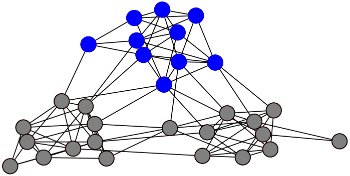
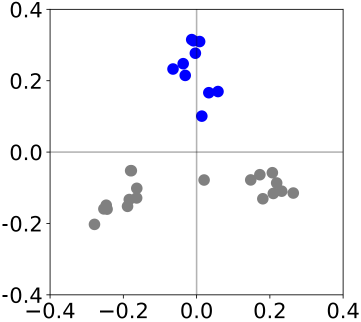
Graph Mining
SD212
7. Graph embedding

Thomas Bonald

2018 – 2019

Graph embedding





Framework

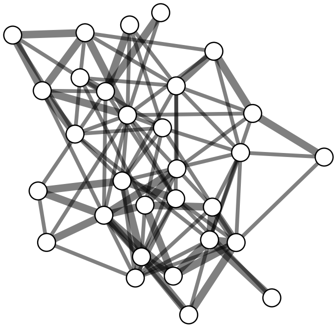
A weighted, **undirected** graph of n nodes and m edges

Connected and **not** bipartite, without self-loops

Weighted adjacency matrix A

Vector of node weights $w = A1$

Diagonal matrix of node weights $D = \mathrm{diag}(w)$

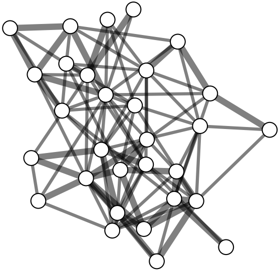


Outline

1. Random walks
2. Laplacian matrix
3. Mechanics
4. Spectral embedding
5. Extensions

Random walk in discrete time

- $P_{ij} = A_{ij}/w_i$, probability of moving from i to j
- A Markov chain with transition matrix $P = D^{-1}A$



Random walk in discrete time

- $P_{ij} = A_{ij}/w_i$, probability of moving from i to j
- A Markov chain with transition matrix $P = D^{-1}A$

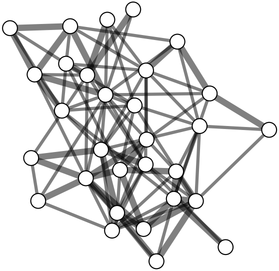
Time evolution

$$\forall t \geq 0, \quad \pi_{t+1} = \pi_t P$$

Equilibrium

$$\pi P = \pi \quad \Longrightarrow \quad \pi \propto w^T$$

Convergence at **exponential rate**



Spectral analysis

Spectral decomposition

$$PV = V\Lambda, \quad V^T DV = I$$

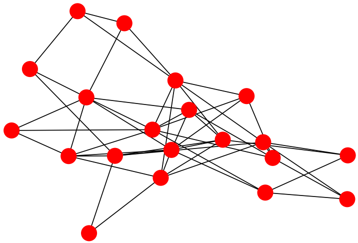
where

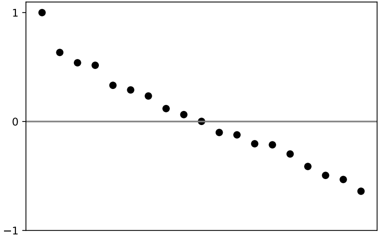
- $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 = 1 > \lambda_2 \geq \ldots \geq \lambda_n > -1$
- $V = (v_1, \ldots, v_n)$ with $v_1 \propto 1$

Note: If the graph has k connected components, then

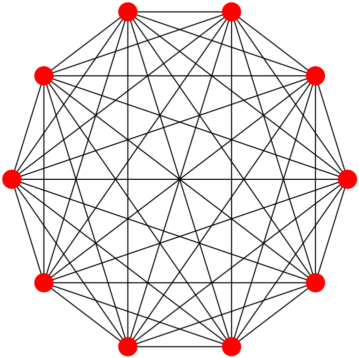
$$\lambda_1 = \ldots = \lambda_k = 1 > \lambda_{k+1}$$

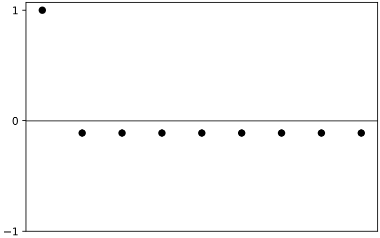
Example



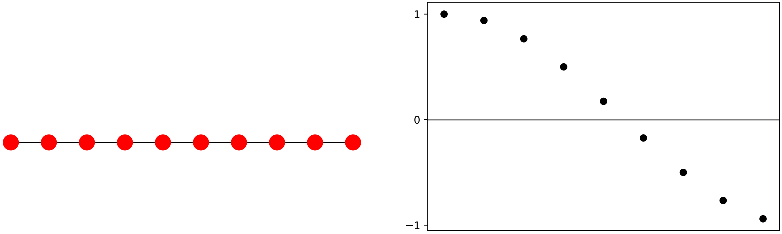


Complete graph

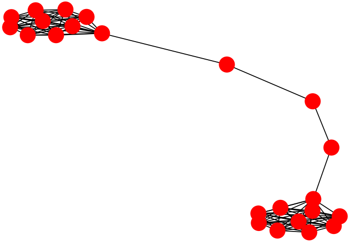


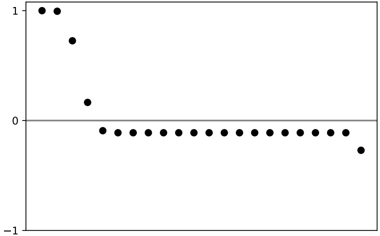


Line

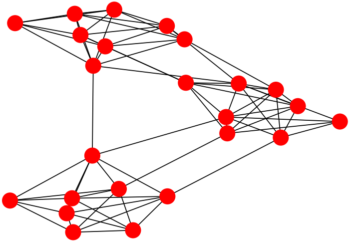


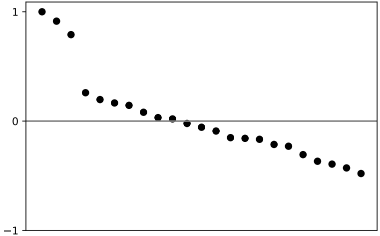
Barbell





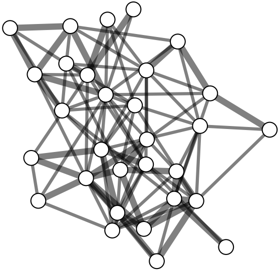
Stochastic block model





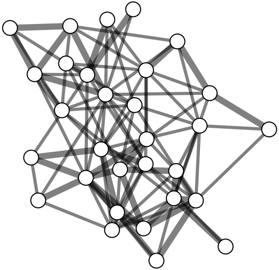
Random walk in continuous time

- Transition rate A_{ij} from node i to node j
- A Markov chain with generator matrix $Q = A - D$



Random walk in continuous time

- Transition rate A_{ij} from node i to node j
- A Markov chain with generator matrix $Q = A - D$



Time evolution

$$\forall t \geq 0, \quad \frac{d\pi_t}{dt} = \pi_t Q$$

Equilibrium

$$\pi Q = 0 \quad \Longrightarrow \quad \pi \propto 1^T$$

Convergence at **exponential rate**

Spectral analysis

Spectral decomposition

Let $L = D - A$:

$$LV = V\Lambda, \quad V^T V = I$$

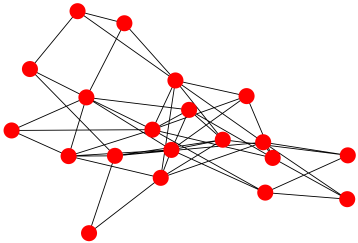
where

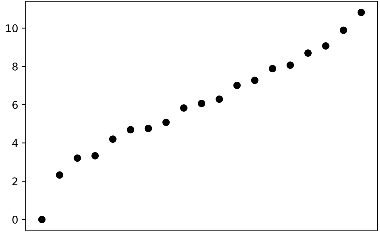
- $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 = 0 < \lambda_2 \leq \ldots \leq \lambda_n$
- $V = (v_1, \ldots, v_n)$ with $v_1 \propto 1$

Note: If the graph has k connected components, then

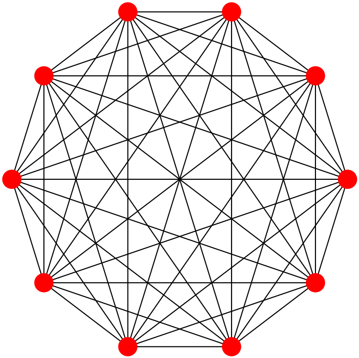
$$\lambda_1 = \ldots = \lambda_k = 0 < \lambda_{k+1}$$

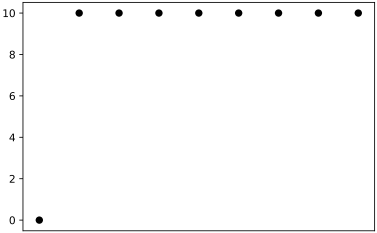
Example



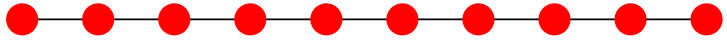


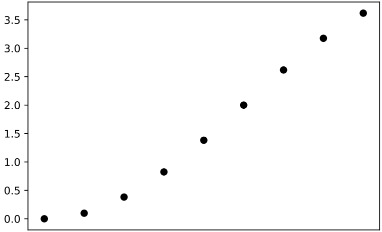
Complete graph



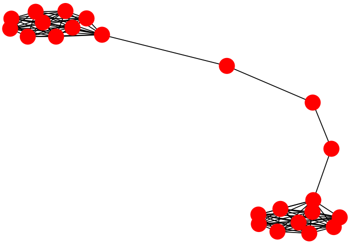


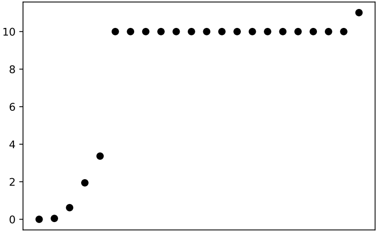
Line



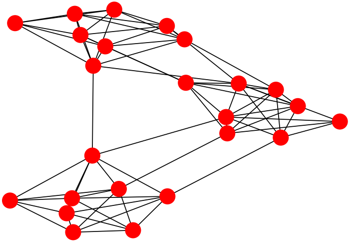


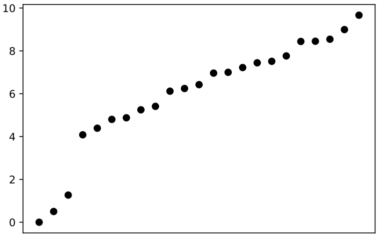
Barbell





Stochastic block model





Outline

1. Random walks
2. **Laplacian matrix**
3. Mechanics
4. Spectral embedding
5. Extensions

Laplacian matrix

Definition

The matrix $L = D - A$ is called the **Laplacian matrix**.

Proposition

This is a **positive semi-definite** matrix:

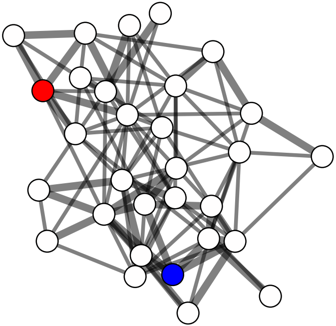
$$\forall x \in \mathbb{R}^n, \quad x^T L x = \sum_{i<j} A_{ij}(x_i - x_j)^2$$

Heat equation

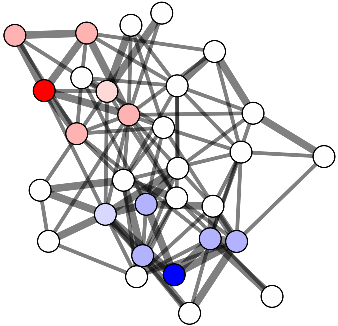
- Fix the temperature of some nodes $S \subset \{1, \ldots, n\}$
- Interpret the weight A_{ij} as the **thermal conductivity**
- Then for each node $i \notin S$,

$$\frac{dT_i}{dt} = \sum_j A_{ij}(T_j - T_i) = -(LT)_i$$

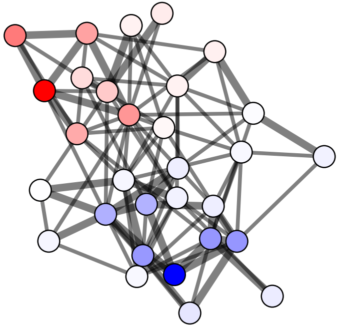
Example



Example



Example



Equilibrium: The Dirichlet problem

For each node $i \notin S$,

$$(LT)_i = 0$$

with boundary condition T_j for all $j \in S$.

Proposition

The **unique solution** to the Dirichlet problem is given by:

$$\forall i \notin S, \quad T_i = \sum_{j \in S} P^S_{ij} T_j$$

where P^S_{ij} is the probability that a random walk starting from i first hits the boundary S in j.

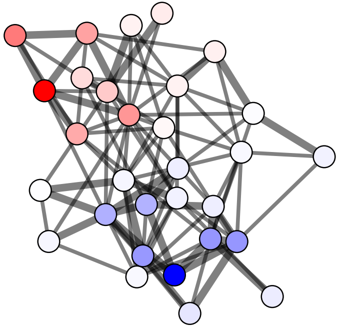
Illustration

Diffusion in discrete time

For each node $i \notin S$,

$$(LT)_i = 0 \quad \Leftrightarrow \quad T_i = (PT)_i$$

The solution of the Dirichlet problem follows from the iteration $T_i \leftarrow (PT)_i$ (outside the boundary S)



A particular case

Let $S = \{s, t\}$ with $T_s = 0$ (source) and $T_t = 1$ (target)

Proposition

The **unique solution** to the Dirichlet problem is given by:

$$T_i = \frac{(e_i - e_s)^T L^+ (e_t - e_s)}{(e_t - e_s)^T L^+ (e_t - e_s)}.$$

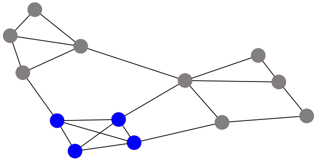
where $L^+ = V\Lambda^+ V^T$ is the **pseudo-inverse** of the Laplacian

Outline

1. Random walks
2. Laplacian matrix
3. **Mechanics**
4. Spectral embedding
5. Extensions

1D embedding

How to position the nodes of the graph on a line?

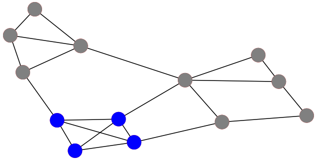


Mechanics

Consider n particles with a **spring** of strength A_{ij} between i and j

- Let $u_1, \ldots, u_n$ be the **location** of these particles
- By **Hooke's law**, the force between i and j is $A_{ij}|u_i - u_j|$
- We deduce the **potential energy**:

$$E = \frac{1}{2}\sum_{i<j} A_{ij}(u_i - u_j)^2 = \frac{1}{2}u^T L u$$



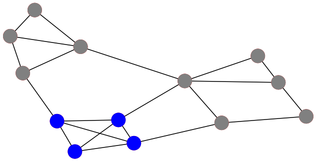
Fiedler vector

Proposition

The second eigenvector v_2 of the Laplacian matrix **minimizes** the energy $\frac{1}{2}u^T L u$ under the constraints:

$$u \perp 1, \quad ||u|| = 1.$$

The energy is then $\frac{1}{2}\lambda_2$.



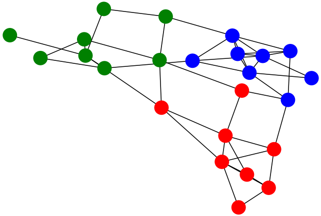
Energy minima

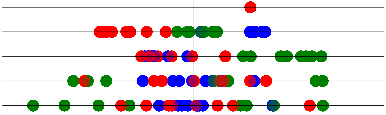
Theorem

For all $k = 1, \ldots, n$,

$$\lambda_k = \min_{\substack{u:||u||=1 \\ u \perp v_1, \ldots, u \perp v_{k-1}}} u^T L u$$

and the minimum is attained for $u = v_k$.





Harmonic oscillator

Assume each particle has **unit** mass

- Force exerted on particle i in state s:

$$\sum_j A_{ij}(s_j - s_i) = -(Ls)_i$$

- By Newton's law,

$$-Ls = \ddot{s}$$

- Eigenmodes:

$$s = ue^{i\omega t} \implies Lu = \omega^2 u$$

Harmonic oscillator (2)

Assume the particles have masses $w_1, \ldots, w_n$

- Force exerted on particle i:

$$\sum_j A_{ij}(s_j - s_i) = -(Ls)_i$$

- By Newton's law,

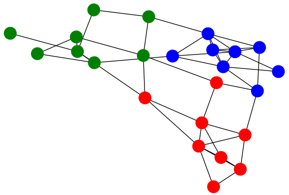
$$-Ls = D\ddot{s}$$

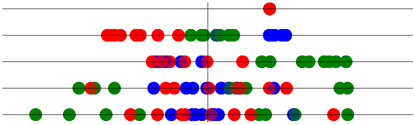
- Eigenmodes:

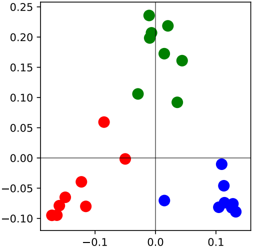
$$s = ue^{i\omega t} \quad \Longrightarrow \quad Lu = \omega^2 Du \quad \Leftrightarrow \quad Pu = (1 - \omega^2)u$$

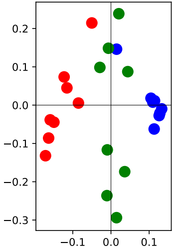
Outline

Spectral embedding

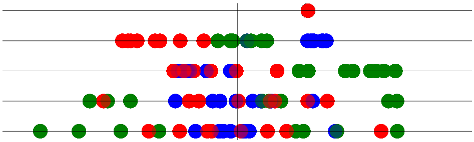




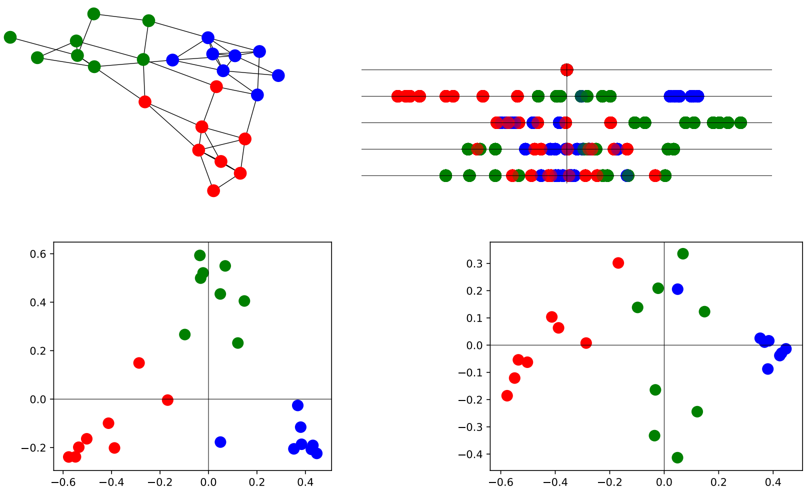




Scaling



Scaled spectral embedding



Gram matrix

- Spectral decomposition of the Laplacian:

$$L = V \Lambda V^T \quad V^T V = I$$

- Gram matrix of the spectral embedding:

$$V V^T = I$$

- After scaling

$$X = V \sqrt{\Lambda^+} \quad \Longrightarrow \quad X X^T = V \Lambda^+ V^T = L^+$$

The Gram matrix is the **pseudo-inverse** of the Laplacian

Back to the Dirichlet problem

- Fix the position of some nodes $S \subset \{1, \ldots, n\}$
- At equilibrium,

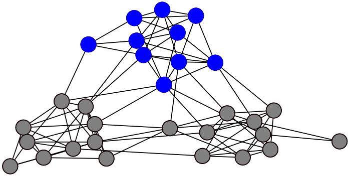
$$\forall i \notin S, \quad (LT)_i = 0$$

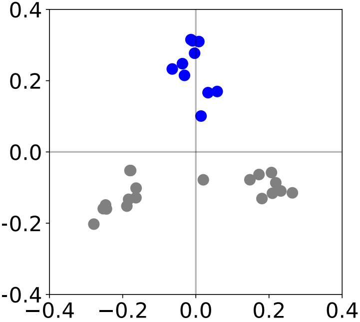
The potential energy $\frac{1}{2} u^T L u$ is minimized for $u = T$

Solution to the Dirichlet problem

For $S = \{s, t\}$ with $T_s = 0$ (source) and $T_t = 1$ (target),

$$\forall i, \quad T_i = \frac{(x_i - x_s)^T (x_t - x_s)}{||x_t - x_s||^2}$$



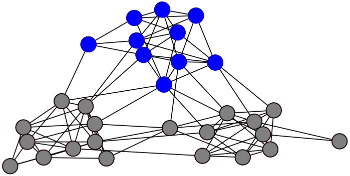


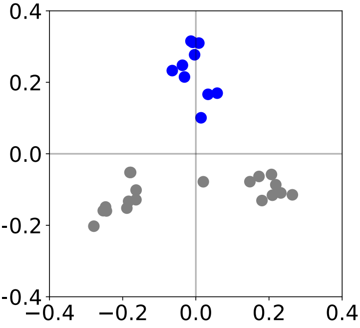
Proximity in the embedding space

For $S = \{s, t\}$ with $T_s = 0$ (source) and $T_t = 1$ (target),
for s a **virtual** particle located at the origin and $t = j$

$$\forall i, \quad T_i = \frac{{x_i}^T x_j}{||x_j||^2}$$

This is a measure of the **proximity** of node i to node j.

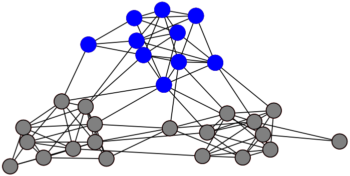


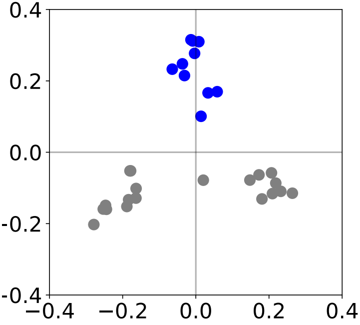


Cosine similarity

The **cosine similarity** is a symmetric measure of proximity:

$$\forall i, j, \quad \cos(x_i, x_j) = \frac{{x_i}^T x_j}{||x_i||||x_j||}$$





Spectral embedding

Embedding

Parameter: k, dimension of the embedding

1. Form the Laplacian $L = D - A$
2. Compute $v_1, \ldots, v_{k+1}$, the eigenvectors of L associated with the $k+1$ lowest eigenvalues, $\lambda_1 = 0 < \lambda_2 \leq \ldots \leq \lambda_{k+1}$
3. Return $X = \left(\frac{v_2}{\sqrt{\lambda_2}}, \ldots, \frac{v_{k+1}}{\sqrt{\lambda_{k+1}}}\right)^T$

Note: Related to the random walk in **continuous time**, $\pi \propto 1^T$

Impact of weighting

- Spectral decomposition of the transition matrix:

$$P = V\Lambda V^T \quad V^T D V = I$$

- Gram matrix of the spectral embedding:

$$VV^T = D^{-1}$$

- After scaling

$$Y = V\sqrt{(I-\Lambda)^+} \implies {y_i}^T y_j = (x_i - \bar{x})^T (x_j - \bar{x})$$

with $\bar{x}$ the center of mass of the embedding X:

$$\bar{x} = \sum_{i=1}^{n} w_i x_i$$

Normalized spectral embedding

Embedding

Parameter: k, dimension of the embedding

1. Form the transition matrix $P = D^{-1}A$
2. Compute $v_1, \ldots, v_{k+1}$, the eigenvectors of P associated with the $k+1$ highest eigenvalues, $\lambda_1 = 1 > \lambda_2 \geq \ldots \geq \lambda_{k+1}$
3. Return $Y = \left(\frac{v_2}{\sqrt{1-\lambda_2}}, \ldots, \frac{v_{k+1}}{\sqrt{1-\lambda_{k+1}}}\right)^T$

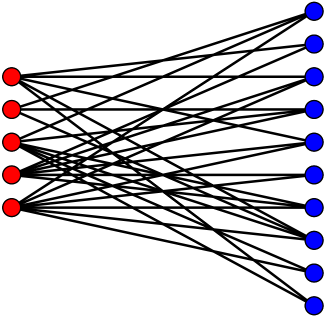
Note: Related to the random walk in **discrete time**, $\pi \propto w^T$

Outline

1. Random walks
2. Laplacian matrix
3. Mechanics
4. Spectral embedding
5. **Extensions**

Bipartite graphs

Let B be the biadjacency matrix



Spectral analysis

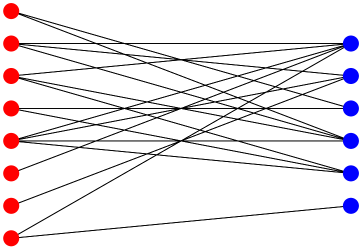
Spectral decomposition

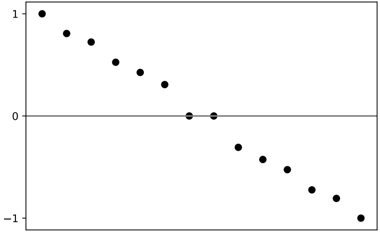
$$P\begin{pmatrix}U\\V\end{pmatrix} = \begin{pmatrix}U\\V\end{pmatrix}\Lambda, \quad \begin{pmatrix}U\\V\end{pmatrix}^T D\begin{pmatrix}U\\V\end{pmatrix} = I$$

where

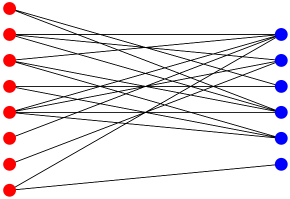
- $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 = 1 > \lambda_2 \geq \ldots > \lambda_n = -1$
- The spectrum is **symmetric**
- The eigenvectors $\begin{pmatrix}u_1\\v_1\end{pmatrix}, \ldots, \begin{pmatrix}u_n\\v_n\end{pmatrix}$ correspond to the **generalized singular vectors** of B

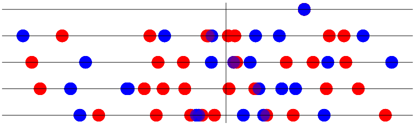
Example



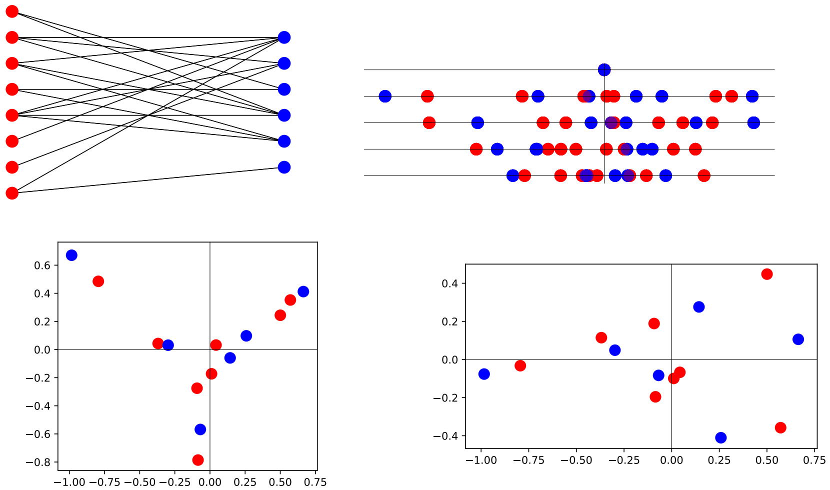


Mechanics

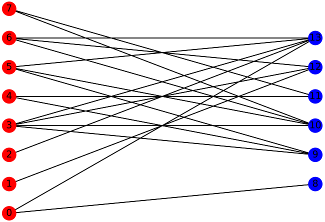


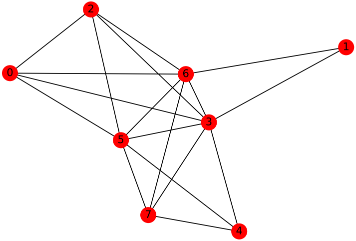


Spectral co-embedding

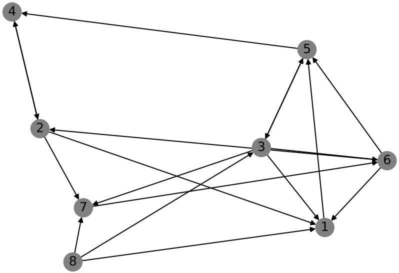


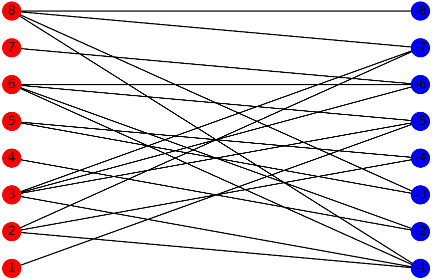
The co-citation graph



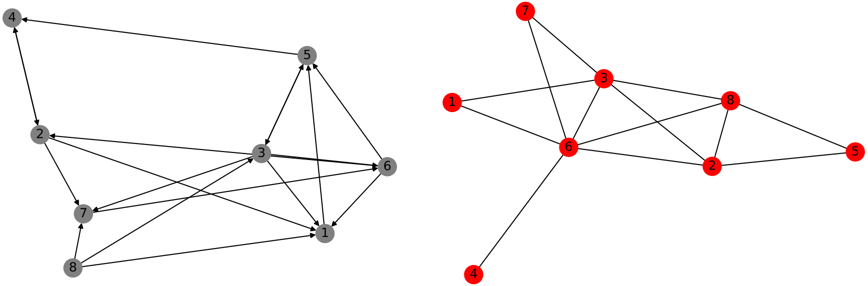


Directed graphs





The co-citation graph



Summary

- Graph embedding, a technique to transform **graph data** into **vector data**
- The Laplacian matrix, related to the **heat equation**
- Spectral decomposition of the Laplacian, interpretable as the eigenmodes of a **mechanical system**
- Cosine similarity in the embedding space, to measure the proximity between nodes, related to the **Dirichlet problem**