

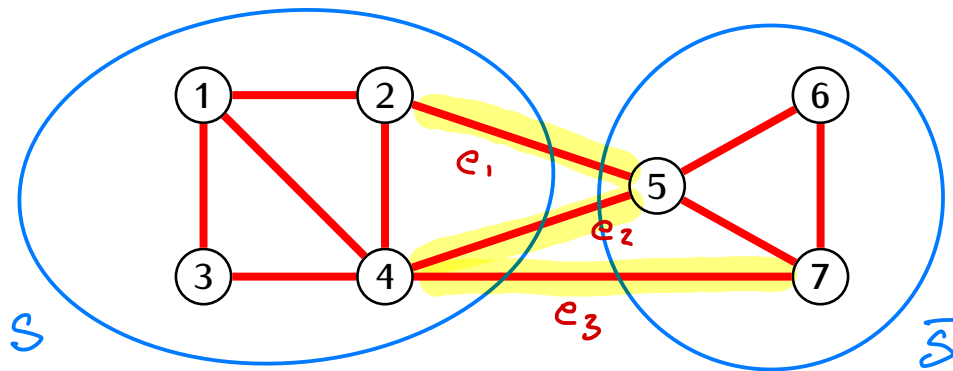
Notation. If S is a finite set then

$$|S| := (\text{the number of elements of } S)$$

Definition

Let G be a graph with the set of vertices V . Let $S \subseteq V$ and let $\bar{S} = V \setminus S$. Then

$$E(S, \bar{S}) = \left(\begin{array}{l} \text{the set of edges of } G \\ \text{with one end in } S \\ \text{and the other end is } \bar{S} \end{array} \right)$$



$$E(S, \bar{S}) = \{e_1, e_2, e_3\}$$

$$|E(S, \bar{S})| = 3$$

Partitioning problem. For a given connected graph with the set of vertices $V = 1, \dots, N$ and a given number $1 \leq k \leq N$ find $S \subseteq V$ such that $|S| = k$ and that $E(S, \bar{S})$ is as small as possible.

Definition

Let G be a graph with vertices $V = \{1, \dots, N\}$, and let $S \subseteq V$. The *selector vector* of S is the vector $\mathbf{v}_S \in \mathbb{R}^N$ given by

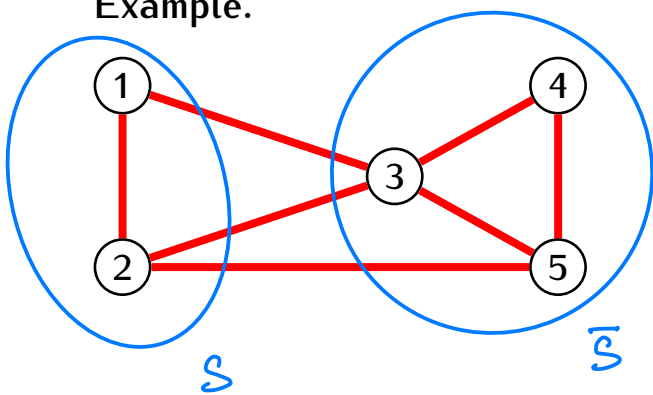
$$\mathbf{v}_S = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \quad \text{where} \quad x_i = \begin{cases} 1 & \text{if } i \in S \\ -1 & \text{if } i \in \bar{S} \end{cases}$$

Proposition

Let G be a graph with vertices $V = \{1, \dots, N\}$, and let L be the Laplacian of G . For $S \subseteq V$ we have:

$$|E(S, \bar{S})| = \frac{1}{4} \cdot \mathbf{v}_S^T L \mathbf{v}_S$$

Example.



$$|E(S, \bar{S})| = 3$$

$$\mathbf{v}_S = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix}$$

$$\mathbf{v}_S^T L \mathbf{v}_S = [1 \ 1 \ -1 \ -1 \ -1] \cdot \begin{matrix} 1 & 2 & 3 & 4 & 5 \\ \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & -1 & 3 \end{bmatrix} \end{matrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}$$

$$= [1 \ 1 \ -1 \ -1 \ -1] \cdot \begin{bmatrix} 2 \\ 4 \\ -4 \\ 0 \\ -2 \end{bmatrix} = 2 + 4 + 4 + 0 + 2 = 12 = 4 \cdot |E(S, \bar{S})|$$

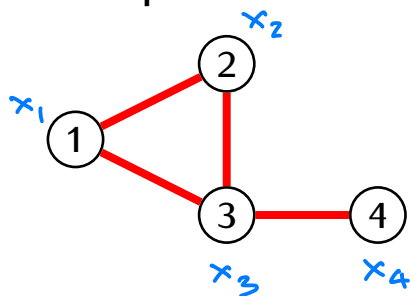
Notation. If i, j are vertices in a graph then we will write $i \sim j$ if there is an edge joining i and j .

Lemma

Let G be a graph with vertices $V = \{1, \dots, N\}$, and let L be the Laplacian of G . For any vector $\mathbf{v} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$ we have

$$\mathbf{v}^T L \mathbf{v} = \sum_{\substack{i < j \\ i \sim j}} (x_i - x_j)^2$$

Example.



$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$


$$\mathbf{v}^T L \mathbf{v} = (x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^2$$

Proof of Lemma. $L = \overset{\substack{\text{degree} \\ \text{matrix}}}{D} - \overset{\substack{\text{adjacency} \\ \text{matrix}}}{A}$ $v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$

$$v^T L v = v^T (D - A) v = v^T D v - v^T A v$$

$$v^T D v = \sum_i \deg(i) x_i^2 = \sum_{\substack{i < j \\ i \sim j}} x_i^2 + x_j^2$$

check



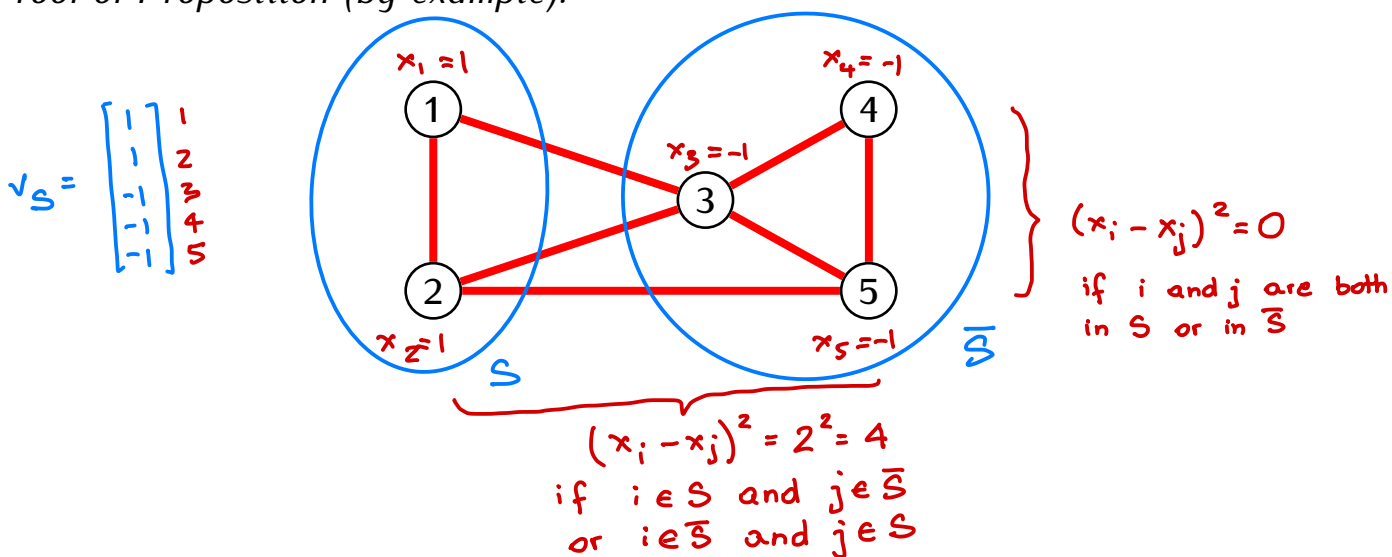
$$v^T A v = \sum_{\substack{i < j \\ i \sim j}} x_i x_j = \sum_{\substack{i < j \\ i \sim j}} x_i x_j + \sum_{\substack{j < i \\ i \sim j}} x_i x_j = 2 \sum_{\substack{i < j \\ i \sim j}} x_i x_j$$

check

This gives:

$$v^T L v = \sum_{\substack{i < j \\ i \sim j}} (x_i^2 + x_j^2 - 2x_i x_j) = \sum_{\substack{i < j \\ i \sim j}} (x_i - x_j)^2$$

Proof of Proposition (by example).



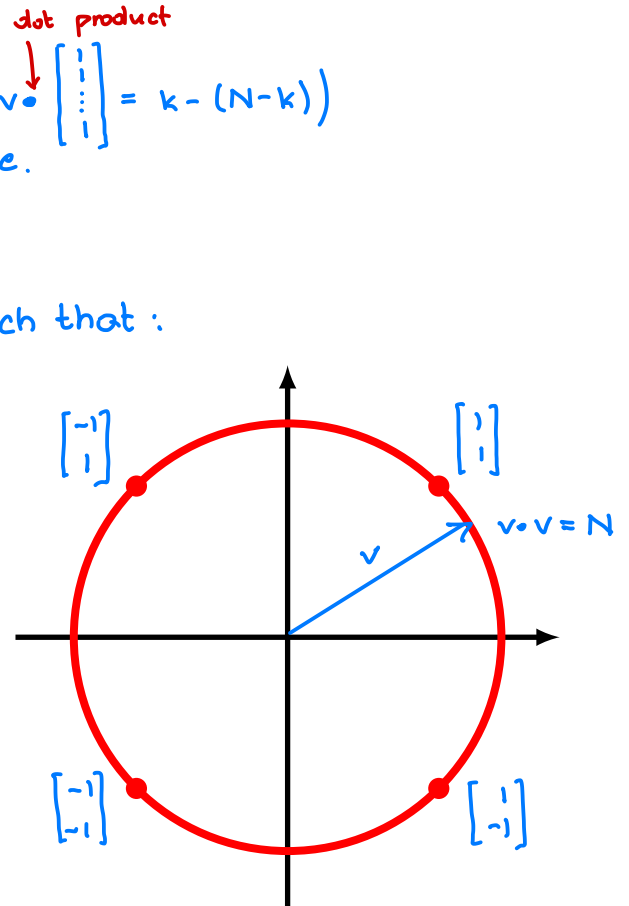
By Lemma: $v_S^T \cdot L \cdot v_S = \sum_{\substack{i < j \\ i \sim j}} (x_i - x_j)^2 = \sum_{\substack{i < j \\ i \sim j \\ i, j \text{ are in different groups}}} 4 = 4 \cdot |E(S, \bar{S})|$

So: $|E(S, \bar{S})| = \frac{1}{4} v_S^T L v_S$

Partitioning problem restated:

Given a connected graph with vertices $\{1, 2, \dots, N\}$ and Laplacian L
find a vector $v = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$ such that:

- hard \rightarrow [(1) $x_i = \pm 1$ for $i=1, 2, \dots, N$
(2) $\sum_i x_i = k - (N-k)$ (equivalently: $v \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = k - (N-k)$)
(3) $v^T L v$ is the smallest possible.



Relaxation:

Find a vector $v = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$ such that:

- (1') $v \cdot v = N$
(2) $v \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = k - (N-k)$
(3) $v^T L v$ is the smallest possible.

Note:

Let v_P = a solution of the partitioning problem
 v_R = a solution of the relaxed problem

Then

1) $v_P^T L v_P \geq v_R^T L v_R$

2) we can use v_R to get an approximated solution of the partitioning problem

Preparation: Eigenvectors of the Laplacian of a graph

Let G be a connected graph with N vertices and L be the Laplacian of G .

1) Since L is a symmetric matrix, it has N orthonormal eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_N$.

orthonormal:

$$\mathbf{u}_i \cdot \mathbf{u}_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

Let

$$\begin{aligned} \lambda_1 &= \text{eigenvalue corresponding to } \mathbf{u}_1 \\ \lambda_2 &= \text{eigenvalue corresponding to } \mathbf{u}_2 \\ \dots & \dots \dots \dots \dots \dots \dots \\ \lambda_N &= \text{eigenvalue corresponding to } \mathbf{u}_N \end{aligned}$$

We can assume that $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$.

2) $\lambda_i \geq 0$ for $i = 1, \dots, N$ (since L can be written in the form BB^T for some matrix B).

3) Since G connected, we have $\lambda_1 = 0$ and $\lambda_i > 0$ for $i = 2, \dots, N$.

4) We can take

$$\mathbf{u}_1 = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Solution of the relaxed problem

Let $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$ - eigenvalues of L

$\frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = u_1 \quad u_2 \quad \dots \quad u_N$ - corresponding orthonormal eigenvectors

Take $v \in \mathbb{R}^N$ that satisfies the conditions (1'), (2), (3)

Since $\{u_1, \dots, u_N\}$ is a basis of \mathbb{R}^N we have:

$$v = \sum_i c_i u_i \quad \text{for some } c_i \in \mathbb{R}$$

Condition (1') gives:

$$N = v \cdot v = \left(\sum_i c_i u_i \right) \left(\sum_j c_j u_j \right) = \sum_{ij} c_i c_j (u_i \cdot u_j) \overset{\substack{\text{by orthonormality} \\ \text{of } \{u_1, \dots, u_N\}}}{=} \sum_i c_i^2$$

Condition (2) gives:

$$k - (N - k) = v \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \left(\sum_i c_i u_i \right) \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \sum_i c_i \left(u_i \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right) \overset{\substack{\text{since } u_i \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = 0 \text{ for } i > 1}}{=} c_1 \cdot \frac{N}{\sqrt{N}} = c_1 \sqrt{N}$$

$$\text{Thus: } c_1 = \frac{k - (N - k)}{\sqrt{N}}$$

Condition (3):

$$\begin{aligned} v^T L v &= \left(\sum_i c_i u_i \right)^T L \left(\sum_j c_j u_j \right) = \left(\sum_i c_i u_i \right)^T \left(\sum_j c_j L u_j \right) \overset{\substack{u_j \text{ is an eigenvector of } L \\ \text{corresp. to } \lambda_j}}{=} \left(\sum_i c_i u_i \right)^T \left(\sum_j c_j \lambda_j u_j \right) \\ &= \sum_{ij} c_i c_j \lambda_j (u_i \cdot u_j) \overset{\substack{\text{by orthonormality} \\ \text{of } \{u_1, \dots, u_N\}}}{=} \sum_i c_i^2 \lambda_i = c_1^2 \cdot 0 + (c_2^2 \cdot \lambda_2 + c_3^2 \cdot \lambda_3 + \dots + c_N^2 \cdot \lambda_N) \\ &\overset{\substack{\lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_N}}{\geq} c_1^2 \cdot 0 + (c_2^2 \cdot \lambda_2 + c_3^2 \cdot \lambda_2 + \dots + c_N^2 \cdot \lambda_2) \\ &= c_1^2 \cdot 0 + (c_2^2 + c_3^2 + \dots + c_N^2) \lambda_2 = w^T L w \\ &\quad \text{for } w = c_1 u_1 + d u_2 \\ &\quad \text{where } d = \sqrt{c_2^2 + \dots + c_N^2} \end{aligned}$$

solution of the relaxed problem continued...

Upshot: To get a vector $v \in \mathbb{R}^N$ that satisfies (1'), (2), (3) we need to take :

$$v = cu_1 + du_2$$

where :

$$c = \frac{k - (N-k)}{\sqrt{N}}, \quad c^2 + d^2 = N \\ (\text{or } d^2 = N - c^2)$$

This gives:

$$d^2 = N - c^2 = N - \frac{(k - (N-k))^2}{N} = \frac{4k(N-k)}{N}$$

$$d = \pm \sqrt{\frac{4k(N-k)}{N}} \quad \uparrow \text{check}$$

We obtain:

1) The solution of the relaxed partitioning problem is given by the vector

$$v_R = \underbrace{\frac{k - (N-k)}{\sqrt{N}}}_c \cdot \underbrace{\frac{1}{\sqrt{N}} \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}}_{w_1} \pm \underbrace{\sqrt{\frac{4k(N-k)}{N}}}_d \cdot u_2$$

2) For this vector we have :

$$v_R^T L v_R = c \cdot 0 + d^2 \cdot \lambda_2 = \frac{4k(N-k)}{N} \cdot \lambda_2$$

Theorem

Let G be a graph with N vertices, and let λ_2 be the second smallest eigenvalue of the Laplacian of G . Then for any set S of vertices of G we have

$$|E(S, \bar{S})| \geq \frac{|S| \cdot |\bar{S}|}{N} \cdot \lambda_2$$

Definition

Let G be a graph. The second smallest eigenvalue λ_2 of the Laplacian of G is called the *algebraic connectivity* of G .

Back to the partitioning problem

Recall: Given a connected graph with the set of vertices $V = \{1, 2, \dots, N\}$ and $0 < k < N$ we want to find $S \subseteq V$ such that $|S| = k$ and $|E(S, \bar{S})|$ is as small as possible (equivalently: $\mathbf{v}_S^T L \mathbf{v}_S$ is as small as possible).

Approximated solution:

The spectral partitioning algorithm

Recall: Given a connected graph with the set of vertices $V = \{1, 2, \dots, N\}$ and $0 < k < N$ we want to find $S \subseteq V$ such that $|E(S, \bar{S})|$ is as small as possible.

Approximated solution:

1. Compute the Laplacian L of the graph.
2. Compute the eigenvector of L corresponding to the second smallest eigenvalue λ_2 :

$$\mathbf{u}_2 = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

3. Let

$$S_+ = \{i_1, \dots, i_k\} \subseteq V$$

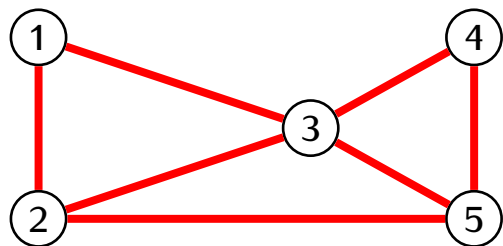
$$S_- = \{j_1, \dots, j_k\} \subseteq V$$

such that

- x_{i_1}, \dots, x_{i_k} are the largest entries of \mathbf{u}_2
- x_{j_1}, \dots, x_{j_k} are the smallest entries of \mathbf{u}_2 .

If $x_{i_1} + \dots + x_{i_k} \geq -(x_{j_1} + \dots + x_{j_k})$ take $S = S_+$. Otherwise take $S = S_-$.

Example.



$$L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & -1 & 3 \end{bmatrix}$$

Definition

Let G be a graph with the set of vertices V . The *Cheeger constant* of G is the number

$$h(G) = \min \left\{ \frac{|E(S, \bar{S})|}{|S|} \mid S \subseteq V, 1 \leq |S| \leq \frac{|V|}{2} \right\}$$

Corollary

If λ_2 is the algebraic connectivity of a graph G then

$$h(G) \geq \frac{1}{2}\lambda_2$$

Theorem (Cheeger inequality)

If λ_2 is the algebraic connectivity of a graph G then

$$\sqrt{2\lambda_2 d_{\max}} \geq h(G)$$

where d_{\max} is the maximal degree of a vertex of G .