Lecture 1: Random Walks

Lecturer: Shayan Oveis Gharan *01/04/22*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

In this note we study random walks and mixing time in weighted undirected non-regular graphs G. The notation that we define here will be used later to study high dimensional random walks on simplicial complexes.

Given an undirected (weighted) graph $G = (V, E)$ with weight function $w : E \to \mathbb{R}_{\geq 0}$, let $A \in \mathbb{R}_{\geq 0}^{V\times V}$ be the adjacency matrix of G.

Consider a natural probability distribution on edges of G where for any edge $\{u, v\}$

$$\pi_1(\{u, v\}) = \frac{w(\{u, v\})}{\sum_e w(e)}.$$

This distribution naturally induces a distribution on vertices; given a random edge e, we randomly drop one of the endpoints of e:

$$\pi_0(v) = \sum_e \mathbb{P}_{\pi_1}[e]\,\mathbb{P}[v|e] = \sum_{e:v\in e} \frac{1}{2}\pi_1(e) = \frac{\sum_{u\sim v} \frac{1}{2} w(\{u, v\})}{\sum_e w(e)} = \frac{d_w(v)}{\sum_u d_w(u)}.$$

Conversely, suppose we want to sample an edge $e = \{u, v\} \sim \pi_1$. One way is to first sample a vertex $v \sim \pi_0$ and then sample an edge $e|v$, i.e., among all edge incident to v choose one proportional to its weight.

$$\mathbb{P}[\{u, v\}|u\}] := \frac{\pi_1\{u, v\}}{\sum_{x\sim u} \pi_1\{u, x\}}$$

We claim this process chooses e with probability $\pi_1(e)$, because,

$$\begin{aligned}
\mathbb{P}[e] &= \mathbb{P}_{\pi_0}[v]\,\mathbb{P}[e|v] + \mathbb{P}_{\pi_0}[u]\,\mathbb{P}[e|u] \\
&= \pi_0(v)\frac{\pi_1(e)}{\sum_{f\sim v} \pi_1(f)} + \pi_0(u)\frac{\pi_1(e)}{\sum_{f\sim u} \pi_1(f)} \\
&= \pi_0(v)\frac{w(e)}{\sum_{f\sim v} w(f)} + \pi_0(u)\frac{w(e)}{\sum_{f\sim u} w(f)} \\
&= \pi_0(v)\frac{w(e)}{d_w(v)} + \pi_0(u)\frac{w(e)}{d_w(u)} = \frac{\pi(e)}{2} + \frac{\pi(e)}{2} = \pi(e).
\end{aligned}$$

Fact 1.1. *Summarizing the above observations,*

- *To choose $v \sim \pi_0$, we can first choose $e \sim \pi_1$ and then drop one endpoint.*
- *To choose $e \sim \pi_1$, we can first choose $v \sim \pi_0$ and then choose $e|v$.*

1.1 Random Walk Operator

We can look at the simple random walk operator on G that says if I am at vertex u, I choose an edge incident to u with probability proportional to its weight. Let

$$\mathbb{P}\left[u \to v\right] = \mathbb{P}\left[\{u,v\}|u\}\right].$$

For a function $f : V \to \mathbb{R}$, we have

$$Pf(v) := \mathbb{E}_{\{u,v\}|v}\left[f(u)\right] = \sum_u \mathbb{P}\left[v \to u\right] f(u).$$

The following reversibility property will also be used crucially.

$$P(u,v)\pi_0(u) = \pi_0(u)\mathbb{P}\left[\{u,v\}|u\right] = \pi_1(\{u,v\})\mathbb{P}\left[u|\{u,v\}\right] = \pi_1(\{u,v\})\mathbb{P}\left[v|\{u,v\}\right] = P(v,u)\pi_0(v) \tag{1.1}$$

We equip the linear space $\mathbb{R}^V$ with the following inner product: For two vectors $f, g : V \to \mathbb{R}$,

$$\langle f, g\rangle = \mathbb{E}_{v\sim\pi_0} f(v)g(v) = \sum_v \pi_0(v)f(v)g(v).$$

This naturally defines a norm, where for any such function f, $\|f\| = \sqrt{\langle f, f\rangle}$.

Using the above fact, it follows that P is self-adjoint with respect to the above inner product, i.e., for any two $f, g : V \to \mathbb{R}$,

$$\begin{aligned}\langle f, Pg\rangle &= \mathbb{E}_{v\sim\pi_0}\left[f(v)Pg(v)\right] \\ &= \mathbb{E}_{v\sim\pi_0}\left[f(v)\mathbb{E}_{\{u,v\}|v}\left[g(u)\right]\right] \\ &= \mathbb{E}_{\{u,v\}\sim\pi_1}\mathbb{E}_{v|\{u,v\}} f(v)g(u) = \mathbb{E}_{\{u,v\}\sim\pi_1}\mathbb{E}_{v|\{u,v\}} f(u)g(v)\end{aligned}$$

Similarly, we can show that $\langle Pf, g\rangle$ is also equal to the RHS.

Fact 1.2. $\lambda_1 = 1$ *and* $\lambda_n \geq -1$.

Proof. First, observe that the all-ones function $\mathbf{1}$ is an eigenfunction,

$$P\mathbf{1} = \mathbf{1}.$$

Second, we show $\lambda_i \leq 1$ for all i. For any eigenfunction $f : V \to \mathbb{R}$, with eigenvalue λ, i.e., $Pf = \lambda f$, we claim that $\lambda \leq 1$. Say $u = \text{argmax}_v |f(v)|$. Then,

$$\lambda f(u) = Pf(u) = \mathbb{E}_{\{u,v\}|u} f(v) \leq \mathbb{E}_{\{u,v\}|u}|f(v)| \leq \mathbb{E}_{\{u,v\}|u}|f(u)| = |f(u)|.$$

In the second inequality we used $|f(v)| \leq |f(u)|$ for all $v \in V$. So, we have $\lambda \leq 1$ as desired. Also, observe from the same inequality that $|\lambda| \leq 1$ as desired. □

Fact 1.3. *The matrix P has n-real eigenvalues with n-orthonormal eigefunctions.*

Proof. The proof is fairly general and holds for any self-adjoint matrix with respect to an inner-product. First, Suppose $Pf = \lambda f$ for $f \in \mathbb{C}^n$ and $\lambda \in C$. And, recall for any function f, $\langle f, f\rangle = \mathbb{E}_u f(u)\overline{f(u)} \in \mathbb{R}$. Then,

$$\|Pf\|^2 = \langle Pf, Pf\rangle = \langle P^2 f, f\rangle = \lambda^2 \|f\|^2.$$

So, $\lambda^2 \in \mathbb{R}$.

On the other hand, suppose P has two eigenfunctions $f, g \in \mathbb{R}^n$ with eigenvalues $\lambda_f \leq \lambda_g$; then,

$$\langle \lambda_f f, g\rangle = \langle Pf, g\rangle = \langle f, Pg\rangle = \langle f, \lambda_g g\rangle.$$

Therefore, since $\lambda_f \neq \lambda_g$ we must have $\langle f, g\rangle = 0$. □

It follows by the spectral theorem that P has $n = |V|$ eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ with corresponding orthonormal eigenfunctions $f_1, \ldots, f_n$ such that for any $1 \leq i < j \leq n$,

$$\langle f_i, f_j\rangle = 0,$$

and for any i, $\|f_i\| = 1$.

1.2 Mixing

Theorem 1.4. *For a weighted graph with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, if $\max\{|\lambda_2|, |\lambda_n|\} < 1$, then for any vertex u,*

$$\lim_{t\to\infty} P^t f(u) = \mathbb{E}f.$$

Letting $f(v) = \mathbf{1}_v$, we have $\mathbb{E}f = \pi_0(v)$. So, the above theorem shows that as $t \to \infty$, the distribution of the walk converges to π_0. In other words, π_0 is the *stationary distribution* of the simple random walk on G.

The mixing time of a random walk is defined as

$$\tau_{\max} = \min\left\{t : \|P^t(u,.) - \pi_0\|_{\mathrm{TV}} = \frac{1}{2}\sum_v |P^t \mathbf{1}_v(u) - \pi_0(v)| \leq \frac{1}{4}, \forall u \in V\right\} \tag{1.2}$$

where τ_{mix} is the first time that the total variation distance of the distribution of the walk started at u, from the stationary distribution is at most a constant.

For a function $f : V \to \mathbb{R}$ define

$$\|f\|_p = (\mathbb{E}_\pi f^p)^{1/p}.$$

More generally, we define the L_p mixing time of the walk as

$$\tau_{\mathrm{mix},p} := \min\left\{t : \left\|\frac{P^t(u,.)}{\pi} - \mathbf{1}\right\|_p \leq \frac{1}{4}, \forall u \in V\right\}$$

Note that by Cauchy-Schwarz inequality

$$\sum_v |P^t(u,v) - \pi(v)| = \mathbb{E}_\pi|P^t(u,.)/\pi - \mathbf{1}| \leq \sqrt{\mathbb{E}_\pi|P^t(u,.)/\pi - \mathbf{1}|^2} = \left\|\frac{P^t(u,.)}{\pi} - \mathbf{1}\right\|_2$$

So, the L_2-mixing of a chain implies L_1 mixing (but not vice versa). And similarly, for any $p > 1$, L_p mixing implies L_1 mixing.

Theorem 1.5. *Let G be a weighted graph with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, such that $\lambda^* := \max\{|\lambda_2|, |\lambda_n|\} < 1$. For any $\epsilon > 0$, any function $f \in V \to \mathbb{R}$ and $t \geq \frac{\log \epsilon}{1-\lambda^*}$*

$$\left\|P^t f - \mathbb{E}f\right\| \leq \epsilon \left\|f\right\|.$$

Proof. Let $f_1, \ldots, f_n$ be the orthonormal eigenfunctions of P with corresponding eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Recall $\lambda_1 = 1$ and $f_1 = \mathbf{1}$. Also note that P^t have exactly the same eigenfunctions with eigenvalues $\lambda_1^t, \ldots, \lambda_n^t$. Therefore, we can write $f = \sum_i \langle f, f_i \rangle f_i$. First, by (**??**),

$$\begin{aligned} P^t f - \mathbb{E} f &= P^t \left(\sum_{i=1}^n \langle f, f_i \rangle f_i \right) - \langle f, \mathbf{1} \rangle \mathbf{1} \\ &= \sum_{i=1}^n \lambda_i^t f_i \langle f, f_i \rangle - \langle f, \mathbf{1} \rangle \mathbf{1} = \sum_{i=2}^n \lambda_i^t \langle f, f_i \rangle f_i \end{aligned}$$

Therefore, using the orthonormality of f_i's,

$$\begin{aligned} \left\| P^t f - \mathbb{E} f \right\|^2 &= \left\langle \sum_i \lambda_i^t \langle f, f_i \rangle, \sum_j \lambda_i^t \langle f, f_j \rangle \right\rangle \\ &= \sum_{i=2}^n \lambda_i^{2t} \langle f, f_i \rangle^2 \\ &\leq \lambda^{*2t} \left\| f \right\|^2 \end{aligned}$$

where in the last inequality we used $|\lambda_i| \leq \lambda^*$ for all $i > 1$. By the above inequality, for $t \geq \frac{\log \epsilon}{1 - \lambda^*}$ we have

$$\left\| P^t f(u) - \mathbb{E} f \right\| \leq \lambda^{*t} \left\| f \right\| = (1 - (1 - \lambda^*))^t \left\| f \right\| \leq e^{-(1-\lambda^*)t} \left\| f \right\| = \epsilon \left\| f \right\|$$

where in the second inequality we used $1 - x \leq e^{-x}$. □

Corollary 1.6. *Suppose for any function $f : V \to \mathbb{R}$, $\epsilon > 0$ and $t \geq \frac{\log \epsilon^{-1}}{1-\lambda^*}$ $\|P^t f - \mathbb{E} f\| \leq \epsilon \|f\|$ then,*

$$\tau_{mix,1} \leq \tau_{mix,2} \leq \max_u \frac{\log \epsilon^{-1} \pi(u)^{-1}}{1 - \lambda^*}$$

Note that in particular, if we want to bound the mixing time starting at a given vertex u, we don't need the max on the right hand side.

Proof. Fix a vertex $u \in V$ and let $f := \mathbf{1}_u / \pi_0(u)$. Then, $\mathbb{E} f = 1$.

$$\left\| P^t f - \mathbb{E} f \right\|^2 = \left\| P^t f - (\mathbb{E} f) \mathbf{1} \right\| = \left\| P^t f \right\|^2 - 1.$$

On the other hand,

$$\begin{aligned} \left\| P^t f \right\|^2 &= \langle P^t f, P^t f \rangle = \mathbb{E}_v P^t f(v) \cdot P^t f(v) \\ &\underset{f(w) \neq 0, \forall w \neq u}{=} \mathbb{E}_{v \sim \pi_0} \frac{P^t(v, u)}{\pi(u)} \cdot \frac{P^t(v, u)}{\pi(u)} \\ &\underset{(1.1)}{=} \mathbb{E}_{v \sim \pi_0} \left(\frac{P^t(u, v)}{\pi(v)} \right)^2 \end{aligned}$$

Putting these together, $\|P^t f - \mathbb{E} f\|^2 = \left\| \frac{P^t(u,.)}{\pi(.)} - 1 \right\|^2$. Finally, since $\|f\|^2 = \frac{1}{\pi(u)}$, letting ϵ of Theorem 1.5 $\epsilon \pi(u)$ proves the claim. □

Theorem 1.5 allows us to estimate the probability of any event. Say $A \subseteq V$ and we want to estimate $\pi_0(A)$. For $f = \mathbf{1}_A$ we have $\mathbb{E}f = \pi_0(A)$ and $\langle f, f\rangle \leq \pi_0(A) < 1$. So, we can use $P^t f(u)$, for a vertex $u \in V$, as an estimate of $\mathbb{E}f$. By Cauchy-Scwharz inequality,

$$\mathbb{E}_v|P^t f(v) - \mathbb{E}f| \leq \sqrt{\mathbb{E}_u|P^t f(u) - \mathbb{E}f|^2} = \left\|P^t f - \mathbb{E}f\right\|.$$

So, for $t = \frac{\log \epsilon^{-1} \log \pi(u)^{-1}}{1-\lambda^*}$

$$|P^t f(u) - F| \leq \frac{1}{\pi_0(u)} \mathbb{E}_v|P^t(v) - F| \underset{\textit{Theorem 1.5}}{\leq} \epsilon \left\|f\right\|$$

Having this, we can simply estimate $\pi_0(A)$ within an additive error of ϵ (with probability $1-\delta$) by averaging k independent $f(X_t)$, where X_t is the t-th state that the simple random walk started at u lands on. Here, $k = O(\frac{1}{\epsilon^2} \log \delta)$ and t is as defined above.