18.317 Combinatorics, Probability, and Computations on Groups 5 October 2001

Lecture 12

Lecturer: Igor Pak *Scribe: Michael Korn*

More About Mixing Times

Definition 1 *For a probability distribution Q, we define the total variation distance, denoted $\|Q-U\|$, by*

$$\|Q-U\| = \max_{X\subset G}\left(Q(X) - \frac{|X|}{|G|}\right) = \max_{X\subset G}\left(\frac{|X|}{|G|} - Q(X)\right) = \frac{1}{2}\sum_{g\in G}\left|Q(g) - \frac{1}{|G|}\right|$$

Definition 2

$$\tau_4 = \min\left\{t : \|Q^t - U\| < \frac{1}{4}\right\}$$

Theorem 3

$$\frac{1}{96}\tau_2 < \tau_4 < 2\tau_2$$

where $\tau_2 = \min\{t : \mathrm{sep}(t) < 1/2\}$, as we defined in the previous lecture.

Proof: First we shall prove that $\tau_4 < 2\tau_2$. In the previous lecture, we showed that $\|Q^t - U\| < \mathrm{sep}(t)$. In particular, $\|Q^{2\tau_2} - U\| < \mathrm{sep}(2\tau_2)$. By sub-multiplicativity, we have that $\mathrm{sep}(2\tau_2) < \mathrm{sep}(\tau_2)^2 < 1/4$. It then follows that $\tau_4 < 2\tau_2$.

Claim 4 *Let $X = \left\{g \in G : Q(g) < \frac{1}{4|G|}\right\}$. Assume $\|Q-U\| < 1/4$. Then $|X| < \frac{|G|}{3}$.*

Proof: The proof of this is one line. Here is that line:

$$\frac{1}{4} > \|Q-U\| \geq \frac{|X|}{|G|} - Q(X) > \frac{|X|}{|G|} - \frac{1}{4}\frac{|X|}{|G|} = \frac{3}{4}\frac{|X|}{|G|} \quad \Rightarrow \quad |X| < \frac{|G|}{3}$$

■

It follows that $|\overline{X}| > \frac{2}{3}|G|$.
Now let us apply this claim. We still assume $\|Q-U\| < 1/4$. Let $P = Q * Q$. For all $g \in G$ we have

$$P(g) = \sum_{ab=g} Q(a)Q(b) \geq \sum_{\substack{a\in\overline{X}\\ b\in\overline{X}\\ ab=g}} Q(a)Q(b).$$

For at least $\frac{2}{3}$ of the elements a in G we have $a \in \overline{X}$. For at least $\frac{2}{3}$ of the elements a in G, we have $a^{-1}g \in \overline{X}$. Therefore, for at least $\frac{1}{3}$ of the elements a in G, both statements hold. So the latter summation in the above equation is a sum over at least $\frac{|G|}{3}$ terms. In these terms, $Q(a)$ and $Q(b)$ are both at least $\frac{1}{4|G|}$, since a and b are in $\overline{X}$. Hence,

$$P(g) > \frac{|G|}{3}\frac{1}{4|G|}\frac{1}{4|G|} = \frac{1}{48|G|}.$$

Now take Q to be Q^{τ_4}, which makes $P = Q^{2\tau_4}$. (Notice that Q^{τ_4} satisfies the condition $\|Q^{\tau_4} - U\| < 1/4$.) The above argument gives us that $Q^{2\tau_4}(g) > \frac{1}{48|G|}$ for all g. Hence $\mathrm{sep}(2\tau_4) < \frac{47}{48}$. So

$$\mathrm{sep}(48(2\tau_4)) < \left(1 - \frac{1}{48}\right)^{48} < \frac{1}{e} < \frac{1}{2}.$$

It follows that $96\tau_4 > \tau_2$. This proves the second half of the theorem. ■

Stopping Times

In this section, we introduce the notion of a *stopping rule* or *stopping time* for a random walk. The intuitive notion of a stopping rule is as follows. There is a machine which observes a random walk as it progresses. It is the job of the machine to say "STOP" at some time. The time at which the machine chooses to say "STOP" will depend on the sequence of states that it has observed. The machine cannot see future states of the random walk; it can only say "STOP" based upon what it has seen up to that point. This leads us to the following definition.

Definition 5 *A function $\varkappa : \{X_t\} \to \mathbb{N}$ is called a* stopping time*, provided it satisfies the following condition: If $\varkappa(\{X_t\}) = k$, and $X'_a = X_a$ for all $a \le k$, then $\varkappa(\{X'_t\}) = k$.*

So formally we treat $\varkappa$ as a function of an infinite walk, whose output is the time (a non-negative integer) at which the machine says "STOP". The condition in the definition ensures that the function does not depend on states of the walk beyond the point at which the machine says "STOP". Now we generalize this concept to allow for randomness in the machine's behavior. The machine still sees the random walk as before, but now it also has access to a random real number in $[0, 1]$ at each time step.

Definition 6 *A function $\varkappa : \{X_t\} \times [0,1]^\infty \to \mathbb{N}$ is called a* randomized stopping time*, provided it satisfies the following condition:*
If $\varkappa(\{X_t\}, \{p_t\}) = k$, $X'_a = X_a$ for all $a \le k$, and $p'_a = p_a$ for all $a \le k$, then $\varkappa(\{X'_t\}, \{p'_t\}) = k$.

From here on, all stopping times will be assumed to be randomized unless otherwise stated.

Definition 7 *We define the* stopping state *to be $\eta = X_\varkappa$.*

Definition 8 *$\varkappa$ is* uniform *if $\Pr(\eta = g) = \frac{1}{|G|}$ for all $g \in G$.*
$\varkappa$ is strong uniform *if $\Pr(\eta = g \mid \varkappa = k) = \frac{1}{|G|}$ for all $g \in G$, and for all k for which $\Pr(\varkappa = k) > 0$.*

Let us look at an example of a uniform stopping rule. Let $G = \mathbb{Z}_m$, and let $\{\pm 1, 0\}$ be our generating set. Suppose $\varkappa$ picks a random element of G, and tells the random walk to stop as soon as it gets to that element. (In other words, the machine uses its first random number to determine what it wants the stopping state to be, and it lets the walk progress until it happens to land on that number.) Clearly this is a uniform stopping rule. It is not strong uniform, however. If the stopping time is small, then the stopping state is more likely to be near 0, rather than near $m/2$ (especially if the stopping time is less than $m/2$).

Theorem 9 *(Aldous-Diaconis)*
Let $\varkappa$ be strong uniform. Then:
(1) $\text{sep}(t) \leq \Pr(\varkappa > t)$ *for all t, and*
(2) There exists a $\varkappa$ which yields equality in part (1) for all t.

This will be proved in the next lecture.

Corollary 10 $\tau_3 = \min E(\varkappa)$, *where the minimum ranges over all strong uniform $\varkappa$.*

Proof: Recall that $\tau_3 = \sum_{t=0}^{\infty} \text{sep}(t)$. Let $\varkappa$ be the optimal value mentioned in part (2) of the theorem. Then $\text{sep}(t) = \Pr(\varkappa > t)$. So

$$\tau_3 = \sum_{t=0}^{\infty} \text{sep}(t) = \sum_{t=0}^{\infty} \Pr(\varkappa > t) = \sum_{t=0}^{\infty} t \cdot \Pr(\varkappa = t) = E(\varkappa)$$

The fact that this is the minimum value follows by the same argument, except with $\leq$ signs. ■

Let us apply this corollary to a specific example.

Example Let $G = \mathbb{Z}_m$, where $m = 2^r$. Let our generating set be $\{\pm 1, 0\}$, with starting point 0. (So we have a random walk on a cycle of length 2^r, where on each step we move forward, backward, or not at all, with probability $1/3$ of each. Now we need to define a stopping rule which is strong uniform. Here is the rule we will use.

> Keep walking until we reach either $\frac{m}{4}$ or $\frac{3m}{4}$.
> Continue walking until we reach $\frac{m}{8}$, $\frac{3m}{8}$, $\frac{5m}{8}$, or $\frac{7m}{8}$.
> Continue walking until we reach $\frac{m}{16}$, $\frac{3m}{16}$, $\frac{5m}{16}$, $\frac{7m}{16}$, $\frac{9m}{16}$, $\frac{11m}{16}$, $\frac{13m}{16}$, or $\frac{15m}{16}$.
>
> $\vdots$
>
> Continue walking until we reach an odd-numbered point.
>
> Continue walking until we encounter a step of $+1$ or 0.
>
> STOP.

Claim 11 *This stopping rule is strong uniform.*

Proof: Our walk begins at 0. By symmetry, we reach $\frac{m}{4}$ and $\frac{3m}{4}$ each with probability $1/2$, regardless of the time it takes us to get there. From that point, now we have to go a distance $\frac{m}{8}$ either forward or backward, and again both happen with equal probability, regardless of the time it takes to get there. So now it is equally likely that we stand at any of those four points. When we've finished all but the last stage of the protocol, we have an equal chance of being at any odd-numbered point, regardless of the time it has taken us to get there.

In doing the last stage of this procedure, we take some number of -1 steps, followed by either a 0 or $+1$. Let j be the number of -1 steps we take in this stage. If j is even, then right before the final step of the walk, we are at a randomly chosen odd point. When we add either $+1$ or 0, this takes us to a point which is randomly chosen among all possible points. Similarly, if j is odd, we are at a randomly chosen even point right before the final step, but again the final addition of $+1$ or 0 gets us to a point chosen uniformly from among all points. So for any value of j, we end up in a uniform distribution. Thus this rule defines a strong uniform stopping time. ■

What is the expected running time of this procedure? To do the first stage requires us to do a random walk until we've reached a distance of $\frac{m}{4}$ from the start. This takes expected time $O\left(\left(\frac{m}{4}\right)^2\right)$. The next stage is similar; it takes expected time $O\left(\left(\frac{m}{8}\right)^2\right)$. So the whole procedure takes time $O\left(\left(\frac{1}{16}+\frac{1}{64}+\frac{1}{256}+\cdots\right)m^2\right) = O\left(\frac{1}{12}m^2\right) = O\left(m^2\right)$.

By the corollary, $\tau_3 \leq E(\varkappa) = O(m^2)$. So for the mixing time for for the random walk in this example we have **mix**$= O(m^2)$.