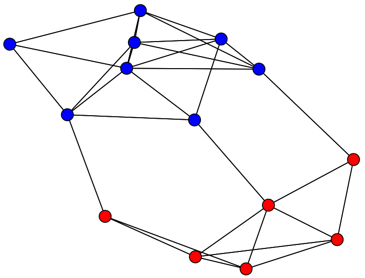
Graph Learning
SD212
6. Spectral Embedding

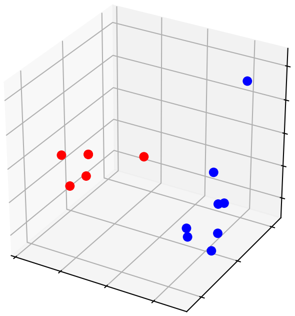
Thomas Bonald

2023 – 2024

Motivation

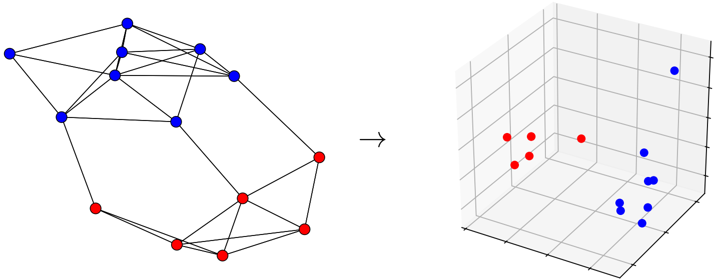
Representation of a graph in a **vector space** of low dimension





Outline

- Laplacian matrix
- Transition matrix
- Spectral embedding
- Algorithms
- Extensions



Laplacian matrix

Definition

$$L = D - A$$

Properties

- Symmetric
- Positive semi-definite
- Discrete differential operator

$$\boxed{L = \nabla^T \nabla}$$

with ∇ the $m \times n$ **incidence matrix** of the graph

Spectral analysis

Theorem

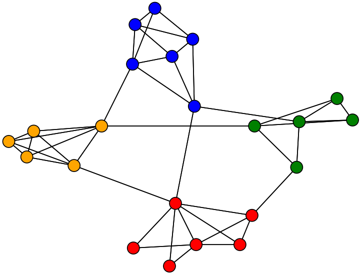
$$L = U \Lambda U^T$$

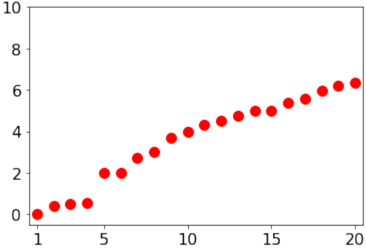
where

- $U = (u_1, \ldots, u_n)$ with $U^T U = I$
- $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 = 0 \le \lambda_2 \le \ldots \le \lambda_n$

Note: $u_1 \propto 1$

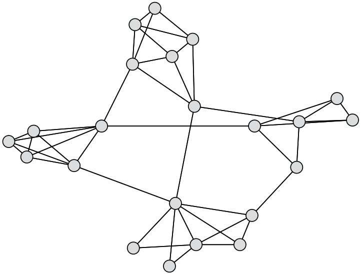
Example



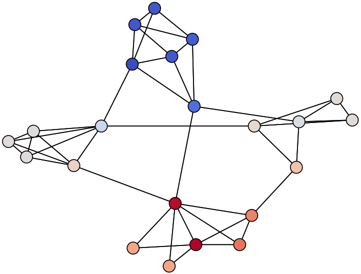


$$\lambda_1 = 0 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

Eigenvectors

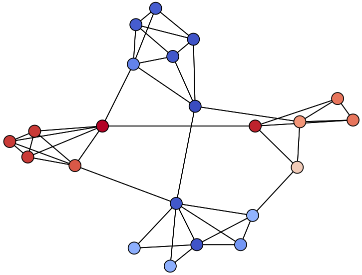


u_1



u_2

Eigenvectors



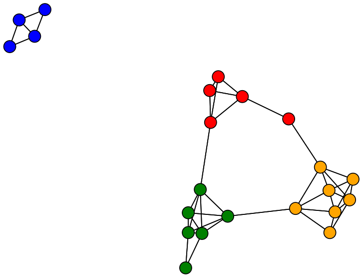
u_3 u_4

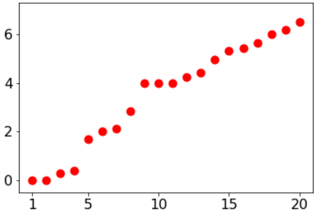
Connected components

Proposition

The **multiplicity** of the eigenvalue $\lambda = 0$ of the Laplacian matrix L is equal to the number of connected components of the graph.

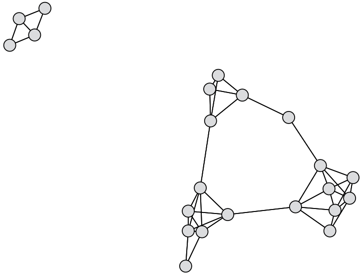
Example



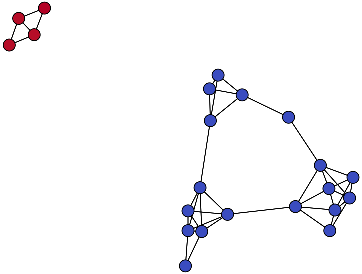


$$\lambda_1 = \lambda_2 = 0 < \lambda_3 \ldots \leq \lambda_n$$

Eigenvectors



u_1

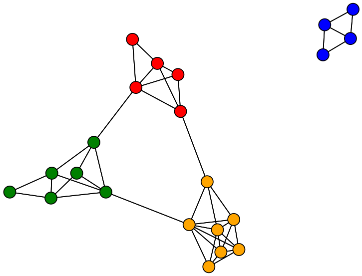


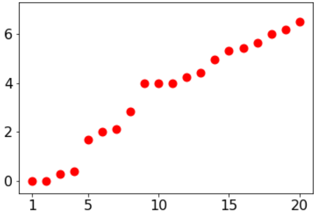
u_2

Heat diffusion

Dynamics

$$\forall t \geq 0, \quad T(t) = e^{-Lt} T(0) \quad \text{with} \quad e^{-Lt} = U e^{-\Lambda t} U^T$$





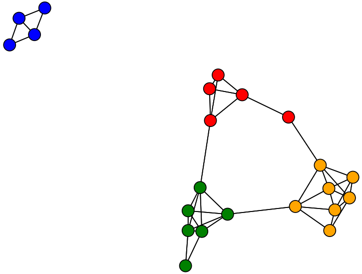
$$\lambda_1 = \lambda_2 = 0 < \lambda_3 \ldots \leq \lambda_n$$

Regularization

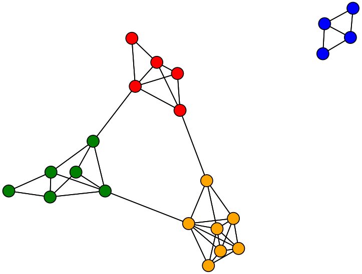
Principle

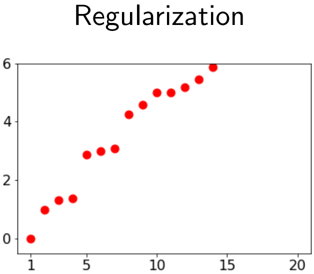
When the graph is disconnected, use

$$A' = A + \frac{11^T}{n} \quad D' = D + I$$



Example





$$\lambda_1' = 0 < \lambda_2' \ldots \leq \lambda_n'$$

Outline

1. Laplacian matrix
2. **Transition matrix**
3. Spectral embedding
4. Algorithms
5. Extensions

Transition matrix

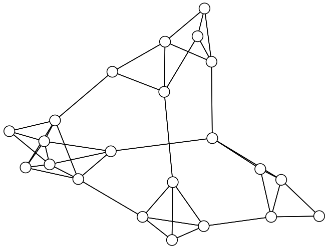
Definition

$$P = D^{-1}A$$

Property

P is a stochastic matrix:

$$P \geq 0 \quad \text{and} \quad P1 = 1$$



Spectral analysis

Theorem

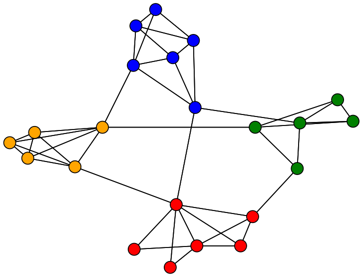
$$P = V\Gamma V^T D$$

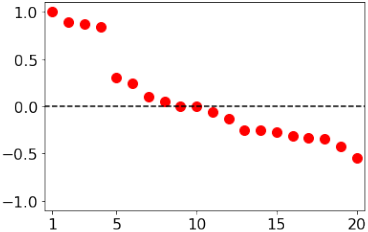
where

- $V = (v_1, \ldots, v_n)$ with $V^T D V = I$
- $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$ with $\gamma_1 = 1 \geq \gamma_2 \geq \ldots \geq \gamma_n \geq -1$

Note: $v_1 \propto 1$

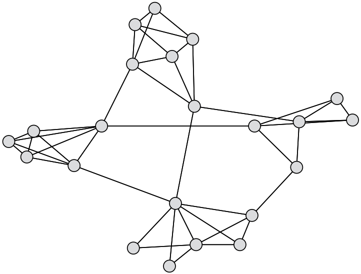
Example



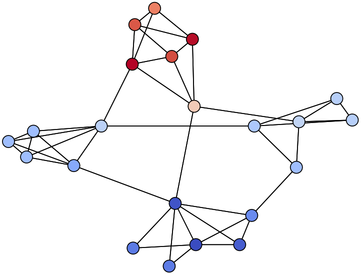


$$\gamma_1 = 1 \geq \gamma_2 \geq \ldots \geq \gamma_n \geq -1$$

Eigenvectors

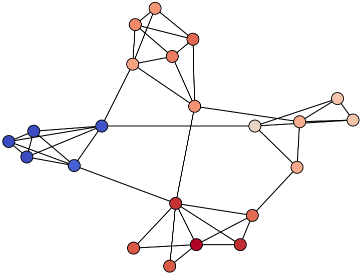


v_1

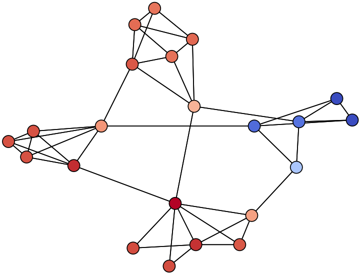


v_2

Eigenvectors



v_3



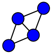
v_4

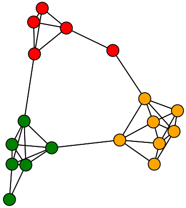
Connected components

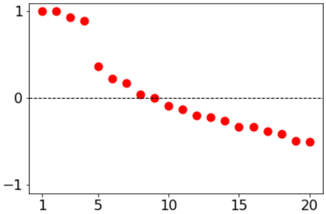
Proposition

The **multiplicity** of the eigenvalue $\gamma = 1$ of the transition matrix P is equal to the number of connected components of the graph.

Example







$$\gamma_1 = \gamma_2 = 1 > \ldots \geq \gamma_n \geq -1$$

Heat diffusion

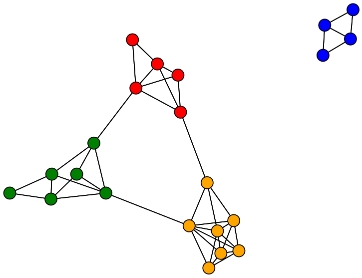
Dynamics

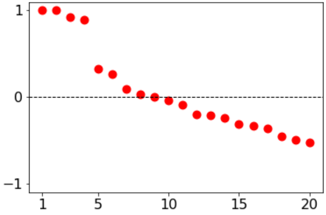
$$\forall t \geq 0, \quad T(t) = P^t T(0) \quad \text{with} \quad P^t = V\Gamma^t V^T D$$

Heat diffusion

Dynamics

$$\forall t \geq 0, \quad T(t) = P^t T(0) \quad \text{with} \quad P^t = V\Gamma^t V^T D$$





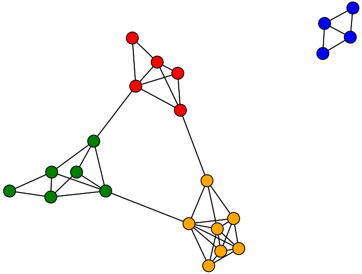
$$\gamma_1 = 1 \geq \gamma_2 \geq \ldots \geq \gamma_n \geq -1$$

Regularization

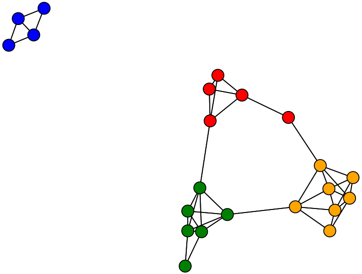
Principle

When the graph is disconnected, use

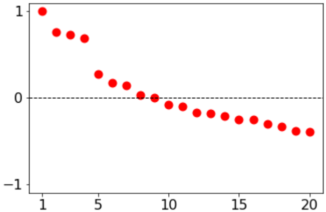
$$A' = A + \frac{11^T}{n} \quad D' = D + I$$



Example



Regularization



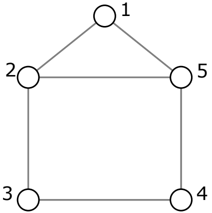
$$\gamma'_1 = 1 > \gamma'_2 \geq \ldots \geq \gamma'_n \geq -1$$

Exercise

The top eigenvalues of the transition matrix are:

$$\gamma_1 = 1, \gamma_2 = \frac{1}{3}, \gamma_3 = 0$$

Give the corresponding **right eigenvectors**.



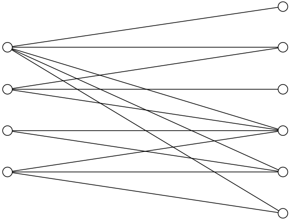
Case of bipartite graphs

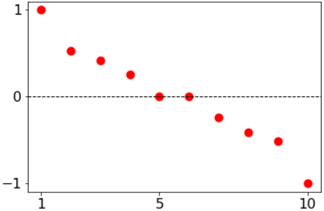
Proposition

The transition matrix of a bipartite graph has a **symmetric** spectrum:

$$\gamma \text{ eigenvalue} \quad \Longleftrightarrow \quad -\gamma \text{ eigenvalue}$$

Example

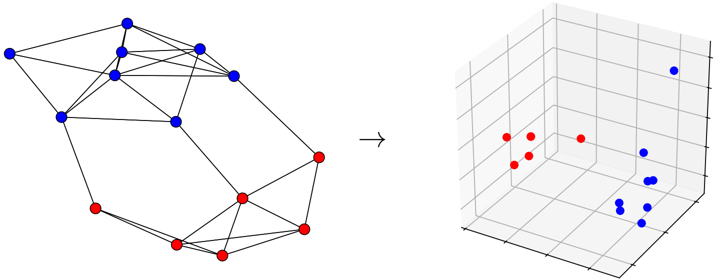




$$\gamma_1 = 1 > \gamma_2 \geq \ldots \geq \gamma_{n-1} > \gamma_n = -1$$

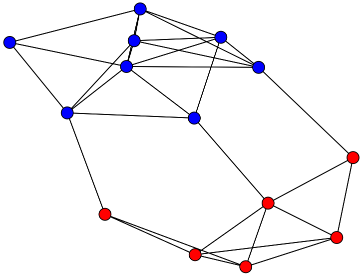
Outline

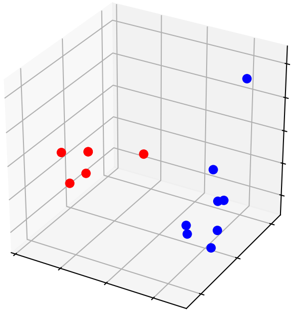
1. Laplacian matrix
2. Transition matrix
3. **Spectral embedding**
4. Algorithms
5. Extensions



An optimization problem

$$\min_X \sum_{i<j} A_{ij} ||X_i - X_j||^2$$





Laplacian matrix

Lemma

$$\mathrm{tr}(X^T L X) = \sum_{i<j} A_{ij} ||X_i - X_j||^2$$

Spectral embedding

Definition

Embedding $X = (u_2, \dots, u_{K+1})$ given by the first K eigenvectors (except the first) of the **Laplacian matrix** L

Theorem

The spectral embedding is optimal:

$$X = \arg \min_{X : X^T 1 = 0, X^T X = I_K} \mathrm{tr}(X^T L X)$$

Spectral embedding

Definition

Embedding $X = (u_2, \dots, u_{K+1})$ given by the first K eigenvectors (except the first) of the **Laplacian matrix** L

Theorem

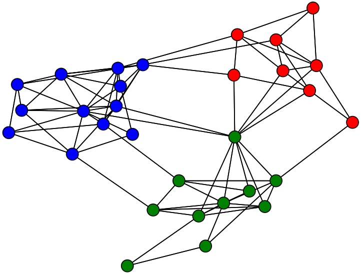
The spectral embedding is optimal:

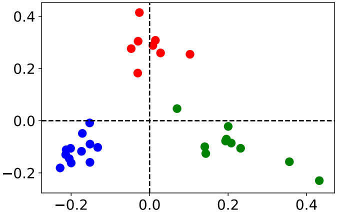
$$X = \arg \min_{X : X^T 1 = 0, X^T X = I_K} \mathrm{tr}(X^T L X)$$

Note: The embedding is centered:

$$\sum_{i=1}^{n} X_i = 0$$

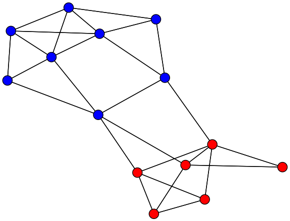
Example



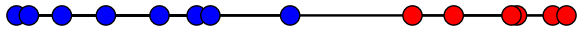


A mechanical system

Nodes = **particles**, edges = (attractive) **springs**
Put nodes on a **line** at positions $x_1, \ldots, x_n \in \mathbb{R}$

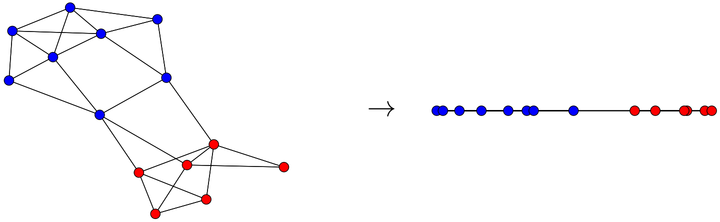


$\rightarrow$



A mechanical system

Nodes = **particles**, edges = (attractive) **springs**
Put nodes on a **line** at positions $x_1, \ldots, x_n \in \mathbb{R}$

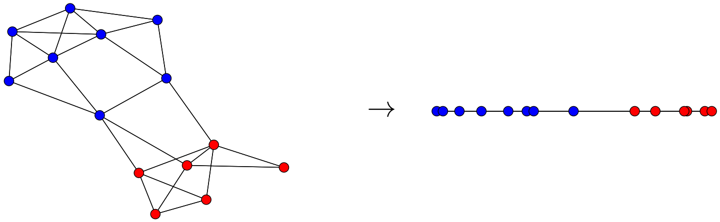


Potential energy

$$E = \frac{1}{2} \sum_{i<j} A_{ij}(x_i - x_j)^2 = \frac{1}{2} x^T L x$$

A harmonic oscillator

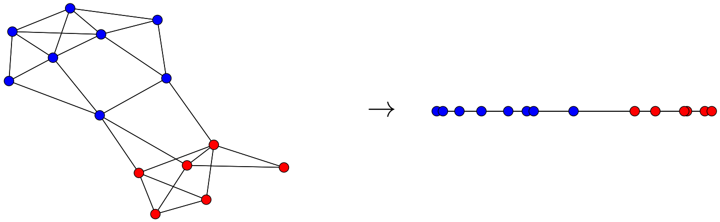
Let the system **evolve**, assuming **unit** masses, starting from positions $x_1, \ldots, x_n \in \mathbb{R}$



$$\forall i, \quad \ddot{x}_i = \sum_j A_{ij}(x_j - x_i) \quad \Longleftrightarrow \quad \ddot{x} = -Lx$$

A harmonic oscillator

Let the system **evolve**, assuming **unit** masses, starting from positions $x_1, \ldots, x_n \in \mathbb{R}$

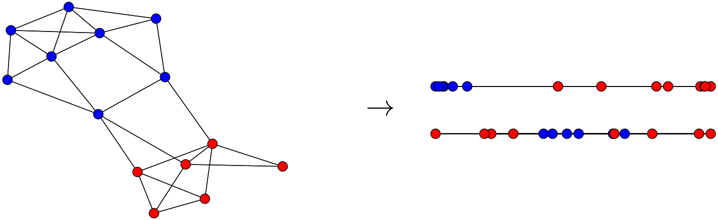


$$\forall i, \quad \ddot{x}_i = \sum_j A_{ij}(x_j - x_i) \quad \iff \quad \ddot{x} = -Lx$$

Eigenvectors of $L \to$ **eigenmodes**
Eigenvalues of $L \to$ **levels of energy**

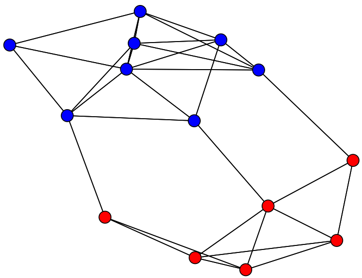
Eigenmodes

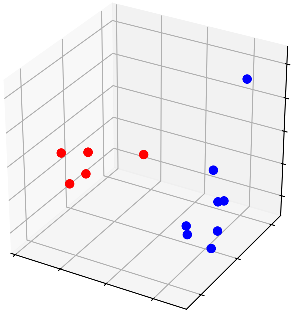
The most interesting **eigenmodes** are those of **lowest energy** (equivalently, of lowest eigenfrequency)



Back to the optimization problem

$$\min_{X:X^T 1=0, X^T X=I} \sum_{i<j} A_{ij} ||X_i - X_j||^2$$





Spectral embedding

Definition

Embedding $X = (v_2, \dots, v_{K+1})$ given by the K leading eigenvectors (except the first) of the **transition matrix** P

Theorem

The spectral embedding is optimal:

$$X = \arg \min_{X : X^T d = 0, X^T D X = I_K} \mathrm{tr}(X^T L X)$$

Spectral embedding

Definition

Embedding $X = (v_2, \dots, v_{K+1})$ given by the K leading eigenvectors (except the first) of the **transition matrix** P

Theorem

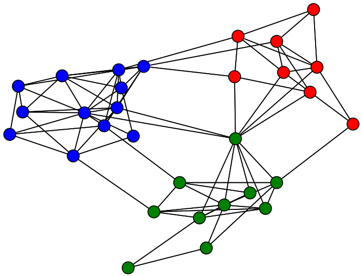
The spectral embedding is optimal:

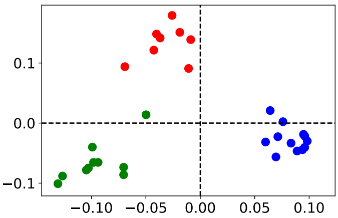
$$X = \arg \min_{X : X^T d = 0, X^T D X = I_K} \mathrm{tr}(X^T L X)$$

Note: The **weighted** embedding is centered:

$$\sum_{i=1}^{n} d_i X_i = 0$$

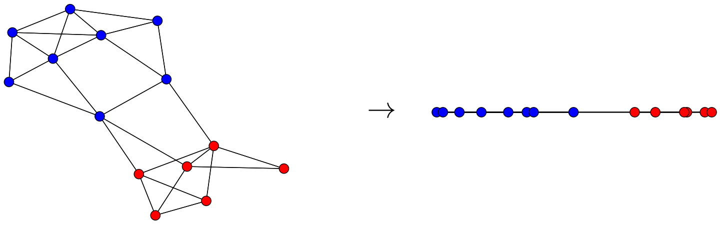
Example





Back to the mechanical system

Nodes = **particles**, edges = (attractive) **springs**
Put nodes on a **line** at positions $x_1, \ldots, x_n \in \mathbb{R}$

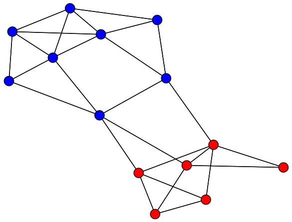


Potential energy

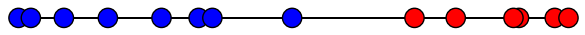
$$E = \frac{1}{2} \sum_{i<j} A_{ij} (x_i - x_j)^2 = \frac{1}{2} x^T L x$$

Harmonic oscillator

Let the system **evolve**, with **masses** equal to the **degrees**, starting from positions $x_1, \dots, x_n \in \mathbb{R}$



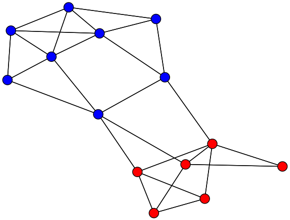
$\rightarrow$



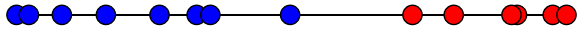
$$\forall i, \quad D\ddot{x}_i = \sum_j A_{ij}(x_j - x_i) \quad \Longleftrightarrow \quad \ddot{x} = -(I - P)x$$

Harmonic oscillator

Let the system **evolve**, with **masses** equal to the **degrees**, starting from positions $x_1, \ldots, x_n \in \mathbb{R}$



$\rightarrow$

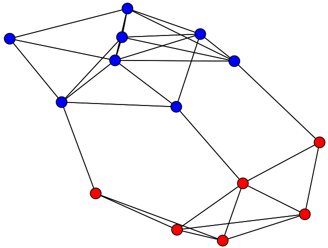


$$\forall i, \quad D\ddot{x}_i = \sum_j A_{ij}(x_j - x_i) \quad \Longleftrightarrow \quad \ddot{x} = -(I - P)x$$

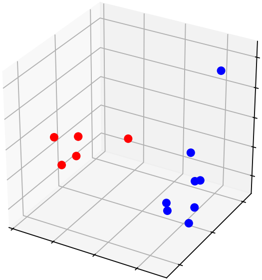
Eigenvectors of P $\rightarrow$ **eigenmodes**
1 - eigenvalues of P $\rightarrow$ **levels of energy**

Outline

1. Laplacian matrix
2. Transition matrix
3. Spectral embedding
4. **Algorithms**
5. Extensions



$\rightarrow$



Algorithms

Need to compute the first **eigenvectors** of some matrix M
(either the Laplacian L or the normalized Laplacian $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$)

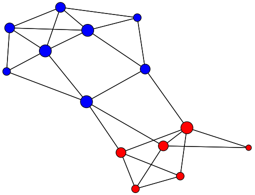
Lanczos' algorithm
Power iteration
Lanczos 1950

$$x \leftarrow \frac{Mx}{||Mx||}$$

Halko's algorithm
Random projection
Power iteration
QR decomposition
Halko 2009

Back to regularization

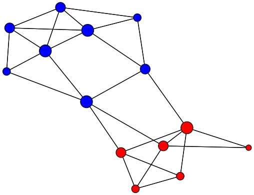
$$A' = A + \frac{11^T}{n}$$



The adjacency matrix becomes **dense**...

Back to regularization

$$A' = A + \frac{11^T}{n}$$

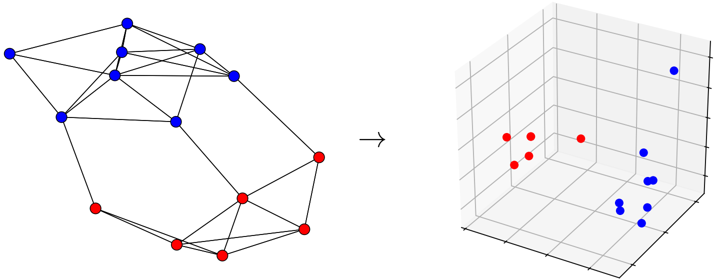


The adjacency matrix becomes **dense**...
but with a nice **sparse + low rank** structure:

$$A'x = Ax + \frac{1^T x}{n} 1$$

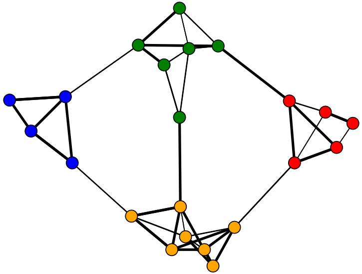
Outline

1. Laplacian matrix
2. Transition matrix
3. Spectral embedding
4. Algorithms
5. **Extensions**



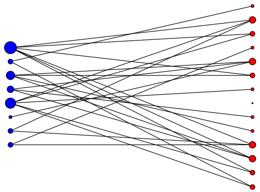
Weighted graphs

$$\min_X \sum_{i<j} A_{ij} ||X_i - X_j||^2$$

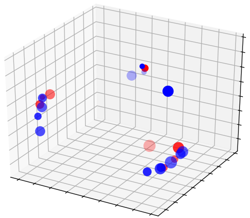


Case of bipartite graphs

Co-embedding of nodes in the **same** space

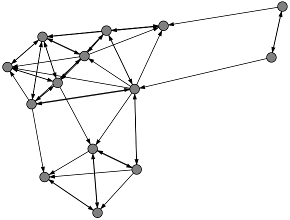


$\rightarrow$

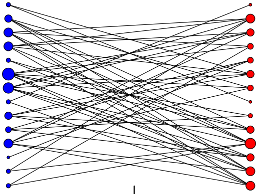


Case of directed graphs

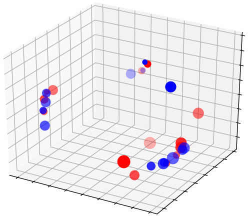
Idea: See the directed graph as a bipartite graph,
with biadjacency matrix A



$\rightarrow$



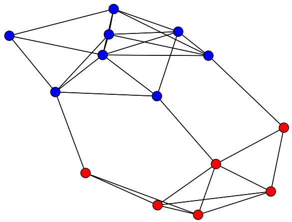
$\downarrow$



Summary

Spectral embedding

- Based on the spectral decomposition of the **Laplacian matrix** $L = D - A$ or the **transition matrix** $P = D^{-1}A$
- Eigenmodes of a **mechanical system**
- **Fast** and **scalable** algorithms
- Applicable to **weighted**, **bipartite** and **directed** graphs



$\rightarrow$