Assignment 4

Probability Theory II
(EN.553.721, Spring 2025)

Assigned: March 31, 2025 Due: 11:59pm EST, April 14, 2025

Solve any four out of the five problems. If you solve more, we will grade the first four solutions you include. Each problem is worth an equal amount towards your grade.

Submit solutions in LaTeX. Write in complete sentences. Include and justify all steps of your arguments, but avoid writing excessive explanation that is not contributing to your solution.

Keep in mind the late submission policy: you may use a total of five late days for homework submissions over the course of the semester without penalty. If you need an extension beyond these, *you must ask me 48 hours before the due date of the homework* and have an excellent reason. After you have used up these late days, further late assignments will be penalized by 20% per day they are late.

Problem 1 (Harmonic analysis and Markov chains)**.** Let (Y_n) be a Markov chain taking values in a countable S and initialized from some probability measure μ_0 over S with transition kernel $p : S \times S \to [0, 1]$. Suppose that S is irreducible.

1. Let $f : S \to \mathbb{R}$ satisfy, for all $y \in S$,

$$f(y) \geq \sum_{z \in S} p(y, z) f(z) = \mathbb{E}_y f(Y_1)$$

   and have that $f(Y_n)$ is integrable (i.e., in L^1) for all $n \geq 0$. Show that $(f(Y_n))_{n \geq 0}$ is a supermartingale, a submartingale if the inequality above is reversed, and a martingale if the inequality is an equality. (This generalizes part of Problem 3 from Homework 2.) In these cases, f is called *superharmonic*, *subharmonic*, and *harmonic*, respectively. For $U \subseteq S$, write $\partial U = \{y \notin U : p(x, y) > 0 \text{ for some } x \in U\}$. Then, $f : U \cup \partial U \to \mathbb{R}$ is called *super/sub/harmonic on* U if it satisfies the above for all $y \in U$ (a weaker condition than having these properties on all of S).

2. Let $U \subseteq S$ be finite and non-empty. Suppose $f : U \cup \partial U \to \mathbb{R}$ is harmonic on U. Show that

$$\max_{x \in U} f(x) \leq \sup_{x \in \partial U} f(x).$$

3. Again, let $U \subseteq S$ be finite and non-empty. For any $f : \partial U \to \mathbb{R}$, show that there is at most one $g : U \cup \partial U \to \mathbb{R}$ that is harmonic on U and that satisfies the "boundary condition" $g(x) = f(x)$ for all $x \in \partial U$.

   (HINT: If g and h are two different such functions, consider using Part 2 on $g - h$ and $h - g$.)

4. Once more, let $U \subseteq S$ be finite and non-empty. Let $T := \min\{n \geq 0 : Y_n \notin U\}$ be the *exit time* from U. Suppose that $f : \partial U \to \mathbb{R}$ is bounded. Show that the function

$$g(x) := \mathop{\mathbb{E}}_{x} f(Y_T)$$

   is well-defined for $x \in U \cup \partial U$, is harmonic on U, and satisfies $g(x) = f(x)$ for all $x \in \partial U$. That is, averaging the value of f at an exit time gives the unique solution of the associated boundary value problem from Part 3.

If you are familiar with basic notions of continuous harmonic analysis and/or the Poisson partial differential equation, you should compare Parts 2 and 3 to the corresponding analytic principles.

Problem 2 (Recurrence and transience of trees)**.** A tree is a simple connected graph with no cycles. All trees in this problem are *locally finite*, meaning each vertex has finite degree (though the total number of vertices may be infinite). The *simple random walk (SRW)* on a tree is the Markov chain taking values in the tree where each successive state is a uniformly random neighbor of the current state.

1. Let G_d be the infinite d-regular tree, the infinite tree where every vertex has degree $d \geq 2$. (For example, $d = 2$ gives the integer graph $\mathbb{Z}$ we have discussed in class.) Fix some $x \in G_d$. Let W_n be the number of closed walks in G_d that start at x, follow an edge at every step, traverse a total of n edges, and end back at x. Show that $W_n = 0$ if n is odd. If n is even, let $W_{n,k}$ be the number of such walks that visit x exactly k times (note that we must have $k \geq 2$). Show that $W_{n,k} \leq 2^{n-k+1} d^{k-1} (d-1)^{n/2-k+1}$.

   (HINT: Group the steps of a walk into three types: (1) steps from x to one of its neighbors, (2) steps from vertices other than x that increase the distance to x, and (3) steps from vertices other than x that decrease the distance to x. How many possible patterns of step types are there for a fixed n and k, and how many walks are there with each pattern?)

2. Derive that, if $d \geq 3$, then there is a constant $C_d > 0$ such that $W_n \leq C_d (2\sqrt{d-1})^n$. Conclude that all $x \in G_d$ are transient for the SRW on G_d for any $d \geq 3$. (In contrast, by Pólya's theorem from class, all $x \in G_2 = \mathbb{Z}$ are recurrent for the SRW on G_2.)

3. Let $x \in G_d$ and pick some neighbor y of x. If we remove the edge between x and y, then G_d is split into two trees, H and H', whose vertices partition the vertices of G_d. Show that, if $d \geq 3$, then the SRW started from any $z \in G_d$ almost surely either visits H only finitely many times or visits H' only finitely many times. (Of course, the SRW must visit one of the two infinitely many times, since all vertices belong to H or H'.)

4. Consider $G_2 = \mathbb{Z}$. In class we showed that the SRW on $\mathbb{Z}$ is recurrent. Consider now picking some $\beta \in (\frac{1}{2}, 1)$, and taking a Markov chain on $\mathbb{Z}$ with the "biased" transition kernel
$$p(x, y) = \left\{ \begin{array}{ll} \beta & \text{if } y = x + 1, \\ 1 - \beta & \text{if } y = x - 1, \\ 0 & \text{otherwise} \end{array} \right\}.$$
Show that every $x \in \mathbb{Z}$ is transient for a Markov chain with this transition kernel with any initialization.

Problem 3 (More on stationary measures)**.** Let (Y_n) be a Markov chain on a countable S such that S is irreducible and every $x \in S$ is recurrent. Recall that we studied the stationary measure $\mu_x(y) := \mathbb{E}_x[\sum_{n=0}^{T_x - 1} \mathbb{1}\{Y_n = y\}]$, where $T_x := \min\{n \geq 1 : Y_n = x\}$.

1. We say that a random variable A *stochastically dominates* a random variable B if, for all $t \in \mathbb{R}$, $\mathbb{P}[A \geq t] \geq \mathbb{P}[B \geq t]$. Write $A \succeq B$ for this relation (not to be confused with positive semidefiniteness of matrices). Show that, if $A \succeq B$, and $f : \mathbb{R} \to \mathbb{R}$ is strictly increasing, then $\mathbb{E} f(A) \geq \mathbb{E} f(B)$ provided that both expectations are well-defined.

2. Show that, for all $x, y \in S$ distinct, $\mathbb{P}_y[T_x < T_y] > 0$.

3. Conclude that $\sum_{n=0}^{T_x - 1} \mathbb{1}\{Y_n = y\}$ is stochastically dominated by a suitable geometric random variable, and therefore that $\mu_x(y) < \infty$ for all $x, y \in S$.

4. Suppose $S = \{0, 1, \dots, M\}$ for some $M \geq 1$; in particular, S is finite. Suppose that $p, q : S \times S \to [0, 1]$ are transition kernels such that, for all $x \in S$, $q(x, \cdot) \succeq p(x, \cdot)$ (note that both sides here, for each fixed $x \in S$, are probability measures). Let μ be the unique stationary probability measure of p and ν that of q. Show that $\nu \succeq \mu$ (i.e., that stochastic domination of transition kernels implies stochastic domination of stationary distributions).

   (**Hint:** Write $\nu^{(k)}(x) = \sum_y \nu(y) p^{(k)}(y, x)$. Show that $\nu \succeq \nu^{(k)}$ for all k, and use the convergence theorem from class.)

Problem 4 (More on stationary probability measures)**.** As in Problem 3, let (Y_n) be a Markov chain on a countable S such that S is irreducible and every $x \in S$ is recurrent. The same μ_x and T_x discussed there will appear below.

1. Suppose that, for some $x \in S$, $\mathbb{E}_x[T_x] < \infty$. Show that then in fact $\mathbb{E}_y[T_y] < \infty$ for all $y \in S$. Generally, such states in a Markov chain are called *positive recurrent*, and positive recurrence, like recurrence, is "contagious under accessibility" (but you do not need to prove that beyond this special case). You may use the result of Problem 3 that the μ_x are measures on S.

   (**Hint:** Use the uniqueness theorem for stationary measures, which shows that the various μ_x are the same up to constant rescaling.)

2. Let $h(x,y) := \mathbb{E}_x[T_y]$, the *expected hitting time* of y starting from x. Note that $h(x,x) = \mathbb{E}_x[T_x]$ is the quantity appearing in Part 1. Show that, for all $x, y \in S$,

$$h(x,y) = 1 + \sum_{\substack{z \in S \\ z \neq y}} p(x,z)h(z,y).$$

3. Deduce that, if π is a stationary probability measure, then, for all $y \in S$,

$$\sum_{x \in S} \pi(x)h(x,y) = 1 + \sum_{\substack{x \in S \\ x \neq y}} \pi(x)h(x,y).$$

4. Suppose that, as in Part 1, all states in S are positive recurrent. Let μ_x be as in Problem 3. Recall that in this case, from the existence theorem from class,

$$\hat{\mu}_x(y) := \frac{\mu_x(y)}{\mu_x(S)} = \frac{\mu_x(y)}{\mathbb{E}_x[T_x]}$$

is a stationary probability measure. Recall also that the uniqueness theorem for stationary measures implies that there is a unique stationary probability measure (since there is a unique stationary measure up to scaling, and the scaling is fixed by asking for a probability measure), so all $\hat{\mu}_x$ must be equal. Show more concretely that, for all $x, y \in S$,

$$\hat{\mu}_x(y) = \frac{1}{\mathbb{E}_y[T_y]},$$

whereby also the unnormalized stationary measures are given by

$$\mu_x(y) = \frac{\mathbb{E}_x[T_x]}{\mathbb{E}_y[T_y]}.$$

(**Hint:** Use Part 3 to show that the first result holds for any stationary probability measure.)

Problem 5 (Continuous state space)**.** Consider a process (Y_n) taking values in $S = \mathbb{R}$ formed by $Y_0 = y_0$ for some (deterministic) $y_0 \in \mathbb{R}$ and setting $Y_{n+1} = \alpha Y_n + Z_{n+1}$ for all $n \geq 0$, for some $\alpha \in \mathbb{R}$, where $Z_1, Z_2, \ldots \sim \mathcal{N}(0,1)$ are i.i.d.

1. Show that (Y_n) is a Markov chain. Describe its transition kernel (recall that in this setting this is a function $p : \mathbb{R} \times \mathcal{B} \to [0,1]$ for $\mathcal{B}$ the Borel σ-algebra of $\mathbb{R}$).

2. Show that this transition kernel has a stationary probability measure if and only if $|\alpha| < 1$. In this case, describe such a probability measure μ_α.

3. If $|\alpha| < 1$, show that $\mathsf{Law}(Y_n)$ converges weakly to μ_α, for any initial y_0.

4. If $|\alpha| < 1$ and $f : \mathbb{R} \to \mathbb{R}$ is smooth and bounded, show that, as $n \to \infty$, $\frac{1}{n}\sum_{i=1}^n f(Y_i)$ converges to $\int f d\mu_\alpha$ in probability. Note that this statement, an ergodic theorem like we saw in class, is like a law of large numbers, except the Y_i are not i.i.d. but rather arise from a Markov chain.

   (**Hint:** Estimate the mean and variance.)