

COL863: Special Topics in Theoretical Computer Science
Rapid Mixing in Markov Chains
II semester, 2016-17
Minor I

Total Marks 100

Due: On moodle at 11:55AM, 6th February 2017

Problem 1.1 (20 marks) *Ex 4.4 of LPW Edition 2.*

We know that \exists a coupling $(X, Y) \parallel \Pr[X \neq Y] = \|\mu - \nu\|_{TV} \forall \mu, \nu$
Specifically, let (X^i, Y^i) be such a coupling for μ^i, ν^i . Since each X^i is independent and so is each Y^i , we can define $(\mathbf{X} \equiv X^1, X^2, \dots, X^n, \mathbf{Y} \equiv Y^1, Y^2, \dots, Y^n)$ as a general coupling on μ, ν .
We also know that \forall couplings (X, Y) on μ, ν , $\|\mu - \nu\|_{TV} \leq \Pr[X \neq Y]$
Hence,

$$\begin{aligned} \|\mu - \nu\|_{TV} &\leq \Pr[X \neq Y] \\ \Pr[\mathbf{X} \neq \mathbf{Y}] &\leq \sum_{i=0}^n \Pr[X_i \neq Y_i] \\ \sum_{i=0}^n \Pr[X_i \neq Y_i] &= \sum_{i=0}^n \|\mu - \nu_i\|_{TV} \\ \therefore \|\mu - \nu\|_{TV} &\leq \sum_{i=0}^n \|\mu - \nu_i\|_{TV} \end{aligned}$$

All that is left to show is that (\mathbf{X}, \mathbf{Y}) is a valid coupling.

$$\begin{aligned} \Pr[\mathbf{X} = x] &= \sum_{\mathbf{Y}} \Pr[\mathbf{X} = x, \mathbf{Y}] \\ &= \sum_{\mathbf{Y}} \prod_{j=0}^n \Pr[X_j = x_j, Y_j] \\ &= \prod_{j=0}^n \sum_{\mathbf{Y}} \Pr[X_j = x_j, Y_j] \\ &= \prod_{j=0}^n \mu_j \\ &= \mu \end{aligned}$$

Problem 1.2 (40 marks) *You have a deck that contains n cards. The shuffling process is*
Make two independent choices of cards and interchange them.

Note that the two choices may be the same, in which case there is no change in the configuration of the deck. Use a coupling based method to prove an upper bound on $t_{\text{mix}}(\epsilon)$ for this chain.

The markov chain used in this answer has a state space consisting of all possible permutations of cards - referred to as deck. The transition matrix is defined by the given shuffling process. I argue that the given process is equivalent to another. Let us call the deck of cards with given shuffling process as X . We will couple it with another deck of cards with its own shuffling process - say Y .

To bound $t_{\text{mix}}(\epsilon)$, we will use the coalescence time (τ_c) of (X, Y) . We know that

$$d(t) \leq \max_{x,y} \mathbb{P}[\tau_c > t]$$

Notice that the process of picking two cards is the same as choosing two positions in the deck or one card and one position. Hence, their mixing times are also the same. We create a coupling, (X, Y) where X, Y denote the state of the deck, as follows:

If $X_t \neq Y_t$, pick a card, say, c uniformly at random. Then, pick a position, say p , uniformly at random. Swap the c with the card at p in both decks.

Let δ_t denote the number of positions at which the decks X, Y agree at time t . I claim that δ_t is a coupon collection problem. We will first show that δ_t is increasing (or constant), and then we'll find the probabilities associated with the chain for δ_t . That will allow us to find τ_c for our couple (X, Y) and hence bound $d(t)$. Notice that:

- If the chosen card is at different places, then, if the cards at the chosen position are the same in both decks, then $\delta_{t+1} = \delta_t$. If the card at the chosen position are different, then $\delta_{t+1} = \delta_t + 1$.
- If the chosen card is at the same place in both decks, then the position picked is irrelevant to δ_t . Hence, $\delta_{t+1} = \delta_t$

While that's nice, we aren't done yet. We can not use the coupon collector's formula for $\mathbb{P}[\tau_c > t]$ because we don't know if δ_t follows the same distribution as the coupon collector from LPW. $\mathbb{P}[\delta_{t+1} \geq \delta_t]$ = probability that the chosen card is at different places and that the position chosen has differing cards in that position.

$$= \left(1 - \frac{\delta_t}{n}\right) * \left(1 - \frac{\delta_t}{n}\right)$$

because the probability of choosing a card that is at different places is the probability of choosing a card that is not agreed upon between the two decks. Similarly for choosing a position which has different cards.

Unfortunately, δ_t does not follow the same distribution as the coupon collector in LPW. But we can still bound $d(t)$ as follows:

$$d(T) \leq \max_{x,y} P_{x,y}[\tau_c > T] \leq \max_{x,y} \frac{E_{x,y}[\tau_c]}{T}$$

The maxima will have $\delta_0 = 0$ since the chain for δ_t is markov itself. We will now calculate $E_{x,y}[\tau_c]$. Let t_i be the time taken for the coupon collecting chain to go from state i to state $i + 1$. Clearly, $\tau_c = \sum_{i=\delta_0}^n t_i$. When we take the max over x, y , we would have $\delta_0 = 0$. Hence, $\max_{x,y} E_{x,y}[\tau_c] = \max_{x,y} E_{x,y}[\sum_{i=0}^{i=n} t_i]$ Since time taken for transition from state i to $i + 1$ follows a geometric distribution with parameter $(1 - \frac{i}{n}) * (1 - \frac{i}{n})$,

$$E[t_i] \leq \left(\frac{n}{n-i}\right)^2$$

$$\implies d(T) \leq \frac{\sum_{i=0}^{n-1} \left(\frac{n}{n-i}\right)^2}{T}$$

This is tricky, but notice that $0 \leq (1/k)^2 \leq \frac{1}{k^2-k} = 1/(k-1) - 1/k$ for $n \geq 2$. The RHS is a telescopic sum, and hence $\sum_{i=2}^{\inf} \frac{1}{i^2}$ is $\leq 1 + 1/n = \mathcal{O}(1)$. Our tricky sum, $\sum_{i=1}^n \frac{1}{i^2}$ must hence be less than some constant.

$$\implies d(T) \leq \frac{cn^2}{T}$$

$$\implies d\left(\frac{c}{\epsilon}n^2\right) \leq \epsilon$$

$$\therefore t_{\text{mix}}(\epsilon) \leq \frac{c}{\epsilon}n^2 = \mathcal{O}(n^2)$$

This bound has one drawback. It is not exactly a coupon collector chain, which is why $t_{\text{mix}}(\epsilon)$ is $\mathcal{O}(n^2)$. Had we made it a coupon collector chain with $P[\delta_{t+1} = \delta_t + 1] = 1 - \frac{\delta_t}{n}$, we would have had an upper bound that is $\mathcal{O}(n \log n)$

Notice that it is a valid coupling because both decks follow the same transition matrix since they pick a card and a position uniformly at random and perform a swap. Hence, our bounds can be used safely. I was unable to come up with a coupling where both X and Y follow the same markov chain and result in the desirable $P[\delta_{t+1} = \delta_t + 1]$

Problem 1.3 (40 marks) Again: You have a deck that contains n cards. This time let us call the n positions in the deck $0, 1, \dots, n-1$. The (lazy) shuffling process is

With probability $1/2$ do nothing. Otherwise pick i uniformly at random from $0, 1, \dots, n-1$ and interchange the cards in positions i and $(i+1) \bmod n$.

Use a coupling based method to prove an upper bound on $t_{mix}(\epsilon)$ for this chain.

We couple markov chains whose state space is defined as all possible permutations of cards and the transition matrix is as defined by the shuffling procedure. Consider the coupling (X, Y) . We will define the coupling in a little bit. Before that, let δ_t be the number of matching positions between the two decks X, Y , i.e., the number of positions which have the same card in X and Y . τ_c , the coupling time is the time at which $\delta_t = n$. Define our coupling as follows:

- With probability $\frac{1}{2}$ decide which deck to modify.
- If X , then pick a position, say p , uniformly at random. If the either p or $p+1 \bmod n$ has the same card in Y , then swap p and $p+1 \bmod n$ in both decks. If it is not agreed upon, then swap p and $p+1 \bmod n$ only in X while leaving Y alone.
- If Y , then do the same as in X except that if either p or $p+1 \bmod n$ matches, then don't swap any cards.

This is a valid coupling because probability that X will make a swap is $\frac{1}{2}$. The swap is chosen randomly. For Y , the probability of a swap happening is $\frac{1}{2}[\frac{\delta_t}{n}] + \frac{1}{2}[1 - \frac{\delta_t}{n}] = \frac{1}{2}$. Again, the swap is chosen randomly. Hence, the distribution over X and Y are the same as the given shuffling procedure. Notice that in this coupling, δ_t never decreases. It would decrease only if a matching card got swapped in one deck but not the other. The coupling is built to prevent such a case. All that is left is to analyse the probability that δ_t increases. Instead of treating δ_t as a markov chain like we did in Q2, we will look at distance between the positions of the same card in the two decks. If we treat positions on the deck as a cycle, consider the cyclic distance between the positions of card c in X and Y . Let $\Delta_{v,t}$ be the cyclic distance between card v in the two decks at time t . $\Delta_{v,t}$ follows the distribution:

$$\Delta_{v,t+1} = \begin{cases} \Delta_{v,t} & \text{with probability } 1 - \frac{2}{n} \\ \Delta_{v,t} + 1 & \text{with probability } \frac{1}{n} \\ \Delta_{v,t} - 1 & \text{with probability } \frac{1}{n} \end{cases}$$

However, the absorbing states for $\Delta_{v,t}$ are n and 0 . Let T_v be the time taken for card v to get matched in both decks. Then, $\tau_c = \max_v T_v$. Notice that $\Delta_{v,t+1}$ is a lazy gambler's ruin chain with probabilities that are not $\frac{1}{2}$. We will deal with the chain soon. We know that

$$d(T) = \max_{x,y} P_{x,y}[\tau_c > T] = \frac{E_{x,y}[\tau_c]}{T} \implies d(T) = \frac{E_{x,y}[\tau_c]}{T} = \frac{E_{x,y}[\max_v T_v]}{T} \leq \frac{\sum_v E_{x,y}[T_v]}{T} = \frac{n E_{x,y}[T_v]}{T}$$

(note: I dropped the $\max_{x,y}$ for brevity.)

We must now calculate $E_{x,y}[T_v]$. Recall that T_v is the time that the chain takes to reach absorbing state 0 or n . We want to calculate $E_k[T_v]$, where k is the initial cyclic distance of card v between X, Y . Let $e_k = E_k[T_v]$. Clearly, $e_0 = e_n = 0$. Also,

$$e_k = \frac{1}{n}e_{k-1} + \frac{1}{n}e_{k+1} + \left(1 - \frac{2}{n}\right)e_k \implies 2e_k = n + e_{k-1} + e_{k+1}$$

Drawing inspiration from the fact that this is essentially a slowed down Gambler's ruin, it is easy to verify that

$$e_k = \frac{n * k * (n - k)}{2}$$

e_k is maximized at $k = n/2$, and has a value of $\frac{n^3}{8}$

$$\therefore d(T) \leq \frac{n^4}{8T} \implies d\left(\frac{n^4}{8\epsilon}\right) \leq \epsilon$$

$$\therefore t_{\text{mix}}(\epsilon) \leq \frac{n^4}{8\epsilon}$$

This bound is slightly weak and a stronger analysis of $\max_v T_v$ will probably result in a better bound.
