Lecture 22

Monday, November 20, 2023 10:05 AM

$A =$ similarity matrix

$n \times n$

$\underbrace{a_{ij} \geq 0}_{\text{non-negative}} \quad \underbrace{a_{ij} = a_{ji}}_{\text{symmetric}}$

eigenvalues can be negative

$|D| = n$

$\vec{x}_i \in \mathbb{R}^d$

$K(\vec{x}_i, \vec{x}_j) = a_{ij}$

$A \equiv$ weighted adjacency matrix

$\Delta \equiv$ degree matrix

$$\begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{bmatrix}$$

diagonal matrix

$$d_i = \underbrace{\sum_{j=1}^{n} a_{ij}}_{\text{sum of row } i}$$

$M \equiv$ normalized adjacency matrix

$= \Delta^{-1} A$

$$= \underbrace{\begin{bmatrix} 1/d_1 & & 0 & \\ & 1/d_2 & & \\ & & \ddots & \\ 0 & & & 1/d_n \end{bmatrix}}_{\Delta^{-1}} \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \end{bmatrix}}_{A}$$

$$= \begin{bmatrix} a_{11}/d_1 & a_{12}/d_1 & \cdots & a_{1n}/d_1 \\ a_{21}/d_2 & a_{22}/d_2 & \cdots & a_{2n}/d_2 \\ \vdots & & & \end{bmatrix}$$

$$\begin{pmatrix} | \\ \vdots \\ | \end{pmatrix} = 1 \begin{pmatrix} | \\ \vdots \\ | \end{pmatrix}$$

$$\begin{bmatrix} & \vdots & & \\ u_{21}/d_2 & u_{22}/d_2 & \cdots & u_{2n}/d_2 \\ & \vdots & & \end{bmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{1}{-} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= \underbrace{\{ m_{ij} \}_{i,j=1..n}}_{\text{Probability of transitioning from } i \text{ to } j} \qquad m_{ij} = a_{ij}/d_i$$

= Markov matrix $\equiv$ row stochastic matrix

each row is a prob vector

(no longer symmetric)

$$M \vec{u_i} = \lambda_i \vec{u_i} \qquad |\lambda_i| \leq 1$$

$\frac{1}{\sqrt{n}}(\vec{1})$ is an eigenvector of M with eigenvalue 1

Graph Laplacian Matrix

$$L = \Delta - A$$

Symmetric, square

positive semi-definite

$$\vec{x}^T L \vec{x} \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n$$

all eigenvalue ≥ 0

$\lambda_n = 0 \leftarrow$ smallest eigenvalue is 0, eigenvector is $\frac{1}{\sqrt{n}} \vec{1}$

$$L \cdot \vec{1} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \cdot \vec{1}$$

$$L \cdot \mathbb{1} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \underset{\lambda_n}{0} \cdot \underset{\vec{u}_n}{\mathbb{1}}$$

$$L = \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ & & & \end{bmatrix}$$

$$L = \begin{bmatrix} d_1 - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & d_2 - a_{22} & -a_{23} & \cdots & -a_{2n} \\ & \vdots & & & \end{bmatrix}$$

Normalized Laplacian

$$\underset{\text{(asymmetric)}}{L_a} = \Delta^{-1} L = \Delta^{-1}(\Delta - A)$$

$$= \Delta^{-1}\Delta - \underline{\Delta^{-1}A}$$

$$= \underline{\underline{I}} - M$$

(L_a is simply the Laplacian of the M matrix)

non-symmetric

still behaves like a positive-semidefinite

$\lambda_i \geq 0 \; \forall i$, all real

$\lambda_n = 0$ ← smallest eigenvalue

Symmetric Normalized Laplacian

$$\Delta = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$$

$$L_s = (\sqrt{\Delta})^{-1} L (\sqrt{\Delta})^{-1}$$

$$= \Delta^{-1/2} L \Delta^{-1/2}$$

$$= \Delta^{-1/2} (\Delta - A) \Delta^{-1/2}$$

$$= \Delta^{-1/2} \Delta \Delta^{-1/2} - \Delta^{-1/2} A \Delta^{-1/2}$$

$$= I - \underbrace{\Delta^{-1/2} A \Delta^{-1/2}}$$

$$\sqrt{\Delta}^{-1} = \Delta^{-1/2} = \begin{bmatrix} 1/\sqrt{d_1} & & & 0 \\ & 1/\sqrt{d_2} & & \\ & & \ddots & \\ 0 & & & 1/\sqrt{d_n} \end{bmatrix}$$

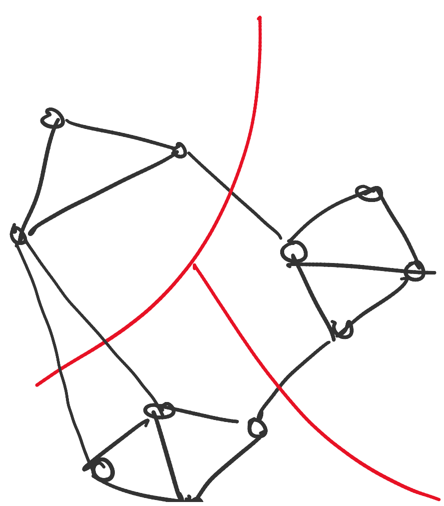
$$L_s = \left\{ \frac{l_{ij}}{\sqrt{d_i}\sqrt{d_j}} \right\}_{i,j=1\ldots n} \quad \text{vs} \quad L_a = \left\{ \frac{l_{ij}}{d_i} \right\}_{i,j=1\ldots n}$$

Symmetric & PSD

Spectral $\rightarrow$ Graph Clustering

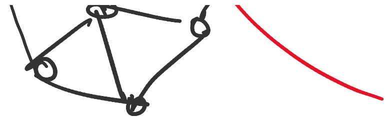
$A = n \times n$

weighted adjacency Matrix



K-way cut

↳ partition the graph into k parts



k parts

1) balance condition

each cluster should have large size or weight

2) edges that are cut should have low weights

$V \equiv$ set of vertices

$$\boxed{\underbrace{W(S,T)}_{\text{cut weight}} = \sum_{i \in S} \sum_{j \in T} a_{ij}} \qquad S,T \subseteq V$$

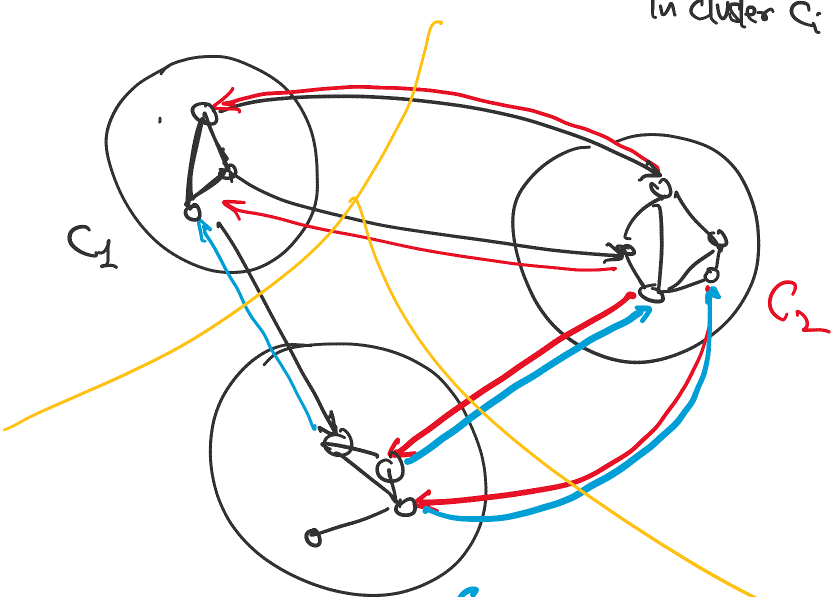
Given V and A, find k clusters

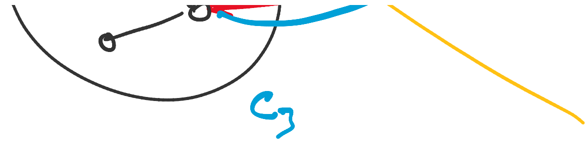
$$? \; \boxed{C_1, C_2, \dots C_k} \subseteq V \qquad c_i \subseteq V$$

$$c_i \cap c_j = \emptyset \qquad \overline{c_i} = V - c_i$$

Objective: $\underbrace{\min J_{rc}}_{\text{ratio cut}} = \sum_{i=1}^{k} \frac{W(c_i, \overline{c_i})}{\underbrace{|c_i|}_{\text{\# of nodes in cluster } c_i}}$

balancing criteria





Objective 2: normalized cut

$$\min \; J_{nc} = \sum_{i=1}^{k} \frac{W(C_i, \overline{C_i})}{Vol(C_i)}$$

$$Vol(C_i) = \underbrace{W(C_i, V)}_{\text{similarity to all vertices (internal + external)}} \quad vs \quad \underbrace{W(C_i, \overline{C_i})}_{\text{external similarity}}$$

Correction

Ratio cut

$$\min \; J_{rc} = \sum_{i=1}^{k} \frac{W(C_i, \overline{C_i})}{|C_i|}$$

Combinatorial objective

$\Downarrow$

Numeric objective

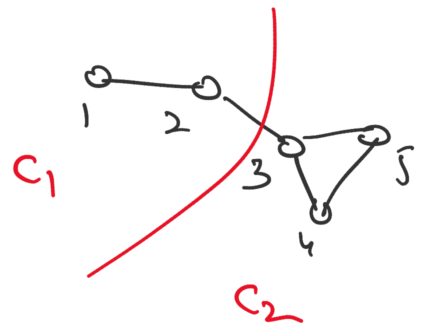
How to characterize $C_i \subseteq V$

$$\vec{c_i} = (c_{i1}, c_{i2}, \ldots, c_{in})^T$$

binary indicator vector

$c_{ij} = 1$ iff $\vec{x_j} \in C_i$

else 0



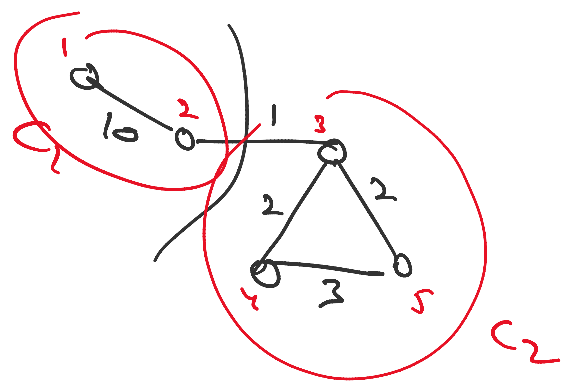
$$\vec{c}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \vec{c}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\underbrace{|C_i|}_{\text{size}} = \vec{c}_i^{\,T}\vec{c}_i = \|\vec{c}_i\|^2$$

$$\vec{c}_1^{\,T} c_1 = 2 \qquad \vec{c}_2^{\,T} c_2 = 3$$

$C_1, C_2 \ldots C_k \Rightarrow \vec{c}_1, \vec{c}_2, \ldots, \vec{c}_k$ all binary indicator vectors

$$\mathrm{vol}(C_i) = \underline{W(C_i, V)}$$
$$= \vec{c}_i^{\,T} \Delta \vec{c}_i$$



$$A = \begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & 10 & 0 & 0 & 0 \\ 2 & 10 & 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 2 & 2 \\ 4 & 0 & 0 & 2 & 0 & 3 \\ 5 & 0 & 0 & 2 & 3 & 0 \end{array}$$

$$\vec{c}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Delta = \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \qquad \vec{c}_2^{\,T} \Delta \vec{c}_2$$

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \end{pmatrix} \Delta \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} = 5+5+5 = 15$$

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \end{pmatrix} \Delta \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 1 \end{pmatrix} = 5+5+5 = 15$$

$$\underline{W(C_i, \bar{C}_i)} = \underbrace{\underbrace{W(C_i, V)}_{\text{all weights}} - \underbrace{W(C_i, C_i)}_{\text{Internal}}}_{\text{external}}$$

$$= \vec{c}_i^{\,T} \Delta \vec{c}_i - \vec{c}_i^{\,T} A \vec{c}_i$$

$$= \vec{c}_i^{\,T} (\Delta - A) \vec{c}_i$$

$$= \vec{c}_i^{\,T} L \vec{c}_i$$

Ratio objective

$$\min_{\vec{c}_1, \vec{c}_2, \ldots \vec{c}_k} J_{rc} = \sum_{i=1}^{k} \frac{\vec{c}_i^{\,T} L \vec{c}_i}{\vec{c}_i^{\,T} \vec{c}_i} \equiv \sum \frac{W(C_i, \bar{C}_i)}{|C_i|}$$

binary indicator vectors!

$\boxed{\text{relax this constraint}}$

allow $\vec{c}_i$ to be any real vector.

Solve optimally (in

→ Solve optimally (in polynomial time)

$$J_{rc} = \sum_{i=1}^{k} \frac{\vec{c_i}^T L \vec{c_i}}{\underbrace{\|\vec{c_i}\|^2}} = \sum \left(\frac{\vec{c_i}}{\|\vec{c_i}\|}\right)^T L \left(\frac{\vec{c_i}}{\|\vec{c_i}\|}\right)$$

$$\min J_{rc} = \sum \vec{u_i} L \vec{u_i}$$

where $u_i = \frac{\vec{c_i}}{\|\vec{c_i}\|} \equiv$ unit vector

Solutions are exactly the eigenvectors of L (Laplacian)

PSD

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \underbrace{\lambda_{n-k+1} \geq \dots \geq \lambda_n = 0}_{\text{pick the } k \text{ smallest eigenvalue}}$$

$$\underbrace{\vec{U}_{n-k+1} \dots \vec{U}_n}_{k \text{ eigenvectors}}$$

$\vec{u_i}$ supposed to be approx to $\vec{c_i}$

real vector → binary vector

run k-means on this!

Spectral Algo

1) A

2) $\underbrace{L = \Delta - A}_{\text{ratio}}$ or $\underbrace{L_a \text{ or } L_s}_{\text{normalized}}$

3) solve

$$L v_i = \lambda_i u_i$$

for the k smallest eigenvalue

4)

$$Y = \begin{bmatrix} | & & | \\ U_{n-k+1} & \dots & U_n \\ | & & | \end{bmatrix}$$

n

$$n \begin{bmatrix} u_{n-k+1} & \cdots & u_n \\ | & & | \end{bmatrix}$$
$$k$$

4a) $\longrightarrow$ normalize the rows of Y (each row is a unit vector) (optional)

5) run k-means on Y

J_{nc}

$$J_{nc} = \sum \frac{w(C_i, \bar{C_i})}{vol(C_i)} = \sum_{i=1}^{k} \frac{\vec{c_i}^T L \vec{c_i}}{\vec{c_i}^T \Delta \vec{c_i}}$$

$$\Delta = \sqrt{\Delta}\sqrt{\Delta}$$

$$\sum u_i^T (\underbrace{\Delta^{-1} L}_{L_a}) \vec{u_i} \qquad \sum \vec{u_i}^T (\underbrace{\Delta^{-1/2} L \Delta^{-1/2}}_{L_s}) \vec{u_i}$$

$$\frac{L}{\Delta^{1/2} \Delta^{1/2}}$$

Markov Clustering

Start with $M = \Delta^{-1} A \equiv$ transition prob matrix for 1 step

1) powers of M^t give us t-step prob matrix for t-step Markov chain

2) damping

MCL:

1) M

2) take a step using M

ie $M^{(t)} = [M^{(t-1)}]^2$

3) Inflate $M^{(t)}$

take each element and inflate by power r

normalize