November 28, 2005
Paul K Horn

Random Walks on Graphs and Directed Graphs

A Result of Lovász and Simonovits

We now have a Cheeger's inequality for both directed and undirected graphs,

$$2h_G \geq \lambda \geq \frac{\alpha_G^2}{2} \geq \frac{h_G^2}{2}$$

We can approximate Cheeger's constant, by algorithmically finding a good cut. Recall that λ is defined in terms of the Raliegh quotient by

$$\lambda = \frac{R(f)}{2}$$

and

$$\alpha_G = \inf_i \frac{F(\partial S_i)}{F(S_i)F(\bar{S}_i)}$$

where

$$F(S) = \sum_{v \in S} \phi(v)$$

and by ordering vertices in terms of f, so that $f(v_1) \leq f(v_2) \leq \cdots \leq f(v_k)$, we take $S_i = \{v_1, \ldots, v_i\}$; the i lowest weighted vertices. For an undirected graph, recall that

$$\phi(v) = \frac{d_v}{\sum_u d_u}.$$

In the directed case it becomes a numerical excercise to solve for u.

1 Approximating the Volume of a Convex Body

Many counting problems can be reduced to computing a volume. In the non-convex case, there are examples that make the volume very hard to compute, but in many problems the desired volume arises from a convex body. Approximating the volume of a covex body can be done via linear programming, and there are polynomial algorithms to solve linear programs, however the efficiency of the algorithms is still important. A method that has been used to improve the efficiency of volume computation is using modified random walks.

To see this connection, we note that computing the volume of a body, and sampling are not so different. Consider putting the body in a box; and sampling n points; then the percentage of points falling inside the body gives an approximation of the volume as n gets large. The question then becomes, how do we appropriately choose a random point in a graph. One method is to use a random walk; start at a point and iteratively choose one of its neighbors for some number of steps. The problen is knowing when to stop. We can find, however, a bound for the convergence of random walks, so the strategy is to choose a k large enough so that after k steps we are sufficiently close to the stationary distribution, where 'sufficiently close' depends on our application, let our random walk run for k steps and take the resulting position to be our random point.

2 A Theorem of Lovász and Simonovits

We consider the following theorem of Lovász and Simonovits [2], which gives a bound on the convergence of random walks and holds in both the directed and undirected case.

Theorem 1 (Theorem (Lovász-Simonovits)). *Let G be a strongly connected directed graph, and $\mathcal{P} = \frac{I+P}{2}$ be the transition probability matrix for a lazy walk. Then*

$$|\mathcal{P}^t(v,u) - \phi(u)| \leq \left(1 - \frac{\beta_{t,v}^2}{8}\right)\sqrt{\frac{\phi(u)}{\phi(v)}}.$$

The $\beta_{t,v}$ are defined in the following, algorithmic, fashion. Fix v, and order the vertices according to $\frac{\mathcal{P}^k(v,u)}{\phi(u)}$. Take

$$S_{j,k,v} = \{j \text{ vertices with largest value of } \frac{\mathcal{P}^k(v,u)}{\phi(u)}\}.$$

Another way of thinking about the collections $S_{j,k,v}$ is we start a random walk at v, and select the j highest weight vertices after going for k steps . Then

$$\beta_{t,v} = \inf_{k'<t} \inf_{j} \frac{F(\partial S_{j,k,v})}{\min\{F(S_{j,k',v}), F(S_{j,k',v})\}}.$$

This theorem has seen use lately in such areas as graph sparsification, and solving linear systems efficiently [3]. Also encoded in the theorem is a hidden Cheeger inequality.

2.1 A Hidden Cheeger Inequality

Encoded in the result of Lovász and Simonovits is the following [1]:

$$2h_G \geq \lambda \geq 1 - \max_{i\neq 0}\left|\frac{1+\rho_i}{2}\right| \geq 1 - \lim_{t\to\infty}(\Delta(t))^{1/t} \geq \frac{\beta_G^2}{8} \geq \frac{h_G^2}{8}.$$

where h_G is the Cheeger constant, the ρ_i are eigenvalue of P, and $\beta_G = \inf_{t,v} \beta_{t,v}$. Note that we sacrifice a factor of four from our original Cheeger inequality, but in return we get the relative pointwise distance metric in the middle.

Previously, we've used eigenvalues to get a bound on the convergence of random walks. Having relative pointwise distance in the middle gives us a method going the other direction. By using random walks, and the cuts gotten from them we get a bound on the eigenvalues of the Laplacian.

2.2 Proving the Theorem of Lovász and Simonovits

We begin with the following lemma

Lemma 1. *Let G be strongly connected, with Perron vector ϕ, with $S \subseteq V$ and $f : V \to \mathbb{R}$. Then*

$$\sum_{v \in S} f \Phi \mathcal{P}(v) = \frac{f \cdot g_1 + f \cdot g_2}{2}$$

where $\Phi = diag(\phi(u))$ and $0 \leq g_i(v) \leq \phi(v)$ with

$$\begin{aligned} \sum_v g_1(v) &= \phi(S) - \phi(\partial S), \\ \sum_v g_2(v) &= \phi(S) + \phi(\partial S). \end{aligned}$$

Proof. We note

$$\begin{aligned}
\sum_{v \in S} f \Phi \mathcal{P}(v) &= \sum_{v \in S} \sum_u f(u)\phi(u)\mathcal{P}(u,v) \\
&= \sum_{u \in S} f(u)\phi(u) \sum_{v \in S} \mathcal{P}(u,v) + \sum_{u \notin S} f(u)\phi(u) \sum_{v \in S} \mathcal{P}(u,v) \\
&= \sum_{u \in S} f(u)\phi(u) \frac{1 + \sum_{v \in S} P(u,v)}{2} + \sum_{u \notin S} f(u)\phi(u) \frac{\sum_{v \in S} P(u,v)}{2} \\
&= \sum_{u \in S} f(u)\phi(u) \frac{2 - \sum_{v \notin S} P(u,v)}{2} + \sum_{u \notin S} f(u)\phi(u) \frac{\sum_{v \in S} P(u,v)}{2} \\
&= \frac{1}{2} \sum_{u \in S} f(u)\phi(u)(1 - \sum_{v \notin S} P(u,v)) + \frac{1}{2} \sum_{u \in S} f(u)\phi(u) + \frac{1}{2} \sum_{u \notin S} f(u)\phi(u) \sum_{v \in S} P(u,v) \\
&= \frac{1}{2}(f \cdot g_1 + f \cdot g_2)
\end{aligned}$$

where

$$g_1(u) = \begin{cases} \phi(u)(1 - \sum_{v \notin S} P(u,v)) & u \in S \\ 0 & \text{else} \end{cases}$$

and

$$g_2(u) = \begin{cases} \phi(u) & u \in S \\ \phi(u) \sum_{v \in S} P(u,v) & \text{else} \end{cases}$$

It remains to show that g_1 and g_2 have the correct sums, but

$$\sum_u g_1(u) = \sum_{u \in S} \phi(u)(1 - \sum_{v \notin S} P(u,v)) = \phi(S) - \phi(\partial S)$$

and

$$\sum_u g_2(u) = \sum_{u \in S} \phi(u) + \sum_{u \notin S} \phi(u) P(u,v) = \phi(S) + \phi(\partial S)$$

as desired. □

For $f : V \to \mathbb{R}$ and for any x we take

$$\hat{f}(x) = \max\{f \cdot g : 0 \leq g(v) \leq \phi(v) \text{ for all} v, \ \sum_v g(v) = x\}$$

Suppose $f(v_1) \geq f(v_2) \geq \cdots \geq f(v_n)$, and

$$\sum_{i=1}^{k} \phi(v_1) \leq x < \sum_{i=1}^{k+1} \phi(v_i)$$

then it is easy to check that

$$\hat{f}(x) = \sum_{i=1}^{k} f(v_i)\phi(v_i) + f(v_{k+1})(x - \sum_{i=1}^{k}(v_i))$$

Proof of Lovasz-Simonovits. Define

$$f_k(u) = \frac{\mathcal{P}^k(v,u) - \phi(u)}{\phi(u)}$$

then

$$f_k \Phi \mathcal{P}(u) = \mathcal{P}^{k+1}(v,u) - \phi(u) = f_{k+1}\Phi(u)$$

We prove the following lemma.

Lemma 2. $\hat{f}_{k+1}(x) < \dfrac{\hat{f}_k(x(1-\beta_{k,v})) + \hat{f}_k(x(1+\beta_{k,v}))}{2}$

Proof.

$$\begin{aligned} \sum_{u \in S} f_{k+1}\Phi\mathcal{P}(v) = \frac{f_k \cdot g_1 + f_k \cdot g_2}{2} \quad &\leq \quad \frac{\hat{f}_k(\phi(S) - \phi(\partial(S))) + \hat{f}_k(\phi(S) + \phi(\partial(S)))}{2} \\ &\leq \quad \frac{\hat{f}_k(x(1-\beta_{k,v})) - \hat{f}_k(x(1+\beta_{k,v}))}{2} \end{aligned}$$

□

We use this to show the following lemma, which implies the result.

Lemma 3.

$$\hat{f}_k \leq \left(1 - \frac{\beta_{k,v}}{8}\right)^k \frac{\min\{\sqrt{x}, \sqrt{1-x}\}}{\sqrt{\phi(u)}}$$

Suppose by induction that we know $|P^t(v,u) - \phi(u)| \leq \left(1 - \frac{\beta_{t,v}^2}{8}\right)^t \sqrt{\frac{\phi(v)}{\phi(u)}}$. Then as $P^k(v,u) - \phi(u) \leq \hat{f}_k(\phi(u))$, the lemma implies the result.

Proof of lemma. We prove this by induction. For $k = 0$, we have $\hat{f}_0(\phi(u)) \leq 1$ and $\hat{f}_0(1) \leq \frac{1}{\sqrt{u}}$ so this holds. By our assumption and the previous lemma we have

$$\hat{f}_{k+1}(x) \leq (1 - \frac{\beta_{k,v}^2}{8})^k \frac{(\min\{\sqrt{x(1-\beta_{k,v})}, \sqrt{1-x(1-\beta k,v)}\} + \min\{\sqrt{x(1+\beta_{k,v})}, \sqrt{1-x(1+\beta_{k,v})}\})}{\phi(u)}$$

To finish the proof, we make the following observation based off of the taylor series for $\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \ldots$ that

$$\begin{aligned} \frac{\sqrt{1-x} + \sqrt{1+x}}{2} &\leq \frac{1 - \frac{x}{2} - \frac{x^2}{8} + 1 + \frac{x}{2} - \frac{x^2}{8}}{2} \\ &= \frac{2 + \frac{x^2}{4}}{2} = 1 + \frac{x^2}{8} \end{aligned}$$

Taking $x = (1 - \beta_{k,v})$ in this we get that

$$\min(\sqrt{\beta_{k,v}}, \sqrt{1-\beta_{k,v}}) + \min(\sqrt{\beta_{k,v}}, \sqrt{1+\beta_{k,v}}) \leq \sqrt{1-\beta_{k,v}} + \sqrt{1+\beta_{k,v}} \leq 1 + \frac{\beta_{k,v}^2}{8}$$

and hence

$$\hat{f}_{k+1}(x) \leq (1 - \frac{\beta_{k,v}^2}{8})^{k+1} \frac{\min\{\sqrt{x}, \sqrt{1-x}\}}{\sqrt{\phi(u)}}$$

as desired. □

This completes the proof of the theorem. □

References

[1] F. Chung, Random walks and cuts in directed graphs, preprint.

[2] L. Lovász, and M. Simonovits, Random walks in a convex body and an improved volume algorithm, *Random Structures and Algorithms* 4 (1993), 359-412

[3] D. Spielman and S.-H. Teng, Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems, *Proceedings of the 96th Annual ACM Symposium on Theory of Computing,* (2004), 91-90.