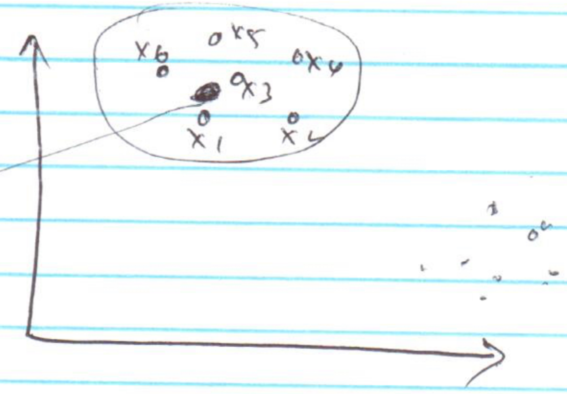
Gradiance 10 – Due Wednesday night

Clustering

$$\begin{array}{cc} x_{11}, & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ \vdots & \\ x_{61} & x_{62} \\ \hline (\bar{x}_{*1}, & \bar{x}_{*2}) \end{array}$$



Generalizing k-means to non-vector data.

1. Use a different distance/similarity measure.
2. Use a different way to find the "center"

"medoid"

Given a set of instances
A medoid is an instance that is closest (on average) to all other instances

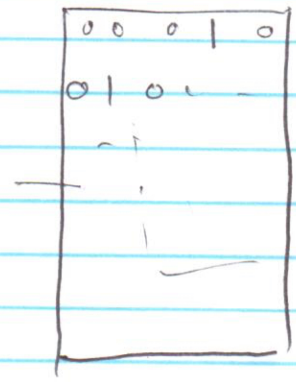
k-medoid

Time series data

Patient 1

Output of k-means : $\mathbf{C} \rightarrow \begin{matrix} c_1 \\ c_2 \\ \vdots \\ c_K \end{matrix} \in \mathbb{R}^D$

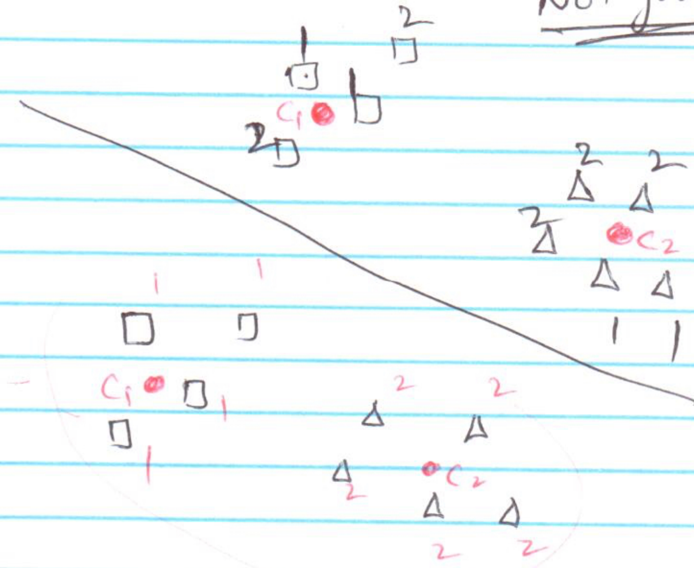
$\underset{N \times K}{R}$



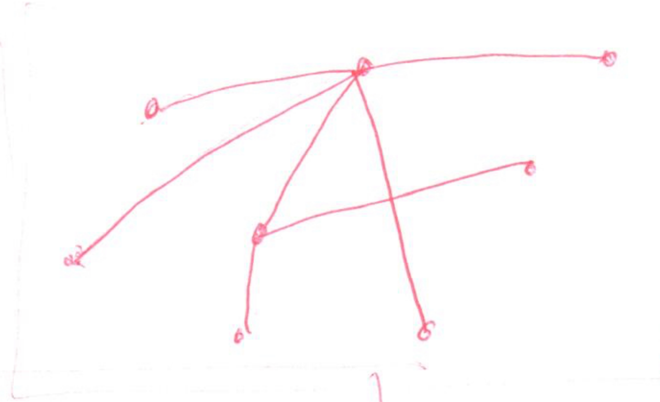
$R_{nk} = 1$

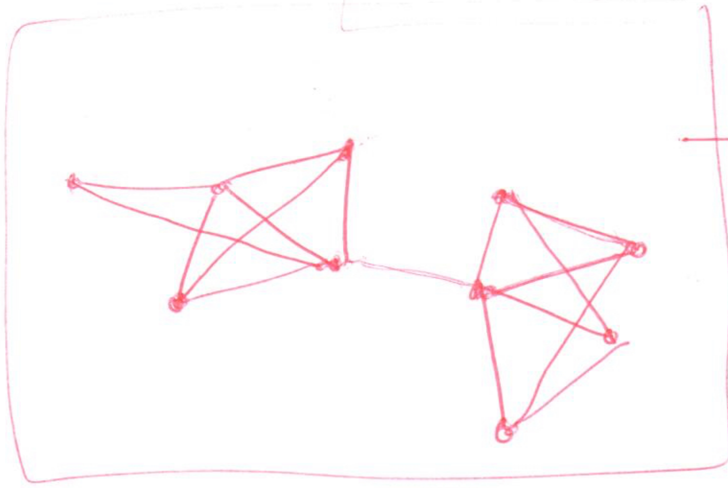
$$J(C, R) = \sum_{n=1}^{N} \sum_{k=1}^{K} R_{nk} \| x_n - c_k \|_2^2$$

Not good



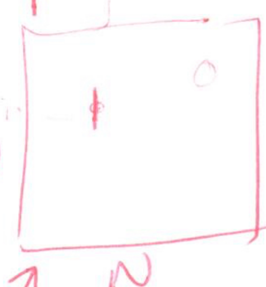
Graph, $G = \{V, E\}$





$\rightarrow$ Graph Partitioning

$= A_N$



Adjacency Matrix.

$A_{ij} \in \{0, 1\}$

Spectral Clustering

$$X = \begin{matrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{matrix} \qquad x_i \in R^D$$

1. Creating a graph G from X — with N vertices.
2. Solve a graph partitioning problem on G
   "Min-cut problem"

Weighted graph: $A_{ij} \in R^+$

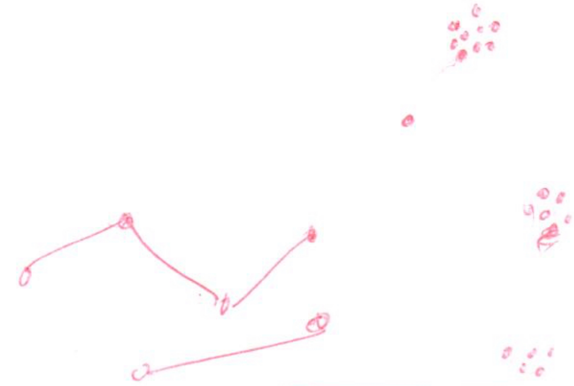
Directed graph. $\boxed{A_{ij} \neq A_{ji}}$

In spectral clustering we create a weighted-undirected-graph

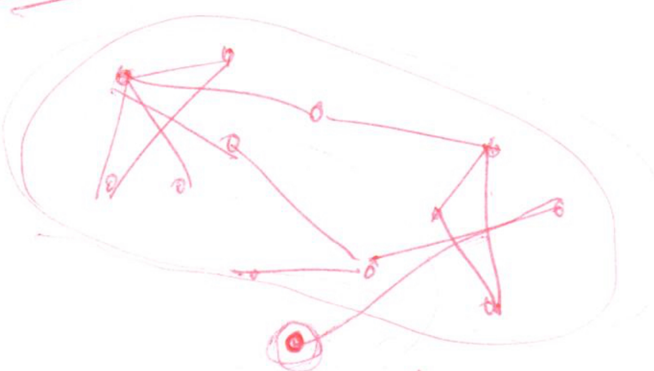
$$W_{(N\times N)} \quad W_{ij} = \begin{cases} Sim(x_i, x_j) & \text{if } x_i \text{ is the nearest neighbour of } x_j \\ 0 & \text{otherwise.} \end{cases}$$

S

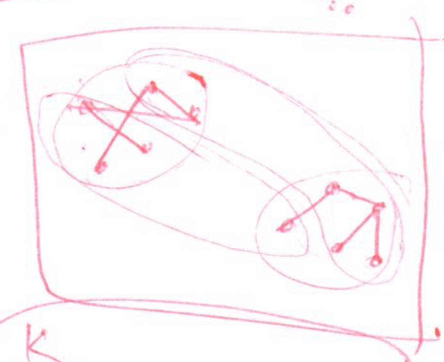
Adjacency matrix of G.



$k=2$



Min-Cut



$$\sum_{k=1}^{K} \frac{1}{2} W(A_k, \bar{A}_k)$$

Trivial Solution

Normalized Min-Cut

$$Vol(A_k) = \sum_{i=1}^{N_k} d_i \longrightarrow \text{weighted degree of node } i$$

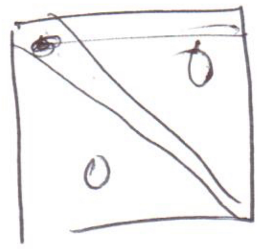
$\boxed{2+2+1+2+1}$

Adjacency matrix of a graph $G \rightarrow W$

$W \rightarrow (N \times N)$ W_{ij} = weight of the edge between node i & node j

Degree matrix

D → Diagonal matrix
$(N \times N)$



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

$$\boxed{L = D - W}$$

$(N \times N)$ → Laplacian of G

① Each row sum to 0

$$\boxed{L} = \boxed{D} - \boxed{W}$$

② L will have one eigen vector which is all 1 and corresponding eigen value of 0

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

A → square symmetric matrix $(N \times N)$

x → vector $(N \times 1)$

We will get N solutions.

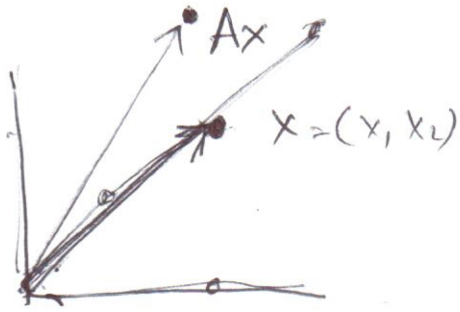
N x's and λ's → scalars

x's → eigenvectors

λ's → eigen values.

$$A \cdot \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

2×2 $\quad$ 2×1



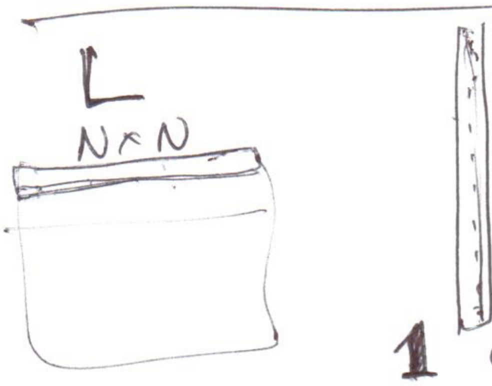
$$\begin{bmatrix} A_{11} x_1 + A_{12} x_2 \\ A_{21} x_1 + A_{22} x_2 \end{bmatrix}$$

~~Ax~~ For a given matrix A:
if we have a vector x, such that Ax only results in stretching
then x is an eigen-vector of A.

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

L $\qquad$ $L\mathbb{1}$ $\qquad$ $L_{i*} * \mathbb{1}$ = row sum of L = 0

$N \times N$



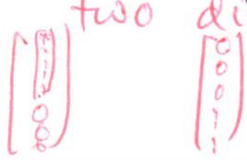
=

$\mathbb{1}$ $(N \times 1)$

$$L \cdot \mathbb{1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \equiv 0 * \mathbb{1}$$

Hence, $\mathbb{1}$ is an eigen vector of L and 0 is the corresponding eigen values.

If G has two disconnected components.



Ax,

$$\underset{(N \times N)}{S} \rightarrow \frac{S + S^T}{(N \times N)}$$

$$\boxed{\underset{(N\times N)}{A}\, \underset{(N\times 1)}{x} = \underbrace{\lambda}_{\text{scalar}} . x} \rightarrow (1)$$

- Each x that satisfies (1) is an eigen vector of A
  - Corresponding λ is called an eigen value.

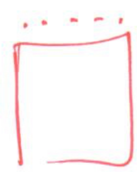
For a Laplacian $x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix}$ will be one of the eigen vectors.

and corresponding eigen value will be $\underline{\underline{0}}$

$L \begin{bmatrix} 1 \\ 1 \\ \vdots \end{bmatrix}$

Spectral Clustering:

- X $\underline{\underline{N \times D}}$
- S (sim ( ))
- W (1-nearest neighbor)
- D
- $L = D - W$
- eig (L) → e, v → (p) NxN



- Use smallest few eigen values.
- Take the eigen vectors corresponding to those p eigen values.

→ run k-means on this data.