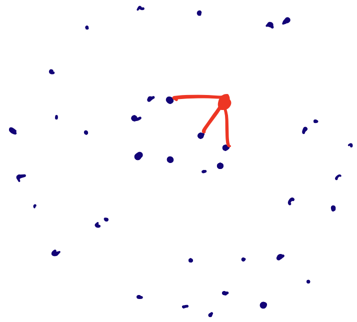
Nov 14, 2024 Spectral Clustering



$S_{ij} = sim(x_i, x_j)$

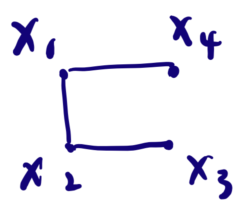
$$\begin{array}{c|cccc} & x_1 & x_2 & x_3 & x_4 \\ \hline x_1 & \times & 0.3 & 0.7 & 0.2 \\ x_2 & 0.3 & \times & 0.4 & 0.8 \\ x_3 & 0.7 & 0.4 & \times & 0.6 \\ x_4 & 0.2 & 0.8 & 0.6 & \times \end{array} \Rightarrow \begin{array}{c|cccc} & x_1 & x_2 & x_3 & x_4 \\ \hline x_1 & 0 & 1 & 0 & 1 \\ x_2 & 1 & 0 & 1 & 0 \\ x_3 & 0 & 1 & 0 & 0 \\ x_4 & 1 & 0 & 0 & 0 \end{array}$$

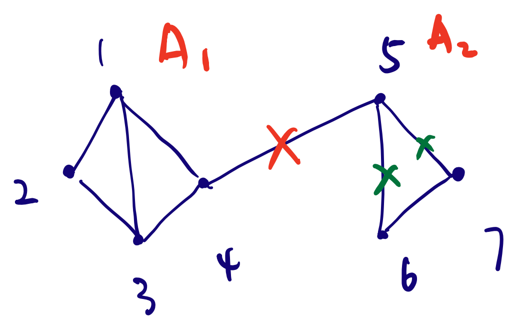
similarity matrix S adjacency matrix W

distance threshold < 0.5

||

graph



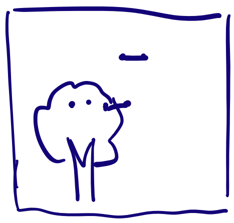


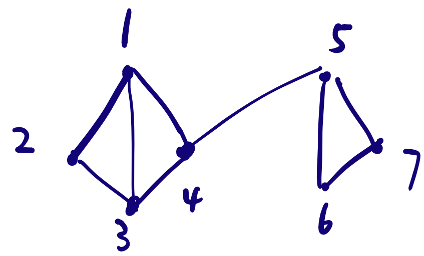
$w(A_1, A_2) = w_{45}$

Clustering = finding k cuts in the graph

$$\min \sum_{k=1}^{K} w(A_k, \bar{A}_k), \quad W(A,B) = \sum_{i \in A, j \in B} w_{ij}$$

graph cut . Image segmentation





adjacency matrix W

$$\begin{array}{c|ccccccc|} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 4 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 5 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 6 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 7 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ \hline \end{array}$$

degree matrix D

$$\begin{bmatrix} 3 & & & & & & \\ & 2 & & & & & \\ & & 3 & & & & \\ & & & 3 & & & \\ & & & & 3 & & \\ & & & & & 2 & \\ & & & & & & 2 \end{bmatrix}$$

L:

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{bmatrix} 3 & -1 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Laplacian matrix $L \triangleq D - W$

properties of Laplacian matrix:

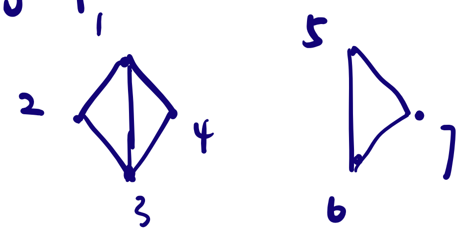
① each row sums to 0

② $L \cdot \mathbb{1} = 0 \cdot \mathbb{1}$

eigen vector $\mathbb{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ with eigen value 0

$$L \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \Sigma \text{ row } 1 \\ \Sigma \text{ row } 2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \cdot \mathbb{1}$$

③ graph has k connected components



L has k eigenvectors $\mathbb{1}_{A_1}, \dots \mathbb{1}_{A_k}$ with eigenvalue 0

$$\mathbb{1}_{A_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbb{1}_{A_2} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$L \cdot \mathbb{1}_{A_1} = 0 \cdot \mathbb{1}_{A_1}$$

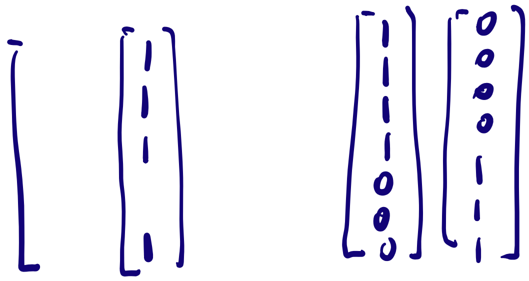
$$L \cdot \boxed{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} = \begin{bmatrix} \boxed{\begin{matrix} 3 & -1 & -1 & -1 \\ -1 & 2 & -1 & \\ -1 & -1 & 3 & -1 \end{matrix}} & \\ & \boxed{\begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{matrix}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0$$

$$L \cdot \mathbb{1}_{A_2} = L \cdot \boxed{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = 0$$

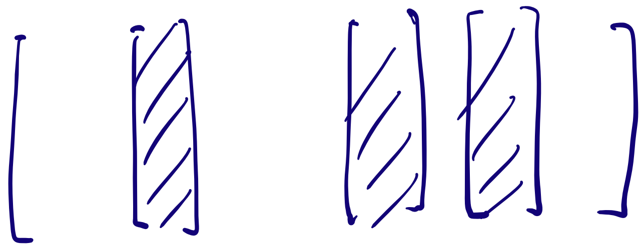
eigenvectors with 0 eigenvalue

$\boxed{L}$ eigenvalues $[x, 0, \cdots 0, 0, x]$

eigenvectors



in pratice $[x, 0 \cdots 0.01, 0.02, x]$



k-means clustering on eigenvectors with small eigenvalues