18.317 Combinatorics, Probability, and Computations on Groups 3 October 2001

Lecture 11

Lecturer: Igor Pak *Scribe: N. Ackerman*

1. Random Walks on Groups

Let G be a finite Group and let S be a set of generators of G. We say that S is symmetric if $S = S^{-1}$. In other words S is symmetric if $\forall s \in S$, $s^{-1} \in S$

Definition 1.1. We define $\Gamma = \Gamma(G, S)$ as the Cayley graph of G with respect to S. This is the graph which has a vertices for each $g \in G$ and edges between each (g, h) such that $g^{-1}h \in S$.

Observe that if S is symmetric then the Cayley graph of G with respect to S, $\Gamma(G, S)$ is unoriented. Also observe that if S contains the identity ($id \in S$) then the Cayley graph of G with respect to S, $\Gamma(G, S)$ has loops.

Definition 1.2. On a walk of a Cayley graph we define x_t as the place you reach after t steps.

We define a random walk as just a random walk on the Cayley graph starting at the identity. In other words at each step you choose (randomly and uniformly) which direction to go. ($x_{t+1} = x_t \cdot s$, $x_0 = \mathrm{id}$, $s \in S$ (uniform in S))

We similarly define a lazy random walk as a random walk, except before each step you choose first whether to move or stay where you are. Then, if you have decided to move, you decide independently where to move ($x_{t+1} = x_t \cdot s^e$, $s \in S, e \in \{0, 1\}$).

We define $Q^t(g) = Pr(x_t = g)$ as the probability that after t steps on the walk you will be at vertex g.

Proposition 1.3. *If the Cayley graph is not bipartite then* $Q^t(g) \to 1/|G|$*, as as* $t \to \infty$*.*

For example, if S contains the identity ($id \in S$) then this proposition is true for the lazy random walk.

Example: If $G = \mathbb{Z}_m$ and $S = \{\pm 1\}$ then the Cayley graph $\Gamma(G, S)$ is bipartite if and only if m is even.

Example: If $G = S_m$ and $S = \{(i, j) | 1 \le i < j \le n\}$ then the Cayley graph $\Gamma(G, S)$ is bipartite.

Definition 1.4. If P and Q are probability distributions on G then we define the convolution of P and Q as

$$P * Q(g) = \sum_{h \in G} P(h) Q(h^{-1} g)$$

Observe that if P is the probability distribution

$$P(g) = \begin{cases} 1/|S|, g \in S \\ 0 \text{ otherwise} \end{cases}$$

then $Q^t = \underbrace{P * P * \cdots * P}_{t \text{ times}}$

Definition 1.5. We then define the separation distance after t steps as

$$\text{sep}(t) = |G| \cdot \max_{g \in G} (1/|G| - Q^t(g))$$

Proposition 1.6.
a) sep*(t+1)* $\leq$ sep*(t)*
b) sep*(t+l)* $\leq$ sep*(t)* $\cdot$ sep*(l)*
c) sep*(t)*$\sim c\rho^t$ *as* $t \to \infty$, $0 \leq \rho \leq 1$,

where $f(x) \sim g(x)$ *means that* $f(x) = g(x \cdot (1 + o(1)))$

Proof. a) Observe that sep(t) $< \epsilon$ is equivalent to saying that $Q^t = (1-\epsilon)U + \epsilon N$ where U is the uniform distribution, and N is some other distribution. Therefore we know that because $Q^{t+1} = Q^t * P$,

$$Q^{t+1} = ((1-\epsilon)U + \epsilon N) * P = (1-\epsilon)U * P + \epsilon N * P$$

But we know $U * P$ is still the uniform distribution, and $\min_{g \in G} N * P \geq \min_{g \in G} N$ by the construction of P. So $\min_{g \in G} Q^{t+1}(g) \geq \min_{g \in G} Q^t(g)$. And so finally sep(t+1) $\leq$ sep(t).

b) Let $Q^t = (1-\epsilon)U + \epsilon N_1$ and $Q^l = (1-\delta)U + \delta N_2$. We know that this is equivalent to sep(t) $< \epsilon$ and sep(l) $< \delta$

We then have

$$Q^{t+l} = ((1-\epsilon)U + \epsilon N_1) * ((1-\delta)U + \delta N_2)$$

and, after we condense terms it is easy to see that

$$Q^{t+l} = (1 - \delta\epsilon)U + \delta\epsilon N_1 * N_2$$

And so we have that sep(t+l) $<$ sep(t) $\cdot$ sep(l) (because sep(t+1)$< \delta\epsilon$ and δ and ϵ are arbitrary).

c) Let $A = (a_{gh})_{g,h \in G}$ be a matrix such that $a_{gh} = P(hg^{-1})$. We then let $A^t = A * \cdots * A$. Then observe that

$$Q^t = \left[A^t \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right]_g$$

We then have

$$A^t \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1/|G| \\ 1/|G| \\ \vdots \\ 1/|G| \end{pmatrix} + \lambda_1^t(v_1) + \cdots + \lambda_n^t(v_n)$$

Where v_i and λ_i are eigenvectors and eigenvalues. And so

$$Q^t(g) = \lambda_1^t(w_1) + \cdots + \lambda_n^t(w_n)$$

Now, if we let $q_t = \min_{g \in G} Q^t(g)$ then $q_t = 1/|G| + w_1\lambda_1^t + \cdots$ and

$$1/|G| - q_t = w_1\lambda_1^t + \sigma C$$

Also observe that if we do this same thing for the lazy random walk we have $\lambda'_i = 1/2(1 + \lambda_i)$.

□

Definition 1.7. We define the <u>relaxation time</u> as: $\tau_1 = 1/(1 - \lambda_1)$
We then define the <u>mixing time</u> as minimum time for the separation time to be less than one half: $\tau_2 = \min\{t : \text{sep(t)} \leq 1/2\}$.
Finally, we define the <u>optimal time</u> as: $\tau_3 = \sum_{t=0}^{\infty} \text{sep(t)}$

Proposition 1.8. $1/2\tau_3 < \tau_2 < 2\tau_3$

Proof. Now, we can see from the definitions that
$\tau_3 \geq \underbrace{1/2 + 1/2 + \cdots + 1/2}_{\tau_2} \geq 1/2\tau_2$
Now we also know from the definitions that
$\tau_3 \leq \underbrace{1 + 1 + \cdots + 1}_{\tau_2} + \underbrace{1/2 + 1/2 + \cdots + 1/2}_{\tau_2} + \underbrace{1/4 + 1/4 + \cdots + 1/4}_{\tau_2} + \cdots$
$= \tau_2(1 + 1/2 + 1/4 + \cdots) = 2\tau_2$
And this completes the proof. □

Proposition 1.9. $\tau_2 < \tau_1 \log(|G|)$

Proof. First, note that

$$|1/|G| - Q^t(g)| < \sum w_i\lambda_i^t < |G|\lambda_1^t$$

(Note that $\lambda = \rho$ if $s = s^{-1}$).

Likewise, from the definitions we see that:
$|G| \cdot |1/|G| - Q^t(g)| < |G|^2(1 - 1/\tau_1)^t \leq 1/e < 1/2 \cdot q_t, t = 2\log(|G|)$

□

Example If $G = \mathbb{Z}_m$ and $S = \{\pm 1\}$ then

$$A = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & \ddots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix}$$

One can show that $\lambda_j = \cos(2\pi j/m)$, so

$$\lambda_1 = 1 - \frac{(2\pi)^2}{m^2} + o\left(\frac{1}{m^4}\right) = 1 - \frac{c}{m^2} + o\left(\frac{1}{m^4}\right).$$

Now Proposition 1.9 implies that $\mathbf{mix} = o(m^2 \log m)$. In the future we will show that $\mathbf{mix} = \theta(m^2)$.