Graph Partitioning II: Spectral Methods

Class	Algorithmic Methods of Data Mining
Program	M. Sc. Data Science
University	Sapienza University of Rome
Semester	Fall 2015
Lecturer	Carlos Castillo http://chato.cl/

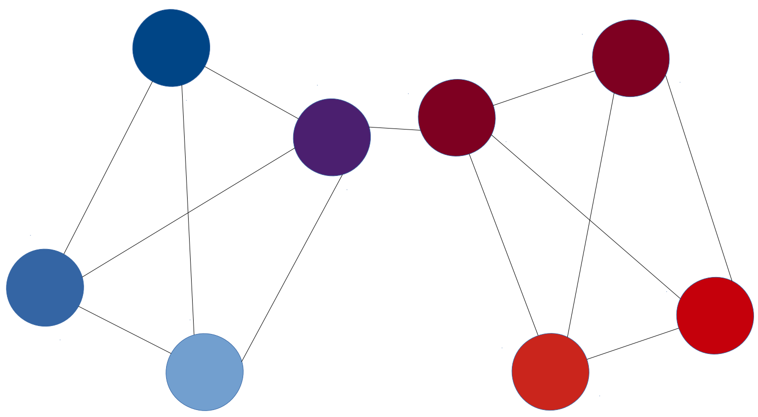
Sources:

- Evimaria Terzi: “Clustering: graph cuts and spectral graph partitioning” [link]
- Daniel A. Spielman: “The Laplacian” [link]
- Jure Leskovec: “Defining the graph laplacian” [link]

Graph projections

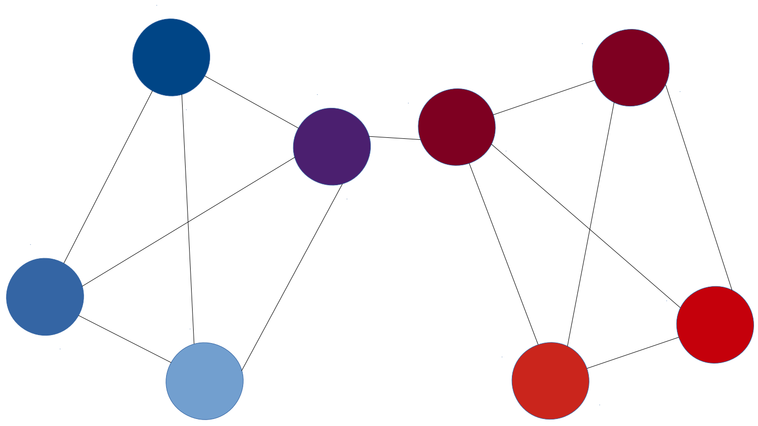
Graphs are nice, but ...

- They describe only local relationships
- We would like to understand a global structure
- Our objective is transforming a graph into a more familiar object: a cloud of points in R^k



Graphs are nice, but ...

- They describe only local relationships
- We would like to understand a global structure
- Our objective is transforming a graph into a more familiar object: a cloud of points in R^k



Distances should be somehow preserved

Simple 3D graph projection

- 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
- Start a BFS from yet another node, which has $z=1$, and nodes visited have ascending z
- Project node n to position (x_n, y_n, z_n)

How do you think this works in practice?

Eigenvectors of adjacency matrix

Adjacency matrix of G=(V,E)

- Matrix of size |V| x |V| such that:

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

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

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

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

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=YuKsZstrua4

Properties of adjacency matrix

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

- 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=YuKsZstrua4

Understanding the adjacency matrix

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

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

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

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

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=YuKsZstrua4

Spectral graph theory

- The study of the eigenvalues and eigenvectors of a graph matrix $Ax = \lambda x$
  - Adjacency matrix
  - Laplacian matrix (next)
- Suppose graph is d-regular (every node has degree d), what is the value of:

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

An easy eigenvector of A

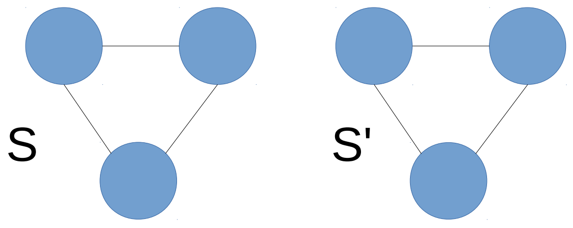
- Suppose graph is d-regular, i.e. all nodes have degree d:

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

- So $[1, 1, \dots, 1]^T$ is an eigenvector of eigenvalue d

Disconnected graphs

- Suppose graph is disconnected (still regular)

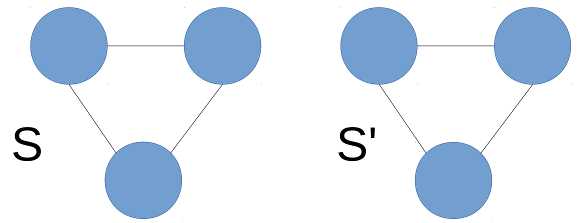


- Then matrix has block structure:

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

Disconnected graphs

- Suppose graph is disconnected (still regular)

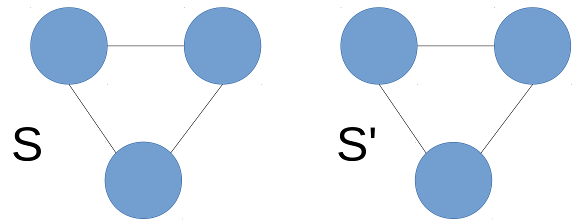


$$\text{Let } x_i^S = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \in S' \end{cases}$$

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

Disconnected graphs

- Suppose graph is disconnected (still regular)

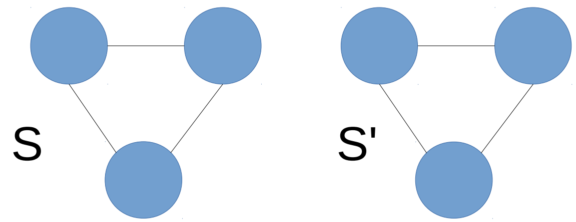


$$\text{Let } x_i^S = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \in S' \end{cases}$$

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

Disconnected graphs

- Suppose graph is disconnected (still regular)



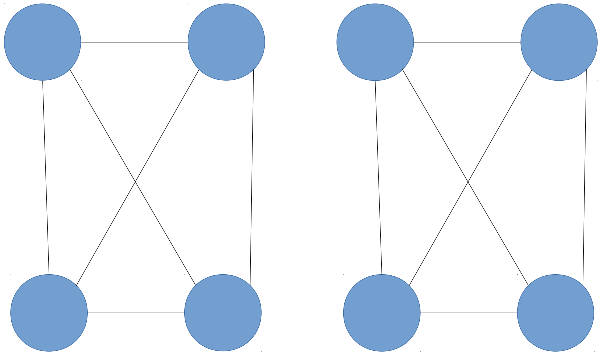
$$Ax^S = dx^S$$

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

- *What happens if there are more than 2 connected components?*

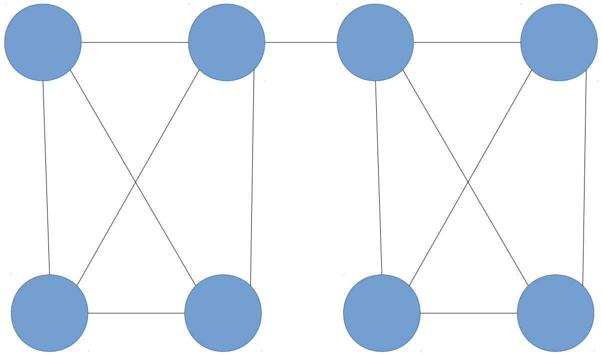
In general

Disconnected graph



$$\lambda_1 = \lambda_2$$

Almost disconnected graph



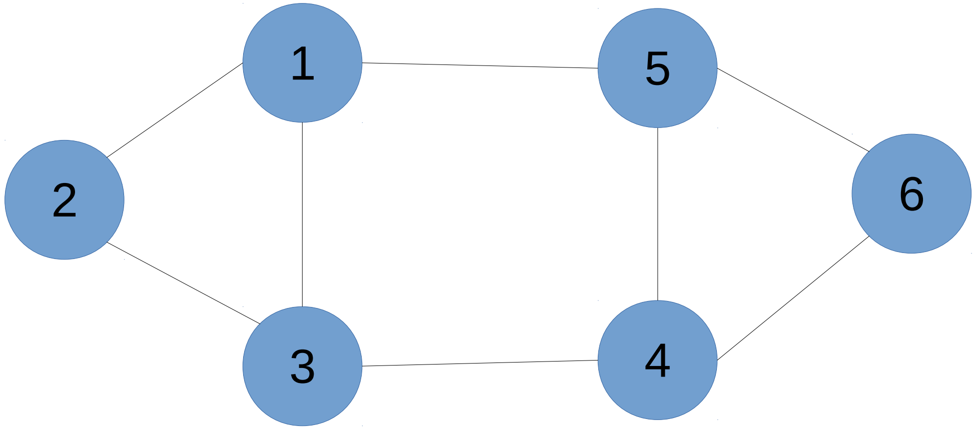
$$\lambda_1 \approx \lambda_2$$

Small "**eigengap**"

Graph Laplacian

Adjacency matrix

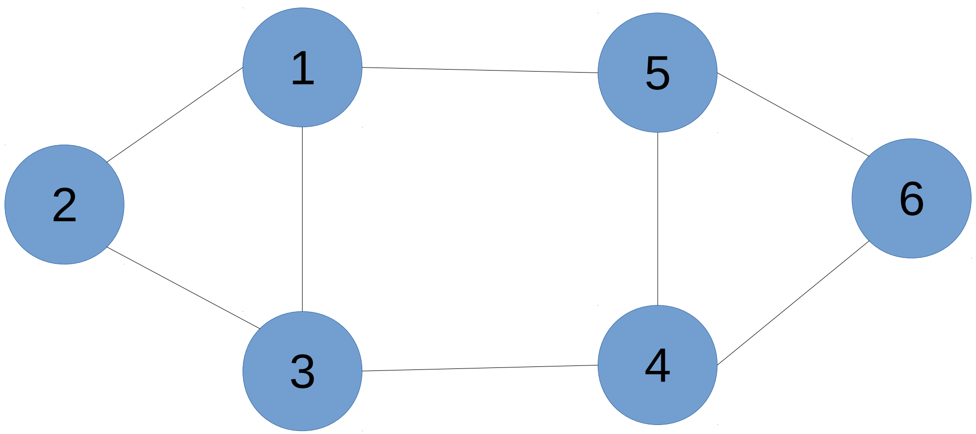
$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

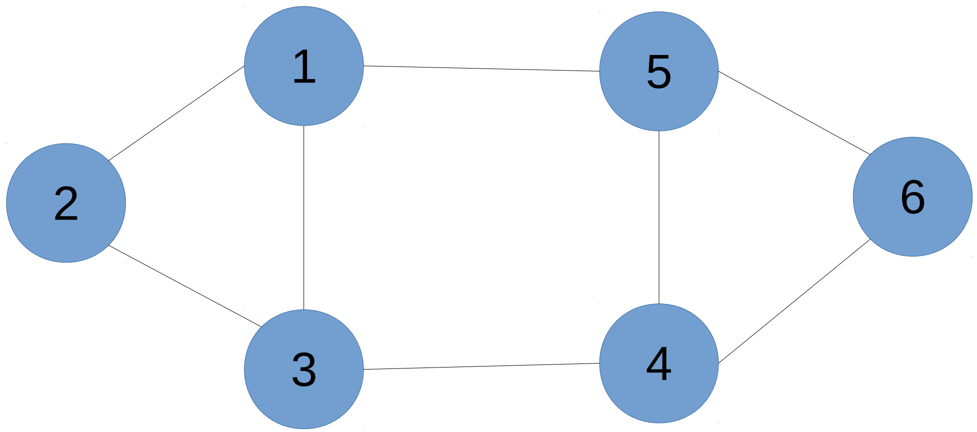
Degree matrix

$$D_{ij} = \begin{cases} d(i) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$



$$D = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

Laplacian matrix



$$L = D - A$$

$$L = \begin{bmatrix} 3 & -1 & -1 & 0 & -1 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & 0 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{bmatrix}$$

Laplacian matrix L = D - A

- Symmetric
- Eigenvalues non-negative and real
- Eigenvectors real and orthogonal

$$L\vec{1} = \begin{bmatrix} 3 & -1 & -1 & 0 & -1 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & 0 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = ?$$

Constant vector is eigenvector of L

- The constant vector x=[1,1,...,1]^T is again an eigenvector, but this time has eigenvalue 0

$$Lx = \begin{bmatrix} 3 & -1 & -1 & 0 & -1 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & 0 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0x$$

If the graph is disconnected

- If there are two connected components, the same argument as for the adjacency matrix applies, and $\lambda_1 = \lambda_2 = 0$
- In general, the multiplicity of eigenvalue 0 is the number of connected components

$$x^TLx$$

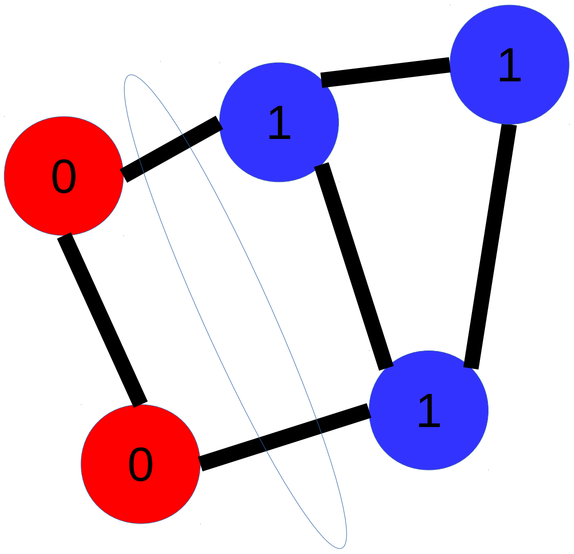
Important property of $x^T Lx$

$$
\begin{aligned}
x^T L x &= \sum_{i=1}^{n} \sum_{j=1}^{n} L_{ij} x_i x_j \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} (D_{ij} - A_{ij}) x_i x_j \\
&= \sum_{i=1}^{n} D_{ii} x_i^2 - \sum_{(i,j)\in E} 2 x_i x_j \\
&= \sum_{(i,j)\in E} x_i^2 + x_j^2 - 2 x_i x_j = \sum_{(i,j)\in E} (x_i - x_j)^2
\end{aligned}
$$

Think of this quantity as the "stress" produced by the assignment of node labels x

$x^T Lx$ and graph cuts

- Suppose (S, S') is a cut of graph G
- Set $x_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \in S' \end{cases}$



$$|c(S, S')| = 2$$

$$x^T Lx = \sum_{(i,j)\in E} (x_i - x_j)^2 = \sum_{(i,j)\in c(S,S')} 1^2 = |c(S, S')|$$

For symmetric matrices

- Important fact for symmetric matrices

$$\lambda_2 = \min_x \frac{x^T M x}{x^T x}$$

https://en.wikipedia.org/wiki/Rayleigh_quotient

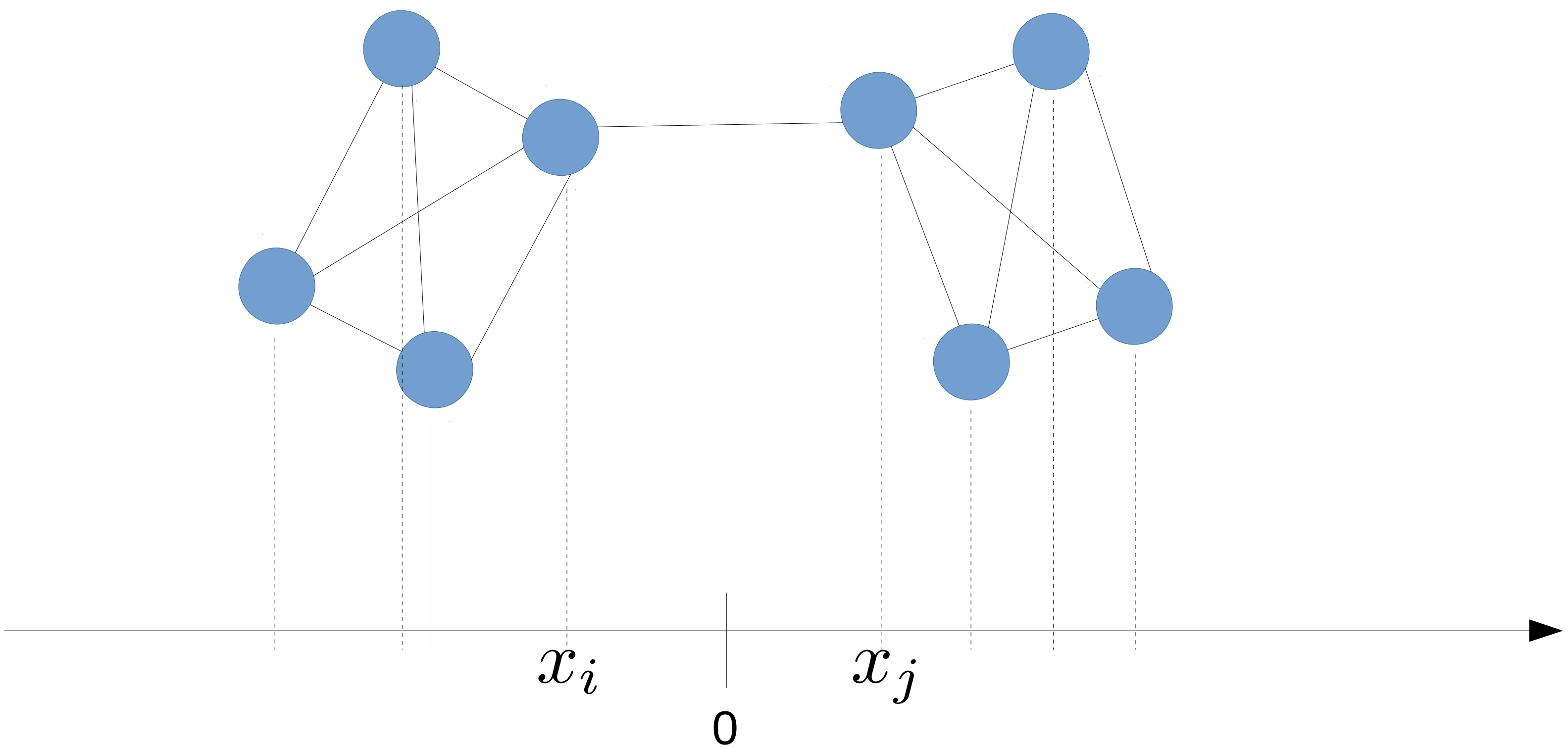
Second eigenvector

- Orthogonal to the first one: $x \cdot \vec{1} = 0 \Rightarrow \sum_i x_i = 0$
- Normal: $\sum_i x_i^2 = 1$

$$\lambda_2 = \min_x \frac{x^T L x}{x^T x} = \min_{x: \sum x_i = 0} \frac{x^T L x}{\sum x_i^2} = \min_{x: \sum x_i = 0} \sum_{(i,j) \in E} (x_i - x_j)^2$$

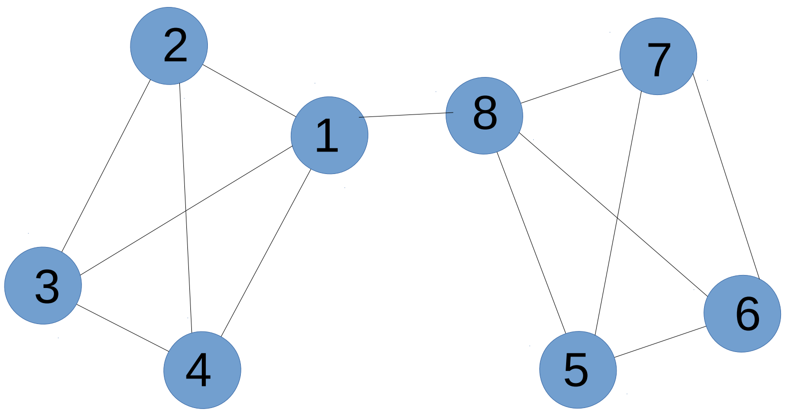
What does this mean?

$$\lambda_2 = \min_{x:\sum x_i = 0} \sum_{(i,j)\in E} (x_i - x_j)^2$$



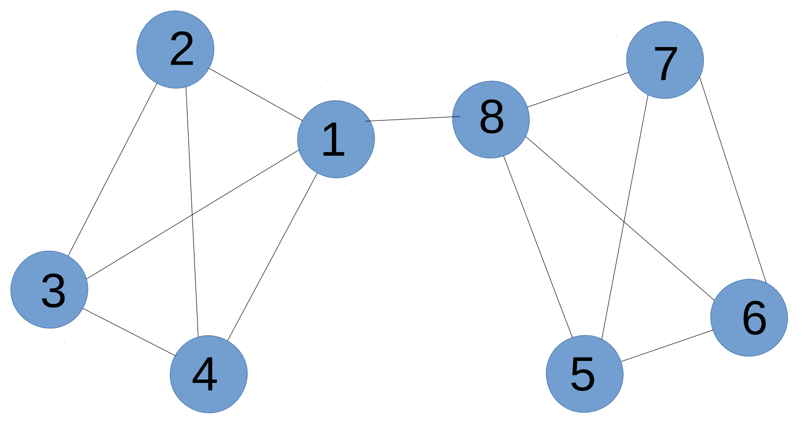
Nodes should be placed at both sides of 0 because $\sum x_i = 0$

Example Graph 1



$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & 0 & 0 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ -1 & 0 & 0 & 0 & -1 & -1 & -1 & 4 \end{bmatrix}$$

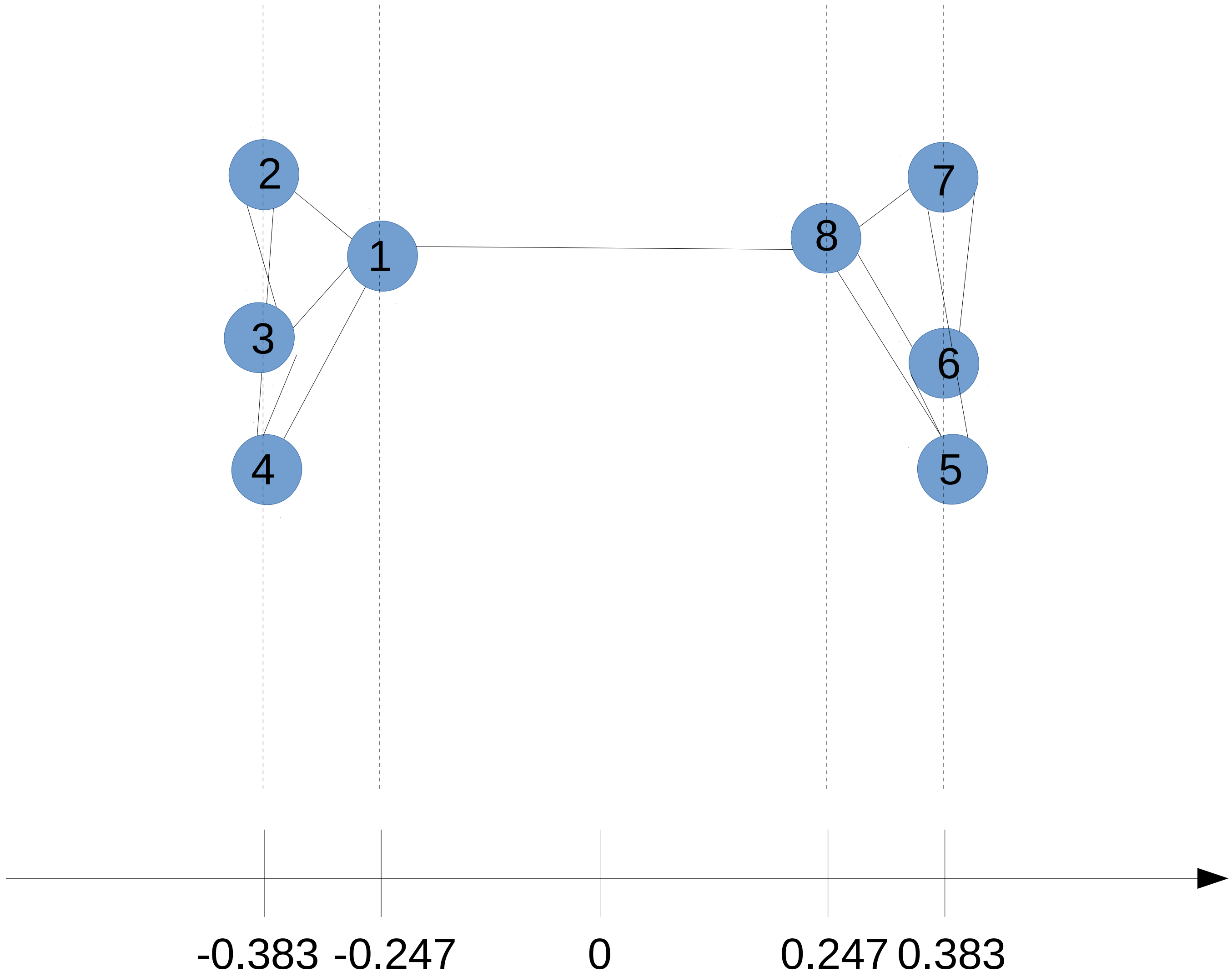
Example Graph 1



$$\lambda_1 = 0$$
$$\lambda_2 = 0.354$$

$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & 0 & 0 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ -1 & 0 & 0 & 0 & -1 & -1 & -1 & 4 \end{bmatrix} \quad v_2 = \begin{bmatrix} 0.247 \\ 0.383 \\ 0.383 \\ 0.383 \\ -0.383 \\ -0.383 \\ -0.383 \\ -0.247 \end{bmatrix}$$

Example Graph 1, projected

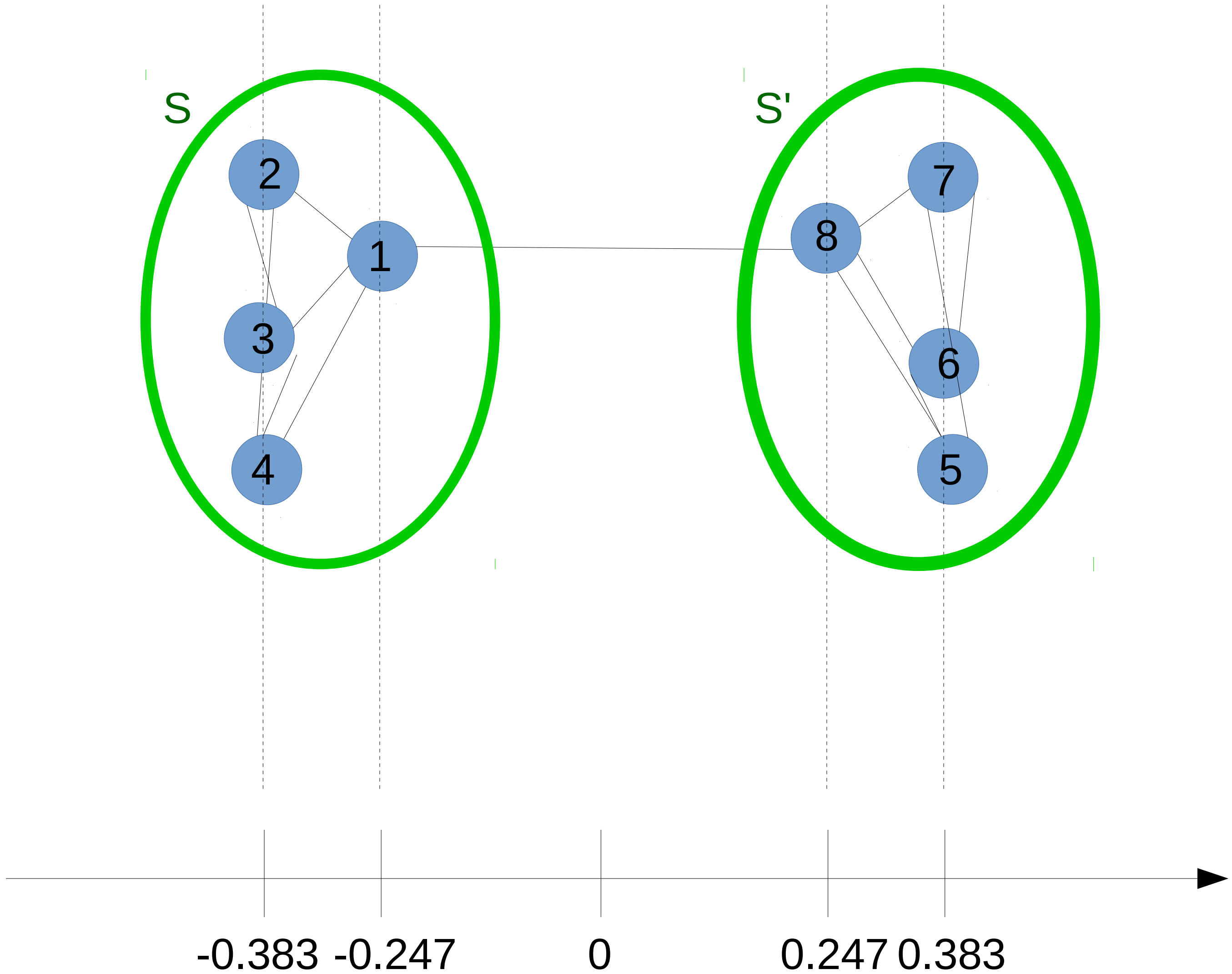


$$\lambda_1 = 0$$

$$\lambda_2 = 0.354$$

$$v_2 = \begin{bmatrix} 0.247 \\ 0.383 \\ 0.383 \\ 0.383 \\ -0.383 \\ -0.383 \\ -0.383 \\ -0.247 \end{bmatrix}$$

Example Graph 1, communities

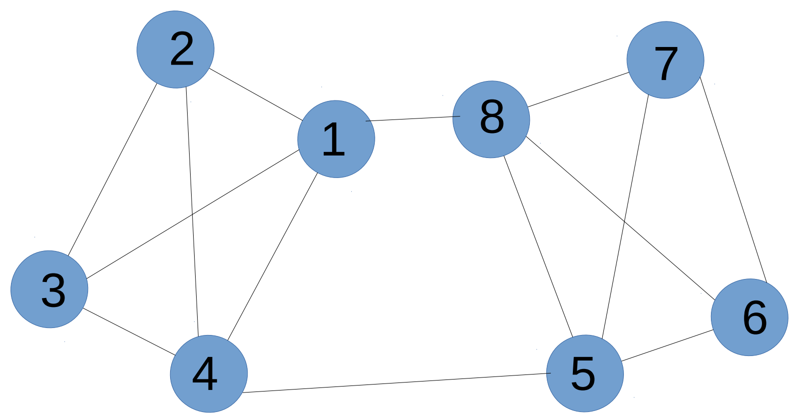


$$\lambda_1 = 0$$

$$\lambda_2 = 0.354$$

$$v_2 = \begin{bmatrix} 0.247 \\ 0.383 \\ 0.383 \\ 0.383 \\ -0.383 \\ -0.383 \\ -0.383 \\ -0.247 \end{bmatrix}$$

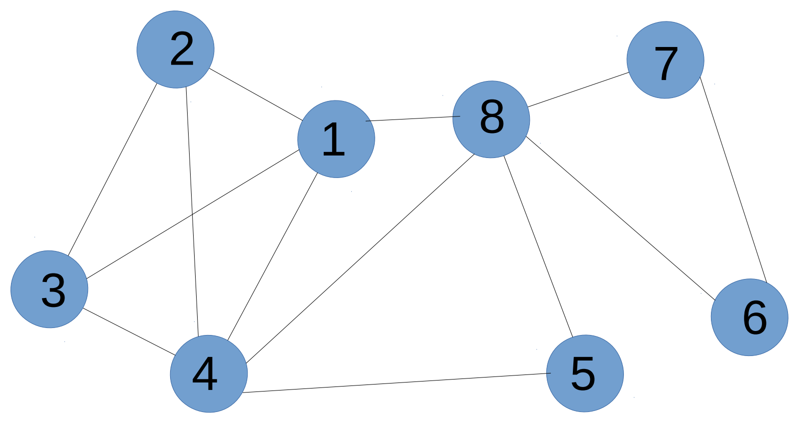
Example Graph 2



$$\lambda_1 = 0$$
$$\lambda_2 = 0.764$$

$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & 0 & 0 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 4 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 4 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ -1 & 0 & 0 & 0 & -1 & -1 & -1 & 4 \end{bmatrix} \quad v_2 = \begin{bmatrix} 0.263 \\ 0.425 \\ 0.425 \\ 0.263 \\ -0.263 \\ -0.425 \\ -0.425 \\ -0.263 \end{bmatrix}$$

Example Graph 3

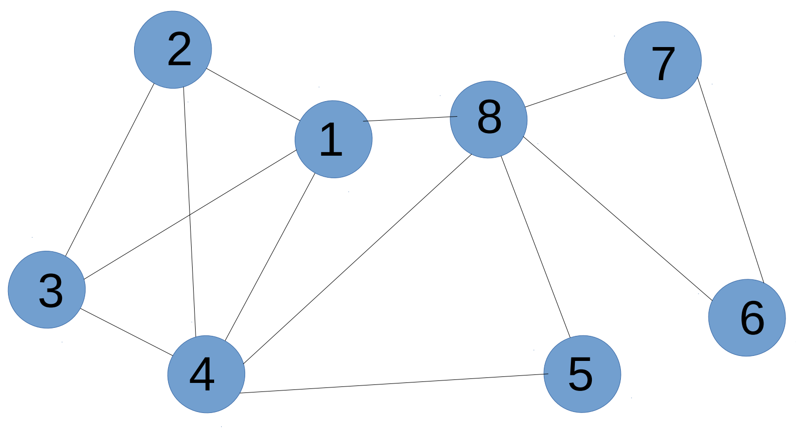


$$\lambda_1 = 0$$

$$\lambda_2 = 0.748$$

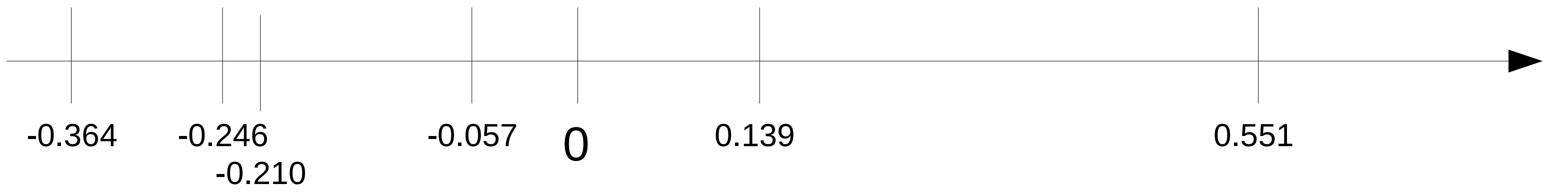
$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & 0 & 0 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 5 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & -1 & -1 & -1 & 5 \end{bmatrix} \quad v_2 = \begin{bmatrix} -0.246 \\ -0.364 \\ -0.364 \\ -0.210 \\ -0.057 \\ 0.551 \\ 0.551 \\ 0.139 \end{bmatrix}$$

Example Graph 3, projected

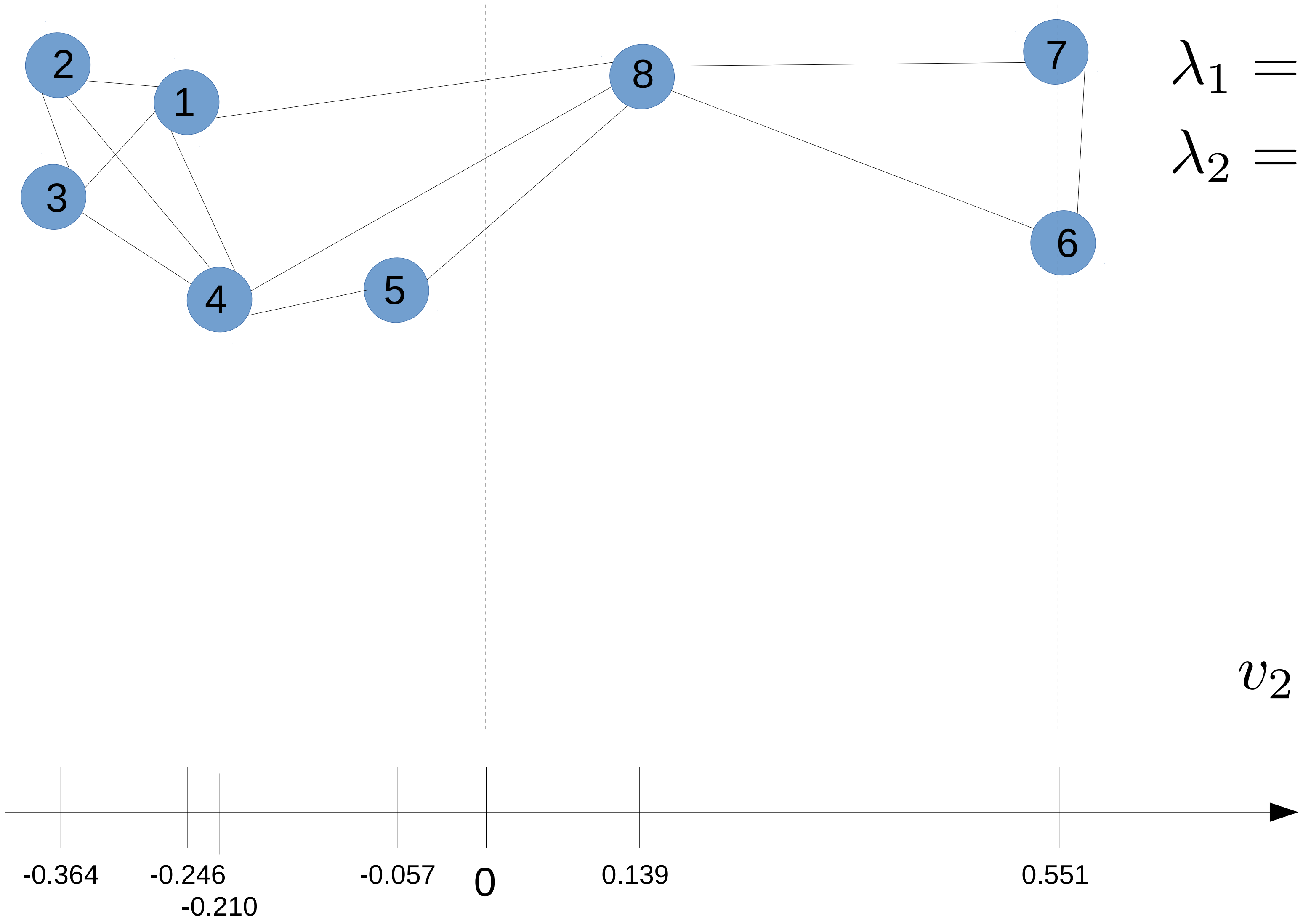


$$\lambda_1 = 0$$
$$\lambda_2 = 0.748$$

$$v_2 = \begin{bmatrix} -0.246 \\ -0.364 \\ -0.364 \\ -0.210 \\ -0.057 \\ 0.551 \\ 0.551 \\ 0.139 \end{bmatrix}$$



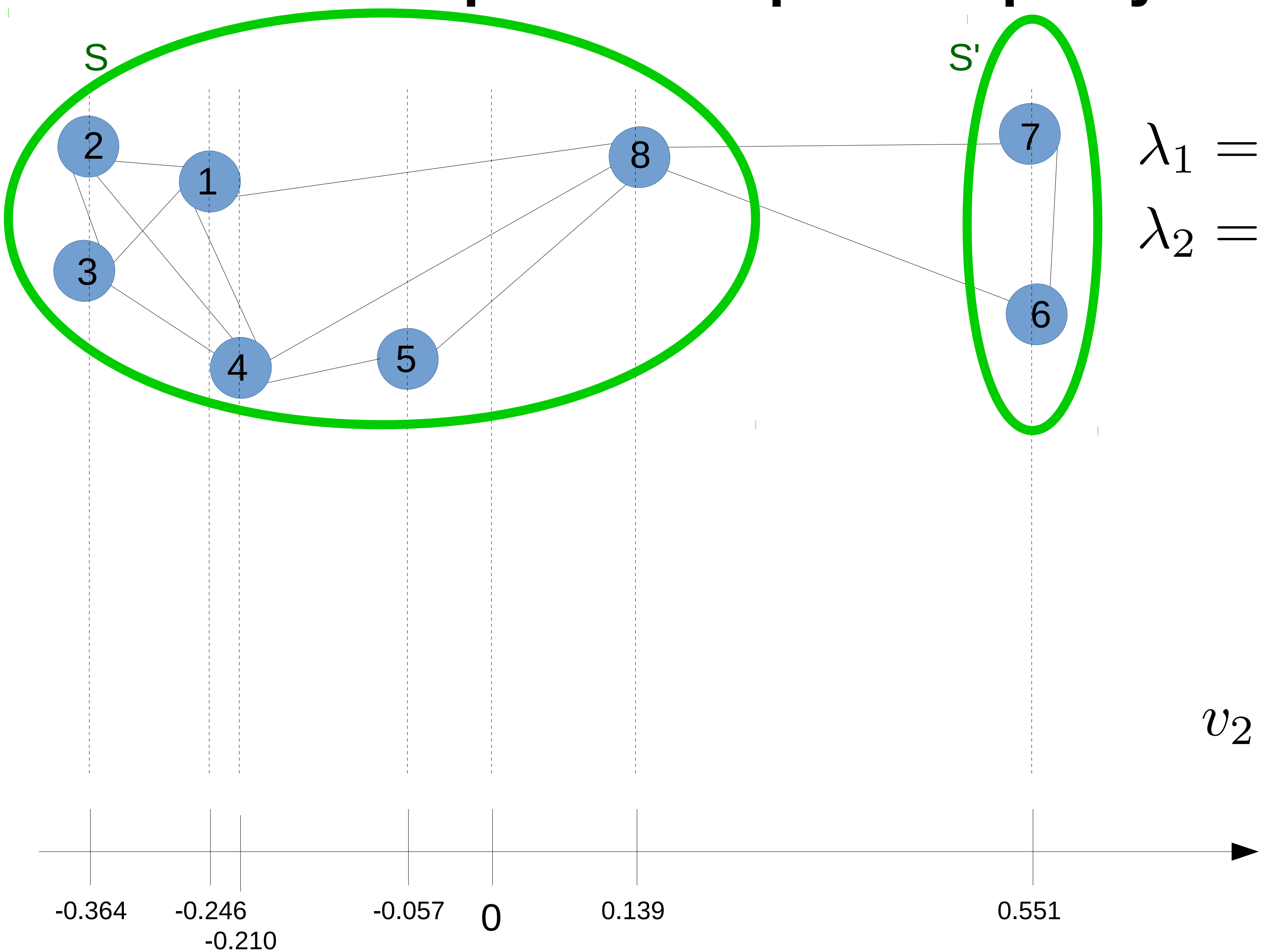
Example Graph 3, projected
how to partition?



$$\lambda_1 = 0$$
$$\lambda_2 = 0.748$$

$$v_2 = \begin{bmatrix} -0.246 \\ -0.364 \\ -0.364 \\ -0.210 \\ -0.057 \\ 0.551 \\ 0.551 \\ 0.139 \end{bmatrix}$$

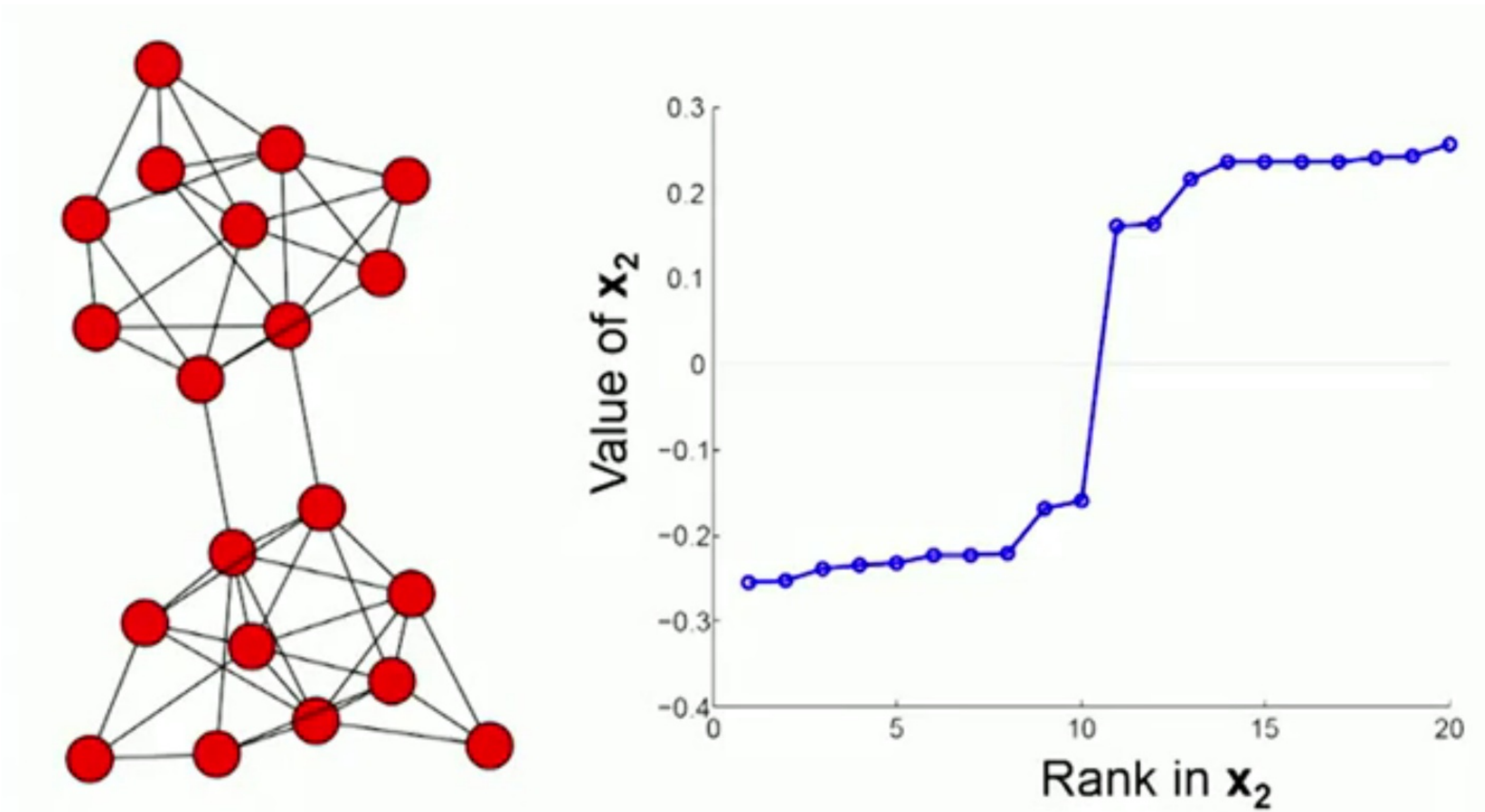
Example Graph 3, projected



$$\lambda_1 = 0$$
$$\lambda_2 = 0.748$$

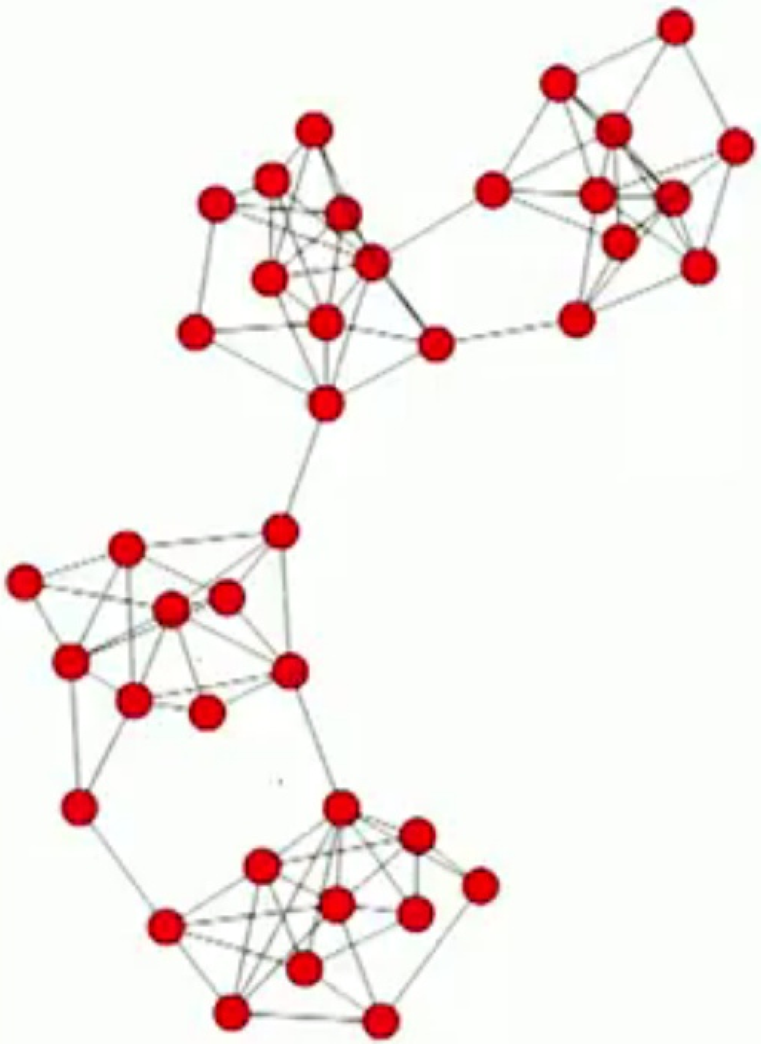
$$v_2 = \begin{bmatrix} -0.246 \\ -0.364 \\ -0.364 \\ -0.210 \\ -0.057 \\ 0.551 \\ 0.551 \\ 0.139 \end{bmatrix}$$

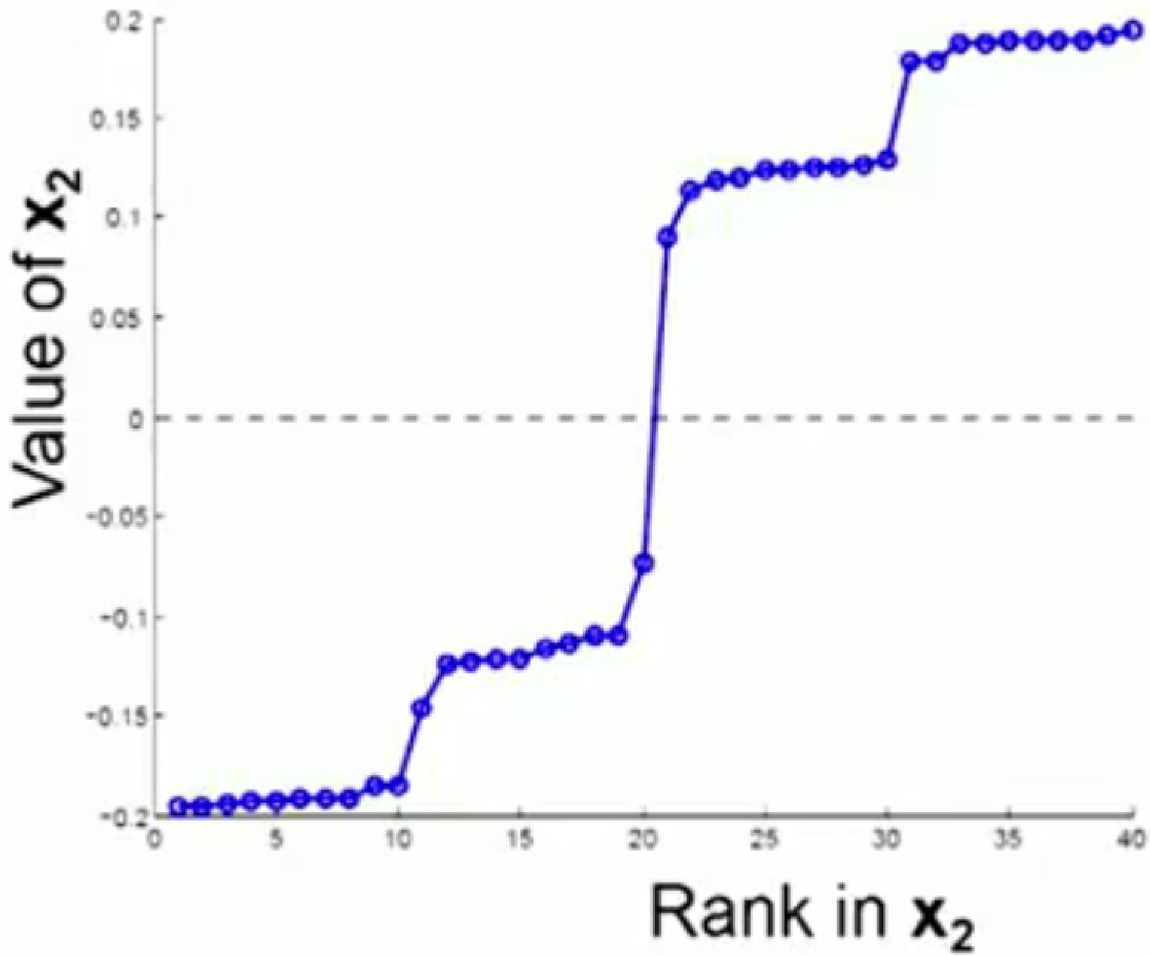
A more complex graph



https://www.youtube.com/watch?v=jpTjj5PmcMM

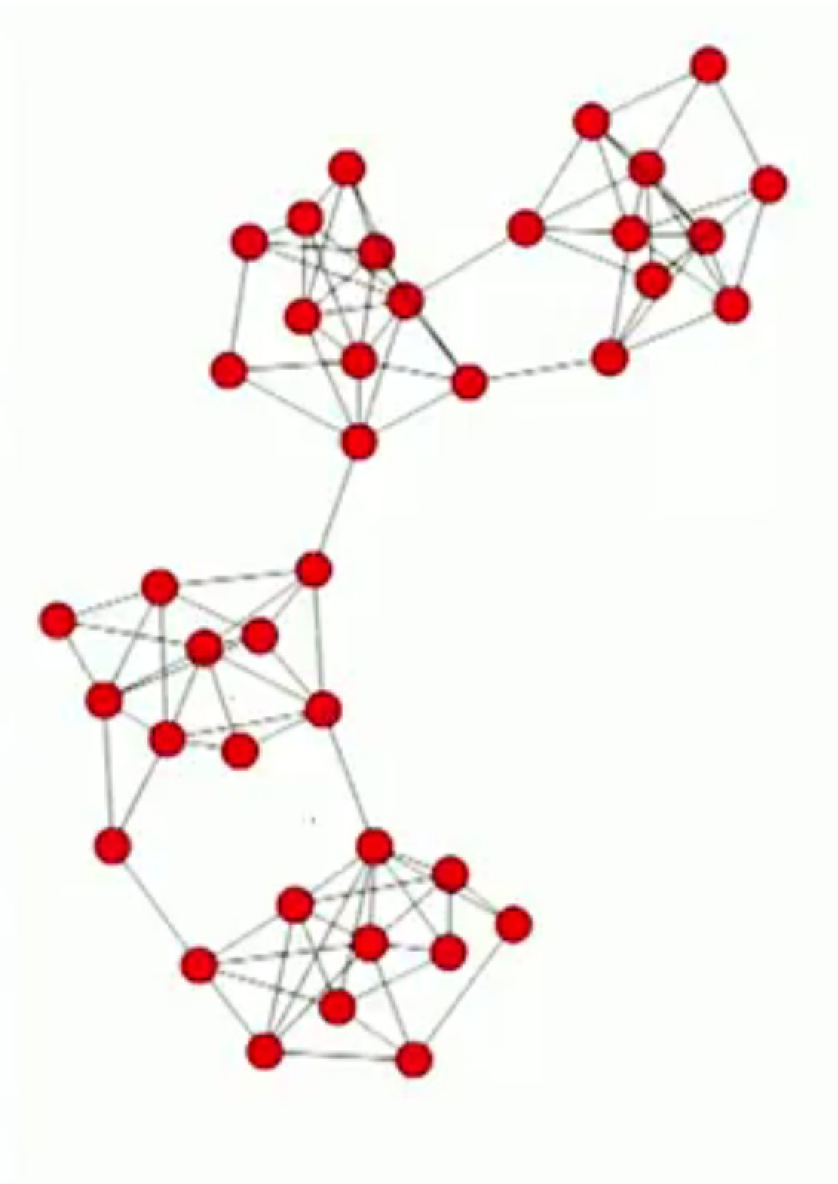
A graph with 4 “communities”

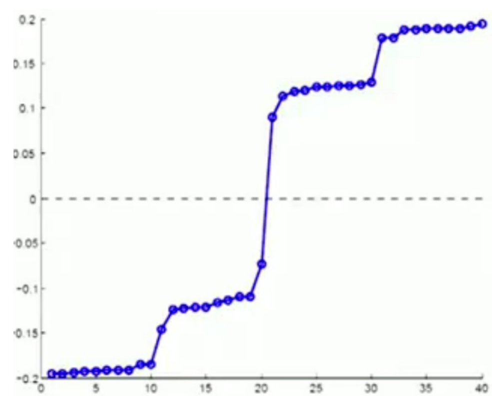




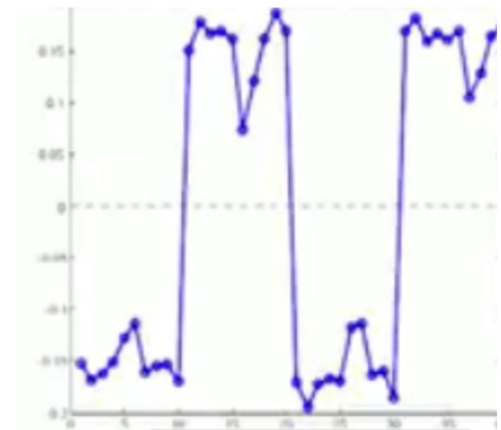
https://www.youtube.com/watch?v=jpTjj5PmcMM

Other eigenvectors



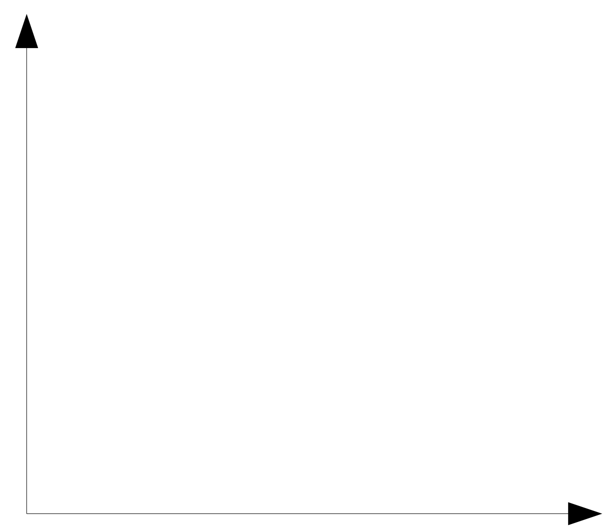


Second eigenvector



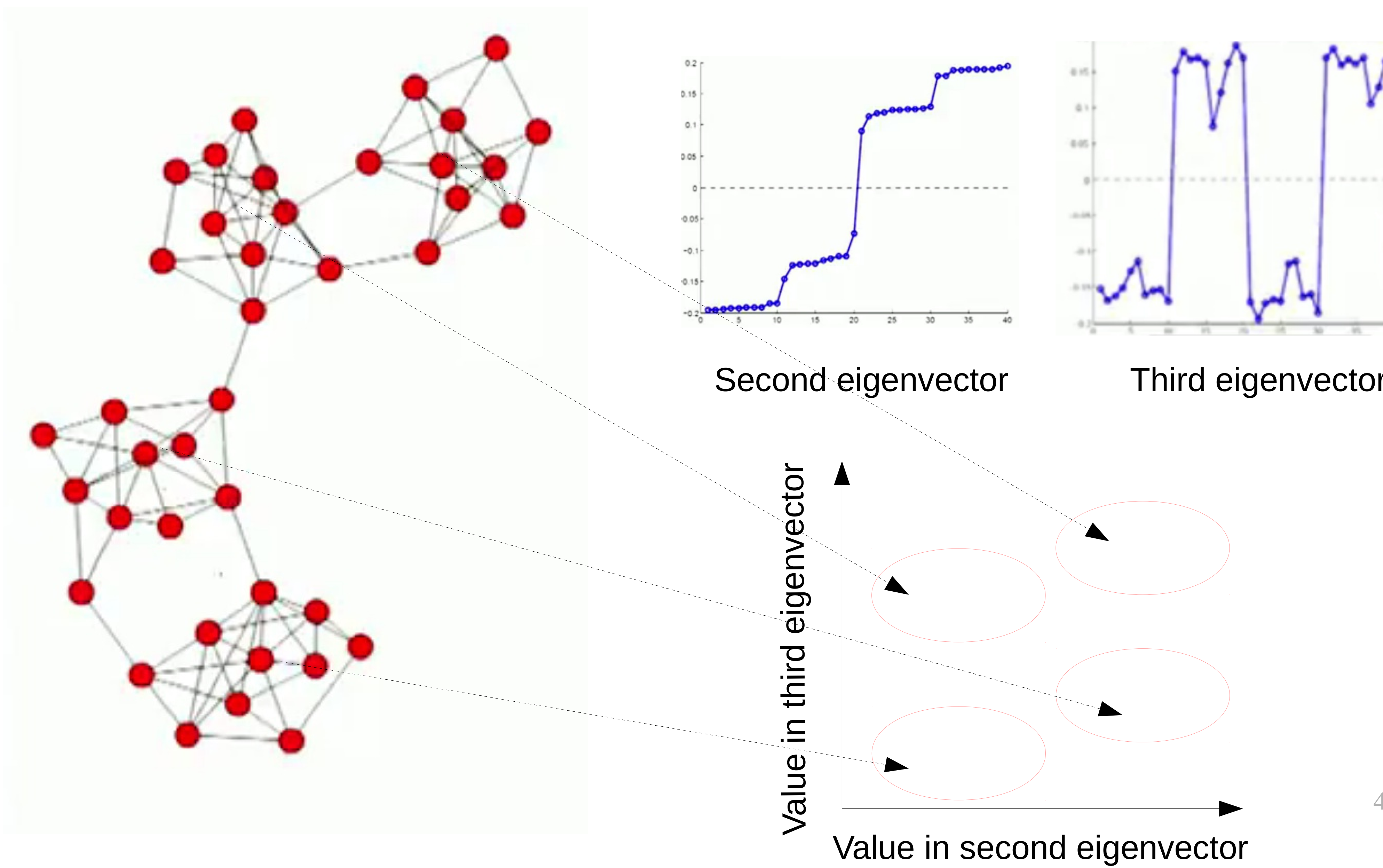
Third eigenvector

Value in third eigenvector

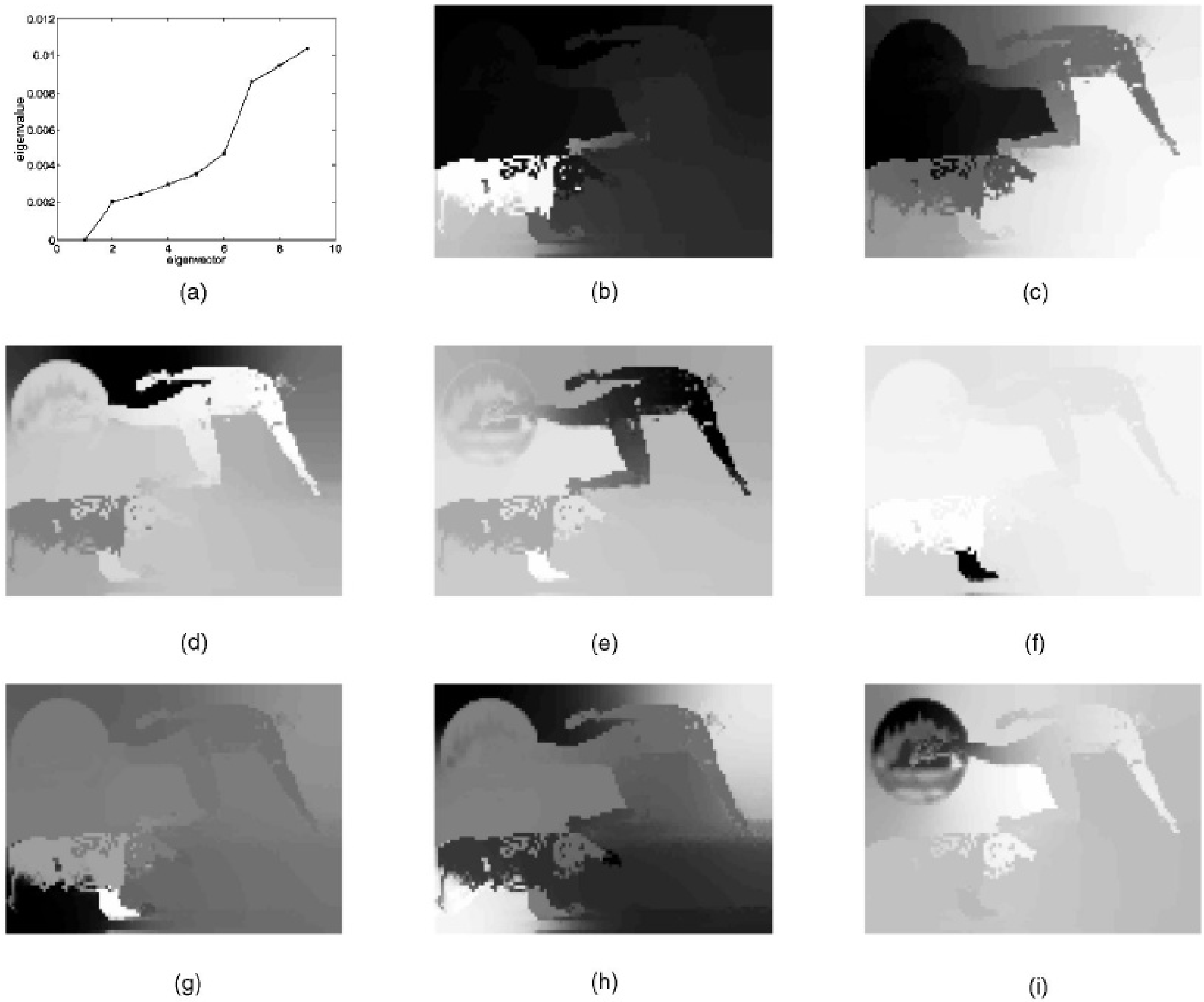


Value in second eigenvector

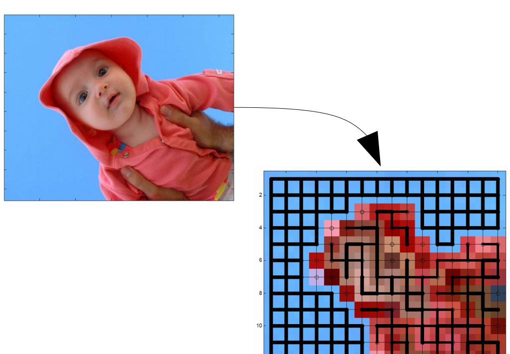
Other eigenvectors



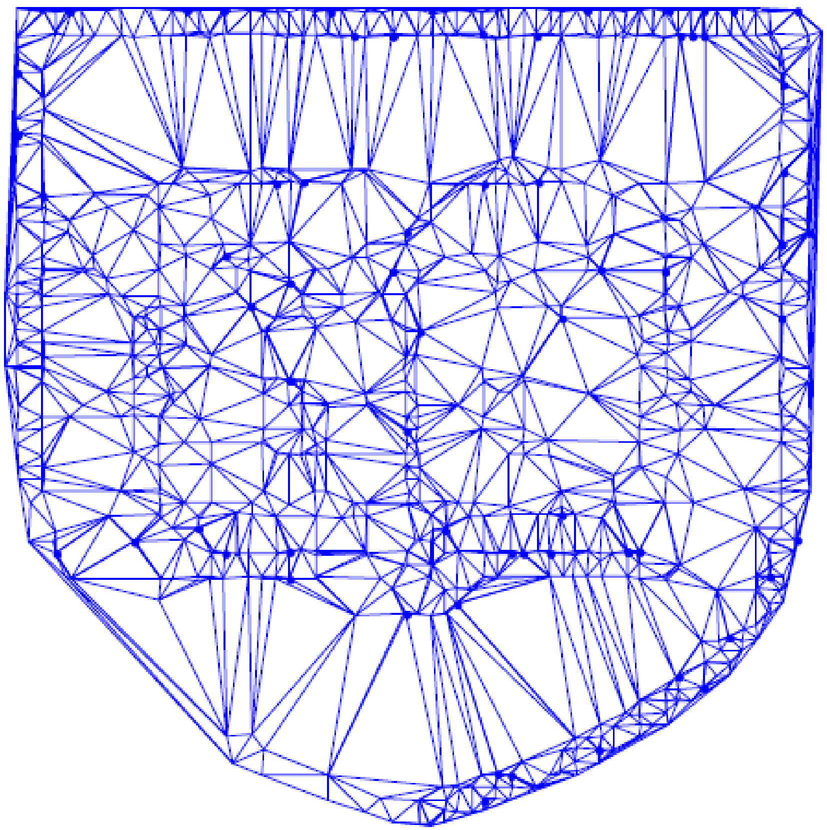
Interesting application: image segmentation [Shi & Malik 2000]

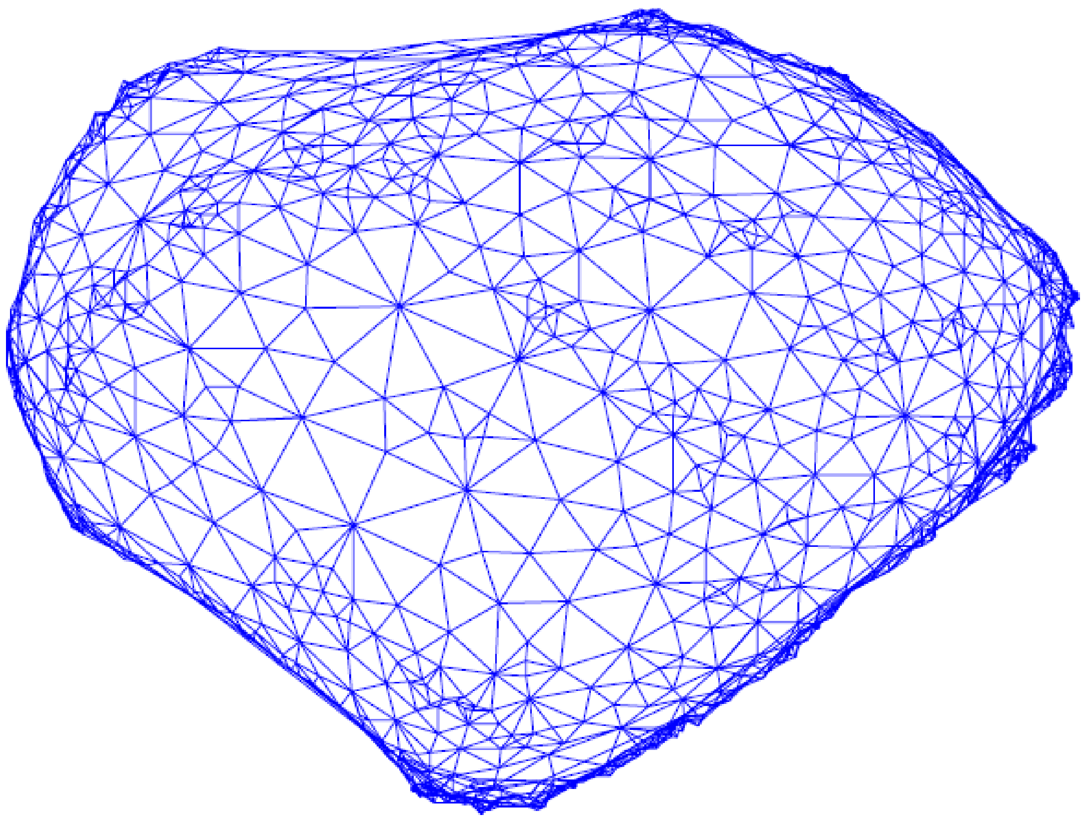


Transform into grid graph with edge weights proportional to pixel similarity



Lots of examples, including one of embedding streets in Rome





http://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf