Lecture 35 - Mixing time of the Riffle Shuffle

Scribe: Csaba Biro

On this lecture we consider a model of the Riffle Shuffle of a deck of cards and we will analyse the mixing time. The Riffle Shuffle is a popular way to shuffle a deck of cards. The shuffler divides the deck into two part. The top part of the deck is placed in the left hand, the remaining stays in the right hand. Cards are then alternatively interleaved from the left and right hands, but not neccessarily perfectly, so several cards may be dropped from each hand at a time. This process is repeated a couple of times.

A freedom left to decide how the deck is divided (with what probabilities) and how many cards are dropped each time. A (hopefully) realistic model is the following:

- Cut the deck with c cards in left hand with probability distibution $\frac{1}{2^n}\binom{n}{c}$ (binomial distribution).
- Drop cards from left hand or right hand with probability proportional to the number of cards in each hand.

So in our Markov Chain, the states will be permutations of cards and the transition probabilities are the probabilities that a permutation is reached from another in one shuffle. The stationary distribution is uniform.

Now we will find a lower bound for the total variation distance. Suppose that the cards are numbered form 1 to n and the starting state is the identity. A *rising sequence* is a maximal increasing subsequence of numbers in a permutation. What is the maximum number of disjoint rising sequences after k shuffles? For $k = 0$ it is 1, for $k = 1$, it at most 2. At each shuffle, we can at most double the number of sequences, so the maximum number of disjoint sequences is 2^k.

Suppose that $A \subset \Omega$ is the set that could be reached after k steps of the Markov Chain. Note, that A is *exactly* the set of permutations in which there are at most 2^k disjoint rising sequences. This is because of the previous paragraph and because every sequence that has at most 2^k rising sequences can be reached after k steps.

Recall that

$$||P^k(x,\cdot) - \pi||_{TV} = \frac{1}{2}\sum_{y\in\Omega}|P^k(x,y) - \pi(y)| = \sup_{B\subset\Omega}|\pi(B) - P^k(x,B)|.$$

Set $B = \Omega\backslash A$.

$$||P^k(x,\cdot) - \pi||_{TV} \geq \pi(\Omega\backslash A) - P^k(x,\Omega\backslash A) = \pi(\Omega\backslash A) = 1 - \pi(A)$$

It is known that the number of ways to get exactly R rising sequences in a permutation of the numbers $1,\ldots,n$ is

$$F_n(R) = \sum_{j=0}^{R}(-1)^j\binom{n+1}{j}(R-j)^n.$$

So the number of ways to get at most 2^k rising sequences is $T_n(k) = \sum_{n=1}^{2^k} F_n(R)$. Hence

$$||P^k(x,\cdot) - \pi||_{TV} \geq 1 - \pi(A) = 1 - \frac{T_n(k)}{n!}.$$

Actual calculations show that if $n = 52$ and $k = 1,2,3$ or 4, then the total variation distance is ≥ 0.99, for $k = 4$ it is about 0.38 and for $k \geq 6$ then it is close to 0. So after 4 shuffles, the deck will be very far from randomness (in uniform distribution) and the first time we may have a chance of an adquately shuffled deck is after 6 shuffles.

In fact it is not so surprising that 5 shuffles is not enough. After 5 shuffles the number of rising sequences is at most 32. But the permutation $(52, 51, \ldots, 1)$ has 52 rising sequences, so this permutation cannot be reached after 5 steps, i. e. its probability is 0. There are several other unreachable permutations which suggests that we must be pretty far from the uniform distribution.[1]

Now we will define the inverse shuffle. Label each card with 0 or 1 with probabilities 1/2. Then take cards labeled by 0 out of the deck and place them to the top, keeping their original order. It is easy to see that the Markov Chain whose transition probabilities (among the permutations of cards) are defined by this inverse shuffle is the reversed chain of the shuffle chain. We can use this chain to study the original one, due to the following lemma.

Lemma 35.1. *Let G be a finite group and S is generating set (it may not contain inverses). Consider the walk on G with S. Then*

$$||P^t - \pi||_{TV} = ||P^{*t} - \pi||_{TV}$$

where $P^*(x, y) = P(y, x)$, i. e. walk with S^{-1} then get the same variation distance as walk with S.

Proof. For all $x, y \in G$ we have $||P^t(x, \cdot) - \pi||_{TV} = ||P^t(y, \cdot) - \pi||_{TV}$, because every vertex is the same on the Cayley graph. Thus

$$\begin{aligned}
||P^t - \pi||_{TV} &= \frac{1}{|\Omega|} \sum_x \frac{1}{2} \sum_y |P^t(x, y) - \pi(y)| \\
&= \frac{1}{|\Omega|} \sum_y \frac{1}{2} \sum_x |P^{*t}(y, x) - \pi(y)| \\
&= \frac{1}{|\Omega|} \sum_y \frac{1}{2} \sum_x |P^{*t}(y, x) - \pi(x)| \\
&= ||P^{*t} - \pi||_{TV}.
\end{aligned}$$

□

References

[1] This paragraph is added by the note-taker.