Spectral Graph Theory

Introduction to Network Science

Instructor: Michele Starnini — https://github.com/chatox/networks-science-course

Universitat Pompeu Fabra Barcelona

Contents

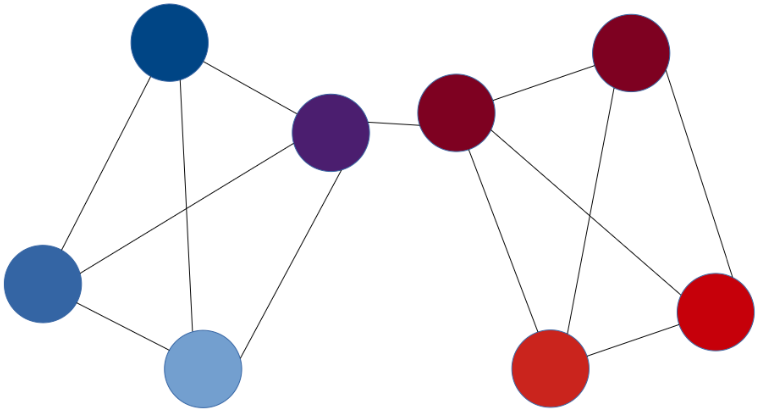
- Graph Embedding
- Linear Algebra recap

Many algorithms are not suitable for graphs

- Many algorithms need a notion of similarity or distance (both are interchangeable)
- **Data mining**: clustering, outlier detection, ...
- **Retrieval/search**: nearest neighbors, ...

Graphs are nice, but ...

- They describe only local relationships
- We would like to understand a global structure
- We will try to transform a graph into a more familiar object: a cloud of points in R^k



Distances should be somehow preserved

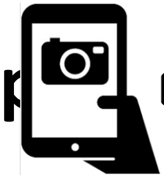
What is a graph embedding?

- A graph **embedding** (or graph **projection**) is a mapping from a graph to a vector space
- If the vector space is $\mathbb{R}^2$ you can think of an embedding as a way of ***drawing*** a graph on paper

Exercise: draw this graph

V = {v1, v2, …, v8}

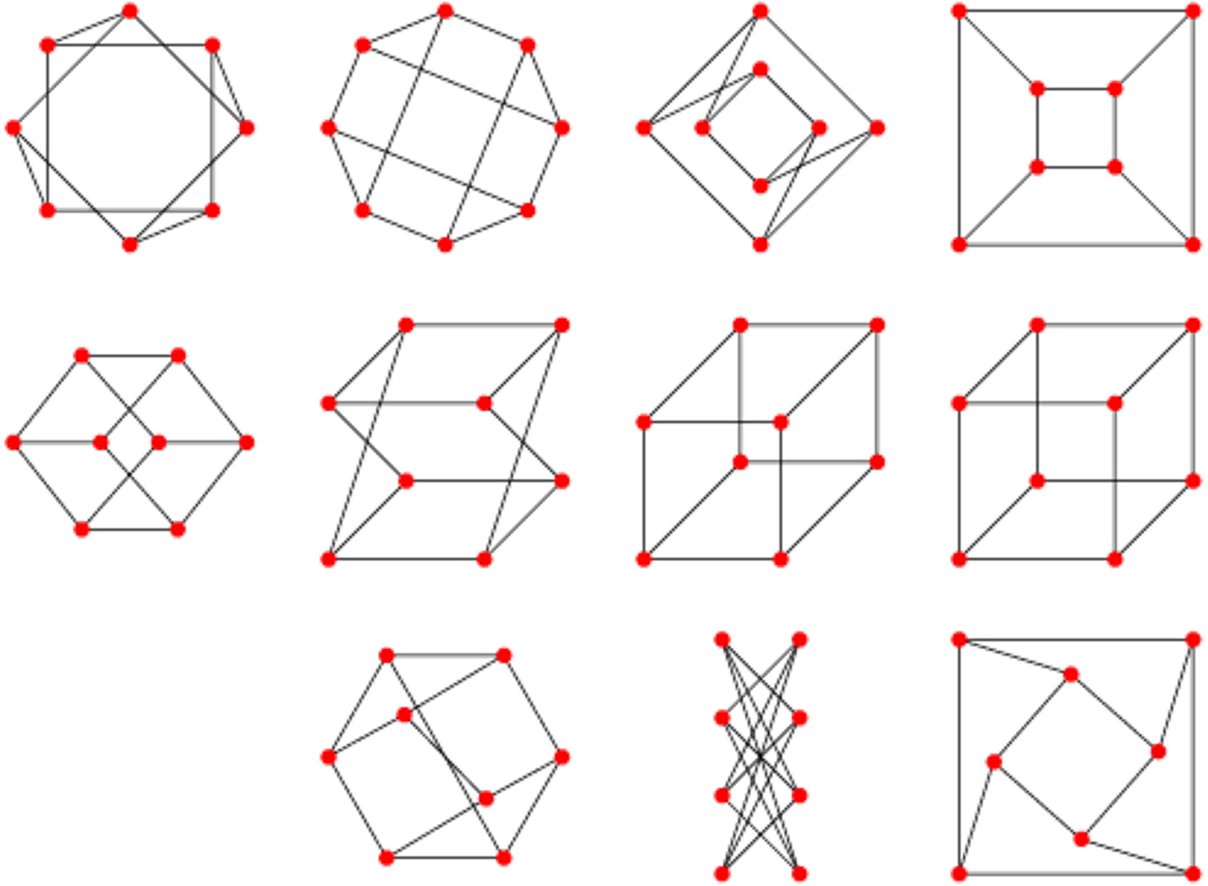
E = { (v1, v2), (v2, v3), (v3, v4), (v4, v1), (v5, v6), (v6, v7), (v7, v8),
(v8, v5), (v1,v5), (v2, v6), (v3, v7), (v4, v8) }

Draw this graph on paper, upload a p  **o**

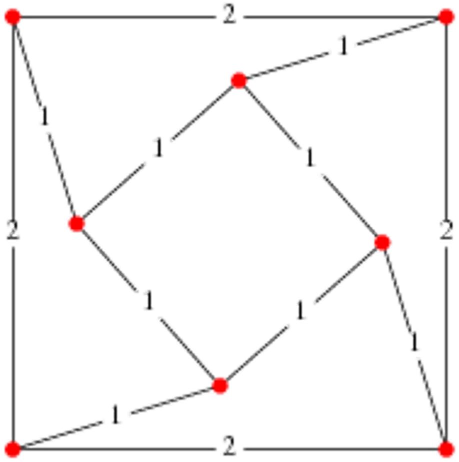
What constitutes a good drawing?

Padlet: https://upfbarcelona.padlet.org/chato/3hwnctvqb7p1z6xx

Example drawings



Look at the 3 neighbors of each corner, they should all be at the same distance

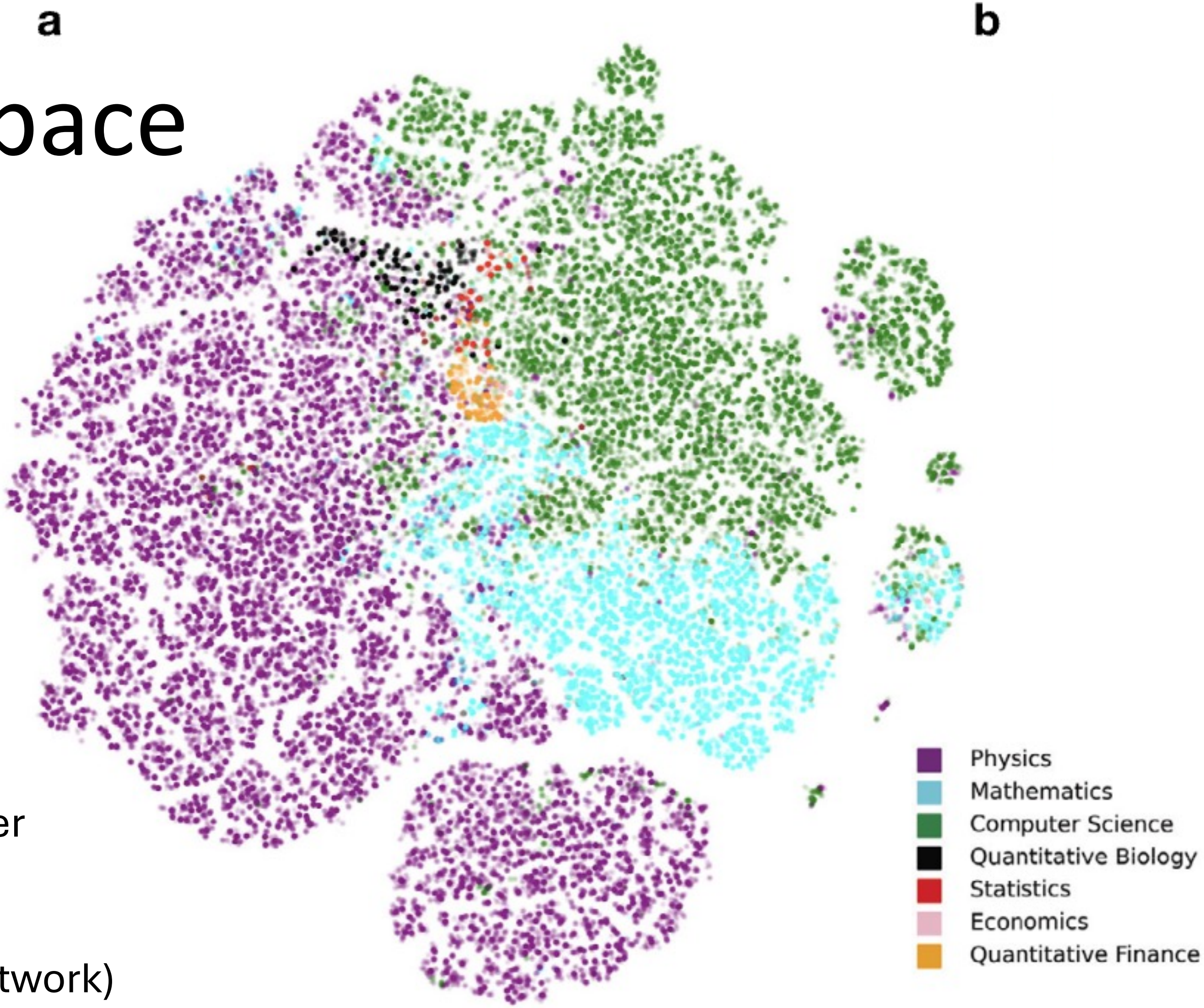


Distances are not preserved well

In a good graph embedding ...

- Pairs of nodes that are **connected** to each other should be **close**
- Pairs of nodes that are not connected should be far
- Compromises will need to be made

Research space



Each point is a scientific paper
Two papers are connected if
they share a common author
(left projection of bipartite network)

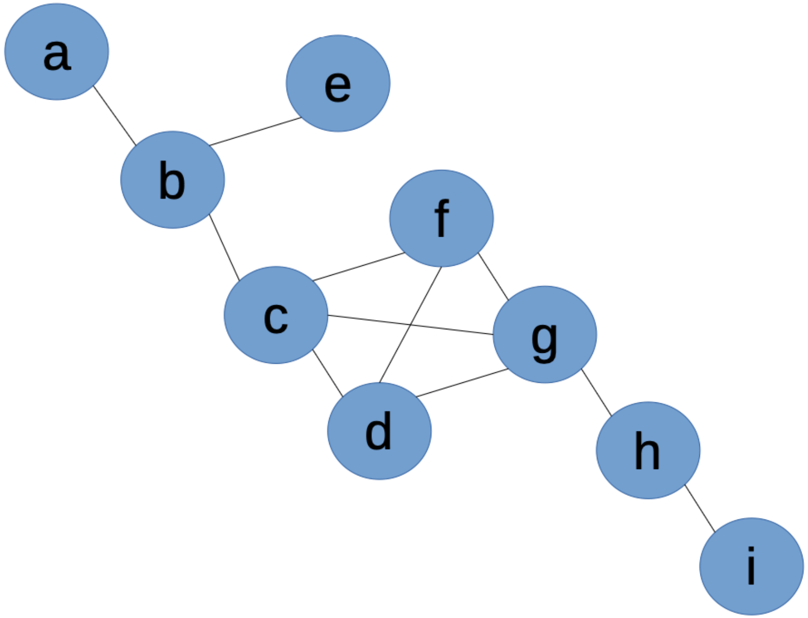
Random projections

Random graph projection (2D)

- Start a BFS from a random node, that has x=1, and nodes visited have ascending x
- Start a BFS from another random node, which has y=1, and nodes visited have ascending y
- Project node i to position (x_i, y_i)

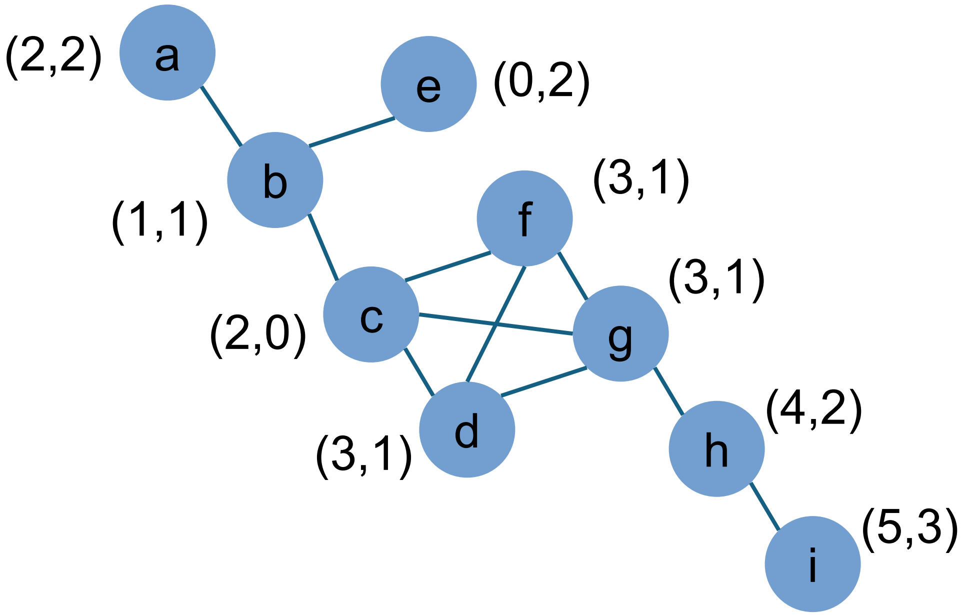
Exercise: random projection

- Given this graph
- Pick a random node *u*
  - Distances from *u* are the x positions
- Pick a random node v
  - Distances from *v* are the y positions
- Draw the graph in an $\mathbb{R}^2$ plane

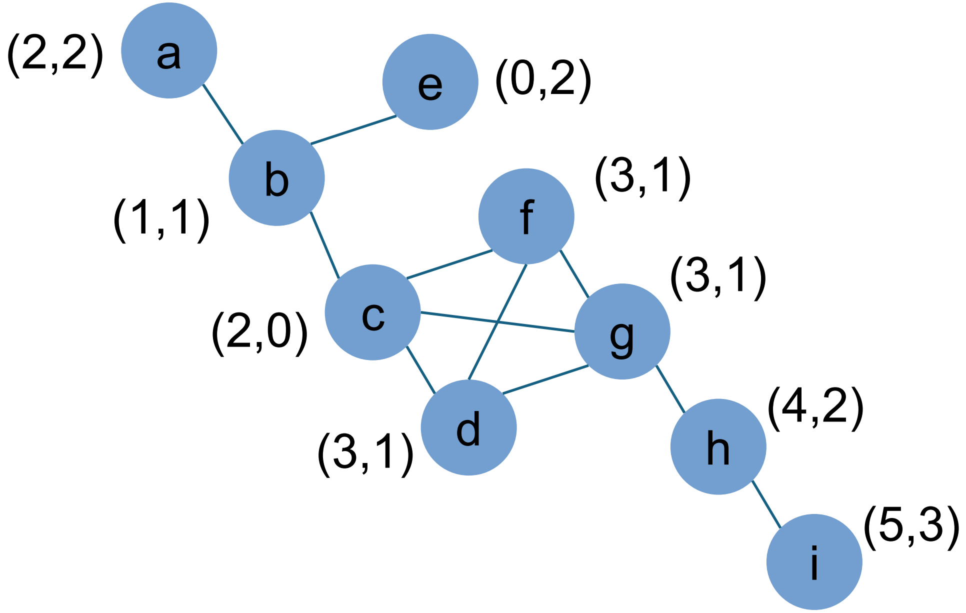


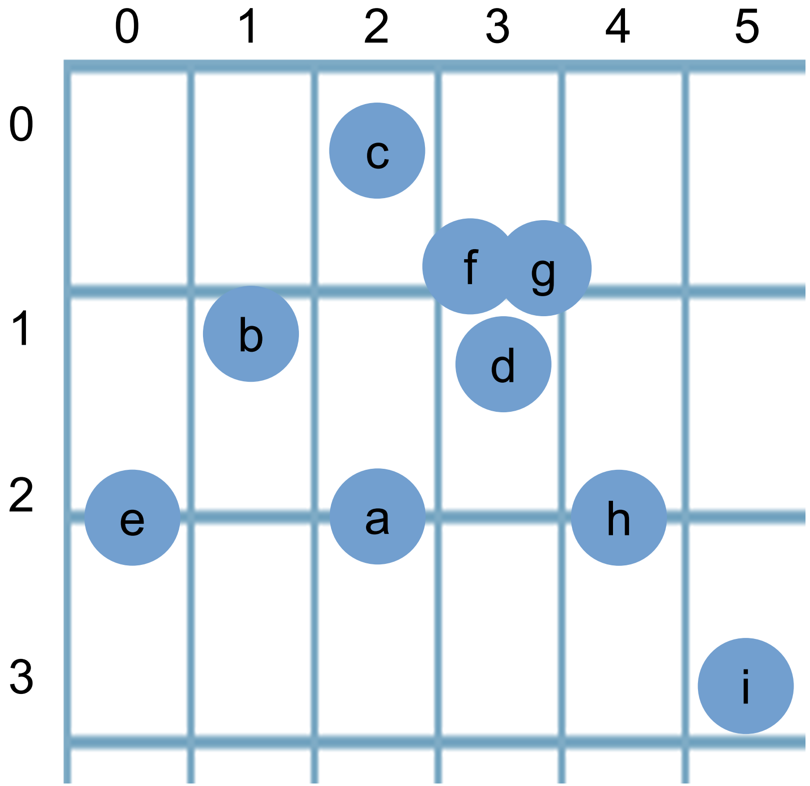
Padlet: https://upfbarcelona.padlet.org/chato/9pd56scbpko5svdj

Answer (Random nodes: e, c)



Answer (Random nodes: e, c)





Spectral Graph Theory

- Understand structural properties of graphs from the characteristics of the adjacency matrix

- Spectral characteristics of a matrix is linear algebra: eigenvalues & eigenvectors

- Spectral graph theory will tell you how to embed a graph

Refresher about eigenvectors/eigenvalues

Eigenvectors of symmetric matrices

- In general $Av = \lambda v$ means A has an eigenvector v of eigenvalue λ
- In **symmetric** matrices ($A=A^T$), **eigenvectors are orthogonal**

Suppose v_1, v_2 are eigenvectors of eigenvalues λ_1, λ_2 with $\lambda_1 \neq \lambda_2$

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle A v_1, v_2 \rangle = \langle v_1, A^T v_2 \rangle$$
$$= \langle v_1, A v_2 \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

For any real matrix

$$\langle Ax, y \rangle = \langle x, A^T y \rangle$$

- Therefore:

$$(\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle = 0 \wedge (\lambda_1 - \lambda_2) \neq 0 \Rightarrow \langle v_1, v_2 \rangle = 0$$

In symmetric matrices

- The **multiplicity** of an eigenvalue λ is the dimension of the space of eigenvectors of eigenvalue λ
- Every n x n symmetric matrix has n eigenvalues counted with multiplicity
- Hence, it has an orthonormal basis of eigenvectors

Eigenvectors of the adjacency matrix
(of an unweighted graph)

Adjacency matrix (binary graph)

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

- How many non-zeros are in every **row** of A?

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

https://www.youtube.com/watch?v=AR7iFxM-NkA

Adjacency matrix of G=(V,E)

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

- **What is *Ax*?** Think of *x* as a set of labels/values:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$y_i = \sum_{j:(i,j)\in E} x_j$$

Ax is a vector whose i^{th} coordinate contains the sum of the x_j who are in-neighbors of i

https://www.youtube.com/watch?v=AR7iFxM-NkA

Spectral graph theory ...

- Studies the eigenvalues and eigenvectors of a graph matrix
  - Adjacency matrix $Ax = \lambda x$
  - Laplacian matrix (next)
- Suppose graph is d-regular: $k_i = d \; \forall i$
- Multiply its adjacency by 1
- Look at the result, what does it imply?

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = ?$$

An eigenvector of a d-regular graph

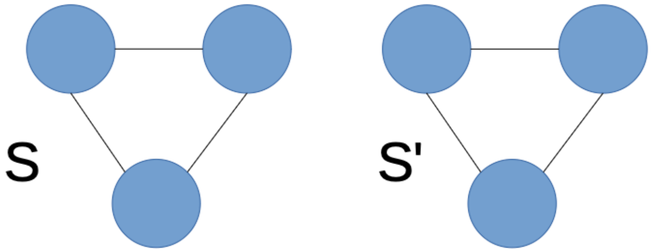
- Suppose graph is d-regular, i.e. all nodes have degree d:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} d \\ d \\ \vdots \\ d \end{bmatrix} = d \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

- **Hence, $[1, 1, \ldots, 1]^T$ is an eigenvector of eigenvalue d**

Disconnected graphs

- Suppose the graph is regular **and disconnected**

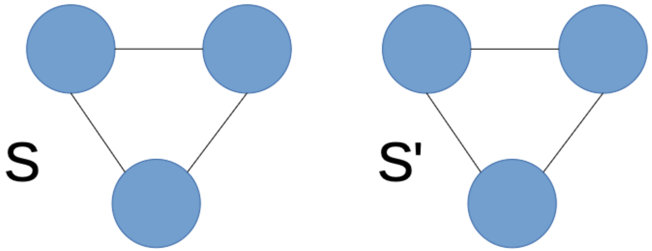


- Then its adjacency matrix has block structure:

$$A = \begin{bmatrix} S & 0 \\ 0 & S' \end{bmatrix}$$

Disconnected graphs

- Suppose the graph is regular **and disconnected**

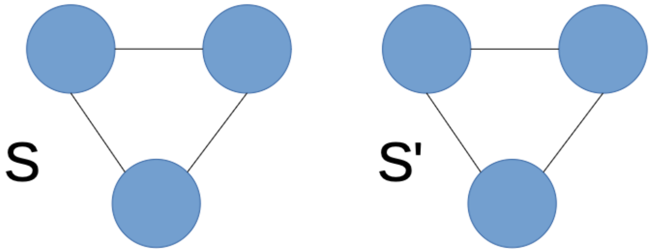


$$\text{Let } x_i^S = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \in S' \end{cases}$$

$$\begin{bmatrix} S & 0 \\ 0 & S' \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = ?$$

Disconnected graphs

- Suppose the graph is regular **and disconnected**

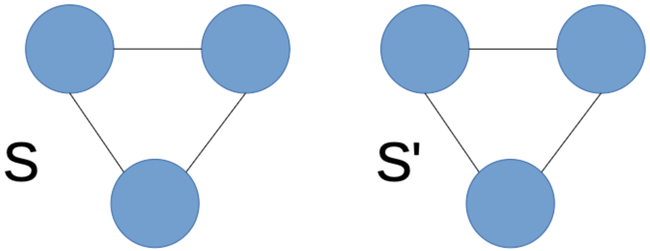


Let $x_i^S = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \in S' \end{cases}$

$$\begin{bmatrix} S & 0 \\ 0 & S' \end{bmatrix} \overset{x^S}{\begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} = \begin{bmatrix} d \\ \vdots \\ d \\ 0 \\ \vdots \\ 0 \end{bmatrix} = d \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Disconnected graphs

- Suppose the graph is regular **and disconnected**



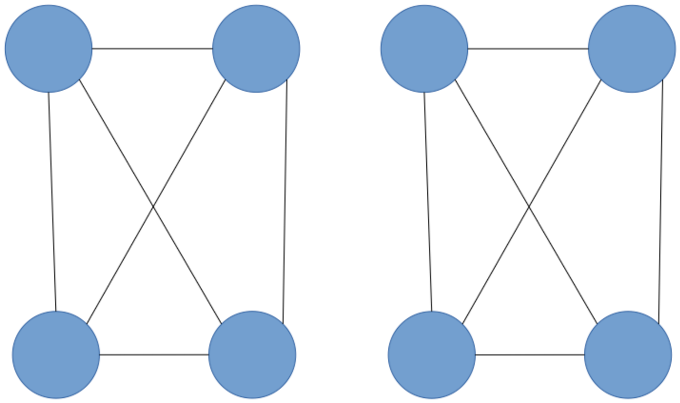
$$Ax^S = dx^S$$

$$Ax^{S'} = dx^{S'}$$

- What is the multiplicity of eigenvalue d?
- What happens if there are more than 2 connected components?

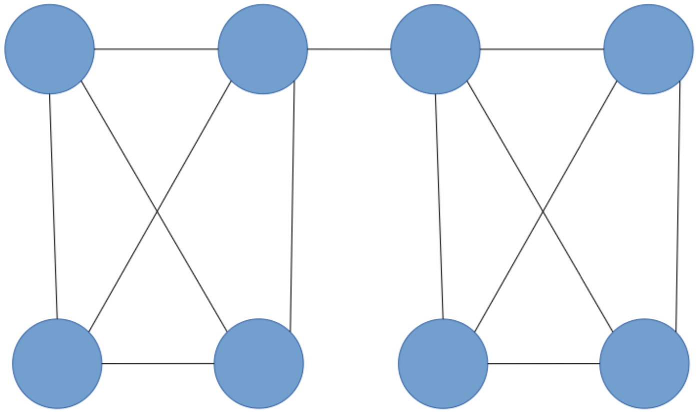
In general

Disconnected graph



$$\lambda_1 = \lambda_2$$

Almost disconnected graph



Communities!

$$\lambda_1 \approx \lambda_2$$

Small "**eigengap**"

Summary

Things to remember

- Graph Embedding: why is it useful
- Eigenvalues, Eigenvectors of Graphs
- Eigenvalues multiplicity and graph structure

Sources

- E. Terzi (2013). Graph cuts — The part on spectral graph partitioning
- URLs cited in the footer of slides