CAS CS 599 B: Mathematical Methods for TCS

Lecturer: Mark Bun | Spring 2022

Lecture Notes 15:

Random Walks

Reading.

- Spielman Chapter 10

A random walk is a process that begins at a vertex in a graph, and in each time step, moves to an adjacent vertex. Last time, we saw that the second eigenvalue of the normalized Laplacian governs the behavior of one step of a random walk. Today we will see how it governs a random walk's convergence rate, or "mixing time."

1 Random Walks

Let $G = (V, E)$ be an undirected graph. We will primarily be interested in the expected behavior of a random walk, i.e., the distribution of a walk after a certain number of time steps.

Denote a probability distribution on V by its probability mass function $p : V \to [0, 1]$. Let p_t denote the distribution of a random walk at time step t. We will usually think of starting the walk deterministically from a fixed vertex, so $p_i = \mathbf{1}_{\{v_0\}}$ for some $v_0 \in V$. To sample from the distribution of random walk at time step $t + 1$:

1. Sample a vertex $u \sim p_t$
2. Sample a random vertex $v \sim u$.

Algebraically, this update rule is given by:

$$p_{t+1}(v) = \sum_{(u,v)\in E} \frac{1}{\deg(u)} \cdot p_t(u).$$

It's a useful exercise to check that

$$p_{t+1} = W p_t = W^t p_0$$

where $W = AD^{-1}$ is the *walk matrix* of the graph G.

The walk matrix has a special (right) eigenvector with positive entries,

$$\pi(u) = \frac{\deg(u)}{\sum_{v\in V} \deg(v)} = \frac{\deg(u)}{\|d^{1/2}\|}$$

with eigenvalue 1:

$$W\pi(v) = \sum_{(u,v)\in E} \frac{1}{\deg(u)} \cdot \frac{\deg(u)}{\|d^{1/2}\|} = \frac{\deg(v)}{\|d^{1/2}\|}.$$

This is the *stationary* or *stable* distribution, as it is the distribution on vertices that remains fixed after taking one step of the random walk.

We'd like to establish conditions under which a random walk converges to its stationary distribution, i.e., $p_t \to \pi$ as $t \to \infty$. For intuition, suppose $\phi_1, \ldots, \phi_n$ is a (not necessarily orthonormal) basis of eigenvectors for W with corresponding eigenvalues $1 = w_1 \geq w_2 \geq \cdots \geq w_n \geq -1$. Then if $p_0 = \alpha_1 \phi_1 + \cdots + \alpha_n \phi_n$, we have

$$W^t p_0 = \alpha_1 w_1^t \phi_1 + \cdots + \alpha_n w_n^t \phi_n.$$

As long as $w_1 > w_2$ and $w_n > -1$, this sum will be dominated by the first term $\alpha_1 \phi_1$, with all other terms vanishing exponentially fast in t.

However, things can go (seriously) wrong when either of these eigenvalue conditions fail. Fortunately, these eigenvalue conditions exactly correspond to conditions under which we expect convergence to fail.

Theorem 1 (Consequence of Perron-Frobenius)**.** *Let W be the walk matrix of a graph G with eigenvalues $1 = w_1 \geq w_2 \geq \cdots \geq w_n \geq -1$. Then*

- *$w_2 = 1$ iff G is disconnected.*
- *If G is connected, then $w_n = -1$ iff G is bipartite.*

I won't prove this, but to understand why it's true, you can read Chapter 4.5 in Spielman. Note that the Perron-Frobenius theorem doesn't directly apply to W since it's not symmetric, so one instead uses the fact that $W = D^{1/2}(D^{-1/2}AD^{-1/2})D^{-1/2}$ is similar to the symmetric normalized adjacency matrix $D^{-1/2}AD^{-1/2}$.

If G is disconnected, then a random walk starting in one component can never reach a different component. If G is bipartite, then a walk will always oscillate between the two halves of the bipartition, so the distribution at time t will depend on the parity of t. Perhaps miraculously, the Perron-Frobenius theorem tells us that these are the only two possible failure modes.

2 Lazy Random Walks

Having to treat bipartite graphs differently is rather annoying, so instead of studying simple random walks, one often studies *lazy random walks*, where the transition operator is given by

$$\widetilde{W} = \frac{1}{2}I + \frac{1}{2}W = \frac{1}{2}I + \frac{1}{2}AD^{-1}.$$

That is, one step of a lazy random walk stays at the same vertex with probability $1/2$ and transitions to a random neighbor with probability $1/2$.

Claim 2. *The eigenvalues of $\widetilde{W}$ are $1 = \omega_1 \geq \omega_2 \geq \cdots \geq \omega_n \geq 0$ where $\omega_i = 1 - \nu_i/2$.*

Here, we recall that $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n \leq 2$ are the eigenvalues of the normalized Laplacian $N = I - D^{-1/2}AD^{-1/2}$.

Proof. Let f be an eigenvector of N with eigenvalue ν. We will show that $D^{1/2}f$ is an eigenvector of $\widetilde{W}$ with eigenvector $1-\nu/2$. To see this, we calculate:

$$\begin{aligned}\widetilde{W}D^{1/2}f &= \frac{1}{2}D^{1/2}f + \frac{1}{2}AD^{-1/2}f \\ &= \frac{1}{2}D^{1/2}f + \frac{1}{2}D^{1/2}(I-N)f \\ &= D^{1/2}f - \frac{1}{2}\nu D^{1/2}f.\end{aligned}$$

□

Now the heuristic calculation from before goes through, using the ω's in place of the w's, as long as $\omega_2 < 1 \iff \nu_2 > 0 \iff G$ is connected.

3 Convergence Rate

Theorem 3. *For any start vertex s, letting $p_0 = \mathbf{1}_{\{s\}}$, and any vertex v and $t \geq 1$, we have*

$$|p_t(v) - \pi(v)| \leq \sqrt{\frac{\deg(v)}{\deg(s)}} \cdot \omega_2^t.$$

Proof. Let $p_0 = \alpha_1 D^{1/2}f_1 + \cdots + \alpha_n D^{1/2}f_n$, where $D^{1/2}f_1, \ldots, D^{1/2}f_n$ are the eigenvectors for $\widetilde{W}$ corresponding to eigenvalues $1 = \omega_1 \geq \omega_2 \geq \cdots \geq \omega_n \geq 0$. Then

$$p_t = \alpha_1 D^{1/2}f_1 + \sum_{i=2}^{n} \omega^t \alpha_i D^{1/2}f_i = \pi + \sum_{i=2}^{n} \omega^t \alpha_i D^{1/2} f_i.$$

Write

$$\begin{aligned}p_t(v) &= \langle \mathbf{1}_{\{v\}}, p_t\rangle \\ &= \langle \mathbf{1}_{\{v\}}, \pi + \sum_{i=2}^{n} \omega_i^t \alpha_i D^{1/2} f_i\rangle \\ &= \pi(v) + \sqrt{\deg(v)} \sum_{i=2}^{n} \omega_i^t \alpha_i f_i(v),\end{aligned}$$

so our goal is to bound the magnitude of the term on the right.

To do this, we observe that the coefficients in our basis decomposition of p_0 satisfy

$$\alpha_i = \langle f_i, D^{-1/2}\mathbf{1}_{\{s\}}\rangle = (\deg(s))^{-1/2} f_i(s)$$

since the f_i's are orthonormal. Plugging this in gives

$$\sqrt{\deg(v)} \sum_{i=2}^{n} \omega_i^t \alpha_i f_i(v) = \sqrt{\frac{\deg(v)}{\deg(s)}} \sum_{i=2}^{n} \omega_i^t f_i(s) f_i(v).$$

The sum on the right is bounded from above by

$$\omega_2^t \sum_{i=2}^{n} |f_i(s)| \cdot |f_i(v)|.$$

By Cauchy-Schwarz, this is at most

$$\omega_2^t \sqrt{\sum_{i=1}^{n} f_i(s)^2} \sqrt{\sum_{i=1}^{n} f_i(v)^2} \leq \omega_2^t$$

since the orthogonal matrix with columns $f_1, \ldots, f_n$ also has orthonormal rows. $\square$