Spectral Clustering

Part 2: Weighted Graph Laplacians

Ng Yen Kaow

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
  - Karger's algorithm finds an optimal solution in $\mathrm{O}(n^2 m \log n)$ time with probability $1/\binom{n}{2}$
- Recall that for the unnormalized $L = D - W$, $x^{\mathrm{T}} L x$ = 4 times the number of adjacent vertices of different values in x
  - We showed in Part 1 that an x which minimizes $x^{\mathrm{T}} L x$ can be approximated from eigendecomposition
  - In fact, it added an (ineffective) balance requirement

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 $|V|$)
- As in minimum cut, let $x_i = \begin{cases} 1 & \text{if } v_i \in S \\ -1 & \text{if } v_i \in \bar{S} \end{cases}$
  - In which case, $|S| = |\bar{S}|$ implies $\sum_i x_i = 0$, which implies $x \perp \mathbf{1}$ (or $x \perp b\mathbf{1}$ for any constant value b)

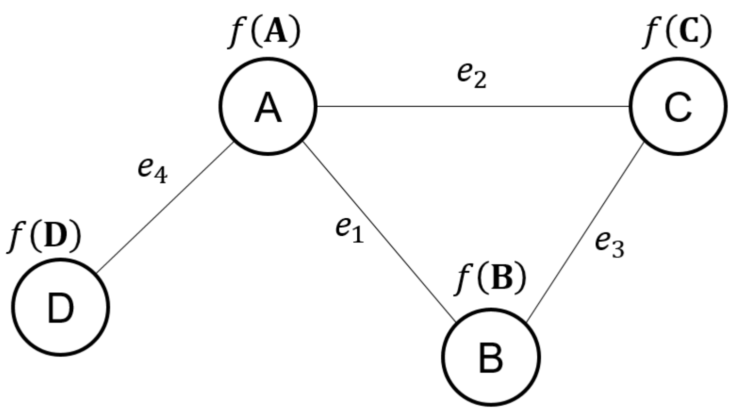
Constrained optimization problem

- Minimize $x^{\mathrm{T}}Lx$ where $L = D - A$ subject to $x_i \in \{1, -1\}$ and $x^{\mathrm{T}}\mathbf{1} = 0$
  - $x_i \in \{1, -1\}$ and $x^{\mathrm{T}}\mathbf{1} = 0$ (that is, $x \perp \mathbf{1}$) together ensures balance in the partition
- Problem is NP-hard

Constrained optimization problem

- Minimize $x^{\mathrm{T}}Lx$ where $L = D - A$ subject to $x_i \in \{1, -1\}$ and $x^{\mathrm{T}}\mathbf{1} = 0$

- Example of cuts (shown with Rayleigh quotient for comparison later)



Group 1	Group 2	$x^T Lx$	$\frac{x^T Lx}{x^T x}$
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

Relaxed Rayleigh quotient version

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

    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

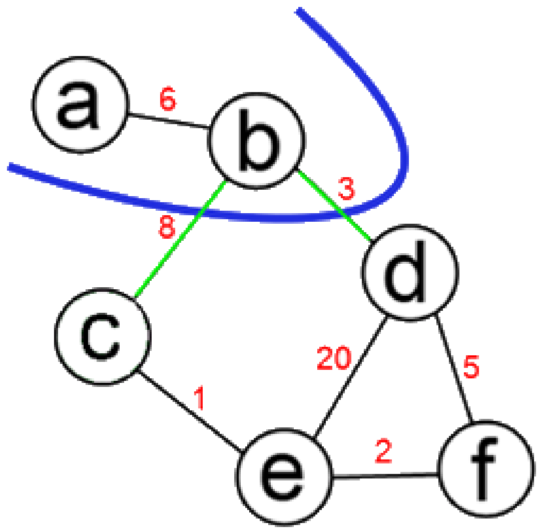
- As expected $\mu_4 = b\mathbf{1}$ $(b = -0.5)$ gives the trivial solution
- As expected $\lambda_3 \leq 2$, the optimal solution under constraint, since λ_3 is minimal among all $\frac{x^{\mathrm{T}}Lx}{x^{\mathrm{T}}x}$ for x orthogonal to μ_4

Introducing weights into problems

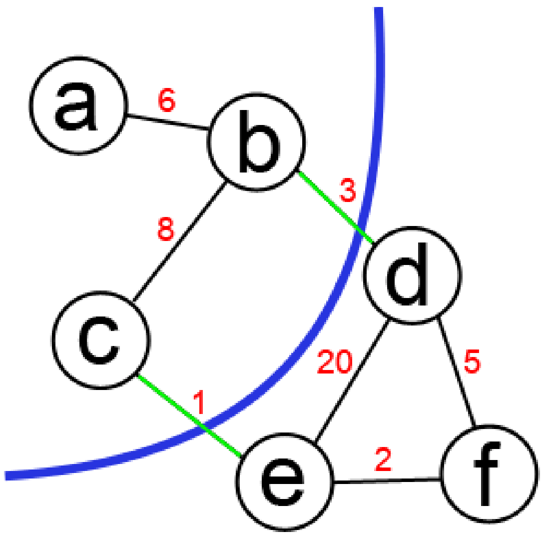
- Unweighted (undirected) graphs
  - Unbalanced version
    - (Unweighted) Minimum Cut Problem
  - Balanced version
    - Minimum Bisection Problem (NP-hard)
- Weighted (undirected) graphs
  - Unbalanced version
    - (Weighted) Minimum Cut Problem $O(|V||E|)$
  - Balanced versions
    - Ratio Cut Problem (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 $\mathrm{cut}(S, \bar{S}) = \sum_{i \in S, j \in \bar{S}} w_{ij}$ is minimized



$\mathrm{cut}(A, B) = 11$



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

(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 algorithm
  - Edmonds-Karp algorithm
  - Current best algorithm runs in $O(|V||E|)$ time
    - No point in using spectral clustering
    - Just as an example try anyway
    - First, define the 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 $x \in \mathbb{R}^m$,

$$x^{\mathrm{T}}(D' - W)x = \frac{1}{2}\sum_{1\le i,j\le m} w_{ij}\left(x_i - x_j\right)^2$$

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

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,

$$x^{\mathrm{T}}(D' - W)x = \frac{1}{2}\sum_{1 \le i,j \le m} w_{ij}\left(x_i - x_j\right)^2$$

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

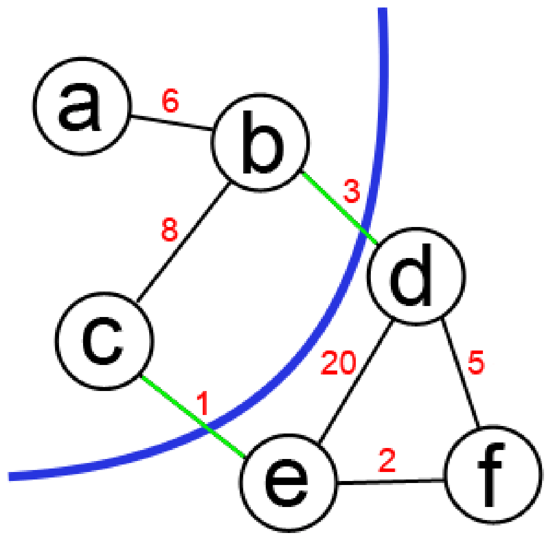
$$= 4\,\mathrm{cut}(A, B)$$

Constrained optimization problem

- Minimize $x^{\mathrm{T}}Lx$ where $L = D' - W$ subject to $x_i \in \{1, -1\}$

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

Group 1	Group 2	$x^T L x$	$\frac{x^T L x}{x^T x}$
A	B C D E F	24	4.00
A B C D E	F	28	4.67
A B	C D E F	44	7.33
A B C E	D F	100	16.67
A B C D	E F	104	17.33
A B C	**D E F**	**16**	**2.67**
A B D	C E F	132	22.00



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

Relaxed Rayleigh quotient version

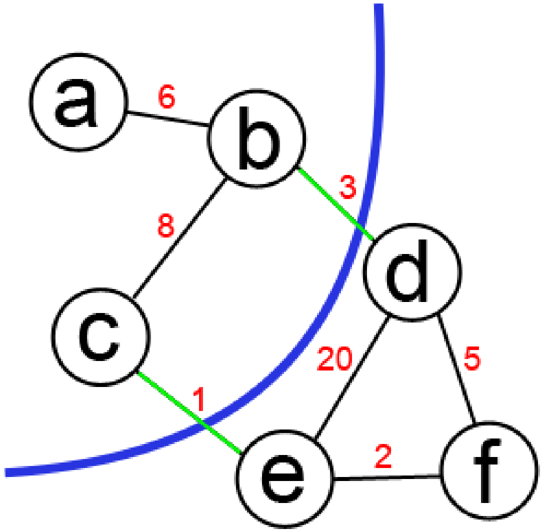
- Minimize $x^{\mathrm{T}}Lx$ where $L = D' - W$
  subject to $x^{\mathrm{T}}x = 1$

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
46.04	23.36	11.07	7.28	**2.25**	0.00

Eigenvectors

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
-0.0136	-0.2879	-0.0224	-0.6854	-0.5291	-0.4082
0.0907	0.8331	0.0189	0.1460	-0.3306	-0.4082
-0.0371	-0.4557	0.1779	0.6912	-0.3390	-0.4082
-0.7519	0.0242	-0.3924	0.0007	0.3368	-0.4082
0.6488	-0.1212	-0.5194	0.0226	0.3570	-0.4082
0.0631	0.0074	0.7374	-0.1750	0.5049	-0.4082



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

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

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

- Then, $x^{\mathrm{T}}x = |S|\frac{|\bar{S}|}{|S|} + |\bar{S}|\frac{|S|}{|\bar{S}|} = |V| = \text{const}$

- $\sum_i x_i = \sum_{i \in S} \sqrt{\frac{|\bar{S}|}{|S|}} - \sum_{v_i \in \bar{S}} \sqrt{\frac{|S|}{|\bar{S}|}} = |S|\sqrt{\frac{|\bar{S}|}{|S|}} - |\bar{S}|\sqrt{\frac{|S|}{|\bar{S}|}} = 0$

  $\Rightarrow x \perp \mathbf{1}$ (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 $x^{\mathrm{T}}Lx = |V|\mathrm{ratio}(S,\bar{S})$

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

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

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

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

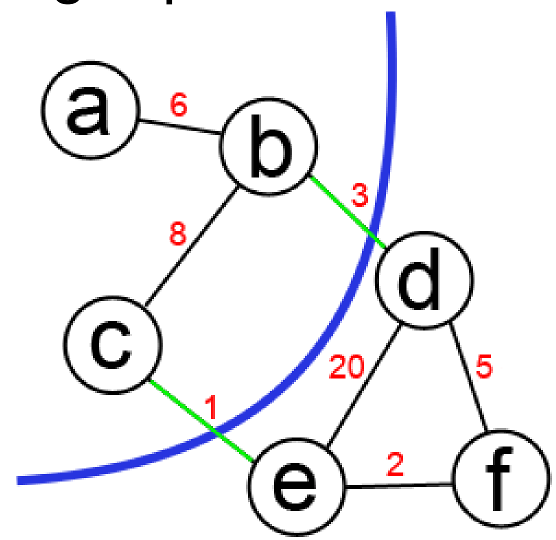
$$= \mathrm{cut}(S,\bar{S})\left(\frac{|S| + |\bar{S}|}{|\bar{S}|} + \frac{|S| + |\bar{S}|}{|S|}\right)$$

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

Constrained optimization problem

- Minimize $x^{\mathrm{T}}Lx$ where $L = D' - W$ subject to $x_i \in \{1, -1\}$ and $x^{\mathrm{T}}\mathbf{1} = 0$
  - $x_i \in \{1, -1\}$ and $x^{\mathrm{T}}\mathbf{1} = 0$ enforce balance
  - However, problem is NP-hard
- Example of cuts with $x^T Lx$ and Rayleigh quotient

Group 1	Group 2	$x^T Lx$	$\frac{x^T Lx}{x^T x}$
A	B C D E F	24	4.00
A B C D E	F	28	4.67
A B	C D E F	44	7.33
A B C E	D F	100	16.67
A B C D	E F	104	17.33
A B C	**D E F**	**16**	**2.67**
A B D	C E F	132	22.00



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

Relaxed Rayleigh quotient version

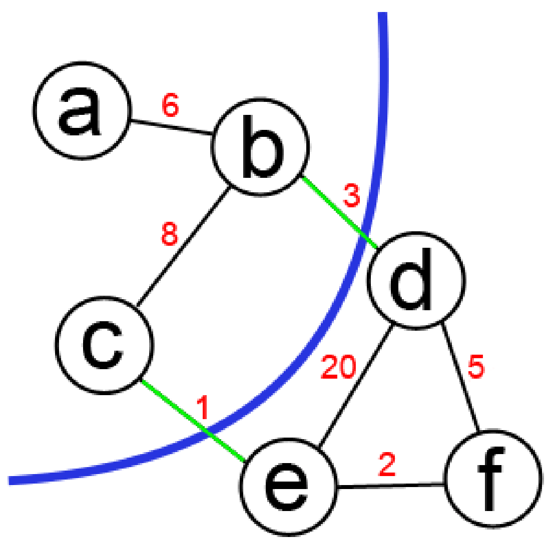
- Minimize $x^{\mathrm{T}}Lx$ where $L = D' - W$ subject to $x^{\mathrm{T}}x = 1$ and $x^{\mathrm{T}}\mathbf{1} = 0$
  - Balance no longer enforced

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
46.04	23.36	11.07	7.28	**2.25**	0.00

Eigenvectors

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
-0.0136	-0.2879	-0.0224	-0.6854	-0.5291	-0.4082
0.0907	0.8331	0.0189	0.1460	-0.3306	-0.4082
-0.0371	-0.4557	0.1779	0.6912	-0.3390	-0.4082
-0.7519	0.0242	-0.3924	0.0007	0.3368	-0.4082
0.6488	-0.1212	-0.5194	0.0226	0.3570	-0.4082
0.0631	0.0074	0.7374	-0.1750	0.5049	-0.4082



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

Relaxed Rayleigh quotient version

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
46.04	23.36	11.07	7.28	**2.25**	0.00

Eigenvectors

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
-0.0136	-0.2879	-0.0224	-0.6854	-0.5291	-0.4082
0.0907	0.8331	0.0189	0.1460	-0.3306	-0.4082
-0.0371	-0.4557	0.1779	0.6912	-0.3390	-0.4082
-0.7519	0.0242	-0.3924	0.0007	0.3368	-0.4082
0.6488	-0.1212	-0.5194	0.0226	0.3570	-0.4082
0.0631	0.0074	0.7374	-0.1750	0.5049	-0.4082

The eigenvalue system is exactly the same as in (Weighted) Minimum Cut – only the optimal solution is different

- As expected $\mu_6 = b\mathbf{1}$ ($b = -0.4082$) gives the trivial solution
- As expected $\lambda_5 \leq 2.67$, the optimal solution under constraint, since λ_5 is minimal among all $\frac{x^{\mathrm{T}} L x}{x^{\mathrm{T}} x}$ 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 $x\mathbf{1} = 0$ balance requirement can better exploit the fact that $\mu_{k-1} \perp \mu_k$ where $\mu_k = \mathbf{1}$ to claim optimality
- However note that even with $x^{\mathrm{T}}\mathbf{1} = 0$, 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 $x^{\mathrm{T}}x = 1$ and $x^{\mathrm{T}}\mathbf{1} = 0$, both of which are fulfilled automatically for a eigenvalue system
- Further constraints can be added to the eigenvalue system
  - However, the resultant eigenvalue system will no longer be standard
  - Demonstrate with Graph Partitioning Problem in Part 3