Spectral Clustering

Part 2: Weighted Graph Laplacians

Ng Yen Kaow

Recap

- An intuition from the Laplacian function (in continuous space) gave us the **graph Laplacian matrix** (in graph space)

- Subsequently people found out that the graph Laplacian possesses several properties that lend it to solve graph cutting problems
  - The basic graph cutting problem is known as minimum cut in the literature

Minimum Cut Problem

- The minimum cut of an undirected graph $G = (V, E)$ is a partition of V into two groups S and $\bar{S}$ so that the number of edges between S and $\bar{S}$ is minimized
  - Solvable in polynomial time $O(|E||V| + |V|^2 \log|V|)$ (Nagamochi *et al* 1992, Stoer-Wagner 1995)

- We saw in Part 1 that given a graph Laplacian $L, x^\top Lx =$ 4 times the weight sum of the edge weights to remove in a partitioning
  - An x which minimizes $x^\top Lx$ can be approximated from eigendecomposition
  - In fact, it adds a balance requirement which is required in a minimum bisection

Minimum Bisection Problem

- The minimum bisection of an undirected graph $G = (V, E)$ is a partition of V into two groups S and $\bar{S}$ so that the number of edges between S and $\bar{S}$ is minimized, under the constraint that $|S| = |\bar{S}|$ (or $\big||S| - |\bar{S}|\big| = 1$ for odd $|V|$)

- If we let $x_i = \begin{cases} 1 & \text{if } v_i \in S \\ -1 & \text{if } v_i \in \bar{S} \end{cases}$ (as in minimum cut)

  Then $|S| = |\bar{S}|$ implies $\sum_i x_i = 0$ (or 1 or -1)
  - This condition is partially ensured by the eigendecomposition

$\sum_i x_i = 0$ condition

- As in minimum cut, let $x_i = \begin{cases} 1 & \text{if } v_i \in S \\ -1 & \text{if } v_i \in \bar{S} \end{cases}$
  - As stated in Part 1, eigenvectors of L are orthogonal
  - Furthermore, the vector $\mathbf{1}$ (that is, $\forall i, x_i = 1$) is a eigenvector (since it minimizes $x^\top L x$)

  $\Rightarrow$ Hence $x \perp \mathbf{1} = 0$ for all eigenvectors x of L

  That is, $x \perp \mathbf{1} = (x_1 \quad \cdots \quad x_n)\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_i x_i = 0$

- We can restate minimum bisection as a constrained optimization problem

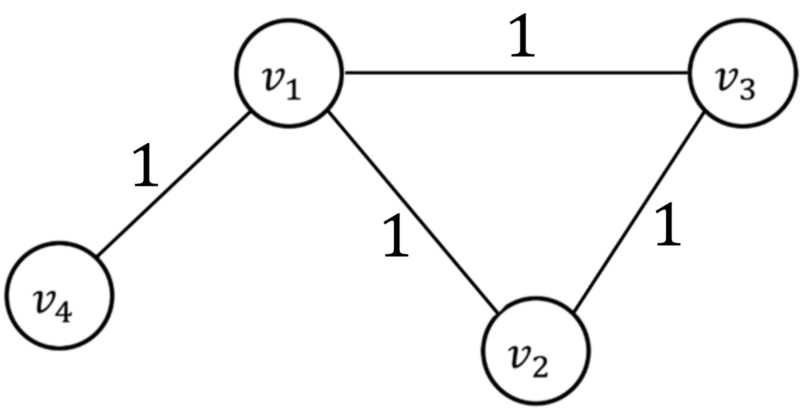
Constrained optimization problem

- Minimize $x^{\top}Lx$ where $L = D - A$
  subject to $x_i \in \{1, -1\}$ and $x^{\top}\mathbf{1} = 0$
  - The constraints $x_i \in \{1, -1\}$ and $x^{\top}\mathbf{1} = 0$ (that is, $x \perp \mathbf{1}$) would together ensures balance in the partition

- Problem (minimum bisection) is NP-hard
- In contrast, eigendecomposition of a $|V| \times |V|$ matrix takes $O(|V|^3)$ time

Constrained optimization problem

- Minimize $x^\top L x$ where $L = D - A$ subject to $x_i \in \{1, -1\}$ and $x^\top \mathbf{1} = 0$
  - Recall the exhaustive search we performed in Part 1



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

Group 1	Group 2	$x^\top L x$	$\frac{x^\top L x}{x^\top x}$
A	B C D	12	3
B	A C D	8	2
C	A B D	8	2
D	A B C	4	1
A B	C D	12	3
A C	B D	12	3
A D	**B C**	**8**	**2**
A B C D	∅	0	0

Under the balance constraint

Constrained optimization problem

- Minimize $x^\top L x$ where $L = D - A$ subject to $x_i \in \{1, -1\}$ and $x^\top \mathbf{1} = 0$
  - Recall the exhaustive search we performed in Part 1

Exercise: Modify the program you wrote in Part 1 to output only balanced partitions

Group 1	Group 2	$x^\top L x$	$\frac{x^\top L x}{x^\top x}$
A	B C D	12	3
B	A C D	8	2
C	A B D	8	2
D	A B C	4	1
A B	C D	12	3
A C	B D	12	3
A D	**B C**	**8**	**2**
A B C D	∅	0	0

Under the balance constraint: rows A B / C D, A C / B D, A D / B C

Relaxed Rayleigh quotient version

- Minimize $x^\top Lx$ where $L = D - A$
  subject to $x^\top x = 1$ and $x^\top \mathbf{1} = 0$
  - $x^\top x = 1$ (or any constant)
    - Allows problem to be solved as minimization of $\frac{\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}}$
      - The (standard) Rayleigh quotient is scale invariant so limiting $x^\top x$ to any constant does not change its value
      - By the min-max theorem, λ_{k-1} is minimal among all $\frac{x^\top Lx}{x^\top x}$ that are orthogonal to μ_k
  - $x^\top \mathbf{1} = 0$
    - Automatically fulfilled by μ_{k-1}
  - Balance no longer ensured

    Both $\frac{[1\ 1\ {-1}\ {-1}]}{\|[1\ 1\ {-1}\ {-1}]\|}$ and $\frac{[1\ 1\ 1\ {-3}]}{\|[1\ 1\ 1\ {-3}]\|}$ fulfill the constraints

Relaxed Rayleigh quotient version

- Eigenvalues

λ_1	λ_2	λ_3	λ_4
4.0000	3.0000	1.0000	0.0000

- Eigenvectors

μ_1	μ_2	μ_3	μ_4
0.8660	0.0000	0.0000	-0.5000
-0.2887	0.7071	-0.4082	-0.5000
-0.2887	-0.7071	-0.4082	-0.5000
-0.2887	0.0000	0.8165	-0.5000

- As expected $\mu_4 = b\mathbf{1}$ $(b = -0.5)$ gives the trivial solution
- Furthermore, $\lambda_3 \leq 2$, the optimal solution under constraint
  - This is as expected since λ_3 is minimal solution among all x orthogonal to μ_4 but without the 1 and -1 restriction

Relaxed Rayleigh quotient version

- Eigenvalues

λ_1	λ_2	λ_3	λ_4
4.0000	3.0000	1.0000	0.0000

- Eigenvectors

μ_1	μ_2	μ_3	μ_4
0.8660	0.0000	0.0000	-0.5000
-0.2887	0.7071	-0.4082	-0.5000
-0.2887	-0.7071	-0.4082	-0.5000
-0.2887	0.0000	0.8165	-0.5000

- **Exercise:** Verify that the eigenvectors μ_1, μ_2, μ_3, and μ_4 are orthogonal by showing that for each i and j, $\mu_i \cdot \mu_2 \approx 0$

Introducing weights into problems

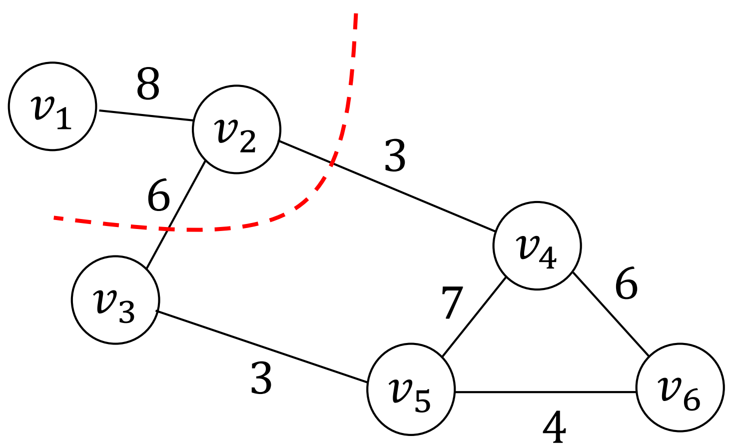
- Unweighted (undirected) graphs
  - Unbalanced version
    - (Unweighted) Minimum Cut Problem
  - Balanced version
    - Minimum Bisection Problem (NP-hard)

Discussed

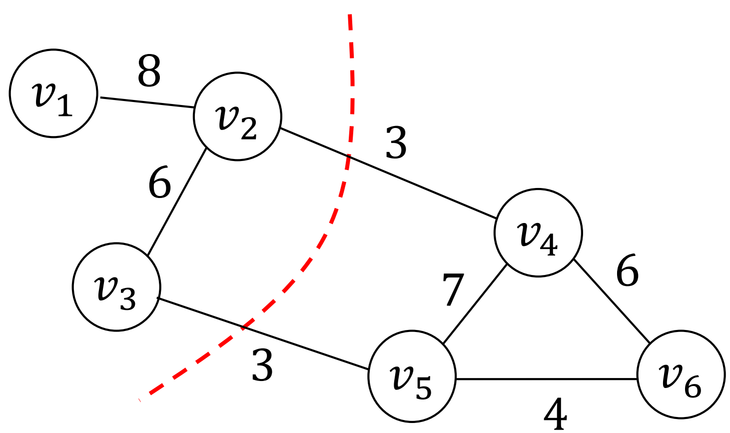
- Weighted (undirected) graphs
  - Unbalanced version
    - (Weighted) Minimum Cut Problem $O(|V||E|)$
  - Balanced versions
    - Ratio Cut Problem (NP-hard)
    - Graph Partitioning Problem (NP-hard)

(Weighted) Minimum Cut Problem

- Given edge weight matrix $W = (w_{ij})$, the minimum cut of an undirected graph $G = (V, E)$ is a partition of V into two groups S and $\bar{S}$ such that $\text{cut}(S, \bar{S}) = \sum_{i \in S, j \in \bar{S}} w_{ij}$ is minimized



$\text{cut}(A, B) = 9$



$\text{cut}(A, B) = 6$

(Weighted) Minimum Cut Problem

- Given edge weight matrix $W = (w_{ij})$, the minimum cut of an undirected graph $G = (V, E)$ is a partition of V into two groups S and $\bar{S}$ such that $\mathrm{cut}(S, \bar{S}) = \sum_{i \in S, j \in \bar{S}} w_{ij}$ is minimized
  - Ford-Fulkerson 1956
  - Edmonds-Karp 1972 (rediscovery of Dinitz 1970)
  - Current best algorithm runs in $O(|V||E|)$ time
    - No point in approximation with spectral clustering
    - Mentioned here only for completeness
- Need graph Laplacian with edge weights

Graph Laplacian with edge weights

- To add weight to the Laplacian
  - Adjacency matrix $A \Longrightarrow$ weight matrix W
  - Degree matrix $D \Longrightarrow$ weighted degree D'
- Laplacian $L = D - A$ becomes $L = D' - W$
- Given edge weights $W = \left(w_{ij}\right)_{m \times m}$, for any vector $x \in \mathbb{R}^m$,

$$x^\top (D' - W) x = \frac{1}{2} \sum_{1 \le i,j \le m} w_{ij} \left(x_i - x_j\right)^2$$

(Proof same as for $x^\top (D - A) x = \frac{1}{2} \sum_{1 \le i,j \le m} a_{ij} \left(x_i - x_j\right)^2$)

Graph Laplacian with edge weights

- To add weight to the Laplacian
  - Adjacency matrix $A \Longrightarrow$ weight matrix W
  - Degree matrix $D \Longrightarrow$ weighted degree D'
- Laplacian $L = D - A$ becomes $L = D' - W$
- Suppose x is a vector of only the values +1 and -1. Then,

$$x^\top (D' - W)x = \frac{1}{2}\sum_{1\le i,j\le m} w_{ij}\left(x_i - x_j\right)^2$$

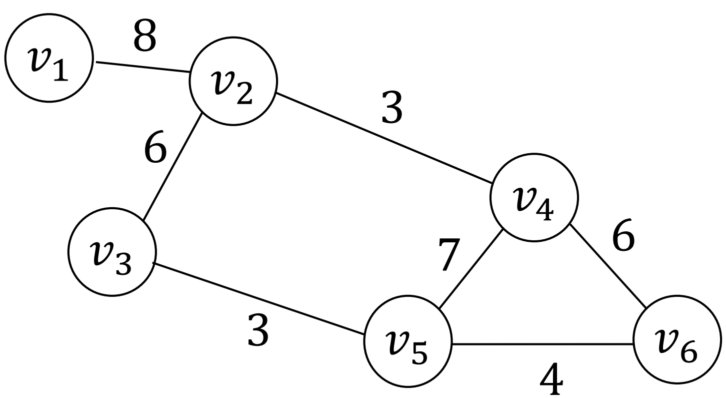
$$= \frac{1}{2}\sum_{1\le i,j\le m} w_{ij}\left(x_i - x_j\right)^2 = 4\sum_{1\le i<j\le m, x_i\neq x_j} w_{ij}$$

$$= 4\,\mathrm{cut}(A, B)$$

Constrained optimization problem

- Minimize $x^\top L x$ where $L = D' - W$ subject to $x_i \in \{1, -1\}$

- Example of cuts with $x^\top L x$ and Rayleigh quotient

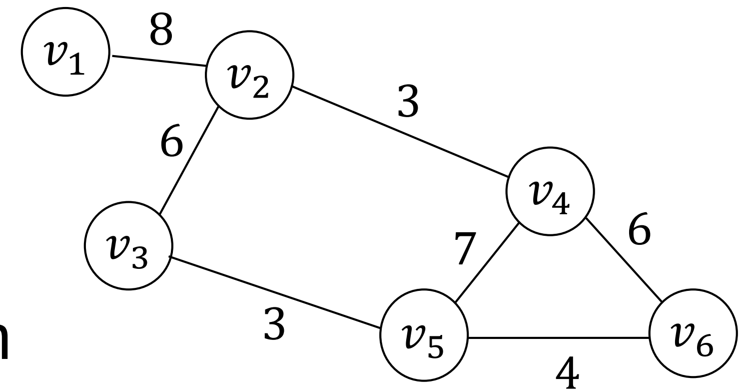


Group 1	Group 2	$x^\top L x$	$\frac{x^\top L x}{x^\top x}$
v_1	$v_2 v_3 v_4 v_5 v_6$	32	5.333
$v_1 v_2 v_3 v_4 v_5$	v_6	40	6.667
$v_1 v_2$	$v_3 v_4 v_5 v_6$	36	6.000
$v_1 v_2 v_3 v_4$	$v_5 v_6$	64	10.667
$v_1 v_2 v_3 v_5$	$v_4 v_6$	56	9.333
$\boldsymbol{v_1 v_2 v_3}$	$\boldsymbol{v_4 v_5 v_6}$	**24**	**4.000**
$v_1 v_2 v_4$	$v_3 v_5 v_6$	76	12.667
⋮	⋮	⋮	⋮

- **Exercise:** Produce a complete list of partitions

Relaxed Rayleigh quotient version

- Minimize $x^\top Lx$ where $L = D' - W$ subject to $x^\top x = 1$



- **Exercise:** Derive W for the graph and obtain the following eigendecomposition

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
25.73	20.49	16.14	8.46	**3.18**	0.00

Eigenvectors

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
-0.1081	0.2775	-0.0777	-0.5096	0.6920	-0.4082
0.4137	-0.7045	0.1720	0.2260	0.2920	-0.4082
-0.2622	0.4073	0.1308	0.7525	0.1237	-0.4082
-0.6924	-0.4100	-0.2506	-0.1448	-0.3193	-0.4082
0.4953	0.2290	-0.6521	-0.0049	-0.3321	-0.4082
0.1538	0.2008	0.6776	-0.3193	-0.4563	-0.4082

Ratio Cut Problem

- Given edge weight matrix $W = (w_{ij})$, the minimum ratio cut of an undirected graph $G = (V, E)$ is a partition of V into two groups S and $\bar{S}$ such that

$$\text{ratio}(S, \bar{S}) = \text{cut}(S, \bar{S}) \left(\frac{1}{|S|} + \frac{1}{|\bar{S}|} \right)$$

  is minimized, where $\text{cut}(S, \bar{S}) = \sum_{i \in S, j \in \bar{S}} w_{ij}$

- Original paper defined $\text{ratio}(S, \bar{S}) = \text{cut}(S, \bar{S}) / |S||\bar{S}|$

$$= \frac{1}{|V|} \text{cut}(S, \bar{S}) \left(\frac{1}{|S|} + \frac{1}{|\bar{S}|} \right)$$

Ratio Cut

- Represent a partition $S, \bar{S}$ of V with $x \in \mathbb{R}^n$, where

$$x_i = \begin{cases} \sqrt{\frac{|S|}{|\bar{S}|}} & \text{if } i \in S \\ -\sqrt{\frac{|\bar{S}|}{|S|}} & \text{if } i \in \bar{S} \end{cases}$$

Unlike earlier formulation, $|x_i|$ **is not a constant – it changes according to the solution**

- Then, $x^{\mathrm{T}}x = |S|\frac{|\bar{S}|}{|S|} + |\bar{S}|\frac{|S|}{|\bar{S}|} = |V| =$ const
- $\sum_i x_i = \sum_{i \in S}\sqrt{\frac{|\bar{S}|}{|S|}} - \sum_{v_i \in \bar{S}}\sqrt{\frac{|S|}{|\bar{S}|}} = |S|\sqrt{\frac{|\bar{S}|}{|S|}} - |\bar{S}|\sqrt{\frac{|S|}{|\bar{S}|}} = 0$

  $\Rightarrow x \perp \mathbf{1}$ (in fact, it can be shown that $x \perp b\mathbf{1}$ for any b)
- For the unnormalized weighted Laplacian $L = D' - W$

$$x^{\mathrm{T}}Lx = |V|\mathrm{cut}(S, \bar{S})\left(\frac{1}{|S|} + \frac{1}{|\bar{S}|}\right) = |V|\mathrm{ratio}(S, \bar{S})$$

Proof for $x^\top Lx = |V|\text{ratio}(S,\bar{S})$

- $x^\top Lx = \frac{1}{2}\sum_{1\le i,j\le m} w_{ij}\left(x_i - x_j\right)^2$

$$= \frac{1}{2}\sum_{i\in S, j\in\bar{S}} w_{ij}\left(\sqrt{\frac{|S|}{|\bar{S}|}} + \sqrt{\frac{|\bar{S}|}{|S|}}\right)^2 + \frac{1}{2}\sum_{i\in S, j\in\bar{S}} w_{ij}\left(-\sqrt{\frac{|S|}{|\bar{S}|}} - \sqrt{\frac{|\bar{S}|}{|S|}}\right)^2$$

$$= \sum_{i\in S, j\in\bar{S}} w_{ij}\left(\frac{|S|}{|\bar{S}|} + \frac{|\bar{S}|}{|S|} + 2\right) = \text{cut}(S,\bar{S})\left(\frac{|S|}{|\bar{S}|} + \frac{|\bar{S}|}{|S|} + 2\right)$$

$$= \text{cut}(S,\bar{S})\left(\frac{|S|}{|\bar{S}|} + \frac{|\bar{S}|}{|S|} + \frac{|S|}{|S|} + \frac{|\bar{S}|}{|\bar{S}|}\right)$$

$$= \text{cut}(S,\bar{S})\left(\frac{|S| + |\bar{S}|}{|\bar{S}|} + \frac{|S| + |\bar{S}|}{|S|}\right)$$

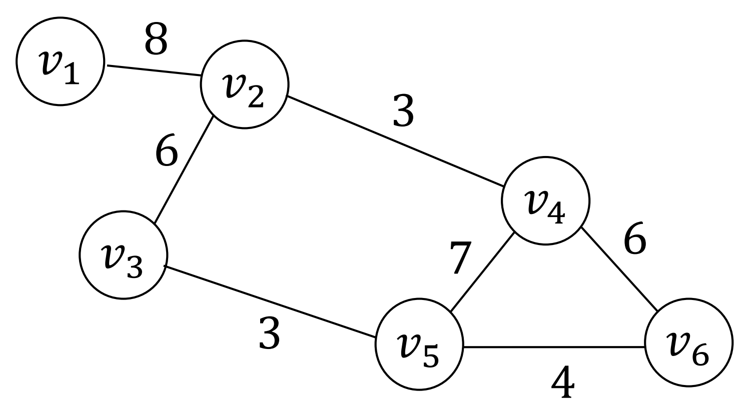
$$= (|S| + |\bar{S}|)\text{cut}(S,\bar{S})\left(\frac{1}{|\bar{S}|} + \frac{1}{|S|}\right) = |V|\text{cut}(S,\bar{S})\left(\frac{1}{|\bar{S}|} + \frac{1}{|S|}\right)$$

Constrained optimization problem

- Minimize $x^\top L x$ where $L = D' - W$ subject to $x_i \in \{\sqrt{|S|/|\bar{S}|}, -\sqrt{|S|/|\bar{S}|}\}$
  - $x_i \in \left\{\sqrt{\frac{|S|}{|\bar{S}|}}, -\sqrt{\frac{|S|}{|\bar{S}|}}\right\} \Rightarrow x^\top x = |V|$ and $x^\top \mathbf{1} = 0$
  - However, problem is NP-hard

- Example of cuts with $x^\top L x$ and Rayleigh quotient

Group 1	Group 2	$x^\top L x$	$\frac{x^\top L x}{x^\top x}$
v_1	$v_2 v_3 v_4 v_5 v_6$	32	5.333
$v_1 v_2 v_3 v_4 v_5$	v_6	40	6.667
$v_1 v_2$	$v_3 v_4 v_5 v_6$	36	6.000
$v_1 v_2 v_3 v_4$	$v_5 v_6$	64	10.667
$v_1 v_2 v_3 v_5$	$v_4 v_6$	56	9.333
$v_1 v_2 v_3$	$v_4 v_5 v_6$	24	4.000
$v_1 v_2 v_4$	$v_3 v_5 v_6$	76	12.667



$\text{cut}(A, B) = 6$

Relaxed Rayleigh quotient version

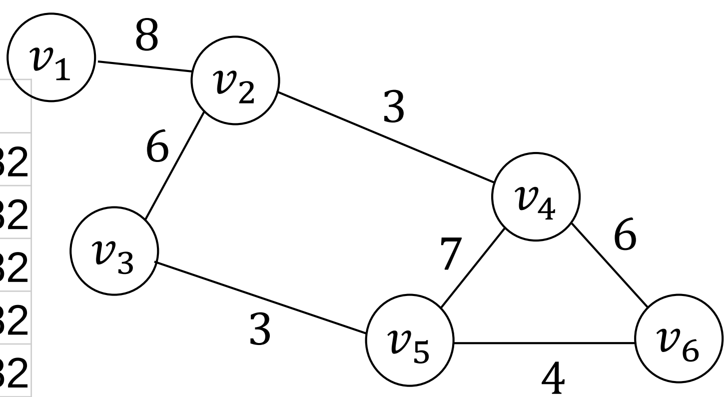
- Minimize $x^{\mathsf{T}}Lx$ where $L = D' - W$ subject to $x^{\mathsf{T}}x = 1$ and $x^{\mathsf{T}}\mathbf{1} = 0$
  - Since $x^{\mathsf{T}}Lx \neq |V|\text{ratio}(S, \bar{S}) \Rightarrow$ balance no longer enforced

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
25.73	20.49	16.14	8.46	**3.18**	0.00

Eigenvectors

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
-0.1081	0.2775	-0.0777	-0.5096	0.6920	-0.4082
0.4137	-0.7045	0.1720	0.2260	0.2920	-0.4082
-0.2622	0.4073	0.1308	0.7525	0.1237	-0.4082
-0.6924	-0.4100	-0.2506	-0.1448	-0.3193	-0.4082
0.4953	0.2290	-0.6521	-0.0049	-0.3321	-0.4082
0.1538	0.2008	0.6776	-0.3193	-0.4563	-0.4082



$\text{cut}(A, B) = 6$

Relaxed Rayleigh quotient version

- Minimize $x^\top Lx$ where $L = D' - W$ subject to $x^\top x = 1$ and $x^\top \mathbf{1} = 0$
  - Since $x^\top Lx \neq |V|\text{ratio}(S, \bar{S}) \Rightarrow$ balance no longer enforced

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
25.73	20.49	16.14	8.46	**3.18**	0.00

Eigenvectors

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
-0.1081	0.2775	-0.0777	-0.5096	0.6920	-0.4082
0.4137	-0.7045	0.1720	0.2260	0.2920	-0.4082
-0.2622	0.4073	0.1308	0.7525	0.1237	-0.4082
-0.6924	-0.4100	-0.2506	-0.1448	-0.3193	-0.4082
0.4953	0.2290	-0.6521	-0.0049	-0.3321	-0.4082
0.1538	0.2008	0.6776	-0.3193	-0.4563	-0.4082

The eigenvalue system is exactly the same as in (Weighted) Minimum Cut

- As expected $\mu_6 = b\mathbf{1}$ $(b = -0.4082)$ provides a trivial solution
- As expected $\lambda_5 \leq 2.67$ since the optimal solution under constraint, since λ_5 is minimal among all $\frac{x^\top Lx}{x^\top x}$ for x orthogonal to μ_6

Comparison of problems

- Unweighted problem ($L = D - A$)
  - Minimum Cut
  - Add balance ⇒ Minimum Bisection
- Weighted problems ($L = D' - W$)
  - (Weighted) Minimum Cut
  - Add balance ⇒ Ratio Cut
- The version of the problem with $x^\top \mathbf{1} = 0$ balance requirement can better exploit the fact that $\mu_{k-1} \perp \mu_k$ where $\mu_k = \mathbf{1}$ which helps optimality
- However note that even with $x^\top \mathbf{1} = 0$, the balance requirement is not ensured

More constraint for balance

- So far, no attempt has been made to maintain the balance of the partition besides $x^\top x = 1$ and $x^\top \mathbf{1} = 0$, constraints which are provided free-of-charge by the eigenvectors of the eigenvalue system
- Further constraints can be added to the eigenvalue system
  - **The Normalized Laplacian**