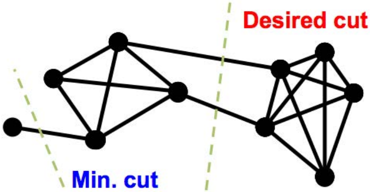
$$\mathbf{s}^* = \mathbf{v}_2 + \frac{|V_2| \quad |V_1|}{|V|}\mathbf{1}$$

  - $\Rightarrow$ where $\mathbf{v}_2$ is the eigenvector of $\mathbf{L}$ with the second smallest eigenvalue

Recovering the partitions

- How do we go back to two groups from the (non-integer) $\mathbf{s}^*$?
  - $\Rightarrow$ Find $\mathbf{s} \in \{\ \ 1, +1\}^{|V|}$ that is most aligned with $\mathbf{s}^*$
- Algorithm: Spectral graph partitioning
  - $\Rightarrow$ 1) Compute Laplacian matrix $\mathbf{L}$ of graph of interest
  - $\Rightarrow$ 2) Find $\mathbf{v}_2$, the eigenvector of $\mathbf{L}$ with the second smallest eigenvalue
  - $\Rightarrow$ 3) Order the entries of $\mathbf{v}_2$ in decreasing order
  - $\Rightarrow$ 4) Assign $s_i = \ \ \ 1$ to the $|V_1|$ top-sorted entries and rest $s_i = +1$
- Minor subtlety $\Rightarrow$ Both $\mathbf{v}_2$ and $\ \ \mathbf{v}_2$ are eigenvectors
  - $\Rightarrow$ Repeat the above procedure for $\ \ \mathbf{v}_2$
  - $\Rightarrow$ Choose partition (from those two) that minimizes the cut

- What if we do not know a priori the sizes of the sought communities?
  - ⇒ We cannot implement the above procedure
  - ⇒ More importantly, the cut might not be the right criterion



- Consider the ratio cut R instead

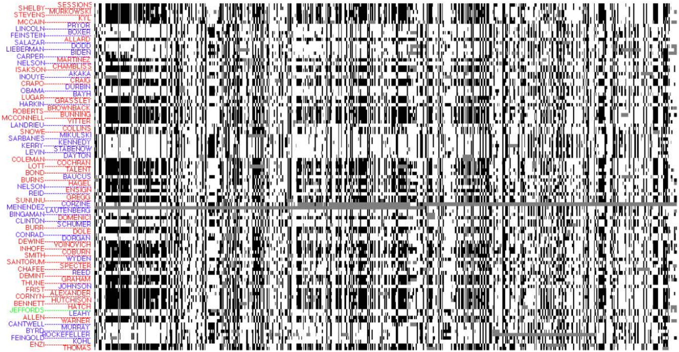
$$R(V_1, V_2) = \frac{C(V_1, V_2)}{|V_1|} + \frac{C(V_1, V_2)}{|V_2|}$$

⇒ Small groups are penalized ⇒ Balanced partition

- The ratio cut criterion can be relaxed in a similar way
- Main difference $\Rightarrow$ Unknown group sizes
- Algorithm: Spectral community detection
  - $\Rightarrow$ 1) Compute Laplacian matrix $\mathbf{L}$ of graph of interest
  - $\Rightarrow$ 2) Find $\mathbf{v}_2$, the eigenvector of $\mathbf{L}$ with the second smallest eigenvalue
  - $\Rightarrow$ 3) Assign $s_i = \quad 1$ if $[\mathbf{v}_2]_i < 0$ and $s_i = +1$ otherwise
- What if we want to detect more than two groups?
- Opt. 1: Apply the above process iteratively

Clustering via other low-dimensional representations

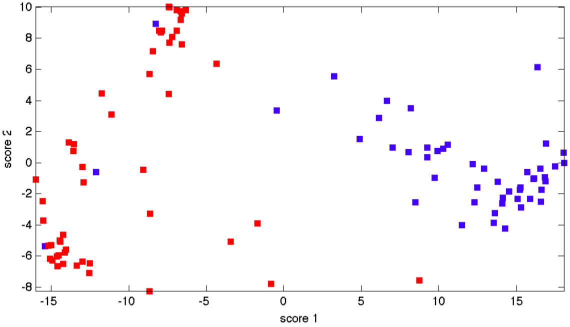
- Votes by $n = 100$ senators in the 2004-2006 US Senate about $m = 542$ bills



- Republicans tend to vote similar to other republicans [El Ghaoui 17]
  - $\Rightarrow$ Same occurs with democrats
  - $\Rightarrow$ How can we capture this? $\Rightarrow$ Notion of covariance

Principal Component Analysis (PCA)

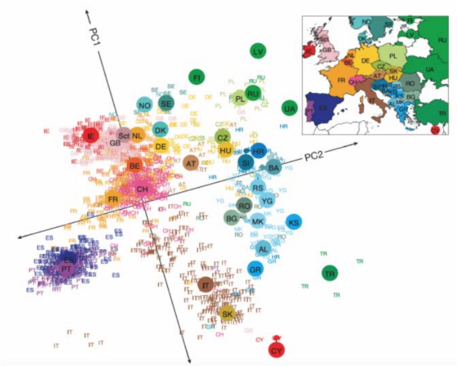
- One can interpret the covariance matrix as a weighted graph
  - ⇒ Large positive values indicate similar voting patterns between senators
- PCA projects senators onto the top eigenvectors of the covariance matrix
  - ⇒ These are the direction of maximum variance



- The parties can be recovered almost perfectly from the 2D representation

Example of more than two groups

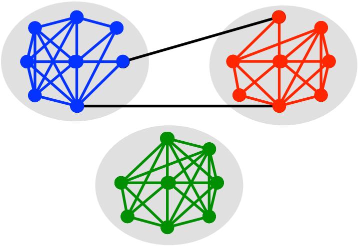
- Genes mirror geography within Europe, Novembre et al., Nature (2008)
- Two-dimensional embedding of ‘gene similarity’ matrix
  - $\Rightarrow$ Consistent with origins of individuals in European map



Novembre, John, Toby Johnson, Katarzyna Bryc, et al. "Genes mirror geography within Europe." *Nature* 456 (2008): 98–101.

Using centrality for community detection

- Communities as connected components when deleting bridges



- How can we automatically detect bridges?
- Can we use the notion of (betweenness) centrality but for edges?

  $\Rightarrow$ For the interested, see Chapter 3.6 of our main text