Exam Advanced Topics in Markov Chains July 9th, 2018

Hints: You can use all results proved in the lecture notes (without proving them yourselves). You can also use claims from exercises in the lecture notes (without solving these exercises). Finally, you can also use a claim you are supposed to prove in one excercise below to solve another excercise (even if you did not prove the claim). Partial solutions also yield points.

Exercise 1 (Book shuffle) Professor David Levin has a shelve with a row of N books on it. When he finishes with a book drawn from the shelve, he does not bother to put it back at the right place but instead places it back at the left end of the row. Let S_N denote the space of all permutations of $\{1, \ldots, N\}$. For $1 \le i \le N$, define $\pi^i \in S_N$ by

$$\pi^{i}(j) := \begin{cases} i & \text{if } j = 1, \\ j - 1 & \text{if } 1 < j \le i, \\ j & \text{if } i < j \le N. \end{cases}$$

This says that after taking the *i*-th book and placing it at the left end of the row, $\pi^i(j)$ is the original position of the book that is now on place j. Let Π_0 and M_0 be random variables with values in S_N and $\{0, \ldots, N\}$ respectively, and let I_1, I_2, \ldots be i.i.d. uniformly distributed on $\{1, \ldots, N\}$ and independent of (Π_0, M_0) . Define

$$\left. \begin{array}{l} \Pi_k := \Pi_{k-1} \circ \pi^{I_k}, \\ M_k := M_{k-1} + \mathbb{1}_{\{I_k > M_{k-1}\}} \end{array} \right\} \quad (k \ge 1).$$

(a) Explain why $(\Pi_k)_{k\geq 0}$, $(M_k)_{k\geq 0}$, as well as the joint process $(\Pi_k, M_k)_{k\geq 0}$ are Markov chains.

(b) Calculate the transition kernel of the Markov chain $(M_k)_{k\geq 0}$. Is $(M_k)_{k\geq 0}$ (viewed as a function of $(\Pi_k, M_k)_{k\geq 0}$) an autonomous Markov chain?

(c) Is $(\Pi_k)_{k>0}$ an autonomous Markov chain?

(d) Let $P(\pi, \pi')$, Q(m, m'), and $R((\pi, m), (\pi', m'))$ denote the transition kernels of $(\Pi_k)_{k\geq 0}$, $(M_k)_{k\geq 0}$, and the joint chain $(\Pi_k, M_k)_{k\geq 0}$. Do the chains evolve independently, in the sense that $R((\pi, m), (\pi', m')) = P(\pi, \pi')Q(m, m')$?

(e) For $m \in \{0, ..., N\}$, let $K(m, \cdot)$ denote the uniform distribution on $\{\pi \in S_N : \pi(m+1) < \cdots < \pi(N)\}$. Let Π have law $K(m, \cdot)$. Prove that for each $1 \le i \le m$, the random variable $\Pi \circ \pi^i$ also has law $K(m, \cdot)$.

Please turn over.

(f) Let Π have law $K(m, \cdot)$. Let I be independent of Π and uniformly distributed on $\{m+1,\ldots,N\}$. Prove that $\Pi \circ \pi^I$ has law $K(m+1, \cdot)$.

(g) Prove the intertwining relation QK = KP.

(h) Assume that the joint chain is started in an initial law such that

$$\mathbb{P}\big[\Pi_0 = \pi \mid M_0 = m\big] = K(m, \pi) \qquad (0 \le m \le N, \ \pi \in S_N).$$
(1)

Prove that

$$\mathbb{P}\big[\Pi_k = \pi \,\big|\, (M_0, \dots, M_k)\big] = K(M_k, \pi) \quad \text{a.s.} \qquad (k \ge 0, \ \pi \in S_N).$$

(i) Let τ be a stopping time w.r.t. the filtration generated by the Markov chain $(M_k)_{k\geq 0}$. Assume that $\tau < \infty$ a.s. Prove that (1) implies that

$$\mathbb{P}\big[\Pi_{\tau} = \pi \mid M_{\tau} = m\big] = K(m, \pi) \qquad (0 \le m \le N, \ \pi \in S_N).$$

Hint: the event $\{\tau = k\}$ is measurable w.r.t. the σ -field generated by (M_0, \ldots, M_k) .

(j) Start the process with Π_0 the identity permutation and $M_0 = 0$. Let $\tau := \inf\{k \ge 0 : M_k = N - 1\}$. Prove that Π_{τ} is uniformly distributed on S_N .

Exercise 2 (Random walk on a tree) For each $h \in \mathbb{Z}$, let S_h denote the space of all functions $x : \{\ldots, h-1, h\} \to \{1, 2\}$ with the property that there exists a $g \leq h$ such that x(i) = 1 for all $i \leq g$. We view an element $x \in S_h$ as an infinite word, e.g., \cdots 11111112122, made up from the alphabet $\{1, 2\}$. For $x \in S_h$ and $\alpha \in \{1, 2\}$, we define $x\alpha \in T_{h+1}$ by $(x\alpha)(i) = x(i)$ $(i \leq h)$ and $(x\alpha)(h+1) = \alpha$, i.e., $x\alpha$ is the word x with an extra letter α added at the right end. For each $x \in S_h$, we define $\tilde{x} \in T_{h+1}$ by $\tilde{x}(i) = x(i)$ $(i \leq h-1)$, i.e., \tilde{x} is the word x from which the last letter has been removed. We let $\mathbb{T} := \{(h, x) : h \in \mathbb{Z}, x \in S_h\}$ and define a probability kernel P on \mathbb{T} by

$$P((h,x),(h-1,\check{x})) := \frac{1}{3}, P((h,x),(h+1,x1)) := \frac{1}{3}, \text{ and } P((h,x),(h+1,x2)) := \frac{1}{3},$$

and P((h, x), (h', x')) := 0 in all other cases. Note that we can view \mathbb{T} as a regular tree in which each vertex has three neighbors, and that P is the transition kernel of a random walk on this tree.

(a) Let $(H_k, X_k)_{k\geq 0}$ be a Markov chain with transition kernel P. Show that $(H_k)_{k\geq 0}$ is a Markov chain and determine its transition kernel.

(b) For each $\theta \in (0, \infty)$, let $f_{\theta} : \mathbb{T} \to (0, \infty)$ be defined a $f(h, x) := \theta^h$. Show that f_{θ} is a positive eigenfunction of P, i.e., there exists a constant λ_{θ} such that $Pf_{\theta} = \lambda_{\theta}f$.

The exercise continues on a second paper.

(c) Using the eigenfunction from part (b), we can transform P into a new probability kernel P_{θ} defined as

$$P_{\theta}((h,x),(h',x')) := \lambda_{\theta}^{-1} f_{\theta}(h,x)^{-1} P((h,x),(h',x')) f_{\theta}(h',x').$$

Let $(H_k^{\theta}, X_k^{\theta})_{k \ge 0}$ denote the Markov chain with transition kernel P_{θ} . Show that θ can be chosen in such a way that $(H_k^{\theta})_{k \ge 0}$ is symmetric random walk on \mathbb{Z} .

(d) Show that if θ is chosen as in part (c), then $\rho(P_{\theta}) = 1$. *Hint:* Define $x_0 \in S_0$ by x(i) = 1 $(i \leq 0)$. Then

$$\rho(P_{\theta}) = \lim_{n \to \infty} \left(P_{\theta}^{2n} \big((0, x_0), (0, x_0) \big) \right)^{1/2n}.$$

You can bound this from below by

$$\mathbb{P}^0 \left[H_{2n}^{\theta} = 0 \text{ and } H_k^{\theta} \ge 0 \ \forall 0 \le k \le 2n \right]$$

and then use Exercise 2.29 from the lecture notes.

(e) Calculate $\rho(P)$.

Solutions

 $\mathbf{Ex} \ \mathbf{1}$

(a) Since $(\Pi_k)_{k\geq 0}$, $(M_k)_{k\geq 0}$, and $(\Pi_k, M_k)_{k\geq 0}$ are defined in terms of i.i.d. random variables by means of a random mapping representation, these processes are Markov chains.

(b) We have

$$\mathbb{P}[M_k = m' | (\Pi_{k-1}, M_{k-1}) = (\pi, m)]$$

= $\mathbb{P}[m + 1_{\{I_k > m\}} = m'] = \begin{cases} \frac{m}{N} & \text{if } m' = m, \\ \frac{N-m}{N} & \text{if } m' = m+1, \\ 0 & \text{otherwise.} \end{cases}$

Since these transition probabilities are a function of m and m' only and do not depend on π , the chain $(M_k)_{k\geq 0}$ is autonomous with transition kernel Q(m,m) = m/N, Q(m,m+1) = (N-m)/N, and Q(m,m') = 0 in all other cases.

(c) By a similar argument as under (b), we see that

$$\mathbb{P}\big[\Pi_k = \pi' \,|\, (\Pi_{k-1}, M_{k-1}) = (\pi, m)\big]$$

can be written as a function of π and π' only, so $(\Pi_k)_{k>0}$ is autonomous too.

(d) If $\Pi_0 = 1$ (the identity permutation) and $M_0 = m$ with 0 < m < N. Then $P(1, \pi^i) = 1/N > 0$ for all $1 \le i \le N$ and Q(m, m) = m/N > 0, but $R((1, m), (\pi^i, m)) = 0$ for all i > m. This proves that the chains do not evolve independently.

(e) To construct a random variable Π with law $K(m, \cdot)$, we can proceed as follows: First, we choose $\Pi(1), \ldots, \Pi(m)$ uniformly at random from $\{1, \ldots, N\}$ without replacement. Next, we let $\Pi(m+1), \ldots, \Pi(N)$ be the remaining elements of $\{1, \ldots, N\}$, ordered from low to high. If Π is constructed in such a way, then clearly any permutation of the first m elements that leaves the remaining N - m elements invariant will preserve the law of Π . In particular, $\Pi \circ \pi^i$ has the same distribution as Π for each $1 \leq i \leq m$.

(f) Let J_1, \ldots, J_N be independent such that J_k is uniformly distributed on $\{k, \ldots, N\}$. Then by the arguments in part (e),

$$\Pi := \pi^{J_1} \circ \cdots \circ \pi^{J_m}$$

has law $K(m, \cdot)$, J_{m+1} is independent of Π , and $\Pi \circ \pi^{J_{m+1}}$ has law $K(m+1, \cdot)$.

(g) Let Π_0 have law $K(m, \cdot)$ and let I_1 be uniformly distributed on $\{1, \ldots, N\}$ and independent of Π_0 . Define

$$\Pi_1 := \Pi_0 \circ \pi^{I_1}, M_1 := m + \mathbb{1}_{\{I_1 > m\}}$$

By parts (e) and (f),

$$\mathbb{P}[\Pi_1 \in \cdot | I_1 \le m] = K(m, \cdot) \quad \text{and} \quad \mathbb{P}[\Pi_1 \in \cdot | I_1 > m] = K(m+1, \cdot),$$

which shows that

$$\mathbb{P}[\Pi_1 \in \cdot \mid M_1] = K(M_1, \cdot) \quad \text{a.s.}$$

It follows that

$$\mathbb{P}[\Pi_1 = \pi] = \sum_{m'} \mathbb{P}[\Pi_1 = \pi \mid M_1 = m'] \mathbb{P}[M_1 = m'] = \sum_{m'} Q(m, m') K(m', \pi)$$

 $(0 \le m \le N, \ \pi \in S_N)$. On the other hand,

$$\mathbb{P}[\Pi_1 = \pi] = \sum_{\pi'} \mathbb{P}[\Pi_1 = \pi \mid \Pi_0 = \pi'] \mathbb{P}[\Pi_0 = \pi'] = \sum_{\pi'} K(m, \pi') P(\pi', \pi).$$

(h) We apply Theorem 3.5 from the lecture notes to the Markov processes $(\Pi_k, M_k)_{k\geq 0}$ and $(M_k)_{k\geq 0}$, which have state spaces $T := S_N \times \{0, \ldots, N\}$ and $S := \{0, \ldots, N\}$. We let $\psi : T \to S$ denote the function $\psi(\pi, m) := m$ and let L denote the probability kernel from S to T defined as

$$L(m, (m', \pi')) := 1_{\{m=m'\}} K(m, \pi').$$

By Theorem 3.5, we need to check that

$$QL = LR,$$

where we recall that R denotes the transition kernel of the joint process $(\Pi_k, M_k)_{k\geq 0}$. Fix $0 \leq m \leq N$ and define Π_0, M_1 , and Π_1 as in part (g). Then, on the one hand,

$$\mathbb{P}\big[(\Pi_1, M_1) = (\pi', m')\big] = \sum_{m''} \mathbb{P}\big[(\Pi_1, M_1) = (\pi', m') \mid M_1 = m''\big] \mathbb{P}\big[M_1 = m''\big] \\ = \sum_{m''} Q(m, m'') L\big(m'', (m', \pi')\big),$$

while on the other hand

$$\mathbb{P}[(\Pi_1, M_1) = (\pi', m')]$$

= $\sum_{(\pi'', m'')} \mathbb{P}[(\Pi_1, M_1) = (\pi', m') | (\Pi_0, m) = (\pi'', m'')] 1_{\{m=m''\}} \mathbb{P}[\Pi_1 = \pi'']$
= $\sum_{(\pi'', m'')} L(m, (m'', \pi'')) R((\pi'', m''), (\pi', m')).$

(i) It suffices to prove that

$$\mathbb{P}[\Pi_k = \pi \mid M_k = m, \ \tau = k] = K(m, \pi) \qquad (k \ge 0, \ 0 \le m \le N, \ \pi \in S_N),$$

since then it follows that

$$\mathbb{P}[\Pi_{\tau} = \pi \mid M_{\tau} = m] = \sum_{k=0}^{\infty} \mathbb{P}[\Pi_{k} = \pi \mid M_{k} = m, \ \tau = k] \mathbb{P}[\tau = k \mid M_{k} = m] = K(m, \pi).$$

By the same argument, summing over all possible values of M_0, \ldots, M_{k-1} , it suffices to prove that

$$\mathbb{P}\big[\Pi_k = \pi \mid (M_0, \dots, M_k), \ \tau = k\big] = K(M_k, \pi) \quad \text{a.s.} \qquad (k \ge 0, \ \pi \in S_N).$$

Since $\{\tau = k\}$ is measurable w.r.t. to the σ -field generated by M_0, \ldots, M_k , it suffices to prove that

$$\mathbb{P}\big[\Pi_k = \pi \mid (M_0, \dots, M_k)\big] = K(M_k, \pi) \quad \text{a.s.} \qquad (k \ge 0, \ \pi \in S_N).$$

But this has already been proved in part (h).

(j) Since $M_{\tau} = N - 1$ a.s., it follows from part (i) that

$$\mathbb{P}\big[\Pi_{\tau} = \pi\big] = \sum_{m} \mathbb{P}\big[\Pi_{\tau} = \pi \,\big| M_{\tau} = m\,\big] \mathbb{P}[M_{\tau} = m] = K(N-1,\pi),$$

so the claim follows by observing that $K(N-1, \cdot)$ is the uniform distribution on S_N . Note that $\tau_m := \inf\{k \ge 0 : M_k = m\}$ is the first time *m* different books have been taken from the shelve. So the claim says that as soon as all but one of the books have been taken from the shelve, the order of the books is completely random.

Ex 2

(a) We observe that

$$\mathbb{P}[H_k = h - 1 \mid (H_{k-1}, X_{k-1}) = (h, x)] = \frac{1}{3},$$

$$\mathbb{P}[H_k = h + 1 \mid (H_{k-1}, X_{k-1}) = (h, x)] = \frac{2}{3}.$$

In particular, these probabilities do not depend on x so $(H_k)_{k\geq 0}$ is an autonomous Markov chain with transition kernel Q given by Q(h, h-1) := 1/3, Q(h, h+1) := 2/3, and Q(h, h') := 0 in all other cases.

(b) We calculate

$$Pf_{\theta}(h,x) = \sum_{(h',x')\in\mathbb{T}} P((h,x),(h',x')) f_{\theta}(h',x') = \frac{1}{3}\theta^{h-1} + \frac{2}{3}\theta^{h+1} = \left(\frac{1}{3}\theta^{-1} + \frac{2}{3}\theta\right) f_{\theta}(h,x),$$

which yields $\lambda_{\theta} = \frac{1}{3}\theta^{-1} + \frac{2}{3}\theta$. (c) We have

$$P_{\theta}((h,x),(h-1,\dot{x})) = \lambda_{\theta}^{-1}(\theta^{h})^{-1}\frac{1}{3}\theta^{h-1} = \frac{1}{3}\lambda_{\theta}^{-1}\theta^{-1},$$

$$P_{\theta}((h,x),(h+1,x1)) = \lambda_{\theta}^{-1}(\theta^{h})^{-1}\frac{1}{3}\theta^{h+1} = \frac{1}{3}\lambda_{\theta}^{-1}\theta,$$

$$P_{\theta}((h,x),(h+1,x2)) = \lambda_{\theta}^{-1}(\theta^{h})^{-1}\frac{1}{3}\theta^{h+1} = \frac{1}{3}\lambda_{\theta}^{-1}\theta.$$

We see from this that $(H_k^{\theta})_{k\geq 0}$ is an autonomous Markov chain with transition kernel Q_{θ} given by $Q_{\theta}(h, h-1) := \frac{1}{3}\lambda_{\theta}^{-1}\theta^{-1}$, $Q_{\theta}(h, h+1) := \frac{2}{3}\lambda_{\theta}^{-1}\theta$, and Q(h, h') := 0 in all other cases. In particular, setting $\theta = 1/\sqrt{2}$ yields $Q_{\theta}(h, h-1) = \frac{1}{2} = Q_{\theta}(h, h+1)$. (d) We use the formula

 $\rho(P_{\theta}) = \lim_{n \to \infty} \left(P_{\theta}^{2n} \big((0, x_0), (0, x_0) \big) \right)^{1/2n}.$

Since P_{θ} is a probability kernel, this limit is ≤ 1 for all $\theta \in (0, \infty)$. On the other hand,

$$P_{\theta}^{2n}((0,x_0),(0,x_0)) = \mathbb{P}^{(0,x_0)}[(H_{2n}^{\theta}, X_{2n}^{\theta}) = (0,x_0)]$$

$$\geq \mathbb{P}^{(0,x_0)}[H_{2n}^{\theta} = 0 \text{ and } H_k^{\theta} \geq 0 \ \forall 0 \leq k \leq 2n] = (Q_{\theta}|_{\mathbb{N}})^{2n}(0,0),$$

which proves that $\rho(P_{\theta}) \geq \rho(Q_{\theta}|_{\mathbb{N}})$ and hence, by Exercise 2.29, $\rho(P_{\theta}) \geq \rho(Q_{\theta})$. Set $\theta := 1/\sqrt{2}$. Then by part (c), Q_{θ} is null-recurrent, and hence by Lemma 2.5 in the lecture notes, $\rho(Q_{\theta}) = 1$. It follows that $\rho(P_{\theta}) = 1$.

(e) Since

$$P_{\theta}^{2n}((0,x_0),(0,x_0)) = \lambda_{\theta}^{-2n} f_{\theta}(0,x_0)^{-1} P^{2n}((0,x_0),(0,x_0)) f_{\theta}(0,x_0) = \lambda_{\theta}^{-2n} P^{2n}((0,x_0),(0,x_0)),$$

we see that $\rho(P) = \lambda_{\theta} \rho(P_{\theta})$. In particular, setting $\theta = 2^{-1/2}$, by part (d) have $\rho(P_{\theta}) = 1$ and hence $\rho(P) = \lambda_{\theta} = \frac{1}{3}2^{1/2} + \frac{2}{3}2^{-1/2} = \frac{1}{3}\sqrt{2}$.

Remark: It can be shown that P_{θ} with $\theta = 2^{-1/2}$ is transient. Using Theorem 2.4 of the lecture notes, this then implie that P_{θ} and hence also P are R-transient.