CS369M: Algorithms for Modern Massive Data Set Analysis **Lecture 15 - 11/11/2009**

Partitioning Algorithms that Combine Spectral and Flow Methods

Lecturer: Michael Mahoney *Scribes: Kshipra Bhawalkar and Deyan Simeonov*

**Unedited notes*

Last time we looked at flow based graph partitioning. Today we will show why it works.

We are given a graph $G = (V, E)$, a cost function $c : E \to \mathbb{R}^+$ and k pairs of nodes (s_i, t_i). Let $x(e)$ indicate if an edge e is cut, $y(i)$ indicate if commodity i is cut and $\mathcal{P}_i, i = 1, 2, ..., k$ be the set of paths between s_i and t_i. Then we want:

$$\min \frac{\sum_{e\in E} c(e)x(e)}{\sum_{i=1}^{k} d(i)y()}, \text{ s.t.}$$

$$\sum_{e\in P} x(e) \geq y(i) \forall P \in \mathcal{P}_i \forall i = 1, 2, ..., k$$

$$y(i) \in \{0, 1\}, x(e) \in \{0, 1\}$$

We relax x and y to be in $[0, 1]$ and, given that for any feasible solution (x, y) and any $\alpha > 0$, $(\alpha x, \alpha y)$ is a feasible solution, we get the following LP:

$$\min \sum_{e\in E} c(e)x(e), \text{ s.t.}$$

$$\sum_{i=1}^{k} d(i)y(i) = 1$$

$$\sum_{e\in P} x(e) \geq y(i) \forall P \in \mathcal{P}_i \forall i = 1, 2, ..., k$$

$$x(e) \geq 0, y(i) \geq 0$$

Our strategy:

- solve the LP
- round the solution.

Recall:
$x \in \mathbb{R}^n, ||x||_p = (\sum_{i=1}^{n} x_i^p)^{\frac{1}{p}}$ - p-norm. Its induced metric is $||x - y||_p$.

Theorem (Bourgain 85)**.** *Any n-point metric space admits an α-distortion embedding in l_p with $\alpha = O(\log n)$.*

Proof idea: Use as coordinates the minimal distances of points to suitably chosen random sets.

Comparison between l_1 and l_2:

- l_2 has good dimension reduction properties, l_1 doesn't.
- There is a "'fast"'(polytime, $O(n^3)$) algorithm to get the best l_2 embedding, for l_1 it is NP.

- l_2 has strong connection to diffusion, low dimensional spaces and manifolds, l_1 has strong connection to cuts, partition in graphs, multicommodity flows.

Connection between l_1 and Cut metrics:

- There exists a representation of l_1 metrics in terms of combination of cut metrics.
- Cut metrics - Extreme rays of the cone of l_1 metrics.
- minimum ratio function over cone $\iff$ minimum over extreme rays.

Definition. Given a graph $G(V, E)$ and $S \subseteq V$, say δ_S is the cut-metric for S, if $\delta_s(x, y)$ is the indicator of x and y being on different sides of S.

Claim. *The set of l_1 metrics is a convex cone, i.e.:*
If $d_1, d_2 \in l_1$, $\alpha_1, \alpha_2 \geq 0$, then $\alpha_1 d_1 + \alpha_2 d_2 \in l_1$.

Does not hold for l_2!
Proof: Line metric is an l_1 metric but $d^{(i)}(x, y) = |x_i - y_i|$ for $x_i, y_i \in \mathbb{R}^n$. If d is an $l_!$ metric then it is a line metric. We can then argue for each dimension.

Claim. *Let d be a finite l_1 metric then we can write d as*

$$d = \sum_{S \subseteq V} \alpha_S \delta_S \quad \alpha_S \in \mathbb{R}, \delta_S - \textit{cut metric}$$

i.e.

$$\begin{aligned} CUT_n &= \{d : d = \sum_{S \subseteq V} \alpha_S \delta_S, \alpha_S \geq 0\} \\ &= \{\textit{positive cone generated by this metric}\} \\ &= \{\textit{all n-points subsets of } \mathbb{R}^n \textit{ under } l_1 \textit{ metric} \end{aligned}$$

Proof: ($CUT_n \subseteq l_1$) $d = \sum_{S \subseteq V} \alpha_S \delta_S \in CUT_n$. Introduce one dimension for each pair of vertices. For a pair of points (i, j) value in the dimension (i, j) is sum of α_S for all cuts $(S, \bar{S})$ over all cuts that put i and j in different sets.

($l_1 \subseteq CUT_n$): Consider a dimension d and sort points along that dimension in increasing values. Let $v_1, v_2, ...v_k$ be the set of distinct values along that dimension. Define k-1 cut metrics $S_i = \{x : x_d \leq v_{i-1}\}$ and let $\alpha_i = v_{i+1} - v_i$. So along that dimention,

$$|x_d - y_d| = \sum_{i=1}^{k} \alpha_i \delta_{S_i}$$

Do this for each dimension. □

Usefulness - you can optimize over l_1 metrics:
Let $C \subseteq \mathbb{R}^n$ - convex cone. $f, g : \mathbb{R}^n \to \mathbb{R}^+$. Then:

$$\min_{x \in C} \frac{f(x)}{g(x)} = \min_{x \in \text{extreme ray of} C} \frac{f(x)}{g(x)}$$

Conductance and Sparsity

Given a graph G(V,E) we define the conductance h_G and sparsity ϕ_G as follows,

$$h_G := min_{S \subseteq V} \frac{E(S, \bar{S})}{min\{|S|, |\bar{S}|\}}$$
$$\phi_G := min_{S \subseteq V} \frac{E(S, \bar{S})}{\frac{1}{n}|S||\bar{S}|}$$

Claim. $\phi_G = \{min_{d \in l_1 metric} \sum_{i,j \in E} d_{ij} s.t. \sum_{i,j} d_{ij} = 1\}$

Idea: Relax to optimization over a larger set i.e. a metric.

$$\lambda^* := min \sum_{i,j \in E} d_{ij}$$
$$s.t. \sum_{i,j} d_{ij} = 1$$
$$d_{ij} \geq 0$$
$$d_{ij} = d_{ji}$$
$$d_{ij} + d_{jk} \geq d_{ik}$$

Clearly $\lambda^* \leq \phi^*$. Less obviously $\phi^* \leq O(\log n)\lambda^*$ (homework 3).

Algorithm: Given G.

- Solve the LP to get metric d,
- Use bourgain embedding result to approximate by l_1 metric (with loss O(log n).
- Round the solution to get a cut.
  - For each dimension covert the l_1 metric along that to a cut metric.
  - Choose the best.

Note: If have l_1 embedding with distortion factor ξ then can approximate the cut upto ξ. (Homework 3).

What else can you relax to?

- l_2 - not convex
- l_2^2 - convex but distortion $Omega(n)$. (Not a metric) (this happens in spectral technique.
  - Can get eigenvalue λ s.t. λ is close to h(G) upto quadratic factor. (Cheegar Inequality).
- l_2^2 + triangle inequality gives a metric
  - This works.
  - Arora, Rao Vazirani.

Various relaxations of spectral cut.
Actual problem:

$$\Phi_G = min_{S \subseteq V} \frac{|E(S, \bar{S})|}{|S||\bar{S}|} = min_{x \in 0,1^V} \frac{\sum_{i,j} A_{ij}|x_i - x_j|}{\sum_{i,j} |x_i - x_j|}$$

Spectral Method: replace l_1 with l_2.

$$d - \lambda_2 = min_{x \in \mathbb{R}^V} \frac{\sum_{i,j} A_{ij}(x_i - x_j)^2}{\sum_{i,j}(x_i - x_j)^2}$$

Leighton, Rao

$$min_{d-metric} \frac{\sum_{i,j} A_{ij} d_{ij}}{\sum_{i,j} d_{ij}}$$

Leighton,Rao approach is bad in expanders, but is good in sparse graphs.

Definition. l_2^2 representation of a graph G is an assignment of a point vector to each node $v_i \in \mathbb{R}^k$ for each i. s.t.

$$|v_i - v_j|_2^2 + |v_j - v_k|_2^2 \geq |v_i - v_k|_2^2$$

It is a unit l_2^2 representation if on unit sphere i.e. $|v_i| = 1, \forall i$.

Thing to note:

- Condition that l_2^2 distance form a metric $\iff$ all triangles are acute.
- $l_2^2 \cap metric$ is a convex cone.
- $d \in l_1 \implies d \in l_2^2$.

Relax to vectors on unit sphere that form a metric,

$$min \sum_{i,j \in E} A_{ij} \|\vec{x_i} - \vec{x_j}\|$$

$$s.t. \sum_{i,j} \|x_i - x_j\|_2^2 = 1$$

$$\|x_i - x_j\|_2^2 + \|x_j - x_k\|_2^2 \geq \|x_i - x_k\|_2^2$$

This it the semi-definite program(SDP) from Arora, Rao, Vazirani.

Theorem: For uniform sparsest cut $d_{ij} = 1 \forall i, j$ SDP integrality gap is $\theta(\sqrt{\log n})$.
For general sparsest cut, SDP integrality gap is $\theta(\sqrt{\log n} \log \log n)$.

Main structure theorem:
Let $v_, v_2, ... v_n$ be points on the unit vall in $\mathbb{R}^n$. s.t. $d_{ij} = \|v_i - v_j\|_2^2$ is a metric and all points are well-seperated. $\sum_{i,j} d_{ij}/n^2 \geq \delta = \Omega(1)$.
then $\exists S, T$ disjoint subsets of V s.t. $|S|, |T| \geq \Omega(n)$

$$min_{i \in S, j \in T} d_{ij} \geq \Omega(1/\sqrt{\log n})$$

Flow methods: Embed scaled version of complete graphs into G. - $O(\log n)$
ARV: embed an arbitrary graph H(iterative construction) such that H is a good expander or we find a cut.
Iterative construction: multiplicative update method $\sim$ online learning.

References

1. Arora, Rao, and Vazirani, CACM article, ”Geometry, flows, and graph-partitioning algorithms”
2. Khandekar, Rao, and Vazirani, ”Graph partitioning using single commodity flows”
3. Orecchia, Schulman, Vazirani, and Vishnoi, ”On Partitioning Graphs via Single Commodity Flows”