# Spectral clustering

INF391 - MINERÍA DE DATOS

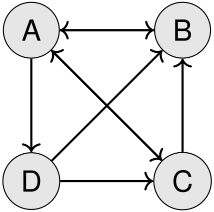
nicolas.torresr@usm.cl
Universidad Técnica Federico Santa María

2020-1

Departamento de Informática
Universidad Técnica Federico Santa María

# Graphs and matrices

- A graph is a set of vertices connected with edges:




- A graph can also be represented by its adjacency matrix $\mathcal{A}$.

| | A | B | C | D |
|---|---|---|---|---|
| A | 0 | 1 | 1 | 1 |
| B | 1 | 0 | 0 | 0 |
| C | 1 | 1 | 0 | 0 |
| D | 0 | 1 | 1 | 0 |

# Eigenvalues and eigenvectors

- The eigenvalues $\lambda$ and the eigenvectors $\nu$ of a square matrix $A$ are defined as follows:

$$A\nu = \lambda\nu [\Rightarrow (A - \lambda I)\nu = 0]$$

- In matrix form:

$$\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- This is solved when:

$$(a_{11} - \lambda)\nu_1 + a_{12}\nu_2 = 0$$
$$a_{21}\nu_1 + (a_{22} - \lambda)\nu_2 = 0$$

# Example of calculating eigenvalues and eigenvectors

- Consider the following example: $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$
- Using $(A - \lambda I)\nu = 0$ we get:

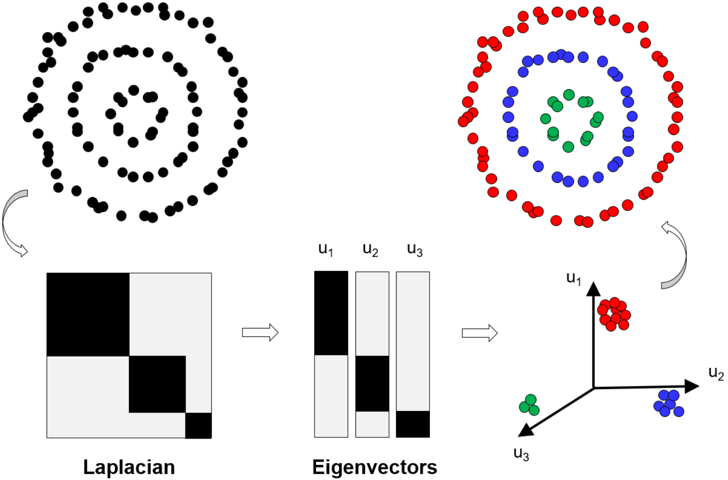
$$\begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- Solved when $det(A) = 0 : (1-\lambda)^2 - 4 = 0$.
- Using $\lambda_1 = 3$ and $\lambda_2 = -1$ we obtain $\nu = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\nu = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

# Spectrum of a graph

- The spectrum of a graph are the eigenvalues of the adjacency matrix $\mathcal{A}$ of the graph.
- The spectrum is considered to capture important structural properties of a graph.
- Some interesting applications of eigenvalues and eigenvectors:
  - Principal Component Analysis.
  - PageRank.
  - Partitioning (i.e. clustering).

# Clustering



# Spectral clustering algorithm

**Algorithm 1** Spectral clustering

1: Compute the affinity (similarity) matrix.
2: Compute the (normalized) Laplacian matrix.
3: Compute the eigenvalues and eigenvectors of Laplacian matrix.
4: Project original objects into a low-dimensional space defined by the $k$ leading eigenvectors.
5: Apply K-Means to find $k$ clusters.

# Affinity Matrix $\mathcal{A}$

$$\mathcal{A}_{i,j} = w_{i,j} \tag{1}$$

where $w_{i,j}$ could be the proximity between $i$ and $j$, or

$$w_{i,j} = \begin{cases} 1 & \text{if exists affinity (edge) between vertices } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

# Degree Matrix $\mathcal{D}$

$$\mathcal{D}_{i,i} = degree(i) \tag{2}$$

where

$$degree(i) = \sum_{j=1}^{n} w_{i,j}$$

# Laplacian Matrix $\mathcal{L}$

$$\mathcal{L} = \mathcal{D} - \mathcal{A} \tag{3}$$

where

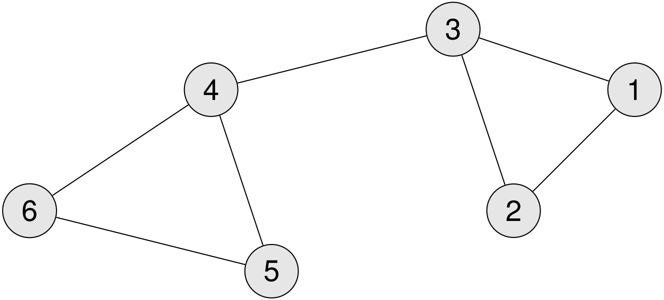
$$\mathcal{L}_{i,j} = \begin{cases} degree(i) & \text{if } i = j \\ -w_{i,j} & \text{otherwise} \end{cases}$$

# Eigenvalues and eigenvectors

$$\mathcal{L}\nu = \lambda\nu \tag{4}$$

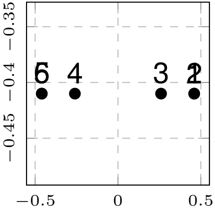
where $\lambda$ is the *eigenvalue* associated with the *eigenvector* $\nu$.

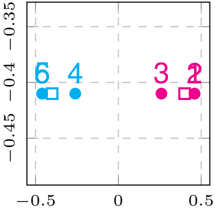
# Step-By-Step Example 1

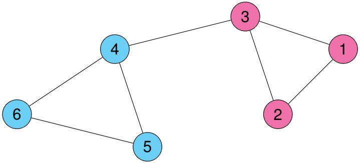


# Step-By-Step Example 1 (cont'd)

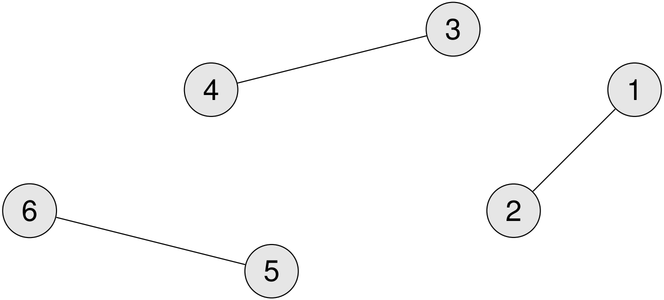
$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix} \mathcal{L} = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{pmatrix} \nu = \begin{pmatrix} 0.2 & 0.0 & 0.0 & 0.8 & 0.5 & -0.4 \\ 0.2 & 0.2 & -0.5 & -0.5 & 0.5 & -0.4 \\ -0.7 & -0.2 & 0.5 & -0.2 & 0.3 & -0.4 \\ 0.7 & -0.2 & 0.5 & -0.2 & -0.3 & -0.4 \\ -0.2 & -0.6 & -0.5 & 0.1 & -0.5 & -0.4 \\ -0.2 & 0.8 & -0.0 & 0.1 & -0.5 & -0.4 \end{pmatrix}$$



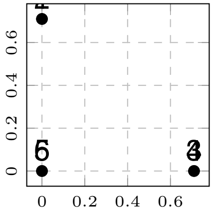



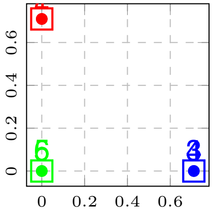
# Step-By-Step Example 2

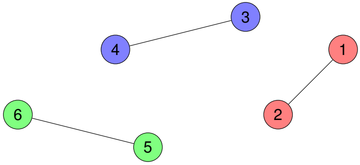


# Step-By-Step Example 2 (cont'd)

$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \mathcal{L} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \nu = \begin{pmatrix} 0.0 & -0.7 & 0.0 & 0.0 & 0.0 & 0.7 \\ 0.0 & 0.7 & 0.0 & 0.0 & 0.0 & 0.7 \\ 0.0 & 0.0 & -0.7 & 0.0 & 0.7 & 0.0 \\ 0.0 & 0.0 & 0.7 & 0.0 & 0.7 & 0.0 \\ -0.7 & 0.0 & 0.0 & 0.7 & 0.0 & 0.0 \\ 0.7 & 0.0 & 0.0 & 0.7 & 0.0 & 0.0 \end{pmatrix}$$

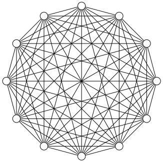


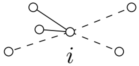



1. Spectral clustering
   - Graph Theory
   - Clustering
   - Examples
   - Graphs

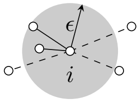
# How to Create the Graph?



- A fully connected graph (complete graph).

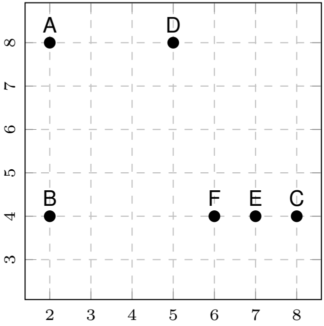



- $k$-nearest-neighbours graph ($k$-NNG): $i \sim j$ **iff** $j$ is one of the $k$ nearest neighbours of $i$.




- $\epsilon$-ball graph: $i \sim j$ **iff** $j \in B(i; \epsilon)$.

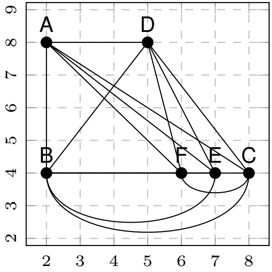
# How to Create the Graph? Example



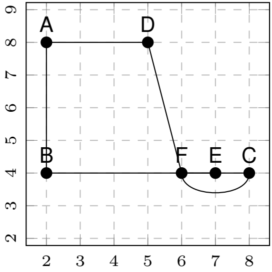

- Fully connected graph.
- $k$-NN-graph with $k = 2$.
- $\epsilon$-ball graph with $\epsilon = 3$.

# How to Create the Graph? Example (cont'd)

**Fully connected graph**




$k$**-NN-graph with** $k = 2$




$\epsilon$**-ball graph with** $\epsilon = 3$