ECE 534 FA26 HW6

November 16, 2026

I. Problems from Prof. Hajek's Lecture Notes

Problem 1. Problem 3.12, Prof. Hajek's lecture notes.

Problem 2. Problem 3.14, Prof. Hajek's lecture notes.

Problem 3. Problem 4.35, Prof. Hajek's lecture notes.

Problem 4. Problem 4.41, Prof. Hajek's lecture notes.

Problem 5. Problem 4.42, Prof. Hajek's lecture notes.

II. Advanced Problem Set on Markov Chains

This section contains advanced problems for students aiming to get an A+ or an A, who may need it for research. Before presenting the problem statement, we first clarify some conventions and list some useful properties.

Conventions. Let $(X_n)_{n\geq 0}$ be a Markov chain on a countable state space S with transition matrix P. A probability vector π is a *stationary distribution* if $\pi P = \pi$. We say π is an *equilibrium* (or *limiting*) distribution if $\lim_{n\to\infty} P^n(x,y) = \pi(y)$ for all $x, y \in S$. For $x \in S$, write $T_x = \inf\{n \geq 1 : X_n = x\}$ for the first return (hitting) time, and $m_x = \mathbb{E}_x[T_x]$ for the mean recurrence time. You may use without proof the following standard facts.

Kac's formula. Let (X_n) be an irreducible Markov chain on a countable state space S. Then (X_n) has a stationary distribution π if and only if it is positive recurrent, i.e. $m_x < \infty$ for some (equivalently, every) $x \in S$. In that case π is unique and

$$\pi(x) = \frac{1}{m_x} = \frac{1}{\mathbb{E}_x[T_x]} \qquad \text{for every } x \in S.$$

No aperiodicity assumption is needed.

Convergence theorem. If (X_n) is irreducible, aperiodic and positive recurrent with stationary distribution π, then $P^n(x,y) \to \pi(y)$ for all $x, y \in S$.

Cesàro means. A sequence $(a_n)_{n\geq 1}$ is *Cesàro convergent* to a if $\frac{1}{n}\sum_{k=1}^n a_k \to a$. Ordinary convergence implies Cesàro convergence to the same limit, but not conversely: $a_n = (-1)^n$ is Cesàro convergent to 0. For Markov chains, Cesàro averaging removes the effect of periodicity. If (X_n) is irreducible and recurrent, then for all $x, y \in S$

$$\frac{1}{n}\sum_{k=1}^{n} P^k(x,y) \longrightarrow \frac{1}{m_y},$$

with $1/m_y := 0$ when $m_y = \infty$. In particular, in the positive recurrent case the Cesàro limit is $\pi(y)$ even when the chain is periodic and $P^n(x,y)$ itself has no limit. Since $\sum_{k=1}^n P^k(x,y) = \mathbb{E}_x\big[\sum_{k=1}^n \mathbf{1}\{X_k = y\}\big]$, this says that the expected long-run fraction of time spent at y is $1/m_y$; the almost sure version, $\frac{1}{n}\sum_{k=1}^n \mathbf{1}\{X_k = y\} \to 1/m_y$, also holds (ergodic theorem for Markov chains). Thus a stationary distribution always describes long-run time averages, whereas an equilibrium distribution additionally describes the law of X_n at a fixed large time.

Abelian and Tauberian theorems. Generating functions such as $U(s) = \sum_{n\geq 0} P^n(x,x)s^n$ encode a sequence through its behavior as $s \uparrow 1$. An *Abelian* theorem passes from the sequence to the generating function. For example, if $a_n \geq 0$ then $\sum_n a_n s^n \uparrow \sum_n a_n \in [0,\infty]$ as $s \uparrow 1$ (Abel's theorem, by monotone convergence), and if $\frac{1}{n}\sum_{k\leq n} a_k \to a$ then $(1-s)\sum_n a_n s^n \to a$. A *Tauberian* theorem goes in the opposite direction. That direction is false in general and needs a side condition, such as nonnegativity of the coefficients. For instance, $a_n = (-1)^n$ has $(1-s)\sum_n a_n s^n = \frac{1-s}{1+s} \to 0$, but this does not tell us that a_n converges. A version that is especially useful for generating functions of Markov chains is:

Karamata's Tauberian theorem. Let $a_n \geq 0$, $\rho \geq 0$, and let L be slowly varying at infinity, i.e. $L(cx)/L(x) \to 1$ as $x \to \infty$ for every $c > 0$ (e.g. $L(x) = \log x$ or $1/\log x$). If

$$\sum_{n\geq 0} a_n s^n \sim (1-s)^{-\rho} L\left(\frac{1}{1-s}\right) \quad (s \uparrow 1), \qquad \text{then} \qquad \sum_{k=0}^{n} a_k \sim \frac{n^\rho L(n)}{\Gamma(1+\rho)} \quad (n \to \infty).$$

The case $\rho = 1$, $L \equiv c$ is the Hardy–Littlewood theorem: for $a_n \geq 0$, $(1-s)\sum_n a_n s^n \to c$ implies $\frac{1}{n}\sum_{k\leq n} a_k \to c$, i.e. Abel convergence implies Cesàro convergence. Note that the conclusion concerns partial sums (Cesàro-type averages), not the individual terms a_n; recovering a_n itself would require a further regularity condition, such as monotonicity.

Krawtchouk polynomials (used in Problem 1). Fix $N \geq 1$. The (binary) Krawtchouk polynomials $K_0, \ldots, K_N$ are defined through the generating function

$$\sum_{j=0}^{N} K_j(x) z^j = (1-z)^x (1+z)^{N-x},$$

that is,

$$K_j(x) = \sum_{i=0}^{j} (-1)^i \binom{x}{i} \binom{N-x}{j-i}, \qquad x \in \{0, 1, \ldots, N\}.$$

K_j is a polynomial of degree j in x; for example $K_0(x) = 1$, $K_1(x) = N - 2x$, and $K_N(x) = (-1)^x$. They are orthogonal with respect to the binomial distribution:

$$\sum_{x=0}^{N} \binom{N}{x} K_i(x) K_j(x) = 2^N \binom{N}{j} \delta_{ij}.$$

Origin in harmonic analysis on the hypercube. The characters of the group $\{0,1\}^N$ (with addition mod 2) are $\chi_u(v) = (-1)^{u\cdot v}$, $u \in \{0,1\}^N$. Summing the characters over all u of Hamming weight j gives a function of v that depends only on the weight $|v|$:

$$\sum_{u:\,|u|=j} (-1)^{u\cdot v} = K_j(|v|).$$

The random walk on $\{0,1\}^N$ that flips a uniformly chosen coordinate has the characters as eigenfunctions, with χ_u having eigenvalue $1 - 2|u|/N$. The Ehrenfest chain of Problem 1 is exactly the image of this walk under $v \mapsto |v|$ (number of balls in urn A). Averaging over each weight class therefore shows that the Ehrenfest transition matrix satisfies

$$PK_j = \left(1 - \frac{2j}{N}\right) K_j, \qquad j = 0, 1, \ldots, N.$$

So the Krawtchouk polynomials diagonalize the Ehrenfest chain. The extreme case $j = N$, with eigenfunction $(-1)^x$ and eigenvalue -1, is the source of the periodicity.

Role in coding theory. The same polynomials are central to the theory of error-correcting codes. A binary code is a subset $C \subseteq \{0,1\}^N$, and distances are measured in the Hamming metric. In that setting the Krawtchouk polynomials are the eigenvalues of the Hamming association scheme (Delsarte), and they act as the "Fourier transform" on weight and distance distributions. Two landmark results illustrate this.

- *MacWilliams identities.* If C is a linear code with A_i codewords of weight i, then its dual code $C^\perp$ has

$$B_j = \frac{1}{|C|} \sum_{i=0}^{N} A_i K_j(i)$$

  codewords of weight j. For example, the $[7, 4]$ Hamming code has nonzero weight counts $A_0 = 1$, $A_3 = 7$, $A_4 = 7$, $A_7 = 1$, and the formula returns the weights of its dual (the simplex code), whose 7 nonzero codewords all have weight 4.
- *Delsarte's linear programming bound.* For any code (linear or not) with minimum distance d, let A_i be the average number of codewords at distance i from a codeword. Then $A_0 = 1$, $A_i \geq 0$, $A_i = 0$ for $1 \leq i < d$, and moreover $\sum_i A_i K_j(i) \geq 0$ for every j. Maximizing $|C| = \sum_i A_i$ subject to these linear constraints gives an upper bound on the size of codes. Using properties of Krawtchouk polynomials, McEliece, Rodemich, Rumsey and Welch (1977) derived an asymptotic upper bound on the rate of binary codes that remained the tightest for nearly 50 years (see [1] for comments on the recent breakthrough).

Krawtchouk polynomials also appear in Lloyd's theorem on the nonexistence of perfect codes. The coding-theoretic and probabilistic roles are two sides of the same fact: both are Fourier analysis on $\{0,1\}^N$, restricted to functions that depend only on Hamming weight.

Now we present the problem statement.

Advanced Problem 1 (The Ehrenfest urn — stationary is not equilibrium)**.** Consider $N \geq 2$ labeled balls distributed between two urns, A and B. At each step a ball is chosen uniformly at random among the N and moved to the other urn. Let X_n be the number of balls in urn A at time n, so that

$$P(k, k-1) = \frac{k}{N}, \qquad P(k, k+1) = \frac{N-k}{N}, \qquad k \in \{0, 1, \ldots, N\}.$$

1) Show that $\pi(k) = \binom{N}{k} 2^{-N}$ is the unique stationary distribution, and that the chain is reversible with respect to π.
2) Show that the chain is irreducible with period 2, and hence has no equilibrium distribution. Compute $\lim_{n\to\infty} P^{2n}(0, y)$ for every y, and verify that the Cesàro averages $\frac{1}{n}\sum_{k=1}^{n} P^k(0, y)$ converge to $\pi(y)$.
3) Compute the mean recurrence times m_0 and, for N even, $m_{N/2}$. Give the asymptotics of $m_{N/2}$ as $N \to \infty$. For $N = 100$ and one move per microsecond, how long (in years) must one wait, on average, between successive visits to the state "urn A is empty"?
4) Starting from $X_0 = 0$, compute $\mathbb{E}_0[X_n]$ and $\mathbb{E}_0[(-1)^{X_n}]$ exactly. Show that $\mathbb{E}_0[X_n] \to \mathbb{E}_\pi[X]$, even though the law of X_n does not converge. Explain this in terms of eigenfunctions of P.
5) Consider the lazy chain $Q = \frac{1}{2}(I + P)$. Show that π is an equilibrium distribution for Q and compute $\mathbb{E}_0[X_n]$ under Q.

REFERENCES

[1] A. Barg, "Comments on the recent improvements of the mrrw bounds," 2026. [Online]. Available: https://arxiv.org/abs/2609.01860