STATS 217: Introduction to Stochastic Processes I

Lecture 22

Last time

- Consider an irreducible transition matrix P on a finite state space S with unique stationary distribution π. A strong stationary time for the starting state x is a stopping time (with respect to the chain and auxiliary randomness) satisfying

$$\mathbb{P}[\tau = t, X_\tau = y \mid X_0 = x] = \mathbb{P}[\tau = t \mid X_0 = x] \cdot \pi(y).$$

- Suppose that τ is a strong stationary time for the starting state x. Then,

$$\mathrm{TV}(X_t \mid X_0 = x, \pi) \leq \mathbb{P}[\tau > t \mid X_0 = x].$$

Riffle shuffles

Recall riffle shuffles from Problem 8 on Homework 6. For a deck of size n:

- Split the deck into two parts according to Binomial$(n, 1/2)$.
- Hold one part in your left hand and the other part in your right hand with the bottom of each deck facing the table.
- Merge the two parts by dropping cards in sequence, where if you have L cards in your left hand and R cards in your right hand at some point, then the probability that the next card comes from your left hand is $L/(L+R)$.

On the homework, you showed that this is an irreducible and aperiodic chain and that the unique stationary distribution is the uniform distribution on all permutations of the cards.

Riffle shuffles

If you attempted this problem, you probably also came up with the following equivalent description:

- Split the deck into two parts according to $\text{Binom}(n, 1/2)$.
- Suppose that there are M 'top' cards and $n - M$ 'bottom' cards. Note that there are $\binom{n}{M}$ possible ways to interleave these cards so that the relative order of the top cards is preserved and the relative order of the bottom cards is also preserved.
- Choose one of these $\binom{n}{M}$ ways to interleave/riffle uniformly at random.

Equivalence of descriptions

- Consider the case when the binomial random variable in both methods is M.
- We know that for the second method, the probability of getting to any riffle is exactly $1/\binom{n}{M}$.
- What about for the first method? Once you know M, then for any valid final outcome, there is exactly one sequence of left and right drops that result in that sequence.
- Consider a sequence of left and right drops. Let I denote the indices corresponding to a left drop and let J denote the indices corresponding to a right drop. Then, the probability of this sequence is

$$\prod_{i \in I} \frac{L_i}{L_i + R_i} \prod_{j \in J} \frac{R_j}{L_j + R_j} = \frac{M(M-1)\ldots,1 \times (n-M)(n-M-1)\ldots 1}{n(n-1)\ldots 1}$$

$$= \binom{n}{M}^{-1}$$

Reverse riffle shuffle

For constructing a strong stationary time, it will be more convenient to work with the inverse riffle shuffle. For a deck of size n:

- Label the cards with independent and uniform bits $b_1, \ldots, b_n \in \{0, 1\}$.
- Move all the cards labelled 0 above all the cards labelled 1 while preserving the relative order within the 0 cards and the 1 cards.

This shuffle is the reverse of the riffle shuffle in a precise sense discussed in Problem 4 of Homework 8.

Here's the idea: the number of cards labelled with 0 has distribution $\text{Binom}(n, 1/2)$. Moreover, conditioned on the value M of this binomial random variable, the location of the M cards labelled 0 is uniform among the $\binom{n}{M}$ possibilities.

Reverse riffle shuffle

- Note that the reverse riffle shuffle is irreducible and aperiodic.
- By the general results of Problem 4 on Homework 8, the uniform distribution on permutations is the unique stationary distribution of the reverse ruffle shuffle, and moreover, the ε-mixing time of the reverse riffle shuffle is the same as that of the riffle shuffle.
- Therefore, it suffices to bound the ε-mixing time of the reverse riffle shuffle.
- We will do this by constructing a suitable strong stationary time for the reverse riffle shuffle.

Strong stationary time for the reverse riffle shuffle

- As a first step towards this, consider what happens after two steps of the reverse riffle shuffle.
- Now, every card i is labelled with a binary string of length 2, denoted by $b_i^2 b_i^1$, where b_i^2 is the bit assigned to it in the second round and b_i^1 is the bit assigned to it in the first round.
- Note that all cards with the string 00 are above all cards with the string 01, which are above all cards with the string 10, which are above all cards with the string 11.
- Within each category (00, 01, 10, 11), the original relative order is preserved.

Strong stationary time for the reverse riffle shuffle

- This suggests the following candidate for a strong stationary time: let τ be the first time when no two cards have the same binary string (of length τ) assigned to them.
- Clearly, τ is a stopping time.
- To check that τ is a strong stationary time, note that given $\tau = t$, we know that each card has a different t-bit string (and that removing the most recent bit leads to at least 2 cards having the same $(t-1)$-bit string). Since the t-bit strings are generated using i.i.d. bits, each resulting permutation must be equally likely by symmetry.

Mixing time of the reverse riffle shuffle

- It remains to estimate $\mathbb{P}[\tau > t]$, or equivalently, $\mathbb{P}[\tau \leq t]$.
- If $\tau \leq t$, then different labels are assigned to all n cards after t rounds. Since each card i is equally likely to get any of the 2^t possible labels, the probability that this happens is

$$\mathbb{P}[\tau \leq t] = 1 \times (1 - 1/2^t) \times (1 - 2/2^t) \times \dots (1 - (n-1)/2^t).$$

- For $2^t = n^2/c^2$ for $c > 0$ (i.e. $t = 2\log_2(n/c)$), this simplifies to

$$\mathbb{P}[\tau \leq t] = e^{-c^2/2}(1 + O(1/n)).$$

- For $\tau_{\text{mix}}(1/4)$, we want $\mathbb{P}[\tau > t] \leq 1/4$, and hence, we want the right hand side to be $3/4$.
- $c = 0.75$ works, and this shows that

$$\tau_{\text{mix}}(1/4) \leq 2\log_2(4n/3).$$

Mixing time of the reverse riffle shuffle

- Our analysis is suboptimal: the well-known paper of Bayer and Diaconis (*Trailing the dovetail shuffle to its lair*) shows that the mixing time is $1.5 \log_2 n + o(\log_2 n)$.
- Here is a numerical computation (to 4 digits of precision) of the total variation distance after t riffle shuffles from the paper of Bayer and Diacnois:

$$t = 1; 1.0000 \qquad t = 2; 1.0000 \qquad t = 3; 1.0000 \qquad t = 4; 1.0000$$

$$t = 5; 0.9237 \qquad t = 6; 0.6135 \qquad t = 7; 0.3341 \qquad t = 8; 0.1672$$

$$t = 9; 0.0854 \qquad t = 10; 0.0429 \qquad t = 11; 0.0215 \qquad t = 12; 0.0108$$