Graph Signal Processing – Part I: Graphs, Graph Spectra, and Spectral Clustering

Qincheng Lu

Ph.D. Student Supervised by Prof. Xiao-Wen CHANG

McGill
School of Computer Science

September 21, 2020

Section 1

Examples

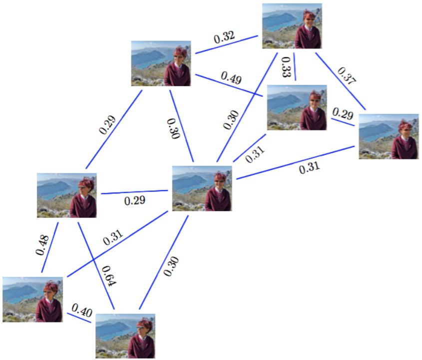


Figure: Images graph

Figure: Minnesota roadmap graph

Graph and Graph Signal

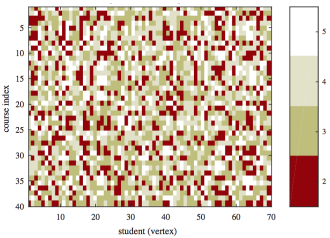


Figure: Marks per student and per course

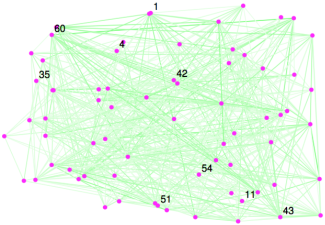


Figure: 2D map with random position

Definition: Graph $\mathcal{G} = \{\mathcal{V}, \mathcal{B}\}$, $\mathcal{B} \subset \mathcal{V} \times \mathcal{V}$
Graph Signal $f \to \mathbb{R}^N$

Classical Discrete Signal Processing

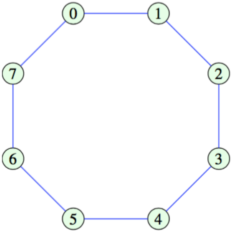


Figure: Time series data

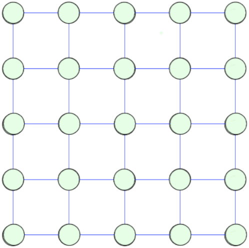


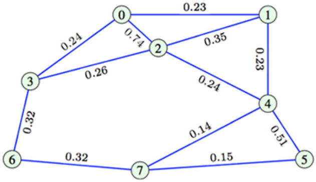
Figure: Digital image data

Adjacency Matrix

1. Adjacency matrix $\mathbf{A}$ for N vertices is an $N \times N$ matrix.
$$A_{mn} \overset{\text{def}}{=} \begin{cases} 1, & \text{if } (m,n) \in \mathcal{B} \\ 0, & \text{if } (m,n) \notin \mathcal{B} \end{cases}$$
2. For undirected graph, if $(n,m) \in \mathcal{B}$ then also $(m,n) \in \mathcal{B}$, $\mathbf{A} = \mathbf{A}^\intercal$.

$$\mathcal{V} = \{0,1,2,3,4,5,6,7\}$$
$$\mathcal{B} \subset \{0,1,2,3,4,5,6,7\} \times \{0,1,2,3,4,5,6,7\}$$
$$\mathcal{B} = \{(0,1),(1,2),(2,0),(2,3),(2,4),(2,7),(3,0),$$
$$(4,1),(4,2),(4,5),(5,7),(6,3),(6,7),(7,2),(7,6)\}.$$

$$\mathbf{A}_{\text{un}} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ \end{array} \begin{array}{c} \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \\ \begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \end{array}$$



$$\mathbf{W} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ \end{array} \begin{array}{c} \begin{bmatrix} 0 & 0.23 & 0.74 & 0.24 & 0 & 0 & 0 & 0 \\ 0.23 & 0 & 0.35 & 0 & 0.23 & 0 & 0 & 0 \\ 0.74 & 0.35 & 0 & 0.26 & 0.24 & 0 & 0 & 0 \\ 0.24 & 0 & 0.26 & 0 & 0 & 0 & 0.32 & 0 \\ 0 & 0.23 & 0.24 & 0 & 0 & 0.51 & 0 & 0.14 \\ 0 & 0 & 0 & 0 & 0.51 & 0 & 0 & 0.15 \\ 0 & 0 & 0 & 0.32 & 0 & 0 & 0 & 0.32 \\ 0 & 0 & 0 & 0 & 0.14 & 0.15 & 0.32 & 0 \end{bmatrix} \\ \begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \end{array}$$

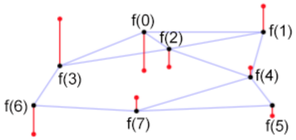
3. Adjacency matrix $\mathbf{A}$ is a special case of the weight matrix $\mathbf{W}$

Laplacian Matrix

1. Degree matrix $\mathbf{D}$ is a diagonal matrix, where $D_{mm} \stackrel{\text{def}}{=} \sum_{n=0}^{N-1} \mathbf{W}_{mn}$
2. Laplacian matrix is defined as $\mathbf{L} \stackrel{\text{def}}{=} \mathbf{D} - \mathbf{W}$

   For undirected graph

   1. Symmetric
   2. Off-diagonal entries are non-positive for non-negative weights
   3. Rows sum up to zero
   4. Eigenvalues are non-negative real numbers
   5. Eigenvectors are real and orthogonal
3. Notion of "smoothness": $\mathbf{f}^\intercal \mathbf{L} \mathbf{f} = \frac{1}{2} \sum_{i,j=0}^{N-1} \mathbf{W}_{ij} (f(i) - f(j))^2$

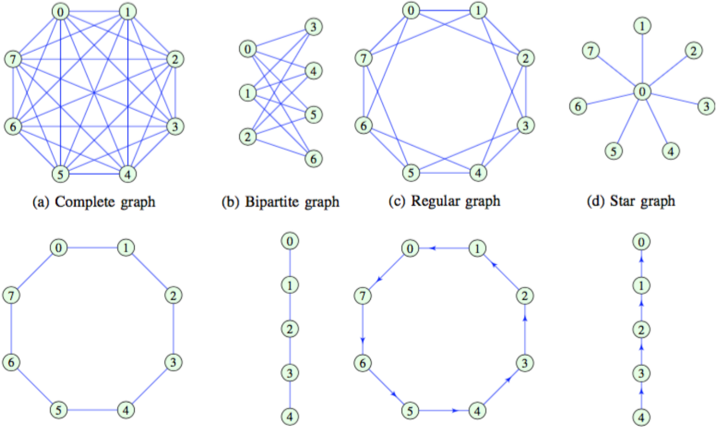


$$\mathbf{L} = \begin{bmatrix} 1.21 & -0.23 & -0.74 & -0.24 & 0 & 0 & 0 & 0 \\ -0.23 & 0.81 & -0.35 & 0 & -0.23 & 0 & 0 & 0 \\ -0.74 & -0.35 & 1.59 & -0.26 & -0.24 & 0 & 0 & 0 \\ -0.24 & 0 & -0.26 & 0.82 & 0 & 0 & -0.32 & 0 \\ 0 & -0.23 & -0.24 & 0 & 1.12 & -0.51 & 0 & -0.14 \\ 0 & 0 & 0 & 0 & -0.51 & 0.66 & 0 & -0.15 \\ 0 & 0 & 0 & -0.32 & 0 & 0 & 0.64 & -0.32 \\ 0 & 0 & 0 & 0 & -0.14 & -0.15 & -0.32 & 0.61 \end{bmatrix}$$

4. Normalized laplacian matrix:

$$\mathbf{L}_N \stackrel{\text{def}}{=} \mathbf{D}^{-1/2}(\mathbf{D} - \mathbf{W})\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$$

Frequently Used Graph Topologies

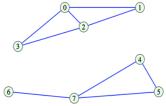


(a) Complete graph (b) Bipartite graph (c) Regular graph (d) Star graph

(e) Circular graph (f) Path graph (g) Directed circular graph (h) Directed path graph

Disconnected Graph

1. Adjacency matrix and Laplacian matrix are of block-diagonal form.
2. The multiplicity of the zero eigenvalue of the Laplacian = the number of disjoint components.



$$\mathbf{A} = \begin{array}{c c} \begin{array}{c} 0\\1\\2\\3\\4\\5\\6\\7 \end{array} & \left[\begin{array}{cccc|cccc} 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 \end{array}\right] \\ & \begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \end{array} \qquad \mathbf{L} = \left[\begin{array}{cccc|cccc} 3 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & 3 \end{array}\right]$$

1. $\mathbf{v_0} = (1, ..., 1, 0, ..., 0)^\intercal$, $\mathbf{v_1} = (0, ..., 0, 1, ..., 1)^\intercal$ are two eigenvectors of eigenvalue $\lambda_0 = 0$ for $\mathbf{L}$

Minimum k-cuts Problem, $k = 2$

- Consider an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{B}\}$ with set of edge weights $\mathbf{W}$
- We want to find $\mathcal{E}$ and $\mathcal{H}$ ($\mathcal{E} \subset \mathcal{V}$, $\mathcal{H} \subset \mathcal{V}$, $\mathcal{E} \cup \mathcal{H} = \mathcal{V}$ and $\mathcal{E} \cap \mathcal{H} = \emptyset$)
- Such that cut $\text{Cut}(\mathcal{E}, \mathcal{H}) = \sum_{m \in \mathcal{E}, n \in \mathcal{H}} \mathbf{W}_{mn}$ is minimized.

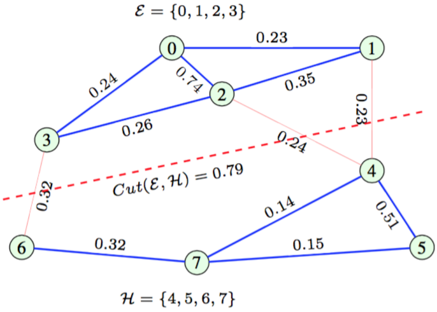


Figure: A cut for a weighted undirected graph

Minimum 2-cuts Problem

- Combinatorial problem: Brute force approach on N vertices takes

$$\binom{N}{1} + \binom{N}{2} + \ldots + \binom{N}{N/2-1} + \binom{N}{N/2}/2 = O(2^N)$$

- Express partition $(\mathcal{E}, \mathcal{H})$ as a vector $\mathbf{x}$: $\mathbf{x}_i \stackrel{\text{def}}{=} \begin{cases} +1, & \text{if } i \in \mathcal{E} \\ -1, & \text{if } i \in \mathcal{H} \end{cases}$

$$\begin{aligned} \mathbf{x}^\intercal \mathbf{L} \mathbf{x} &= \frac{1}{2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \mathbf{W}_{mn} \left(\mathbf{x}(n) - \mathbf{x}(m) \right)^2 \\ &= 4 \sum_{m \in \mathcal{E}, n \in \mathcal{H}} \mathbf{W}_{mn} \\ &= 4 \mathsf{Cut}(\mathcal{E}, \mathcal{H}) \end{aligned}$$

- Fiedler vector $\mathbf{x} = \arg\min_{\mathbf{y} \in \mathbb{R}^N} \mathbf{y}^\intercal \mathbf{L} \mathbf{y}$

Minimum Normalized 2-cuts Problem

- Normalized (ratio) cut $\text{CutN}(\mathcal{E},\mathcal{H}) = (\frac{1}{N_\mathcal{E}} + \frac{1}{N_\mathcal{H}}) \sum_{m\in\mathcal{E}, n\in\mathcal{H}} \mathbf{W}_{mn}$
- where $N_\mathcal{E}$ and $N_\mathcal{H}$ are the respective numbers of vertices in the sets $\mathcal{E}$ and $\mathcal{H}$.
- The normalized indicator $\mathbf{x}$: $\mathbf{x}_i \stackrel{\text{def}}{=} \begin{cases} +1/(N_\mathcal{E} e_x), & \text{if } i \in \mathcal{E} \\ -1/(N_\mathcal{H} e_x), & \text{if } i \in \mathcal{H} \end{cases}$, $||\mathbf{x}||_2^2 = 1$, $e_x^2 = \frac{1}{N_\mathcal{E}} + \frac{1}{N_\mathcal{H}}$
- The indicator $\mathbf{x}$ is orthogonal to the eigenvector of $\mathbf{L}$ for $\lambda_0 = 0$

$$
\begin{aligned}
\frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} &= \frac{1}{||\mathbf{x}||_2^2}\frac{1}{2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\mathbf{W}_{mn}\left(\mathbf{x}(n)-\mathbf{x}(m)\right)^2 \\
&= \frac{1}{||\mathbf{x}||_2^2}\frac{1}{e_x^2}\left(\frac{1}{N_\mathcal{E}}+\frac{1}{N_\mathcal{H}}\right)^2 \sum_{m\in\mathcal{E}, n\in\mathcal{H}}\mathbf{W}_{mn} \\
&= \text{CutN}(\mathcal{E},\mathcal{H})
\end{aligned}
$$

Minimum Normalized 2-cuts Problem (Continue)

- $\frac{\mathbf{x}^\intercal \mathbf{L}\mathbf{x}}{\mathbf{x}^\intercal \mathbf{x}} = \mathsf{CutN}(\mathcal{E}, \mathcal{H})$, with indicator $\mathbf{x}$ normalized to unit energy
- $\min\{\mathbf{x}^\intercal \mathbf{L}\mathbf{x}\}$ subject to $\mathbf{x}^\intercal \mathbf{x} = 1$
- $\mathcal{L}(\mathbf{x}) = \mathbf{x}^\intercal \mathbf{L}\mathbf{x} - \lambda(\mathbf{x}^\intercal \mathbf{x} - 1) \Rightarrow \partial\mathcal{L}(\mathbf{x})/\partial\mathbf{x}^\intercal = \mathbf{0} \Rightarrow \mathbf{L}\mathbf{x} = \lambda\mathbf{x}$
- $\mathbf{x}$ is an eigenvector of $\mathcal{L}$
- $\min\{\mathbf{x}^\intercal \mathbf{L}\mathbf{x}\} = \min\{\lambda\mathbf{x}^\intercal \mathbf{x}\} = \min\{\lambda\}$

Eigenvalue Decomposition of A

1. $\mathbf{Ax}$ is output after a movement of graph signal $\mathbf{x}$ along walks of length one.
2. The output signal from a system on a graph

$$\mathbf{y} = h_0\mathbf{A}^0\mathbf{x} + h_1\mathbf{A}^1\mathbf{x} + \ldots + h_{M-1}\mathbf{A}^{M-1}\mathbf{x} = \sum_{m=0}^{M-1} h_m\mathbf{A}^m\mathbf{x}$$

3. Given $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{-1}$, $\mathbf{A}^m = \mathbf{U}\boldsymbol{\Lambda}^m\mathbf{U}^{-1}$
4. Characteristic polynomial of $\mathbf{A}$

$$\begin{aligned} P(\lambda) &= \mathrm{det}|\mathbf{A} - \lambda\mathbf{I}| = \lambda^N + c_1\lambda^{N-1} + c2\lambda^{N-2} + \ldots + c_N \\ &= (\lambda - \mu_1)^{p_1}(\lambda - \mu_2)^{p_2} \cdots (\lambda - \mu_{N_m})^{p_{N_m}} \\ p_1 + p_2 + \ldots + p_{N_m} &= N, N_m \leq N \end{aligned}$$

5. The minimal polynomial of $\mathbf{A}$

$$P_{\mathsf{min}}(\lambda) = (\lambda - \mu_1)(\lambda - \mu_2) \cdots (\lambda - \mu_{N_m})$$

Eigenvalue Decomposition of L

1. The set of the eigenvalues of the graph Laplacian $\mathbf{L}$ is called graph spectrum
2. Eigenvalues are usually sorted increasingly: $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{N-1}$
3. If $\lambda_1 \neq 0$, λ_1 is called algebraic connectivity
4. The smoothness of an eigenvector $\mathbf{u}_k$ is $\mathbf{u}_k^\intercal \mathbf{L} \mathbf{u}_k = \lambda_k$