Problem Set 7 CS265, Winter 2021-2022
Due: March 2nd (Wednesday) at 11:30 (Pacific Time)

Please follow the homework policies on the course website.

---

1. **(4 pt.) Coin Tosses Modulo k**

   A biased coin with probability of heads $p \in (0, 1)$ is tossed repeatedly. Let X_t denote the number of heads obtained in the first t tosses. Prove that for any integer $k \ge 1$ and $0 \le j \le k - 1$,

   $$\lim_{t \to +\infty} \Pr[(X_t \bmod k) = j] = \frac{1}{k}.$$

   [**HINT:** *If your proof applies the fundamental theorem of Markov chains at some point, make sure to prove that the Markov chain that you defined satisfies all the preconditions.*]

2. **(12 pt.) Fundamental Theorem of Markov Chains: A Special Case**

   Let $X_0, X_1, \ldots$ be a Markov chain over n states (labeled $1, 2, \ldots, n$) with transition matrix $P \in \mathbb{R}^{n \times n}$, i.e., for any $t \ge 0$, $\Pr[X_{t+1} = j | X_t = i] = P_{ij}$. In addition, we assume that $P_{ij} > 0$ for all $i, j \in [n]$, and define $p_{\min} := \min_{i,j \in [n]} P_{ij} > 0$. In this problem, we will prove part of the fundamental theorem of Markov chains for this special case. In particular, we will show that there exists a unique stationary distribution π such that for all $i, j \in [n]$,

   $$\lim_{t \to +\infty} \Pr[X_t = j | X_0 = i] = \pi_j.$$

   (a) **(1 pt.)** As a warmup, show that the assumption $P_{ij} > 0$ for all $i, j \in [n]$ implies that the Markov chain is irreducible and aperiodic. Thus, the assumption that we made is not weaker than the one in the original theorem.

   (b) **(4 pt.)** Suppose that vectors $a, b \in \mathbb{R}^n$ satisfy $\sum_{i=1}^n a_i = 0$ and $\min_{i \in [n]} b_i \ge \epsilon > 0$. Prove that $|\sum_{i=1}^n a_i b_i| \le \sum_{i=1}^n |a_i| b_i - \frac{\epsilon}{2} \sum_{i=1}^n |a_i|$.
   [**HINT:** *It might be useful to define $S^+ := \{i \in [n] : a_i \ge 0\}$ and $S^- := \{i \in [n] : a_i < 0\}$, and then relate $\sum_{i=1}^n a_i b_i$ to $\sum_{i \in S^+} |a_i| b_i$ and $\sum_{i \in S^-} |a_i| b_i$.*]
   [**HINT:** *It is possible to prove a stronger bound with the $\frac{\epsilon}{2}$ factor replaced by ϵ.*]

   (c) **(2 pt.)** Let $a = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$ be a row vector that satisfies $\sum_{i=1}^n a_i = 0$. Prove that $\|aP\|_1 \le (1 - np_{\min}/2)\|a\|_1$.
   [**HINT:** *Use the previous part.*]

   (d) **(3 pt.)** Prove that there exists an n-dimensional row vector $\pi = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_n \end{bmatrix}$ such that: (1) $\pi = \pi P$; (2) $\sum_{i=1}^n \pi_i = 1$.
   [**HINT:** *First prove the existence of a non-zero vector π satisfying $\pi = \pi P$, and then show that the second condition can be satisfied by scaling π. For the first step, you may use the following fact without proof: if λ is an eigenvalue of a square matrix A, λ is also an eigenvalue of A^T. Part 2c might be helpful for the second step.*]

   (e) **(2 pt.)** Let $v = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ be a row vector that satisfies $\sum_{i=1}^n v_i = 1$. Let π be a vector chosen as in Part 2d. Prove that $\lim_{t \to +\infty} vP^t = \pi$. Then, derive that for all $i, j \in [n]$,

   $$\lim_{t \to +\infty} \Pr[X_t = j | X_0 = i] = \pi_j.$$

[**HINT:** *Apply Part 2c to* $(v-\pi),(v-\pi)P,(v-\pi)P^2,\ldots$]

(f) **(0 pt.) [Optional: this won't be graded.]** Extend the proof to the general case, where the Markov chain is irreducible and aperiodic but $P_{ij}>0$ might not hold.

3. **(14 pt.) Coupon Collection on Markov Chains**

Let $X_1, X_2, \ldots$ be an irreducible and aperiodic Markov chain over $n \ge 2$ states (labeled $1, 2, \ldots, n$) with a uniformly distributed initial state and transition matrix $P \in \mathbb{R}^{n\times n}$, i.e., $\Pr[X_1 = i] = 1/n$ for all $i \in [n]$ and $\Pr[X_{t+1} = j | X_t = i] = P_{ij}$ for any $t \ge 1$ and $i, j \in [n]$.[1] Furthermore, we assume that P is *doubly stochastic,* which means that $\sum_{i=1}^{n} P_{ij} = 1$ for every $j \in [n]$, as well as $\sum_{j=1}^{n} P_{ij} = 1$ for $i \in [n]$. (Note that a general transition matrix need only satisfy the second of these).

(a) **(1 pt.)** Prove that the unique stationary distribution of this Markov chain is uniform over $[n]$. Moreover, show that $\Pr[X_t = i] = 1/n$ for every $t \ge 1$ and $i \in [n]$.

(b) **(5 pt.)** Let T denote the earliest time when all the states have been visited at least once and the chain returns to the initial state X_1. Formally, random variable T is defined as

$$T = \min\{t \ge 1 : X_t = X_1 \text{ and } \{X_1, X_2, \ldots, X_t\} = [n].\}$$

Consider the following argument for bounding $\mathbb{E}[T]$:

> *The conclusion of Part 3a says that each X_t is uniformly distributed over $[n]$. Therefore, $X_1, X_2, \ldots$ can be viewed as the coupons that we draw in the coupon collector's problem. Then, T is exactly the number of coupons needed for first collecting the n coupons and then waiting for coupon X_1 to appear again. In expectation, the first part takes $n\sum_{i=1}^{n}\frac{1}{i}$ steps and the second part takes $\frac{1}{1/n} = n$ steps, so $\mathbb{E}[T] = n\sum_{i=1}^{n}\frac{1}{i} + n = \Theta(n\log n)$.*

Disprove the above by doing the following: (1) point out the flaw in the reasoning; (2) construct counterexamples to show that $\mathbb{E}[T]$ might be as small as $O(n)$; (3) construct counterexamples to show that $\mathbb{E}[T]$ might not be upper bounded by any function of n.

More concretely, for Part (2), show that for every $n \ge 2$ there exists a Markov chain with n states satisfying all the assumptions and $\mathbb{E}[T] \le 2n$. (You will also get full credit if $\mathbb{E}[T] \le Cn$ for some other constant C.)

For Part (3), show that for every $M > 0$ there exists a Markov chain with $n = 2$ states satisfying all the assumptions and $\mathbb{E}[T] \ge M$. (You will also get full credit if the number of states is $n = C$ for some other constant C.)

(c) **(3 pt.)** Define $p_{\min} := \min_{i,j\in[n]} P_{ij}$. Prove that if $p_{\min} > 0$, we have $\mathbb{E}[T] \le O\left(\frac{\log n}{p_{\min}}\right)$.

[**HINT:** *It can be proved that* $\mathbb{E}[T] \le 1 + \frac{1}{p_{min}}\left(1 + \sum_{i=1}^{n-1}\frac{1}{i}\right)$.]

(d) **(5 pt.)** Let $\mu := \mathbb{E}[T]$. Show that T has a *sub-exponential tail,* i.e., that there exist constants $c_0, c_1 > 0$ (that do not depend on n or any other parameters of the Markov chain) such that for every $\lambda \ge c_0$:

$$\Pr[T > \lambda\mu] \le e^{-c_1\lambda}.$$

[1] Note that in this problem, the initial state is denoted by X_1 (instead of X_0, which is used in the lecture notes). This slight change makes the analogy to coupon collecting in Part 3b more natural.

You may use the following claim without proof:

Claim 1. *There exists a constant $c_2 > 0$ such that the following holds for any Markov chain with n states that satisfies all the assumptions, and for any $i \in [n]$: Let $M = \lfloor c_2 \mu \rfloor$. Conditioning on $X_1 = i$, with probability at least $1 - 1/e$, there exists $t \in [M]$ such that $\{X_1, X_2, \ldots, X_t\} = \{X_{t+1}, X_{t+2}, \ldots, X_M\} = [n]$.*

[**HINT:** *You might find the argument at the end of the proof of Theorem 1 in Lecture Note #13 useful.*]

(e) **(0 pt.) [Optional: this won't be graded.]** Prove Claim 1.