STA447/2006: Midterm Exam #1

Instructor: Wenlong Mou

Feb 4th, 2026

Student name: _______________

Student ID: _______________

Student signature: _______________

This exam contains 12 pages.

Total marks: 100 pts

Time Allowed: 150 minutes

Question 1. [30 points, 3 points for each question] Mark each of the following statements with T (true) or F (false). *No justification is required.* Your grade will be solely based on your true-or-false choices.

(1) If for some states i, j we have $\sum_{n\geq 0} p_{ij}^{(n)} < \infty$, then j is transient. **Answer:** F

(2) If for some states i, j we have $\sum_{n\geq 0} p_{ij}^{(n)} = \infty$, then j is recurrent. **Answer:** T

(3) Let P be an irreducible and transient Markov chain. Then for any pair $i, j \in S$, with probability 1, the Markov chain starting from i will visit j only finite times. **Answer:** T

(4) Let i, j be a pair of states of an irreducible Markov chain P. If $f_{ij} = 1$. Then P is recurrent. **Answer:** F

(5) There exists a Markov chain that has infinitely many stationary distributions. **Answer:** T

(6) Let P be a Markov chain with stationary measure π. If we have

$$\lim_{n\to+\infty} |p_{ij}^{(n)} - p_{ik}^{(n)}| = 0,$$

for any $i, j, k \in S$, then π can be normalized to become a stationary distribution $\widetilde{\pi}$, with $\lim_{n\to+\infty} p_{ij} = \widetilde{\pi}_j$ **Answer:** F

(7) If state i is aperiodic and $i \to j$, then state j is also necessarily aperiodic. **Answer:** F

(8) Let P be an irreducible Markov chain on a finite state space S, then P has a unique stationary distribution. **Answer:** T

(9) Let P be a reducible Markov chain and $i \in S$ is a transient state. If P has a stationary distribution π, then we must have $\pi_i = 0$. **Answer:** T

(10) Let P be a reducible Markov chain and $i \in S$ is a transient state. If P has a stationary measure π, then we must have $\pi_i = 0$. **Answer:** F

Question 2. [20 pts] Consider a Markov chain on a finite state space $S = \{1, 2, 3, 4, 5\}$, with the transition matrix given by

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & 1/2 & 0 & 1/4 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(1) [5 pts]. Which states are recurrent? Which states are transient? Please explain your reasoning.

Answer: States 1 and 5 are recurrent, while states $2, 3, 4$ are transient.

Rubrics: 1pt for each state.

(2) [9 pts]. Compute f_{25} and $\mathbb{E}_2[N(3)]$, where $N(i)$ is the number of visits to the state i. Explain your reasoning.

Answer: We use f-expansion to compute the relevant quantities. We note that

$$f_{25} = p_{21}f_{15} + p_{23}f_{35} + p_{25} = \frac{1}{4}\cdot 0 + \frac{1}{2}f_{35} + \frac{1}{4}\cdot 1$$
$$f_{35} = p_{32}f_{25} + p_{34}f_{45} = \frac{1}{2}f_{25} + \frac{1}{2}f_{45}$$
$$f_{45} = p_{44}f_{45} + p_{45} = \frac{1}{2}f_{45} + \frac{1}{2}$$

From the last equation: $f_{45} = \frac{1}{2}f_{45} + \frac{1}{2}$, so $f_{45} = 1$.
Substituting back, we have

$$f_{25} = \frac{1}{2}f_{35} + \frac{1}{4}, \qquad f_{35} = \frac{1}{2}f_{25} + \frac{1}{2}.$$

Solving this system of equations, we have

$$f_{25} = \frac{2}{3}, \qquad f_{35} = \frac{5}{6}.$$

To compute $\mathbb{E}_2[N(3)]$, we note that

$$\mathbb{E}_2[N(3)] = \frac{f_{23}}{1 - f_{33}}.$$

Applying f-expansion again, we have

$$f_{23} = p_{21}f_{13} + p_{23} + p_{25}f_{53} = \frac{1}{2},$$
$$f_{33} = p_{32}f_{23} + p_{34}f_{43} = \frac{1}{4}.$$

Therefore, we have $\mathbb{E}_2[N(3)] = \frac{1/2}{1-1/4} = \frac{2}{3}$.

Rubrics: 4 points for computing f_{25} correctly.
5 points for computing $\mathbb{E}_2[N(3)]$ correctly.
For each part, if you use the correct method but make minor calculation mistakes, you can get deduction of 1 point.

(3) [6 pts]. Let $X_0 = 2$, compute the probability that state 4 is never visited but the chain is eventually absorbed by state 5.

Answer: Note that starting from state 4, the chain will eventually be absorbed by state 5 with probability 1. So the desired probability is $f_{25} - f_{24}$. From part (2), we have $f_{25} = 2/3$. To compute f_{24}, we apply f-expansion again:

$$f_{24} = p_{21}f_{14} + p_{23}f_{34} + p_{25}f_{54} = \frac{1}{4} \cdot 0 + \frac{1}{2}f_{34} + \frac{1}{4} \cdot 0 = \frac{1}{2}f_{34},$$
$$f_{34} = p_{32}f_{24} + p_{34} = \frac{1}{2}f_{24} + \frac{1}{2} = \frac{1}{2}f_{24} + \frac{1}{2}.$$

Solving this system of equations, we have

$$f_{24} = \frac{1}{3}, \qquad f_{34} = \frac{2}{3}.$$

Therefore, the desired probability is $f_{25} - f_{24} = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$.

Rubrics: Any other method that correctly derive the desired probability is also acceptable.

If the overall idea is correct but there are minor calculation mistakes, you can get deduction of 1 point.

If you realize that the desired probability can be computed as $f_{25} - f_{24}$, you get 2 points.

Question 3. [15 pts] Recall the independent coupling defined in class. Consider two copies of Markov chains $\{X_n^{(1)}\}_{n\geq 0}$ and $\{X_n^{(2)}\}_{n\geq 0}$ with the same transition matrix P. At each step, the two chains move independently according to P, regardless of their current states. For this problem, we assume that P is irreducible, aperiodic, and recurrent (but not necessarily positive recurrent).

(1) [6 pts]. Show that the joint chain $\{(X_n^{(1)}, X_n^{(2)})\}_{n\geq 0}$ has a stationary measure.

Answer: Let π be a stationary measure of P. We can verify that the product measure $\pi \otimes \pi$ is a stationary measure for the joint chain, i.e., for any $(i,j) \in S \times S$, the stationary measure $\pi \otimes \pi$ takes the value $\pi_i \cdot \pi_j$.

To verify stationarity, for any pair of states (i,j), we have

$$\sum_{(k,l)\in S\times S} (\pi \otimes \pi)_{(k,l)} \cdot p_{(k,l),(i,j)}$$
$$= \sum_{k,l} \pi_k \pi_l p_{ki} p_{lj} = \left(\sum_k \pi_k p_{ki}\right) \cdot \left(\sum_l \pi_l p_{lj}\right) = \pi_i \cdot \pi_j = (\pi \otimes \pi)_{(i,j)}.$$

Rubrics: If you construct the correct stationary measure but do not verify the stationarity condition, you can get 3 points.

An incomplete verification of the stationarity condition can get partial credits.

(2) [9 pts]. Construct a Markov chain P satisfying above conditions, such that the joint chain $\{(X_n^{(1)}, X_n^{(2)})\}_{n\geq 0}$ is transient.

Answer: Let P be a lazy version of 2-dimensional simple random walk, i.e., with the state space being $\mathbb{Z}^2$, we take the transition kernel as follows

$$X_{n+1} = X_n + (\varepsilon_{n+1}^{(1)}, \varepsilon_{n+1}^{(2)}),$$

where $\varepsilon_{n+1}^{(1)}$ and $\varepsilon_{n+1}^{(2)}$ are independent and identically distributed random variables, with $\varepsilon_{n+1}^{(1)} = 0$ with probability $1/2$, and $\varepsilon_{n+1}^{(1)} = \pm 1$ with probability $1/4$ each. The distribution of $\varepsilon_{n+1}^{(2)}$ is the same as that of $\varepsilon_{n+1}^{(1)}$.

The chain is clearly irreducible and aperiodic. Moreover, it is recurrent since the simple random walk on $\mathbb{Z}^2$ is recurrent. However, the joint chain $\{(X_n^{(1)}, X_n^{(2)})\}_{n\geq 0}$ can be viewed as a lazy version of simple random walk on $\mathbb{Z}^4$, which is transient.

Rubrics: Any valid construction gets ful credit.

If you construct a valid Markov chain but fail to explain why the joint chain is transient, you can get 8 points.

If you just use a 2-dimensional simple random walk without laziness, you can get 6 points, since the chain is not aperiodic.

Question 4. [8 pts] Let P be an irreducible Markov chain on a finite state space S. Let T_i be the hitting time to state i. Show that for any state $i \in S$.

$$\mathbb{E}_i\left[T_i^n\right] < +\infty,$$

for any $n \in \mathbb{N}_+$.

Answer: (This result can be proved using the question 1.7 from Lawler, which is listed as a practice question. Here we provide a complete proof.)

Since P is irreducible, for every $j \in S$, there exists $m_j \in \mathbb{N}_+$ such that $p_{ji}^{(m_j)} > 0$. Let $m = \max_{j\in S} m_j$ and $\varepsilon = \min_{j\in S} p_{ji}^{(m_j)}$. Since S is finite, we have $m < +\infty$ and $\varepsilon > 0$, and

$$\mathbb{P}_j(T_i > m) \leq 1 - \varepsilon, \quad \text{for any } j \in S.$$

Therefore, for any $k \in \mathbb{N}_+$, we have

$$\mathbb{P}_i(T_i > km) \leq (1-\varepsilon)^k.$$

And hence

$$\mathbb{E}_i\left[T_i^n\right] = \sum_{t=0}^{+\infty} \mathbb{P}_i(T_i \geq t) \cdot \left((t+1)^n - t^n\right) \leq \sum_{t=0}^{+\infty} (1-\varepsilon)^{\lfloor t/m \rfloor} \cdot (n+1)(t+1)^{n-1} < +\infty.$$

Rubrics: If you can clearly state the conclusion from Lawler's practice question and use it to prove the result, you still get full credit.

If you prove the exponentially decaying tail bound for T_i but fail to use it to conclude the finiteness of $\mathbb{E}_i[T_i^n]$, you can get 5 points.

If you only proved $\mathbb{E}_i[T_i] < +\infty$, you can get 2 points.

Question 5. [15 pts] Consider a Markov chain on the state space $S = \{0, 1, 2, \cdots\}$. For any $i \geq 1$, we define the transition from the state i as

$$p_{i,i+1} = \frac{i}{2i+2}, \quad \text{and} \quad p_{i,i-1} = \frac{i+2}{2i+2},$$

and $p_{i,j} = 0$ for $j \notin \{i-1, i+1\}$. We further let $p_{0,1} = 1$ and $p_{0,j} = 0$ for $j > 1$.

(1) [7 pts]. Let T_i be the hitting time to state i. Compute $\mathbb{E}_0[T_0]$.

Answer: Since the chain is irreducible, the expected return time to state 0 satisfies

$$\mathbb{E}_0[T_0] = \frac{1}{\pi_0},$$

where π is the stationary distribution.

Using detailed balance $\pi_i p_{i,i+1} = \pi_{i+1} p_{i+1,i}$, we get

$$\pi_0 \cdot 1 = \pi_1 \cdot \frac{3}{4} \;\Rightarrow\; \pi_1 = \frac{4}{3}\pi_0,$$

and for $i \geq 1$,

$$\pi_i \frac{i}{2(i+1)} = \pi_{i+1} \frac{i+3}{2(i+2)} \;\Rightarrow\; \pi_{i+1} = \pi_i \frac{i(i+2)}{(i+1)(i+3)}.$$

Thus, for $n \geq 1$,

$$\pi_n = \pi_1 \prod_{k=1}^{n-1} \frac{k(k+2)}{(k+1)(k+3)} = \frac{4}{3}\pi_0 \prod_{k=1}^{n-1} \frac{k}{k+1} \cdot \frac{k+2}{k+3} = \frac{4}{3}\pi_0 \cdot \frac{1}{n} \cdot \frac{3}{n+2} = \pi_0 \frac{4}{n(n+2)}.$$

Hence

$$1 = \sum_{n=0}^{\infty} \pi_n = \pi_0 + \sum_{n=1}^{\infty} \pi_0 \frac{4}{n(n+2)} = \pi_0 \Big(1 + 4 \sum_{n=1}^{\infty} \frac{1}{n(n+2)}\Big).$$

Using

$$\frac{1}{n(n+2)} = \frac{1}{2}\Big(\frac{1}{n} - \frac{1}{n+2}\Big) \;\Rightarrow\; \sum_{n=1}^{\infty} \frac{1}{n(n+2)} = \frac{3}{4},$$

we get

$$1 = \pi_0 \big(1 + 4 \cdot \tfrac{3}{4}\big) = \pi_0 \cdot 4 \;\Rightarrow\; \pi_0 = \frac{1}{4}.$$

Therefore,

$$\mathbb{E}_0[T_0] = \frac{1}{\pi_0} = 4.$$

Rubrics: 2 points for realizing that $\mathbb{E}_0[T_0]$ can be computed as $1/\pi_0$.

5 points for correctly computing π_0.

If the overall idea is correct but there are minor calculation mistakes, you can get deduction of 1 point.

If you get a stationary measure but fail to normalize it to get a stationary distribution, you can get 3 points out of 5 for computing π_0.

(2) [8 pts]. Compute the expectation

$$\mathbb{E}_0\Big[\sum_{t=1}^{T_0} \mathbf{1}_{X_t=1}\Big].$$

Answer: By the construction of stationary measure in lecture, we have that

$$\mu_0(0) = 1, \qquad \mu_0(i) = \mathbb{E}_0\Big[\sum_{t=1}^{T_0} \mathbf{1}_{X_t=i}\Big], \quad \text{for any } i \geq 1,$$

is a stationary measure. And since the Markov chain is positive recurrent as shown in part (1), this stationary measure can be normalized to become a stationary distribution. We denote

$$\widetilde{\mu}(i) = \frac{\mu_0(i)}{\sum_{j\in S} \mu_0(j)}, \quad \text{for any } i \in S,$$

as the normalized stationary distribution. Since the Markov chain is irreducible and positive recurrent, the stationary distribution is unique, and hence $\widetilde{\mu} = \pi$. Therefore, we have

$$\mathbb{E}_0\Big[\sum_{t=1}^{T_0} \mathbf{1}_{X_t=1}\Big] = \mu_0(1) = \pi_1 \cdot \sum_{j\in S} \mu_0(j) = \pi_1/\pi_0 = 4/3.$$

Rubrics: 5 points for correctly applying the construction of stationary measure to derive the formula.

2 points for correctly calculating π_1/π_0.

1 points for showing the uniqueness of stationary distribution to conclude $\widetilde{\mu} = \pi$.

Question 6. [12 pts] During the class, we proved asymptotic convergence of Markov chains using independent coupling. In this problem, we extend the idea to general coupling and prove concrete convergence rates.

Concretely, suppose that we can construct a joint chain $\{(X_n^{(1)}, X_n^{(2)})\}_{n\geq 0}$, such that marginally $\{X_n^{(1)}\}_{n\geq 0}$ and $\{X_n^{(2)}\}_{n\geq 0}$ are both Markov chains with the same transition matrix P (but they are not necessarily independent). Suppose that $X_0^{(1)} = i$ for some $i \in S$, and $X_0^{(2)}$ is drawn from the stationary distribution π. Define the coupling time

$$T = \inf\{n \geq 0 : X_n^{(1)} = X_n^{(2)}\}.$$

(1) [4 pts]. Show that for any $n \geq 0$,

$$\sum_{j\in S} |p_{ij}^{(n)} - \pi_j| \leq 2\mathbb{P}(T \geq n).$$

Hint: you may use the following fact without proof: for any two discrete distributions μ and ν on S, let (X, Y) be a coupling of μ and ν, i.e., marginally $X \sim \mu$ and $Y \sim \nu$. Then we have

$$\sum_{j\in S} |\mu_j - \nu_j| \leq 2\mathbb{P}(X \neq Y).$$

Answer: Under the coupling, we have

$$X_n^{(1)} \sim P^n(i, \cdot), \quad X_n^{(2)} \sim \pi, \quad \text{for any } n \geq 0.$$

Using the hint, we conclude that

$$\sum_{j\in S} |p_{ij}^{(n)} - \pi_j| \leq 2\mathbb{P}(X_n^{(1)} \neq X_n^{(2)}).$$

Moreover, by strong Markov property, starting from the stopping time T, we can modify coupling after first meeting so they move together thereafter, i.e., we can construct a coupling such that $X_{T+k}^{(1)} = X_{T+k}^{(2)}$ for any $k \geq 0$. Therefore, we have

$$\mathbb{P}(X_n^{(1)} \neq X_n^{(2)}) = \mathbb{P}(T \geq n),$$

which concludes the desired result.

Rubrics: If you only apply the hint without explaining why we can relate $\mathbb{P}(X_n^{(1)} \neq X_n^{(2)})$ to $\mathbb{P}(T \geq n)$, you can get 2 points.

(2) [8 pts]. Let us now consider a concrete example. Consider a Markov chain on the state space $S = \{0,1\}^N$ for some positive integer N. The transition probability P is defined as follows: given the current state $x = (x_1, x_2, \cdots, x_N) \in S$, we choose an index i uniformly at random from $\{1, 2, \cdots, N\}$, and update x_i with probability $1/2$ to be either 0 or 1 (independently of the previous value of x_i). All other coordinates remain unchanged.

Let π be the stationary distribution of this Markov chain. Show that for any initial state $x \in S$,

$$\sum_{y \in S} |p_{xy}^{(n)} - \pi_y| \le 2N\left(1 - \frac{1}{N}\right)^n.$$

Answer: We construct a coupling explicitly and apply the results from part (1). In particular, we let $X_0^{(1)} = x$ and $X_0^{(2)} \sim \pi$. At each step, we choose the same index i for both chains, and update the i-th coordinate of both chains to be the same value (either 0 or 1 with probability 1/2 each). It is easy to verify that under this coupling, marginally $\{X_n^{(1)}\}_{n\ge 0}$ and $\{X_n^{(2)}\}_{n\ge 0}$ are both Markov chains with transition matrix P. Moreover, we have

$$\begin{aligned}\mathbb{P}(T \ge n) &= \mathbb{P}\big(\text{at least one coordinate is not updated in the first } n \text{ steps}\big)\\ &\le \sum_{i=1}^{N} \mathbb{P}\big(\text{the } i\text{-th coordinate is not updated in the first } n \text{ steps}\big)\\ &= N\left(1 - \frac{1}{N}\right)^n.\end{aligned}$$

Applying the result from part (1), we conclude the desired result.

Rubrics: Any other valid construction deserves full credit.

5 points for constructing the coupling correctly. 3 points for computing the upper bound for coupling time tail probability.

If you prove an upper bound that is weaker than the desired bound, you get 1 or 2 points out of 3 points for proving the upper bound.

A partially correct construction gets partial credits.