Lecture 8: Intro to Graph Cuts and Spectral Gap

Instructor: Lorenzo Orecchia   Scribe: Erasmo Tani

Introduction

Recall:

Lemma 1. $\lambda_2 = 0$ *iff G is disconnected.*

Ideally we would want to prove some result on the lines of :

Lemma 2. λ_2 *is "almost" zero iff G is "almost disconnected"*

Note that for any $v \perp \vec{1}$:

$$\lambda_2 \leq \frac{v^t L v}{v^T D v} \longrightarrow \text{"small cuts"}$$

the above operation is called rounding, since we round a fractional value to an integral one.

Definition 1 (Graph Conductance)**.** Given a graph $G = (V, E)$, and a subset $S \subseteq V$ the *conductance* of S is the quantity:

$$\phi(S) := \frac{|E(S, \overline{S})|}{\min\{vol(S), vol(\overline{S})\}}$$

the *graph conductance* of G is the the number:

$$\phi(G) := \min_{S \subseteq G} \frac{|E(S, \overline{S})|}{\min\{vol(S), vol(\overline{S})\}}$$

Intuitevely, the conductance is a sort of isoperimetric ratio. It also has an interpretation in terms of random walk. Suppose we are given distribution π over S where, for each $s \in S$ $\pi(s) \propto d_s$, what is the probability of leaving S in a step of the random walk given that the current position is chosen according to π. You could change the denominator of the conductance with any other measure on the vertices of G. This notion plays an important role in community detection and is used a lot, for instance, in image segmentation.

Minimum Conductance Problem

The minimum conductance problem is the problem of finding a minimum conductance cut in the graph.

Example: n-Cycle the conductance of an n-cycle is $\frac{1}{n}$.

Example: Complete Graph Any cut into S and $\overline{S}$ s.t.: $|S| = k$ satisfies:

$$\phi_{K_n} = \frac{k(n-k)}{(n-1)k} = \frac{n-k}{n-1} \approx \frac{1}{2}$$

In lots of real world graphs the structure exhibits a large, well connected core, and a few vertices that live outside of it and are connect to it via few edges. In these cases the conductance is not particularly informative, later in the class we will look at the problem of finding balanced cuts instead.

The minimum conductance problem in NP-Hard and the best possible approximation is $O(\sqrt{\log n})$ and the best known hardness is $\omega(1)$.

We will use the following definition of conductance istead:

$$\overline{\phi}(S) = \frac{|E(S,\overline{S}|}{Vol(S)Vol(\overline{S})} \cdot Vol(V)$$

We have:

$$\phi(S) \leq \overline{\phi}(S) \leq 2\phi(S)$$

We will show that spectral gap is a relaxation of conductance. Recall that the spectral gap can be written as:

$$\lambda_2 = \min_{x \perp \vec{1}} \frac{x^T L x}{x^T D x} = \min \frac{x^T L x}{x^T L(K_G) x} = |E_{K_G}(S,\overline{S})|$$

We have:

$$\overline{\phi}(S) = \frac{|E_G(S,\overline{S})|}{|E_{k_G}(S,\overline{S})|}$$

A cut vector is a vector $\vec{1}_S$ defined by:

$$\vec{1}_S = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases}$$

We then have:

$$\vec{1}_S^T L \vec{1}_S = |\partial(S)|$$

The previous notion of conductance is really an optimization problem over all cut-vectors:

$$\overline{\phi} = \min_{S \subseteq V} \overline{\phi}(S) = \min_{S \subseteq V} \frac{\vec{1}^T L \vec{1}_S}{\vec{1}_S^T L(K_G)\vec{1}_S} \geq \min_{x \in \mathbb{R}^n} \frac{x^T L x}{x^T L(K_G) x} = \lambda_2$$

Theorem 1 (Easy side of Cheeger's Inequality). *We have:*

$$\overline{\phi}_G \geq \lambda_2$$

This reads "If there is a small conductance cut, then the random walks must converge slowly for some distribution". We will prove the following:

Theorem 2 (Cheeger's Inequality).

$$\overline{\phi}_G \leq 2\sqrt{2\lambda_2}$$