CS265, Fall 2023

Class 15: Agenda and Questions

1 Announcements

- HW7 out now (due after break)
- Tomorrow (Friday) is deadline to change to/from CR/NC, and also the withdrawal deadline.
- Have a nice fall break!

2 Questions?

Any questions from the minilectures and/or the quiz? (Coupling, Mixing times)

3 Shuffling I

In this exercise we'll practice coming up with a coupling.

Group Work

Consider the following Markov Chain for shuffling a deck of n cards:

> At each timestep, choose a uniformly random card and move it to the top of the deck. (If you choose the top card, don't do anything).

Let X_t denote the state of the deck after t steps.

1. Convince yourself that this Markov chain is irreducible and aperiodic. What is the stationary distribution of $X_0, X_1, \ldots$?
2. We are going to bound the mixing time of this Markov chain using *couplings.* Come up with a coupling on this Markov chain that you think will "couple" quickly.

   Hint: *You might want to take inspiration from the graph-coloring example we saw, where we tried to make the same choice in both chains.*
3. For the coupling that you came up with, how long is it likely to take for the two chains to couple?

   Hint: *Assuming you came up with the coupling that I think you did, it might be helpful to remember the* coupon collector's problem. *In particular, if you are trying to collect n coupons, the probability that you need more than $2n \log n$ tries to collect them all is $o(1)$.*

4. Come up with a bound on τ_{mix} using your coupling. Can you show that $\tau_{mix} = O(n \log n)$? (You may assume that n is sufficiently large).

Group Work: Solutions

1. This Markov chain is irreducible since you can get from any deck to any other deck you want by choosing cards in the correct order. It is aperiodic since it has a self loop. The stationary distribution is uniform since if you start with a uniformly random deck and put a random card on top, you still have a uniformly random deck.
2. Our coupling will be as follows. Let $X_0, X_1, \ldots,$ be the chain as described above. Say we choose a particular card c_t at step t (for example, maybe c_t is the ace of spades). Then we define $Y_0, Y_1, \ldots,$ by taking the card c_t at time t and putting it on top of the deck.
3. The amount of time it takes for the chains to couple is at most the amount of time for each of the n cards in the deck to be chosen once. Indeed, if each card has been chosen once, then it will be in the same position in the deck in both the X chain and the Y chain. The amount of time that this takes is $O(n \log n)$, by the coupon collector's problem. As per the hint, we have, for any initial deck configurations s, s',

$$\Pr[\text{time to couple starting from } s, s' > 2n \log n] = \Pr[T_{s,s'} > 2n \log n] = o(1).$$

4. As we saw in the mini-lecture,

$$\Delta(t) \leq \max_{s,s'} \Pr[T_{s,s'} > t] = o(1)$$

when $t \geq 2n \log n$. Thus, $\tau_{mix} \leq 2n \log n$. (Assuming that n is sufficiently large so that the $o(1)$ term is at most $1/(2e)$).

4 Shuffling II

In this exercise, we'll see a different way to bound mixing times, other than coupling.

Group Work

Consider this **different shuffling scheme** for shuffling a deck of n cards:

> At each timestep, choose the top card and move it to a uniformly random position in the deck. (Note that it is possible that we choose to keep the top

card on the top, in which case nothing happens).

Let X_t denote the state of the deck after t steps.

1. Convince yourself that this chain is aperiodic and irreducible, and that the stationary distribution is uniform.
2. Let T be the first time at which the original bottom card of the deck is placed randomly somewhere. (That is, if the deck starts out with the ace of spades on the bottom, then time $T-1$ is the first time that the ace of spades is on the top).

   Argue that, at any time $t \geq T$, the deck is completely uniform. (That is, the Markov chain has converged exactly to its stationary distribution).
3. What is $\mathbb{E}[T]$?

   Hint: *Assuming that the ace of spades is originally on the bottom, write*

   $$\begin{aligned} T &= \textit{time it takes for the ace of spades to move to the second-from-bottom position} \\ &\quad + \textit{time it takes for the ace of spades to move from the second-from-bottom position} \\ &\qquad\quad \textit{to the third-from-bottom position} \\ &\quad + \cdots \end{aligned}$$

   and so on, and use linearity of expectation.
4. Notice that Markov's inequality implies that

   $$\Pr[T \geq 2e\mathbb{E}[T]] \leq 1/(2e).$$

   Explain why this implies that the mixing time of $X_0, X_1, \ldots$ is at most $2e\mathbb{E}[T]$.

   Hint: *This is not quite as simple as saying "we just said it was fully mixed at time T," since the formal definition of the mixing time is a bit different.*

   Hint: *Write $P_s^t = \Pr[T \leq t] \cdot \pi + \Pr[T > t] \cdot \sigma$, where π is the uniform (stationary) distribution and σ is some other distribution (the distribution of P_s^t conditioned on $T > t$). Then use that expression in the definition of $\Delta(t)$.*

Group Work: Solutions

1. This chain is irreducible since we can get from any deck to any other deck by building the deck "at the bottom" of the current deck, by iteratively putting the top card to be in the location that we want it. The chain is aperiodic since there is a self-loop. The stationary distribution is uniform since if we start with a uniformly random deck and move the top card to a random place, the deck is still uniformly

random.

2. Whenever $t \geq T$, every card has been placed at least once randomly, and so the overall distribution is random. (Formally, you can show this by induction, with the inductive hypothesis "everything below the original bottom card is in a uniformly random order.")

3. Following the hint, we compute

$$\mathbb{E}[\text{time it takes for ace of spades to move to 2nd-bottom}] = n,$$

because the probability that the ace of spaces moves up one spot is $1/n$. Similarly, the expected time it takes to move from the second-to-bottom to third to bottom is $n/2$. Altogether, we have

$$\mathbb{E}[T] = \sum_{i=1}^{n} \frac{n}{i} = \Theta(n \log n).$$

4. Write

$$P_s^t = \Pr[t > T] \cdot \pi + \Pr[t \leq T] \cdot \sigma,$$

where σ is the distribution of P_s^t conditioned on the event that $t \leq T$. Then we can use the definition of $\Delta(t)$:

$$\begin{aligned}
\Delta(t) &= \max_s \|\pi - P_s^t\| \\
&= \max_s \|\pi - (\Pr[t > T]\pi + \Pr[t \leq T]\sigma)\| \\
&= \max_s \|\Pr[t \leq T](\pi - \sigma)\| \\
&= \Pr[t \leq T]\|\pi - \sigma\| \\
&\leq \Pr[t \leq T].
\end{aligned}$$

Thus, if we choose $t = 2e\mathbb{E}[T] = \Theta(n \log n)$, we have

$$\Delta(t) \leq 1/(2e)$$

by the expression above.

5 Strong Stationary Stopping Times

A random variable T is a **strong stationary stopping time** if:

- The event $T = t$ depends only on $X_1, \ldots, X_t$
- For all s, $\Pr[X_t = s | t \geq T] = \pi(s)$.

That is, you can tell if T has occurred based only on the steps so far, and once T occurs, the chain is completely mixed.

In the previous group work, we essentially showed that:

Theorem 1. *Let $X_0, X_1, \ldots$ be a Markov chain with stationary distribution π and let T be a strong stationary stopping time for this chain. Then*

$$\Delta(t) \leq \Pr[T > t].$$

Group Work

(Bonus, if time.) Consider the following shuffling step for a deck of n cards:

- Assign each card a label "L" or "R," independently and uniformly at random.
- Put all the cards labeled "L" to the left, *preserving their relative order.* Put all the cards labeled "R" to the right, again preserving relative order.
- Put the "L" stack on top of the "R" stack.

You might recognize the as the *inverse* of a standard riffle shuffle. That is, if you do this process in reverse, you cut the deck at a random point and randomly interleave the two parts of the deck.

Use the method of strong stationary stopping times to show that the mixing time of this shuffle (repeated t times) is $O(\log n)$.

Group Work: Solutions

Consider doing this shuffle repeatedly t times. We end up with a string in $\{L, R\}^t$ associated to each element. If two elements have the same string, then their order relative to each other will be the same as in the original deck. On the other hand, if two elements have different strings, then at some point they appeared in different halves of the deck, so their order relative to each other will be random (since it's equally likely that one was L and the other was R as it is to be the other way around).

Thus, a strong stationary stopping time T can be defined as:

> T is the first time that all of the elements have distinct strings.

We can compute the probability that T is large as follows: for any two cards x, y, the probability that those two cards have the same string after t steps is $1/2^t$. Thus, if we choose $t \geq 2 \log n + 10$ (say), the probability that any two cards have the same string is at most

$$n^2 2^{-2 \log n + 10} = 2^{-10} \leq 1/(2e).$$

So our theorem above guarantees that the mixing time of this shuffle is at most $O(\log n)$, as desired.