- Lovasz-Simonovits Theorem for random walks

1 Lovasz-Simonovits Theorem for random walks

We will now study a combinatorial analysis of random walks, a beautiful theorem due to Lovasz-Simonovits that analyzes the progress of a random walk in a more local, iterative manner. The theorem has found applications in algorithms for estimating the volume of a convex body and for local clustering of large graphs.

Consider an arbitrary undirected graph G. We analyze a *lazy random walk* on G: in each step, the random walk stays at the current node with probability $1/2$, and moves to a neighbor chosen at random with probability $1/d$, where d is the degree of the current node. Alternatively, we can view the lazy random walk as the standard random walk on the directed graph $\widehat{G}$ in which each edge (u,v) in G is replaced by two directed edges (u,v) and (v,u), and we also add $d(u)$ self-loops at u, where $d(u)$ is the degree of u in G. This is the graph we will work with in the remainder of this lecture.

We will express the rate of convergence of the random walk (to the stationary distribution) in terms of the conductance ϕ of G, which is defined as

$$\phi = \min_{S \subset V} \frac{e(S, V-S)}{\min\{e(S), e(V-S)\}},$$

where $e(X,Y)$ denotes the number of edges with one endpoint in X and the other in Y, and $e(X)$ denotes the number of edges with one endpoint in X.

Instead of considering the probability on vertices, we will consider the probability on edges. Let $p_t(u)$ denote the probability that the random walk is at u at the start of step t, where we number of steps from 0. For a directed edge e, let $p_t(e)$ denote the probability that the walk goes along edge e in step t. For any edge (u,v), we have $p_t(u,v) = p_t(u)/d(u)$. Note that in stationary distribution $p_t(u,v) = 1/2m$.

The Lovasz-Simonovits Theorem analyzes the progress of the random walk, not by directly bounding the difference of the vector p_t from the stationary distribution, instead by measuring changes to a particular curve that captures the probability on the edges.

In any step, we will order the arcs such that $p_t(e_1) \geq p_t(e_2) \ldots \leq p_t(e_{em})$. We define $I : [0, 2m] \rightarrow [0,1]$ as follows. For each integer $k \in \{0, \ldots, 2m\}$, we define

$$I_t(k) = \sum_{i=1}^{k} p_t(e_i).$$

We extend the domain of I_t to the real interval $[0, 2m]$ by linear interpolation.

Theorem 1. *For any initial probability distribution and every t, we have*

$$I_t(x) \leq \min(\sqrt{x}, \sqrt{2m - x}) \left(1 - \frac{\phi^2}{2}\right)^t + \frac{x}{2m}.$$

2 Proof of the L-S theorem

The proof will proceed in three parts. First we prove some facts about I_t. Second, we show how I_t "decreases" over increasing time. Finally, we will derive the theorem.

2.1 Facts about I_t

Here are some easy observations on I_t.

- $I_t(0) = 0$, $I_t(1) = 1$, and $I_t(x)$ is monotonically nondecreasing with x.
- Since all edges going out of a vertex have the same value of p_t, we can assume that the edges are ordered so that all of the outgoing edges of any vertex are in contiguous order.
- I_t is concave in x; that is, for any $0 \leq x \leq y \leq 2m$ and any t, $I_t((x+y)/2) \geq (I_t(x)+I_t(y))/2$.
- If I_t has reached stationarity, then $I_t(x) = x/2m$ for all x.

Lemma 1. *For all x and t, $I_t(x) \leq I_{t-1}(x)$.*

Proof: Fix t and let $p_t(e_1) \geq p_t(e_2) \geq \ldots \geq p_t(e_{2m})$ denote the order of the edges. Let $e_i = (u_i, v_i)$; note that u_is (resp., v_is) are not necessarily distinct for different i. We prove the claim for all $x = k$, where $\{e_1, \ldots, e_k\}$ form the set of all edges that go out of the vertices in the set $W = \{u_i : 1 \leq i \leq k\}$. It is not too hard to extend the claim to all real x in the interval $[0, 2m]$.

$$\begin{aligned} I_t(k) &= \sum_{i=1}^{k} p_t(u_i, v_i) \\ &= \sum_{i=1}^{k} p_{t-1}(v_i, u_i) \\ &\leq I_{t-1}(k). \end{aligned}$$

□

2.2 Change in I_t

Lemma 1 shows that the curve I_t is monotonically nonincreasing with t; so at the very least, we are not moving away from stationarity. In order to place any bounds on the convergence time, however, we need to see how fast does I_t get closer to the limit.

Lemma 2. *For $x \in [0, m]$, we have*

$$I_t(x) \leq \frac{1}{2}\left(I_{t-1}(x - 2\phi x) + I_{t-1}(x + 2\phi x)\right).$$

For $x \in [m, 2m]$, we have

$$I_t(x) \leq \frac{1}{2}\left(I_{t-1}(x - 2\phi(2m - x)) + I_{t-1}(x + 2\phi(2m - x))\right).$$

Proof: Consider $x \in [0, m]$. We assume that x equals an integer k such that $\{e_1, \ldots, e_k\}$ is the set of *all* outgoing edges from the vertex set $S = \{u_i : 1 \leq i \leq k\}$ where $e_i = (u_i, v_i)$. By the definition of the random walk, we have

$$I_t(k) = \sum_{i=1} k p_t(u_i, v_i) = \sum_{i=1} k - 1 p_{t-1}(v_i, u_i).$$

Let W denote the set $\{(v_i, u_i) : 1 \leq i \leq k\}$. We divide the set W into three groups.

$$\begin{aligned} W_1 &= \{(v_i, u_i) : u_i, v_i \in S, u_i \neq v_i\} \\ W_2 &= \{(v_i, u_i) : u_i \in S, v_i \notin S\} \\ W_3 &= \{(u_i, u_i) : u_i \in S\} \end{aligned}$$

We now show the following two inequalities.

$$\sum_{e \in W_1} p_{t-1}(e) \leq \frac{1}{2} I_{t-1}(k - 2|W_2|) \tag{1}$$

$$\sum_{e \in W_2 \cup W_3} p_{t-1}(e) \leq \frac{1}{2} I_{t-1}(x + 2|W_2|) \tag{2}$$

For a set X of edges, let $p_t(X)$ denote sum of $p_t(e)$ over all edges $e \in X$. Since all edges in W_1 are non-loop outgoing edges for nodes in S, and there are as many self-loops at a node as edges leaving the node, we obtain that $p_{t-1}(W_1)$ is at most $I_{t-1}(2|W_1|)/2$. Since $|W_3| \geq |W_1| + |W_2|$, $2|W_1| \leq k - 2|W_2|$ yielding us Equation 1.

We now consider $p_{t-1}(W_2 \cup W_3)$. Since the number of self-loop edges at a node equals the number of outgoing edges and all outgoing edges at a node carry the same mass, it follows that the mass of edges in W_3 is identical to the mass of edges in W_1 and the edges that form the reverse of edges in W_2. Also, the mass of edges in W_2 is at most the mass of a subset of the self-loops (whose total size is $|W_2|$) around the vertices that are at the tail of edges in W_2. It follows that $p_{t-1}(W_2 \cup W_3)$ is at most

$$\frac{1}{2} I_{t-1}(|W_3| + |W_1| + |W_2| + |W_2| + |W_2|) = \frac{1}{2} I_{t-1}(k + 2|W_2|),$$

thus establishing Equation 2. Since $|W_2| \geq \phi k$ (since $k \leq m$), we obtain from the concavity of I_{t-1} that

$$I_t(x) \leq \frac{1}{2}\left(I_{t-1}(x - 2\phi x) + I_{t-1}(x + 2\phi x)\right).$$

We can extend the claim for all $x \in [0, m]$ (not just the integers k of the type assumed above) and also establish the desired claim for $x \in [m, 2m]$ using a similar argument. $\square$

2.3 Putting together a proof for L-S

Proof of Theorem 1: The proof of the main theorem now follows from Lemma 2 using an induction argument. Define the function R_t as follows. Let $R_0(x) = \min(\sqrt{x}, \sqrt{2m-x}) + x/2m$.

$$\begin{aligned} R_t(x) &= \frac{1}{2}\left(R_{t-1}(x - 2\phi x) + R_{t-1}(x + 2\phi x)\right) \text{ for } x \in [0, m] \\ R_t(x) &= \frac{1}{2}\left(R_{t-1}(x - 2\phi(2m-x)) + R_{t-1}(x + 2\phi(2m-x))\right) \text{ for } x \in [m, 2m]. \end{aligned}$$

We first show using induction on t that $I_t(x) \le R_t(x)$ for all t. The base case is immediate since $I_0(x) \le 1 \le R_0(x)$ for $x \ge 1$ and $I_0(x) \le x \le \sqrt{x} \le R_0(x)$ for $x \in [0, 1]$. The induction step also follows immediately from Lemma 2 and the definition of $R_t(x)$.

We next show the following for all x, which will complete the proof of the theorem.

$$R_t(x) \le \min\{\sqrt{x}, \sqrt{2m-x}\}\left(1 - \frac{\phi^2}{2}\right)^t + \frac{x}{2m}.$$

We only consider the case $x \in [0, m]$; the other case is similar. The induction base $t = 0$ is immediate. For the induction step, we have

$$\begin{aligned} R_t(x) &= \frac{1}{2}\left(R_{t-1}(x - 2\phi x) + R_{t-1}(x + 2\phi x)\right) \\ &\le \frac{1}{2}\left(\sqrt{x - 2\phi x}\left(1 - \frac{\phi^2}{2}\right)^{t-1} + \sqrt{x + 2\phi x}\left(1 - \frac{\phi^2}{2}\right)^{t-1}\right) + \frac{x}{2m} \\ &= \frac{1}{2}\left(1 - \frac{\phi^2}{2}\right)^{t-1}\left(\sqrt{x - 2\phi x} + \sqrt{x + 2\phi x}\right) + \frac{x}{2m}. \\ &\le \frac{1}{2}\left(1 - \frac{\phi^2}{2}\right)^{t-1}\sqrt{x}\left(1 - \phi - \frac{\phi^2}{2} + 1 + \phi - \frac{\phi^2}{2}\right) + \frac{x}{2m}. \\ &= \sqrt{x}\left(1 - \frac{\phi^2}{2}\right)^t + \frac{x}{2m}. \end{aligned}$$

□ Note that one can show convergence to the stationary distribution using the above Theorem, in terms of the conductance of the graph. The bound we get above is related to the one we would obtain if we use the bound on the second smallest eigenvalue of the Laplacian, as obtained from the version of Cheeger's inequality we derived in class. In fact, there is a version of Cheeger's inequality hidden in the above proof that relates the spectral gap to the conductance of the graph.

Our proof of Theorem 1 can, in fact, be extended to the following corollary.

Corollary 1. *For any set W of vertices and $x = \sum_{w \in W} d(w)$, we have*

$$\left|\sum_{w \in W}(p_t(w) - \pi(w))\right| \le \min(\sqrt{x}, \sqrt{2m-x})\left(1 - \frac{\phi(W)^2}{2}\right)^t,$$

where $\phi(W)$ is the conductance of W.

This corollary and its "inverse" have been used to design local clustering algorithms for graphs, whose running time is nearly linear in the size of cluster returned.