# Non Overlapping Communities - Proofs and Theory

Davide Mottin, Konstantina Lazaridou

Hasso Plattner Institute

Graph Mining course Winter Semester 2016

# Acknowledgements

- Most of this lecture is taken from: http://web.stanford.edu/class/cs224w/slides
- Some other material:
  - https://www.ismll.uni-hildesheim.de/lehre/cmie-11w/script/lecture5.pdf

Details!

# $d$ is the largest eigenvalue of $A$

- **G** is **d**-regular connected, **A** is its adjacency matrix
- **Claim:**
  - **d** is largest eigenvalue of **A,**
  - **d** has multiplicity of **1** (there is only **1** eigenvector associated with eigenvalue **d**)
- **Proof: Why no eigenvalue $d' > d$?**
  - To obtain **d** we need $\boldsymbol{x}_i = \boldsymbol{x}_j$ for every $i, j$
  - This means $\boldsymbol{x} = c \cdot (1,1, \dots, 1)$ for some const. $c$
  - **Define:** $\boldsymbol{S}$ = nodes $\boldsymbol{i}$ with maximum possible value of $\boldsymbol{x}_i$
  - Then consider some vector $\boldsymbol{y}$ which is not a multiple of vector $(\mathbf{1}, \dots, \mathbf{1})$. So not all nodes $\boldsymbol{i}$ (with labels $\boldsymbol{y}_i$ ) are in $\boldsymbol{S}$
  - Consider some node $\boldsymbol{j} \in \boldsymbol{S}$ and a neighbor $\boldsymbol{i} \notin \boldsymbol{S}$ then node $\boldsymbol{j}$ gets a value strictly less than $\boldsymbol{d}$
  - So $y$ is not eigenvector! And so $\boldsymbol{d}$ is the largest eigenvalue!

Details!

# Facts about the Laplacian L

**(a)** All eigenvalues are $\geq 0$

**(b)** $x^T L x = \sum_{ij} L_{ij} x_i x_j \geq 0$ for every $x$

**(c)** $L = N^T \cdot N$

- That is, $L$ is positive semi-definite

- **Proof:**
  - **(c)⇒(b):** $x^T L x = x^T N^T N x = (xN)^T (Nx) \geq 0$
    - As it is just the square of length of $Nx$ (it's a quadratic form of a matrix)
  - **(b)⇒(a):** Let $\boldsymbol{\lambda}$ be an eigenvalue of $\boldsymbol{L}$. Then by **(b)** $x^T L x \geq 0$ so $x^T L x = x^T \lambda x = \lambda x^T x \Rightarrow \boldsymbol{\lambda} \geq \boldsymbol{0}$
  - **(a)⇒(c):** is also easy! Do it yourself.

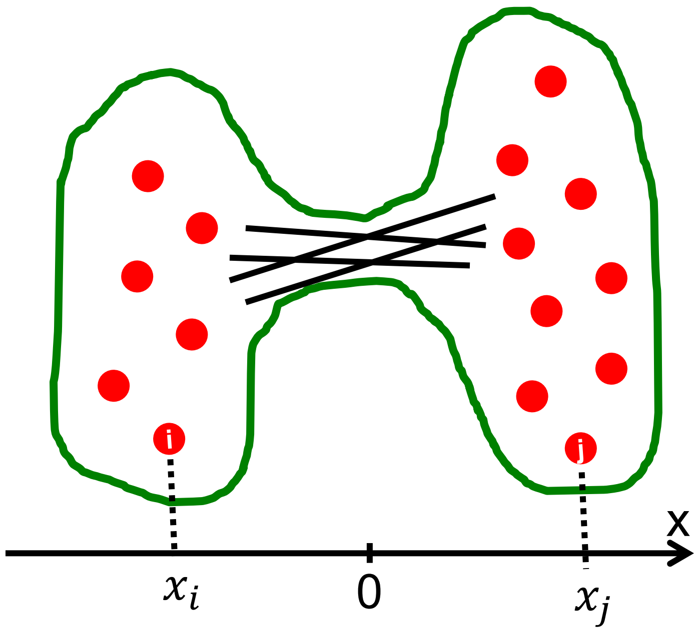
# $\lambda_2$ as optimization problem

- **What else do we know about $x$?**
  - $\boldsymbol{x}$ is unit vector: $\sum_i \boldsymbol{x}_i^2 = \mathbf{1}$
  - $\boldsymbol{x}$ is orthogonal to 1st eigenvector $(\mathbf{1}, \dots, \mathbf{1})$ thus: $\sum_i \boldsymbol{x}_i \cdot \mathbf{1} = \sum_i \boldsymbol{x}_i = \mathbf{0}$

- **Remember:**

$$\lambda_2 = \min_{\substack{\text{All labelings} \\ \text{of nodes } i \text{ so} \\ \text{that } \sum x_i = 0}} \frac{\sum_{(i,j)\in E} (x_i - x_j)^2}{\sum_i x_i^2}$$

**We want to assign values $x_i$ to nodes $i$ such that few edges cross 0.**
**(we want $x_i$ and $x_j$ to subtract each other)**




**Balance to minimize**

Details!

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

- Write $x$ in axes of eigenvectors $w_1, w_2, \ldots, w_n$ of $\boldsymbol{M}$. So, $x = \sum_i^n \alpha_i w_i$
- Then we get: $Mx = \sum_i \alpha_i M w_i = \sum_i \alpha_i \lambda_i w_i$ (where $M w_i = \boldsymbol{\lambda_i w_i}$)
- So, what is $\boldsymbol{x^T M x}$?
  - $x^T M x = (\sum_i \alpha_i w_i)^T (\sum_i \alpha_i \lambda_i w_i) = \sum_{ij} \alpha_i \lambda_j \alpha_j w_i w_j$ (where $w_i w_j = 0$ if $i \neq j$, 1 otherwise)

    $= \sum_i \alpha_i \lambda_i \alpha_i w_i w_i = \sum_i \boldsymbol{\lambda_i \alpha_i^2}$ (since $w_i = 1$)
  - To minimize this over all unit vectors x orthogonal to: (1,1,..1)
  - Well … if you have to find a vector orthogonal to (1,1,…1) that minimizes the above quantity
  - Then, considering that all the eigenvectors are orthogonal the only choice is to pick
    - $\alpha_2 = 1, \alpha_3 = \cdots = \alpha_n = 0$

Details!

# Approx. Guarantee of Spectral

- Suppose there is a partition of **G** into **A** and **B** where $|A| \leq |B|$, s.t. $\boldsymbol{\alpha} = \frac{(\#\, edges\, from\, A\, to\, B)}{|A|}$ then $2\boldsymbol{\alpha} \geq \boldsymbol{\lambda_2}$
  - This is the approximation guarantee of the spectral clustering. It says the cut spectral finds is at most **2** away from the optimal one of score $\boldsymbol{\alpha}$.
- **Proof:**
  - Let: **a=|A|, b=|B|** and **e=** # edges from **A** to **B**
  - Enough to choose some $\boldsymbol{x_i}$ based on **A** and **B** such that: $\lambda_2 \leq \underbrace{\frac{\sum (x_i - x_j)^2}{\sum_i x_i^2}}_{\lambda_2 \text{ is only smaller}} \leq 2\alpha$ (while also $\sum_i x_i = 0$)

Details!

# Approx. Guarantee of Spectral

- Proof (continued):
  - 1) Let's set: $\boldsymbol{x_i} = \begin{cases} -\frac{1}{a} & \boldsymbol{if\ i \in A} \\ +\frac{1}{b} & \boldsymbol{if\ i \in B} \end{cases}$
    - Let's quickly verify that $\sum_i x_i = 0$: $a\left(-\frac{1}{a}\right) + b\left(\frac{1}{b}\right) = \mathbf{0}$
  - 2) Then: $\frac{\sum (x_i - x_j)^2}{\sum_i x_i^2} = \frac{\sum_{i \in A, j \in B}\left(\frac{1}{b}+\frac{1}{a}\right)^2}{a\left(-\frac{1}{a}\right)^2 + b\left(\frac{1}{b}\right)^2} = \frac{e \cdot \left(\frac{1}{a}+\frac{1}{b}\right)^2}{\frac{1}{a}+\frac{1}{b}} = e\left(\frac{1}{a}+\frac{1}{b}\right) \le e\left(\frac{1}{a}+\frac{1}{a}\right) \le$

    $\boldsymbol{e\frac{2}{a} = 2\alpha}$

    e … number of edges between A and B

    Which proves that the cost achieved by spectral is better than twice the OPT cost

Details!

# Approx. Guarantee of Spectral

- **Putting it all together:**

$$2\alpha \geq \lambda_2 \geq \frac{\alpha^2}{2k_{max}}$$

  - where $k_{max}$ is the maximum node degree in the graph
    - Note we only provide the 1st part: $2\alpha \geq \lambda_2$
    - We did not prove $\lambda_2 \geq \frac{\alpha^2}{2k_{max}}$
  - Overall this always certifies that $\lambda_2$ always gives a useful bound