

Last modified 3/9/2016

Lecture 9: The Sparsest Cut Problem

In which we introduce the sparsest cut problem and the Leighton-Rao relaxation.

1 The Uniform Sparsest Cut problem, Edge Expansion and λ_2

Let $G = (V, E)$ be an undirected graph with $n := |V|$ vertices.

We define the uniform sparsity of a cut $(S, V - S)$ as

$$\text{usc}_G(S) := \frac{E(S, V - S)}{|S| \cdot |V - S|}$$

(we will omit the subscript when clear from the context) and the uniform sparsest cut of a graph is

$$\text{usc}(G) := \min_S \text{usc}_G(S)$$

In d -regular graphs, approximating the uniform sparsest cut is equivalent (up to a factor of 2 in the approximation) to approximating the edge expansion, because, for every cut $(S, V - S)$, we have

$$\phi(S, V - S) = \frac{E(S, V - S)}{d \cdot \min\{|S|, |V - S|\}}$$

and, noting that, for every S ,

$$\frac{1}{n}|S| \cdot |V - S| \leq \min\{|S|, |V - S|\} \leq \frac{2}{n}|S| \cdot |V - S|$$

we have, for every S ,

$$\phi(S, V - S) \leq \frac{n}{d} \cdot \text{usc}(S) \leq 2\phi(S, V - S)$$

and so

$$\phi(G) \leq \frac{n}{d} \cdot \text{usc}(G) \leq 2\phi(G)$$

It will be instructive to see that, in d -regular graphs, λ_2 is a relaxation of $\frac{n}{d}\text{usc}(G)$, a fact that gives an alternative proof of the easy direction $\lambda_2 \leq 2\phi(G)$ of Cheeger's inequalities.

If G is d -regular, then λ_2 satisfies

$$\begin{aligned} \lambda_2 &= \min_{\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}, \mathbf{x} \perp \mathbf{1}} \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_v dx_v^2} \\ &= \min_{\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}, \mathbf{x} \perp \mathbf{1}} \frac{n}{d} \cdot \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_{\{u,v\}} (x_u - x_v)^2} \\ &= \min_{\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}} \frac{n}{d} \cdot \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_{\{u,v\}} (x_u - x_v)^2} \end{aligned}$$

where the first identity above comes from the fact that

$$\begin{aligned} \sum_{\{u,v\}} (x_u - x_v)^2 &= \frac{1}{2} \sum_{(u,v) \in V^2} (x_u - x_v)^2 = n \sum_v x_v^2 - \sum_{u,v} x_u x_v \\ &= n \sum_v x_v^2 - \left(\sum_v x_v \right)^2 \\ &= n \sum_v x_v^2 - \langle \mathbf{x}, \mathbf{1} \rangle^2 \\ &= n \sum_v x_v^2 \end{aligned}$$

and the second identity follows by noticing that the cost function is invariant by addition of a multiple of $\mathbf{1}$, and so optimizing over all non-zero vectors gives the same result as optimizing over all vectors orthogonal to $\mathbf{1}$.

On the other hand, the uniform sparsest cut problem can be formulated as

$$\text{usc}(G) = \min_{\mathbf{x} \in \{0,1\}^n - \{\mathbf{0},\mathbf{1}\}} \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_{\{u,v\}} (x_u - x_v)^2}$$

(because the square of a number in $\{-1, 0, 1\}$ is the same as its absolute value) and we see that λ_2 can be considered a continuous relaxation of $\frac{n}{d}\text{usc}(G)$.

2 The Non-Uniform Sparsest Cut Problem

In the non-uniform sparsest cut problem, we are given two graphs $G = (V, E_G)$ and $H = (V, E_H)$, over the same set of vertices; the non-uniform sparsity of a cut $(S, V - S)$ is defined as

$$\text{sc}_{G,H}(S) := \frac{E_G(S, V - S)}{E_H(S, V - S)}$$

and the non-uniform sparsest cut problem is the optimization problem

$$\text{sc}(G, H) := \min_S \text{sc}_{G,H}(S)$$

Note that the non-uniform sparsest cut problem generalizes the sparsest cut problem (consider the case in which H is a clique).

If H is the graph that contains the single edge $\{s, t\}$, then $\text{sc}(G, H)$ is the undirected min-st-cut problem, in which we want to find the cut that separates two vertices s and t and that minimizes the number of crossing edges.

3 The Leighton-Rao Relaxation

We can write the non-uniform sparsity of a set as

$$\text{sc}_{G,H}(S) = \frac{\sum_{\{u,v\} \in E_G} |1_S(u) - 1_S(v)|}{\sum_{\{u,v\} \in E_H} |1_S(u) - 1_S(v)|}$$

The observation that led us to see λ_2 as the optimum of a continuous relaxation of $\frac{n}{d} \text{sc}_{G, K_n}$ was to observe that $|1_S(u) - 1_S(v)| = |1_S(u) - 1_S(v)|^2$, and then relax the problem by allowing arbitrary functions $x : V \rightarrow \mathbb{R}$ instead of indicator functions $1_S : V \rightarrow \{0, 1\}$.

The Leighton-Rao relaxation of sparsest cut is obtained using, instead, the following observation: if, for a set S , we define $d_S(u, v) := |1_S(u) - 1_S(v)|$, then $d_S(\cdot, \cdot)$ defines a semi-metric over the set V , because d_S is symmetric, $d_S(v, v) = 0$, and the triangle inequality holds. So we could think about allowing *arbitrary semi-metrics* in the expression for sc , and define

$$LR(G, H) := \min_{\substack{d : V \times V \rightarrow \mathbb{R} \\ d \text{ semi-metric}}} \frac{\sum_{\{u,v\} \in E_G} d(u, v)}{\sum_{\{u,v\} \in E_H} d(u, v)} \quad (1)$$

This might seem like such a broad relaxation that there could be graphs on which $LR(G, H)$ bears no connection to $\text{sc}(G, H)$. Instead, we will prove the fairly good estimate

$$LR(G, H) \leq \text{sc}(G, H) \leq O(\log |V|) \cdot LR(G, H) \quad (2)$$

The value $LR(G, H)$ and an optimal $d(\cdot, \cdot)$ can be computed in polynomial time by solving the following linear program

$$\begin{aligned} & \text{minimize} && \sum_{\{u,v\} \in E_G} d_{u,v} \\ & \text{subject to} && \sum_{\{u,v\} \in E_H} d_{u,v} = 1 \\ & && d_{u,v} \leq d_{u,z} + d_{z,v} \quad \forall u, v, z \in V \\ & && d_{u,v} \geq 0 \quad \forall u, v \in V \end{aligned} \quad (3)$$

that has a variable $d_{u,v}$ for every unordered pair of distinct vertices $\{u, v\}$. Clearly, every solution to the linear program (3) is also a solution to the right-hand side of the definition (1) of the Leighton-Rao parameter, with the same cost. Also every semi-metric can be normalized so that $\sum_{\{u,v\} \in E_H} d(u, v) = 1$ by multiplying every distance by a fixed constant, and the normalization does not change the value of the right-hand side of (1); after the normalization, the semimetric is a feasible solution to the linear program (3), with the same cost.

4 An L1 Relaxation of Sparsest Cut

In the Leighton-Rao relaxation, we relax distance functions of the form $d(i, j) = |1_S(i) - 1_S(j)|$ to completely arbitrary distance functions. Let us consider an intermediate relaxation, in which we allow distance functions that can be realized by *an embedding of the vertices in an ℓ_1 space*.

Recall that, for a vector $\mathbf{x} \in \mathbb{R}^n$, its ℓ_1 norm is defined as $\|\mathbf{x}\|_1 := \sum_i |x_i|$, and that this norm makes \mathbb{R}^n into a metric space with the ℓ_1 distance function

$$\|\mathbf{x} - \mathbf{y}\|_1 = \sum_i |x_i - y_i|$$

The distance function $d(u, v) = |1_S(u) - 1_S(v)|$ is an example of a distance function that can be realized by mapping each vertex to a real vector, and then defining the distance between two vertices as the ℓ_1 norm of the respective vectors. Of course it is an extremely restrictive special case, in which the dimension of the vectors is one, and in which every vertex is actually mapping to either zero or one. Let us consider the relaxation of sparsest cut to arbitrary ℓ_1 mappings, and define

$$\text{L1sc}(G, H) := \inf_{m, f: V \rightarrow \mathbb{R}^m} \frac{\sum_{\{u, v\} \in E_G} \|f(u) - f(v)\|_1}{\sum_{\{u, v\} \in E_H} \|f(u) - f(v)\|_1}$$

This may seem like another very broad relaxation of sparsest cut, whose optimum might be much smaller than the sparsest cut optimum. The following theorem shows that this is not the case.

Theorem 1 *For every graphs G, H , $\text{sc}(G, H) = \text{L1sc}(G, H)$.*

Furthermore, there is a polynomial time algorithm that, given a mapping $f : V \rightarrow \mathbb{R}^m$, finds a cut S such that

$$\frac{\sum_{\{u, v\} \in E_G} |1_S(u) - 1_S(v)|}{\sum_{\{u, v\} \in E_H} |1_S(u) - 1_S(v)|} \leq \frac{\sum_{\{u, v\} \in E_G} \|f(u) - f(v)\|_1}{\sum_{\{u, v\} \in E_H} \|f(u) - f(v)\|_1} \quad (4)$$

PROOF: We use ideas that have already come up in the proof the difficult direction of Cheeger's inequality. First, recall that for every nonnegative reals a_1, \dots, a_m and positive reals b_1, \dots, b_m we have

$$\frac{a_1 + \dots + a_m}{b_1 + \dots + b_m} \geq \min_i \frac{a_i}{b_i} \quad (5)$$

as can be seen by noting that

$$\sum_j a_j = \sum_j b_j \cdot \frac{a_j}{b_j} \geq \left(\min_i \frac{a_i}{b_i} \right) \cdot \sum_j b_j$$

Let $f_i(v)$ be the i -th coordinate of the vector $f(v)$, thus $f(v) = (f_1(v), \dots, f_m(v))$. Then we can decompose the right-hand side of (4) by coordinates, and write

$$\begin{aligned} & \frac{\sum_{\{u, v\} \in E_G} \|f(u) - f(v)\|_1}{\sum_{\{u, v\} \in E_H} \|f(u) - f(v)\|_1} \\ &= \frac{\sum_i \sum_{\{u, v\} \in E_G} |f_i(u) - f_i(v)|}{\sum_i \sum_{\{u, v\} \in E_H} |f_i(u) - f_i(v)|} \\ &\geq \min_i \frac{\sum_{\{u, v\} \in E_G} |f_i(u) - f_i(v)|}{\sum_{\{u, v\} \in E_H} |f_i(u) - f_i(v)|} \end{aligned}$$

This already shows that, in the definition of ϕ' , we can map, with no loss of generality, to 1-dimensional ℓ_1 spaces.

Let i^* be the coordinate that achieves the minimum above. Because the cost function is invariant under the shifts and scalings (that is, the cost of a function $x \rightarrow f(x)$ is the same as the cost of $x \rightarrow af(x) + b$ for every two constants $a \neq 0$ and b) there is a function $g : V \rightarrow \mathbb{R}$ such that g has the same cost function as f_{i^*} and it has a unit-length range $\max_v g(v) - \min_v g(v) = 1$.

Let us now pick a threshold t uniformly at random from the interval $[\min_v g(v), \max_v g(v)]$, and define the random variables

$$S_t := \{v : g(v) \leq t\}$$

We observe that for every pairs of vertices u, v we have

$$\mathbb{E} |1_{S_t}(u) - 1_{S_t}(v)| = |g(u) - g(v)|$$

and so we get

$$\begin{aligned} & \frac{\sum_{\{u,v\} \in E_G} \|f(u) - f(v)\|_1}{\sum_{\{u,v\} \in E_H} \|f(u) - f(v)\|_1} \\ & \geq \frac{\sum_{\{u,v\} \in E_G} |g(u) - g(v)|}{\sum_{\{u,v\} \in E_H} |g(u) - g(v)|} \\ & = \frac{\mathbb{E} \sum_{\{u,v\} \in E_G} |1_{S_t}(u) - 1_{S_t}(v)|}{\mathbb{E} \sum_{\{u,v\} \in E_H} |1_{S_t}(u) - 1_{S_t}(v)|} \end{aligned}$$

Finally, by an application of (5), we see that there must be a set S among the possible values of S_t such that (4) holds.

Notice that the proof was completely constructive: we simply took the coordinate f_{i^*} of f with the lowest cost function, and then the “threshold cut” given by f_{i^*} with the smallest sparsity. \square

5 A Theorem of Bourgain

We will derive our main result (2) from the L1 “rounding” process of the previous section, and from the following theorem of Bourgain (the efficiency considerations are due to Linial, London and Rabinovich).

Theorem 2 (Bourgain) *Let $d : V \times V \rightarrow \mathbb{R}$ be a semimetric defined over a finite set V . Then there exists a mapping $f : V \rightarrow \mathbb{R}^m$ such that, for every two elements $u, v \in V$,*

$$\|f(u) - f(v)\|_1 \leq d(u, v) \leq \|f(u) - f(v)\|_1 \cdot c \cdot \log |V|$$

where c is an absolute constant. Given d , the mapping f can be found with high probability in randomized polynomial time in $|V|$.

To see that the above theorem of Bourgain implies (2), consider a graph G , and let d be the optimal solution of the Leighton-Rao relaxation of the sparsest cut problem on G , and let $f : V \rightarrow \mathbb{R}$ be a mapping as in Bourgain's theorem applied to d . Then

$$\begin{aligned} LR(G, H) &= \frac{\sum_{\{u,v\} \in E_G} d(u, v)}{\sum_{\{u,v\} \in E_H} d(u, v)} \\ &\geq \frac{\sum_{\{u,v\} \in E_G} \|f(u) - f(v)\|_1}{c \cdot \log |V| \cdot \sum_{\{u,v\} \in E_H} \|f(u) - f(v)\|_1} \\ &\geq \frac{1}{c \cdot \log |V|} \cdot \text{sc}(G, H) \end{aligned}$$