Title: Clustering: Good, Bad, and Spectral

L10: Spectral Clustering

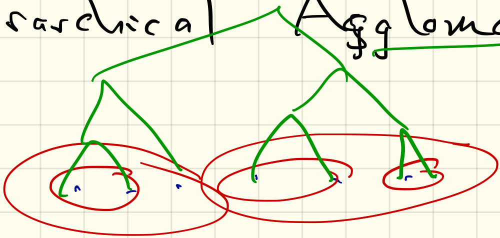
Jeff M. Phillips

February 10, 2020

Clustering

$$X, d \Rightarrow S_1, S_2, \dots S_k \subset X$$

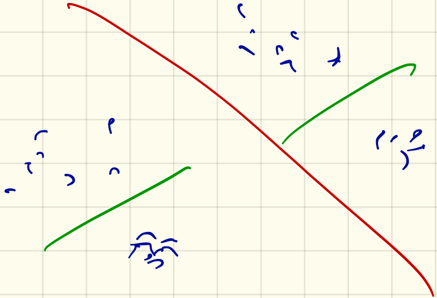
1. Heirarchical Agglomerative Clustering



2. Assignment-based Clustering

3. Spectral Clustering

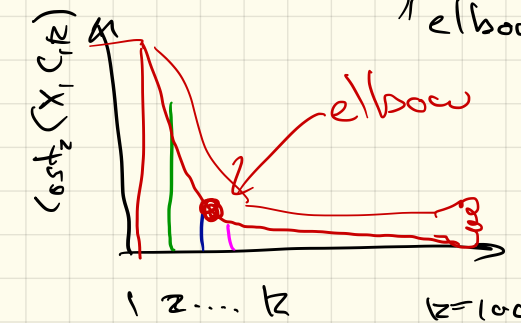
- hierarchy (top down)
+ dimensionality reduction
+ graphs

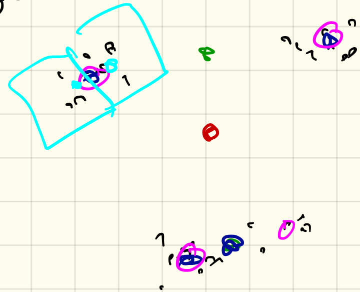


k-means / k-medoid

1. How to choose k?

"elbow" technique





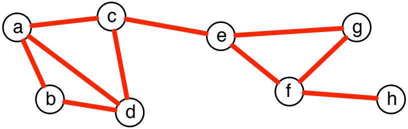
2. k-medoid: minimize

$$\text{Cost}(X, C) = \frac{1}{|X|}\sum_{x\in X} d(x, \phi_C(x))$$

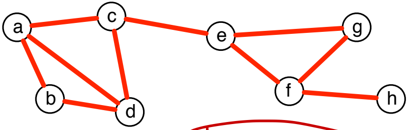
s.t. $C \subset X$

↑ apply to more general X, d

Graphs



Graphs



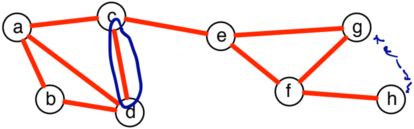
Mathematically: $\boxed{G = (V, E)}$ where

$V = \{a, b, c, d, e, f, g\}$ and

no order

$$E = \Big\{\{a, b\}, \{a, c\}, \{a, d\}, \{b, d\}, \{c, d\}, \{c, e\}, \{e, f\}, \{e, g\}, \{f, g\}, \{f, h\}\Big\}.$$

Graphs



$|V| = n$

Mathematically: $G = (V, E)$ where

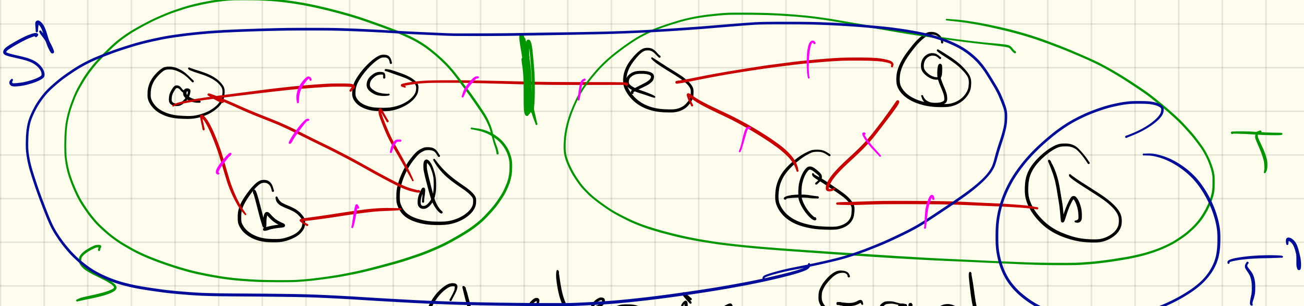
$V = \{a, b, c, d, e, f, g\}$ and

$$E = \Big\{\{a,b\},\{a,c\},\{a,d\},\{b,d\},\{c,d\},\{c,e\},\{e,f\},\{e,g\},\{f,g\},\{f,h\}\Big\}.$$

Matrix-Style: As a matrix with 1 if there is an edge, and 0 otherwise. (For a directed graph, it may not be symmetric).

$n \times n$ matrix

$$G = \begin{array}{c|cccccccc} & a & b & c & d & e & f & g & h \\ \hline a & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ b & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ c & 1 & 0 & 0 & \boxed{1} & 1 & 0 & 0 & 0 \\ d & 1 & 1 & \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ e & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ f & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ g & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ h & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$



Cluster in Graph

Split $V \rightarrow S, T \qquad T = V \setminus S$

Use edges E to measure split

$\text{Cost}(S,T) = \#$ edges in E s.t. $e = \{v_1, v_2\}$ $\quad v_1 \in S$, $v_2 \in T$

$= \text{Cut}(S,T) = 1$

$\text{cut}(S', T') = 1$

Normalize Cut

$Vol(S)$ = # edges with at least 1 endpoint in S.

$Vol(S) = 6$ $\quad Vol(T) = 5$

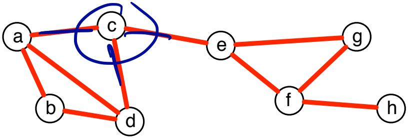
$Vol(S') = 10$ $\quad Vol(T') = 1$

$$NCut(S,T) = \frac{Cut(S,T)}{Vol(S)} + \frac{Cut(S,T)}{Vol(T)}$$

$$= \frac{1}{6} + \frac{1}{5} = 0.367$$

$$NCut(S',T') = \frac{1}{10} + \frac{1}{1} = 1.1$$

Laplacian Matrix



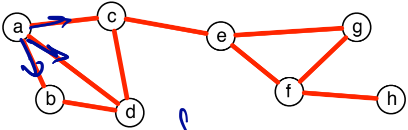
adjacency

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

degree

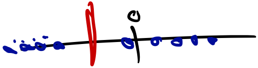
$$D = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

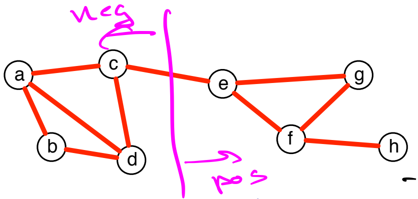
Unnormalized Laplacian Matrix



$$L_0 = D - A = \begin{pmatrix} 3 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 3 & -1 & -1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 3 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}.$$

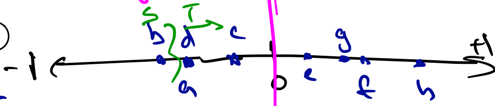
Unnormalized Laplacian Matrix





$[V, \lambda] = \mathrm{eigs}(L_0)$

1. split at 0



2. Scan L→R maintain NCut

eigenvectors of L_0

1st

Fiedler vector

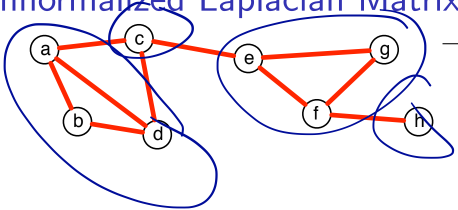
λ	0	0.278	1.11	2.31	3.46	4	4.82
V_a	$1/\sqrt{8}$	−.36	0.08	0.10	0.28	0.25	$1/\sqrt{2}$
b	$1/\sqrt{8}$	−.42	0.18	0.64	−.38	0.25	0
c	$1/\sqrt{8}$	−.20	−.11	0.61	0.03	−.25	0
d	$1/\sqrt{8}$	−.36	0.08	0.10	0.28	0.25	$-1/\sqrt{2}$
e	$1/\sqrt{8}$	0.17	−.37	0.21	−.54	−.25	0
f	$1/\sqrt{8}$	0.36	−.08	−.10	−.28	0.75	0
g	$1/\sqrt{8}$	0.31	−.51	−.36	−.56	0.56	0
h	$1/\sqrt{8}$	0.50	0.73	0.08	0.11	0.11	0

← eigen values

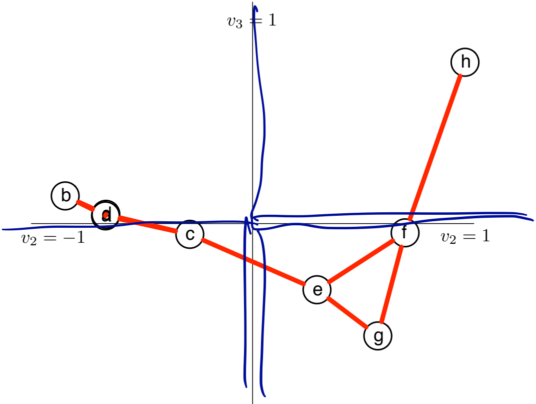
← eigen vectors

→ How to embed G into $\mathbb{R}^1$

Unnormalized Laplacian Matrix



λ	**0.278**	1.11	
V	−.36	0.08	a
	−.42	0.18	b
	−.20	−.11	c
	−.36	0.08	d
	0.17	−.37	e
	0.36	−.08	f
	0.31	−.51	g
	0.50	0.73	h
	v_2	v_3	



Affinity Matrix

$A = \mathbb{R}^{n \times n}$

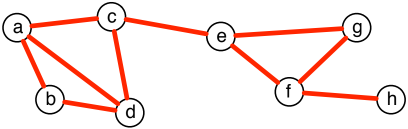
V vertex set
$\parallel$
X data set, similarity
$S: X \times X \to \mathbb{R}$
$s(a,b) = [0,1]$

$$A = \begin{array}{cc} & b \\ g & \begin{bmatrix} & & \\ & \square = s(b,g) & \\ & & \end{bmatrix} \end{array}$$

$$D = \begin{bmatrix} \ddots & & 0 \\ & D_{ii} = \sum_{j=1}^{n} s(i,j) & \\ 0 & & \ddots \end{bmatrix}$$

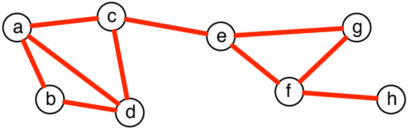
$$L_0 = D - A$$

Laplacian Matrix



$$D^{-1/2} = \begin{pmatrix} 0.577 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.707 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.577 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.577 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.577 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.577 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.707 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Laplacian Matrix

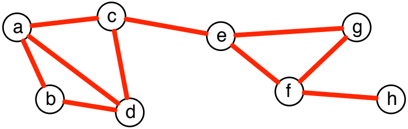


normalized Laplacian

$L = I - D^{-1/2} A D^{-1/2} =$ $D^{-1/2} \, L_0 \, D^{-1/2}$

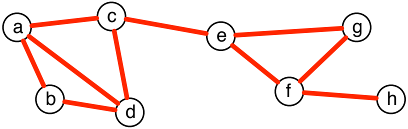
$$\begin{pmatrix} 1 & -0.408 & -0.333 & -0.333 & 0 & 0 & 0 & 0 \\ -0.408 & 1 & 0 & -0.408 & 0 & 0 & 0 & 0 \\ -0.333 & 0 & 1 & -0.333 & -0.333 & 0 & 0 & 0 \\ -0.333 & -0.408 & -0.333 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.333 & 0 & 1 & -0.333 & -0.408 & 0 \\ 0 & 0 & 0 & 0 & -0.333 & 1 & -0.408 & -0.577 \\ 0 & 0 & 0 & 0 & -0.408 & -0.408 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.577 & 0 & 1 \end{pmatrix}.$$

Laplacian Matrix



eigenvectors of L

Laplacian Matrix

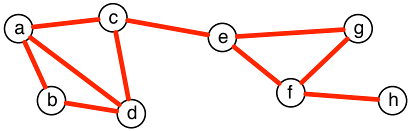


eigenvectors of L

λ	0	**0.125**	0.724	1.00	1.33	1.42	1.66	1.73
V	−.39	0.38	−.09	0.00	0.71	0.26	−.32	0.16
	−.32	0.36	−.27	0.50	0.00	−.51	0.38	−.18
	−.39	0.18	0.36	−.61	0.00	0.03	0.47	−.29
	−.39	0.38	−.09	0.00	−.71	0.26	−.32	0.16
	−.39	−.28	0.48	0.00	0.00	−.57	0.31	0.33
	−.39	−.48	−.29	0.00	0.00	0.05	−.31	−.65
	−.31	−.36	0.27	0.50	0.00	0.51	0.38	−.18
	−.22	−.32	−.61	−.35	0.00	−.07	0.27	0.51

Laplacian Matrix

scaled by $\frac{1}{\lambda_i}$



λ	**0.125**	0.724	
V	0.38	$-.09$	a
	0.36	$-.27$	b
	0.18	0.36	c
	0.38	$-.09$	d
	$-.28$	0.48	e
	$-.48$	$-.29$	f
	$-.36$	0.27	g
	$-.32$	$-.61$	h
	v_2	v_3