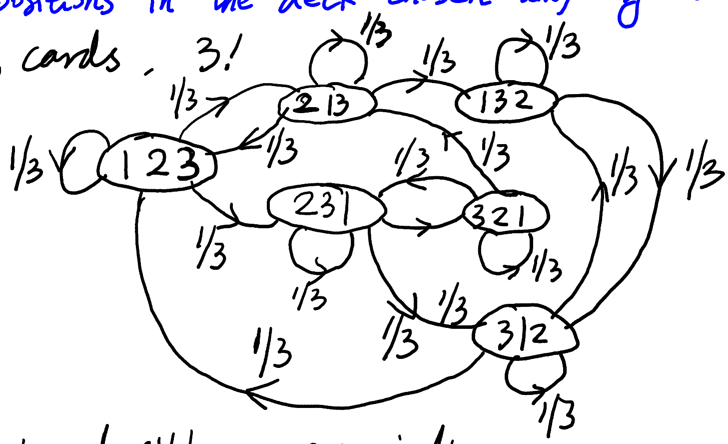
Shuffling

Tuesday, January 11, 2011
9:19 PM

Top-to-Random

policy: take the top card and insert it at one of the n positions in the deck chosen uniformly at random.

3 cards, 3!



irreducible, aperiodic.

but ~~it~~ it is not reversible:

If we insert the top card into (say) the middle of the deck, we cannot bring the card back to the top in one step.

However, notice that every permutation y can be obtained in one step, from exactly n different permutations, since every non-zero transition probability is $\frac{1}{n}$. this implies that $\sum_x P(x,y)=1$.

$$\pi(y)=\sum_x \pi(x)P(x,y)$$

$$\begin{cases} \sum_y P(x,y)=1 \\ \sum_x P(x,y)=1. \end{cases}$$

In fact, π is uniform iff P is doubly stochastic.

Consider the following coupling:

Given two copies X_t and Y_t of the chain in different states, choose a position j uniformly at random from 1 to n and simultaneously move the top card into position j in both chains.

This is a valid coupling, because each chain individually acts as the original shuffling Markov chain.

claim: the mixing times of the original Markov chain and the reverse Markov chain are identical.

Consider the reverse shuffling:
- Pick a card c from the deck uniformly at random.
- Move card c to the top of the deck.

Define the coupling by making both X_t and Y_t choose the same card c (which of course is not necessarily in the same position in both decks) and move it to the top.

Now, the key observation is the following: once a card has been chosen in the coupling, this card will be in the same position in both decks for the rest of time.

T_{xy} is therefore once again dominated by the coupon collector random variable for n coupons. This leads to

$$\tau_{mix} \leq n \ln n + o(n)$$

and

$$\tau(\varepsilon) \leq n \ln n + \lceil n \ln \varepsilon^{-1} \rceil \quad \blacksquare$$

Random Transpositions

Pick two cards i and j uniformly at random with replacement. and switch cards i and j.

irreducible: every permutation can be expressed as a product of transposition.
aperiodic: since we may chose i=j, so the chain has self-loops.
invertible: the random transpositions are invertible.

$$P(x,y) = P(y,x) \Rightarrow \text{uniform stationary distribution.}$$

An equivalent, more convenient description is the following:
- pick card c and position p uniformly at random.
- exchange card c with the card at position p in the deck.

It is easy to define a coupling using this second definition: make X_t and Y_t choose the same c and p at each step. This coupling ensures that the distance between X and Y is non-increasing. More explicitly, writing $d_t = d(X_t, Y_t)$ for the number of positions at which the two decks differ, we have the following case analysis:

① If card c is in the same position in both decks, then $d_{t+1} = d_t$.

② If card c is in different positions in the two decks, there are two possible subcases.

(a) If the card at position p in both decks is the same, then $d_{t+1} = d_t$.

(b) otherwise, $d_{t+1} \le d_t - 1$

Thus, we get a decrease in distance only in case 2(b), and this occurs with probability

$$Pr\{ d_{t+1} < d_t \} = \left(\frac{d_t}{n}\right)^2$$

There, the time for d_t to decrease from value d is stochastically dominated by a geometric random variable with mean $\left(\frac{n}{d}\right)^2$. This implies that $E[T_{xy}] \le \sum_{d=1}^{n} \left(\frac{n}{d}\right)^2$, which is $O(n^2)$

$$Pr\{T_{xy} > cn^2\} < \frac{E[T_{xy}]}{cn^2} = \frac{\sum_{d=1}^{n}\left(\frac{n}{d}\right)^2}{cn^2} = \frac{1}{c}\sum_{d=1}^{n}\frac{1}{d^2} = \frac{1}{2e}$$

$$\Rightarrow \quad \tau_{mix} \le cn^2 \quad \left(\text{where } \frac{1}{c}\sum_{d=1}^{n}\frac{1}{d^2} = \frac{1}{2e} \right)$$

Remarks: Actually, for this shuffle it is known that

$$\tau_{mix} \sim \frac{1}{2} n \ln n.$$

So our analysis in this case is off by quite a bit.

Exercise: Design a better coupling that gives $\tau_{mix} \le O(n \ln n)$.

Background: Random walks on Groups

1. a group is a set G endowed with an associative operation $G \times G \to G$ and an identity $id \in G$ such that

for all $g \in G$,

(i) $id \cdot g = g$ and $g \cdot id = g$

(ii) there exists an inverse $g^{-1} \in G$ for which $g \cdot g^{-1} = g^{-1} \cdot g = id$.

2. Given a probability distribution μ on a group $(G, \cdot)$, we define the random walk on G with increment distribution μ as follows: it is a Markov chain with state space G and which moves by multiplying the current state on the left by a random element of G selected according to μ. Equivalently, the transition matrix P of this chain has entries

$$P(g, hg) = \mu(h)$$

for all $g, h \in G$.

3. (proposition). Let P be the transition matrix of a random walk on a finite group G and let U be the uniform probability distribution on G. Then U is a stationary distribution for P.

proof: Let μ be the increment distribution of the random walk. For any $g \in G$,

$$\sum_{h \in G} U(h) P(h, g) = \frac{1}{|G|} \sum_{k \in G} P(k^{-1}g, g) = \frac{1}{|G|} \sum_{k \in G} \mu(k) = \frac{1}{|G|} = U(g)$$

For the first equality, we re-indexed by setting $k = gh^{-1}$.

4. Let P be the transition matrix of a random walk on a group G with increment distribution μ and let $\hat{P}$ be that of the walk on G with increment distribution $\hat{\mu}$. Let π be the uniform distribution on G. Then for any $t \geq 0$

distribution on G. Then for any $t \geq 0$,

$$\|P^t(id,\cdot)-\pi\|_{TV} = \|\hat{P}^t(id,\cdot)-\pi\|_{TV}$$

proof: Let $(X_t)=(id, X_1, \cdots)$ be a Markov chain with transition matrix P and initial state id. We can write $X_k = g_1 g_2 \cdots g_k$, where the random elements $g_1, g_2, \cdots \in G$ are independent choices from the distribution μ.

Similarly, let (Y_t) be a chain with transition matrix $\hat{P}$, with increments $h_1, h_2, \cdots \in G$, chosen independently from $\hat{\mu}$.

For any fixed elements $a_1, \cdots, a_t \in G$.

$$P\{g_1=a_1, \cdots, g_t=a_t\} = \hat{P}\{h_1=a_t^{-1}, \cdots, h_t=a_1^{-1}\}$$

by the definition of $\hat{P}$.

Summing over all strings such that $a_1 a_2 \cdots a_t = a$ yields

$$P^t(id,a) = \hat{P}^t(id,a^{-1})$$

Hence,

$$\sum_{a\in G}\left|P^t(id,a)-|G|^{-1}\right| = \sum_{a\in G}\left|\hat{P}^t(id,a^{-1})-|G|^{-1}\right| = \sum_{a\in G}\left|\hat{P}^t(id,a)-|G|^{-1}\right| \blacksquare$$