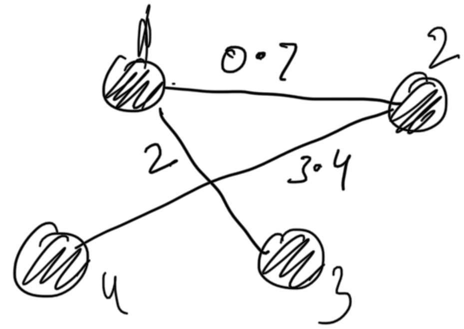
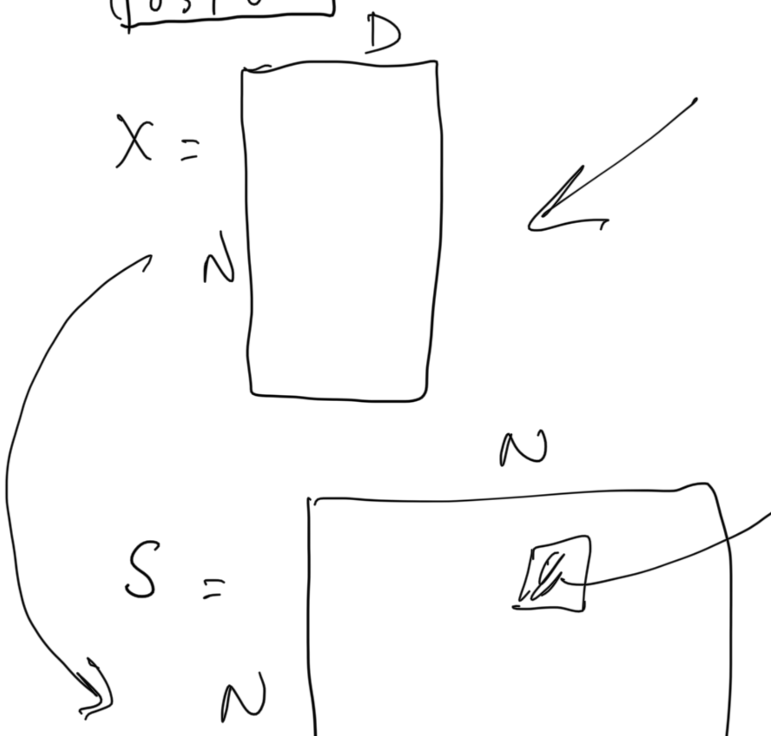
Clustering

Fri Apr 30 -

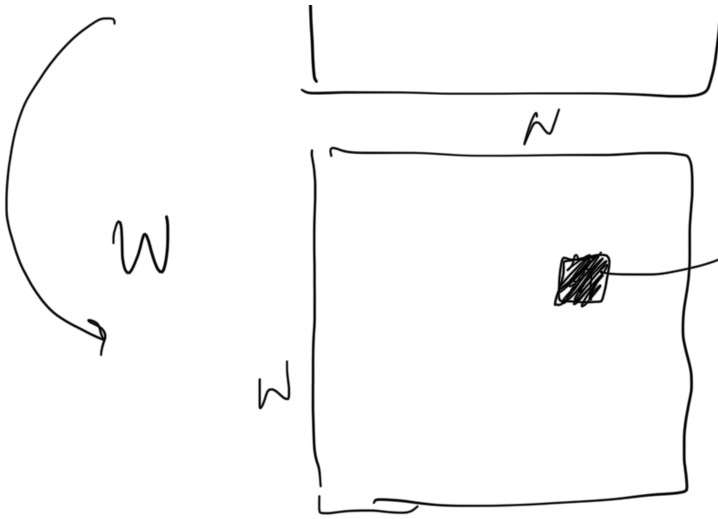
weighted undirected Graphs.



$$
\begin{array}{c|cccc|}
 & 1 & 2 & 3 & 4 \\
\hline
1 & 0 & 0.7 & 2 & 0 \\
2 & 0.7 & 0 & 0 & 3.4 \\
3 & 2 & 0 & 0 & 0 \\
4 & 0 & 3.4 & 0 & 0 \\
\hline
\end{array}
$$



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



$w_{ij} = S[i,j]$

if x_i is the nearest neighbor of x_j

$= 0$

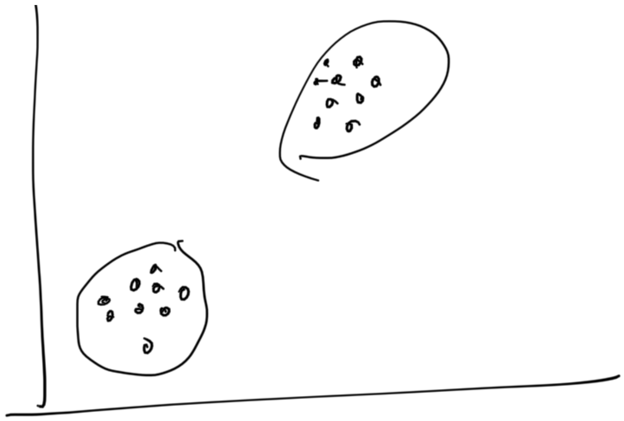
$$X = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}$$

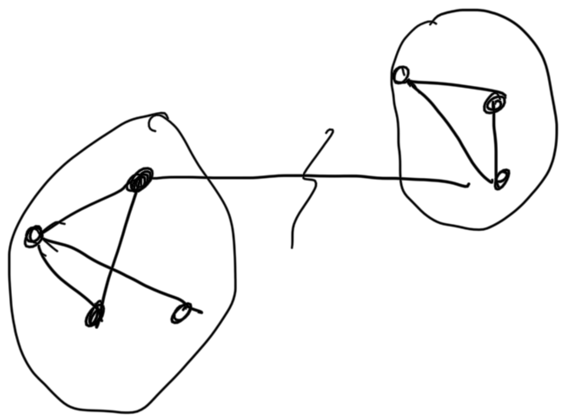
$$\text{sim}(\ ) = \frac{1 - \text{euc}(\ )}{\cancel{\text{euc}(\ )}}$$

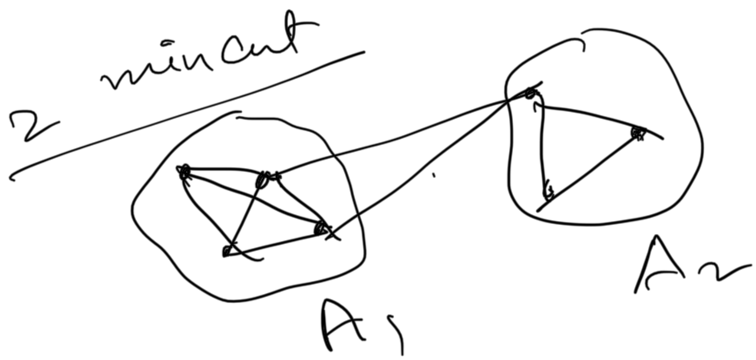
$$S = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & -1 \\ -1 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & & 0 & \\ 0 & & & 0 \end{bmatrix}$$

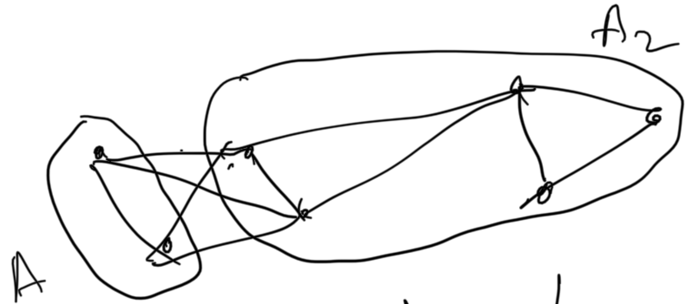
1



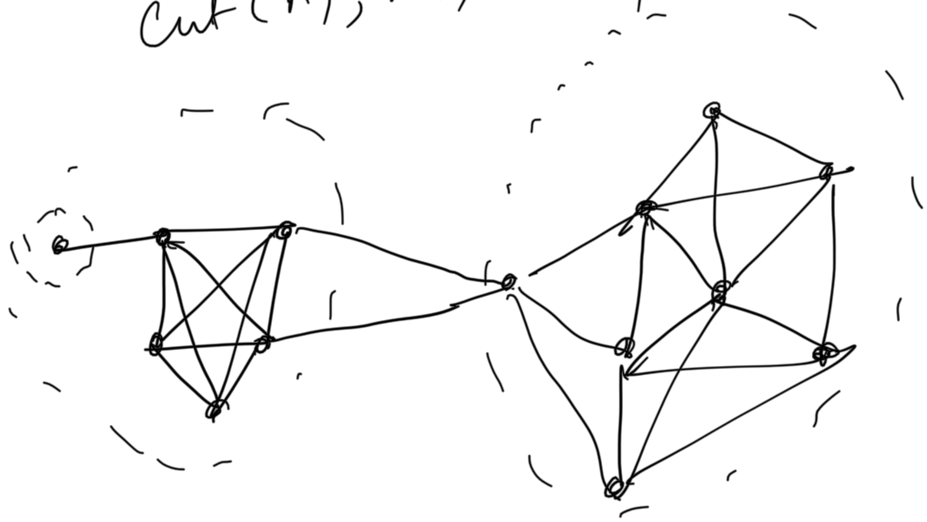



2 min cut
A1
A2

$cut(A_1, A_2) = 2$



$cut(A_1, A_2) = 4$



$W_{N\times N}$ — adjacency matrix

$D_{N\times N}$ — diagonal matrix

$$D_{ii} = \sum_{j=1}^{N} W_{ij}$$

$$L = D - W$$

Laplacian Matrix

$$L = D - W$$

$$\boxed{Lx = \lambda x}$$

$x = (N \times 1)$ λ - scalar

$$x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{1} \qquad \lambda = 0$$

$L\mathbf{1}$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad 0 \times \begin{matrix} 1 \\ 1 \\ 1 \\ 1 \end{matrix}$$

$$
W = \begin{array}{c|c} & \begin{matrix}1 & 2 & 3 & 4 & 5\end{matrix} \\ \hline \begin{matrix}1\\2\\3\\4\\5\end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}
\qquad
D = \begin{bmatrix} 2 & & & & \\ & 2 & & 0 & \\ & & 2 & & \\ & 0 & & 1 & \\ & & & & 1 \end{bmatrix}
$$

$$
L = D - W = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}
$$

$$
\begin{bmatrix} \boxed{\begin{matrix}\neq 0 \\ \neq 0 \\ \neq 0\end{matrix}} \\ 0 \\ 0 \end{bmatrix} \quad \underline{0}
\qquad\qquad
\begin{bmatrix} 0 \\ 0 \\ 0 \\ \boxed{\begin{matrix}\neq 0 \\ \neq 0\end{matrix}} \end{bmatrix} \quad \underline{0}
$$

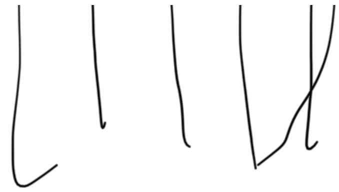
<u>N = 10</u>

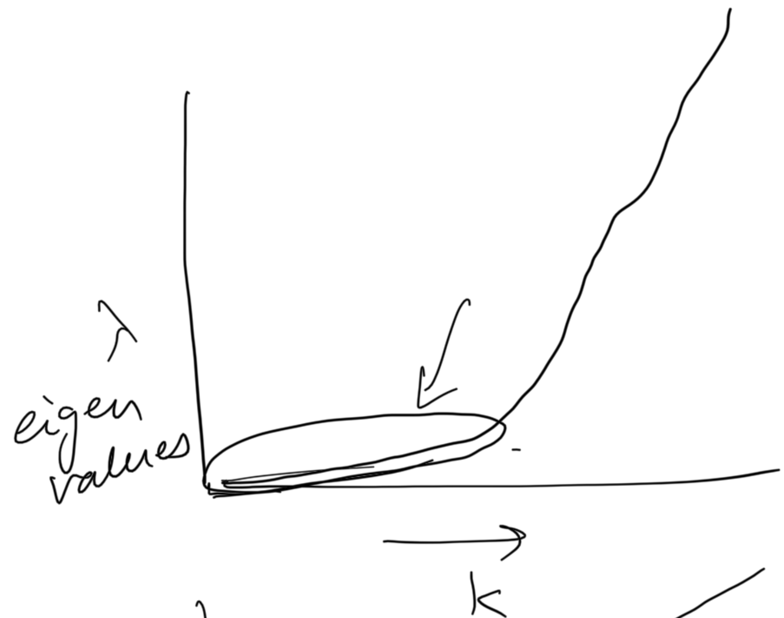
$$
\boxed{\begin{matrix} 0 \\ 0.3 \\ 0.2 \\ 0.25 \end{matrix}}
$$

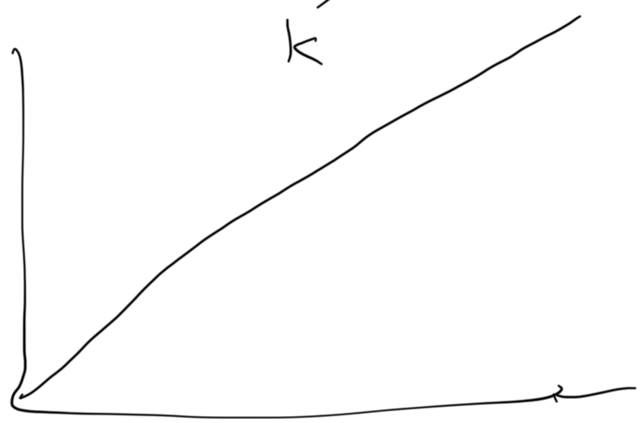
3.4

3.8

3.9
4.7
5.2
8.0







Mon May 3

$X \xrightarrow[\text{sim}(\ ,\ )]{} \text{Graph}$

W adjacency (N×N) matrix

Clustering problem $\Longrightarrow$ Graph partitioning problem

Laplacian Matrix L:

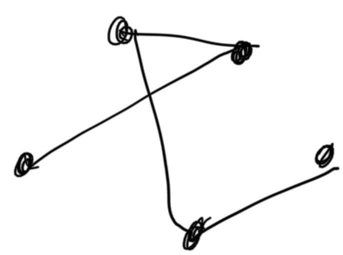
$$L = \underbrace{D}_{\text{degree matrix}} - W$$

$$\boxed{D_{ii} = \sum_{j=1}^{N} w_{ij}}$$

$D_{ij} = 0$ if $i \neq j$

$$\boxed{\underbrace{L}_{N \times N} \underbrace{x}_{N \times 1} = \underbrace{\lambda}_{\text{scalar}} x}$$

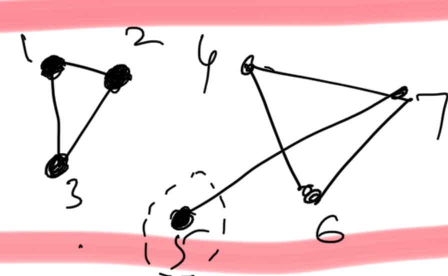
1 - component



1 eigen value = 0

1 eigen vector: $[1, 1, 1, 1, 1]^T$

2 - components

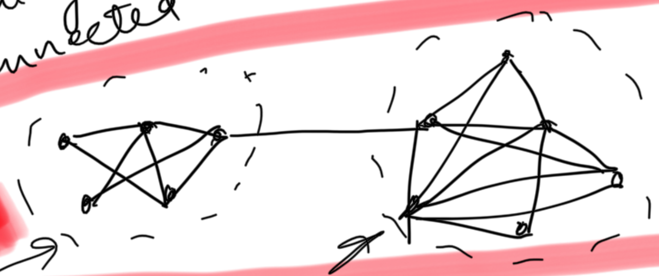


2 eigen values = 0

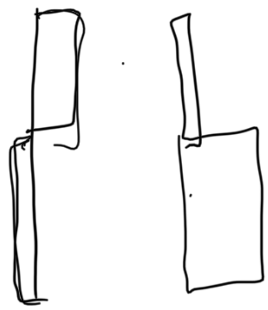
2 eigen vectors: $[1\ 1\ 1\ 0\ 0\ 0\ 0]^T$

$[0\ 0\ 0\ 1\ 1\ 1\ 1]$

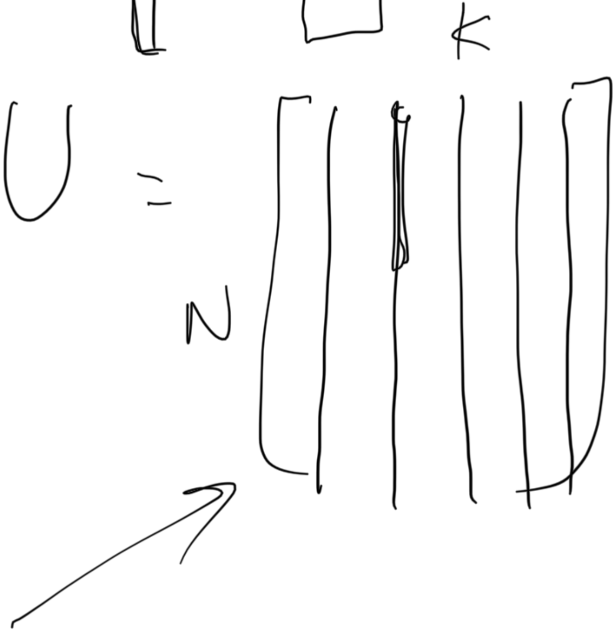
weakly connected components.



2 eigen values which are very small.



$U =$



Clus