Lecture 8

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

Let $G = (V, E)$ be an undirected graph. Unlike past lectures, we will not need to assume that G is regular. We are interested in finding a *sparsest cut* in G, where the sparsity of a non-trivial bipartition $(S, V - S)$ of the vertices is

$$\phi_G(S) := \frac{\frac{1}{|E|} \cdot Edges(S, V - S)}{\frac{2}{V^2} \cdot |S| \cdot |V - S|}$$

which is the ratio between the fraction of edges that are cut by $(S, V - S)$ and the fraction of pairs of vertices that are disconnected by the removal of those edges.

Another way to write the sparsity of a cut is as

$$\phi_G(S) := \frac{|V|^2}{2|E|} \cdot \frac{\sum_{i,j} A_{i,j} |1_S(i) - 1_S(j)|}{\sum_{i,j} |1_S(i) - 1_S(j)|}$$

where A is the adjacency matrix of G and $1_S(\cdot)$ is the indicator function of the set S.

The observation that led us to see $1 - \lambda_2$ as the optimum of a continuous relaxation of ϕ was to observe that $|1_S(i) - 1_S(j)| = |1_S(i) - 1_S(j)|^2$, and then relax the problem by allowing arbitrary functions $x : V \to \mathbb{R}$ instead of indicator functions $1_S : V \to \{0, 1\}$.

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

$$LR(G) := \min_{\substack{d : V \times V \to \mathbb{R} \\ d \text{ semi-metric}}} \frac{|V|^2}{2|E|} \cdot \frac{\sum_{i,j} A_{i,j} d(i, j)}{\sum_{i,j} d(i, j)} \qquad (1)$$

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

$$LR(G) \leq \phi(G) \leq O(\log |V|) \cdot LR(G) \tag{2}$$

Furthermore, we will show that $LR(G)$, and an optimal solution $d(\cdot, \cdot)$ can be computed in polynomial time, and the second inequality above has a constructive proof, from which we derive a polynomial time $O(\log |V|)$-approximate algorithm for sparsest cut.

1 Formulating the Leighton-Rao Relaxation as a Linear Program

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

$$\begin{array}{lll} \text{minimize} & \sum_{i,j} A_{i,j} d_{i,j} & \\ \text{subject to} & & \\ & \sum_{i,j} d_{i,j} = \frac{|V|^2}{2|E|} & \\ & d_{i,k} \leq d_{i,j} + d_{j,k} & \forall i,j,k \in V \\ & d_{i,j} \geq 0 & \forall i \in V \end{array} \tag{3}$$

that has a variable $d_{i,j}$ for every unordered pair of distinct vertices i, j. 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 $\sum_{i,j} d(i,j) = |V|^2/2|E|$ 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.

In the rest of this lecture and the next, we will show how to round a solution to (3) into a cut, achieving the logarithmic approximation promised in (2).

2 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 $d(i,j) = |1_S(i) - 1_S(j)|$ 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

$$\phi'(G) := \inf_{m, f: V \to \mathbb{R}^m} \frac{|V|^2}{2|E|} \cdot \frac{\sum_{i,j} A_{i,j} ||f(i) - f(j)||_1}{\sum_{i,j} ||f(i) - f(j)||_1}$$

This may seem like another very broad relaxation of sparsest cut, whose optimum might bear no correlation with the sparsest cut optimum. The following theorem shows that this is not the case.

Theorem 1 *For every graph G, $\phi(G) = \phi'(G)$.*

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

$$\frac{\sum_{u,v} A_{u,v} |1_S(u) - 1_S(v)|}{\sum_{u,v} |1_S(u) - 1_S(v)|} \leq \frac{\sum_{u,v} A_{u,v} ||f(u) - f(v)||_1}{\sum_{u,v} ||f(u) - f(v)||_1} \tag{4}$$

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

$$\frac{a_1 + \cdots a_m}{b_1 + \cdots b_m} \geq \min_i \frac{a_i}{b_i} \tag{5}$$

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 $f(v) = (f_1(v), \ldots, f_m(v))$. Then we can decompose the right-hand side of (4) by coordinates, and write

$$\frac{\sum_{u,v} A_{u,v}||f(u) - f(v)||_1}{\sum_{u,v} ||f(u) - f(v)||_1}$$

$$= \frac{\sum_i \sum_{u,v} A_{u,v}|f_i(u) - f_i(v)|}{\sum_i \sum_{u,v} |f_i(u) - f_i(v)|}$$

$$\geq \min_i \frac{\sum_{u,v} A_{u,v}|f_i(u) - f_i(v)|}{\sum_{u,v} |f_i(u) - f_i(v)|}$$

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 $x \to f(x)$ is the same as the cost of $x \to af(x) + b$ for every two constants $a \neq 0$ and b) there is a function $g : V \to \mathbb{R}$ 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

$$\frac{\sum_{u,v} A_{u,v}||f(u) - f(v)||_1}{\sum_{u,v} ||f(u) - f(v)||_1}$$

$$\geq \frac{\sum_{u,v} A_{u,v}|g(u) - g(v)|}{\sum_{u,v} |g(u) - g(v)|}$$

$$= \frac{\mathbb{E} \sum_{u,v} A_{u,v}|1_{S_t}(u) - 1_{S_t}(v)|}{\mathbb{E} \sum_{u,v} |1_{S_t}(u) - 1_{S_t}(v)|}$$

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. □

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

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

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 $f : V \to \mathbb{R}$ be a mapping as in Bourgain's theorem applied to d. Then

$$LR(G) = \frac{|V|^2}{|E|} \cdot \frac{\sum_{u,v} A_{u,v} d(u, v)}{\sum_{u,v} d(u, v)}$$

$$\geq \frac{|V|^2}{|E|} \cdot \frac{\sum_{u,v} A_{u,v} ||f(u) - f(v)||_1}{c \cdot \log |V| \cdot \sum_{u,v} ||f(u) - f(v)||_1}$$

$$\geq \frac{1}{c \cdot \log |U|} \cdot \phi(G)$$