Markov chains questions

A few interesting exam questions: 2015 Paper 1 Section 2, 2018 Paper 1 Section 2 and 2018 Paper 2 Section 2, 2016 Paper 2 Section 2.

By a lazy version of the Markov chain with transition P we mean a chain which at each step stays in place with probability $\frac{1}{2}$ and otherwise moves according to P. This is a chain with transition matrix $\frac{P+I}{2}$. For example, if we look at the simple symmetric random walk on $\mathbb{Z}$ a lazy version of it would be the walk which moves by $+1$ or -1 with probability $\frac{1}{4}$ and stays in place with probability $\frac{1}{2}$.

1. For an irreducible finite state space chain, find the sufficient and necessary condition for an invariant distribution to be uniform.

2. Show that all eigenvalues of a transition matrix (i.e. any stochastic matrix) P have modulus at most 1 and that all eigenvalues of a reversible matrix are in $[-1, 1]$. Suppose that P is irreducible and for every x consider the set $T(x) = \{t : P^t(x, x) > 0\}$. Show that $T(x) \subseteq 2\mathbb{Z}$ if and only if -1 is an eigenvalue of P. Show that eigenvalues of the lazy version of a reversible P are in $[0, 1]$.

3. Let X be a simple random walk on a graph G and let Y be a lazy simple random walk on G. Label by $\mathcal{R}_X(t)$ and $\mathcal{R}_Y(t)$ the random variable corresponding to the number of different vertices visited by time t by the walks X and Y, respectively ($\mathcal{R}$ stands for the range). Show that for any starting distribution μ all $t, m \in \mathbb{N}$ we have $\mathbb{P}_\mu(\mathcal{R}_X(t) > m) \geq \mathbb{P}_\mu(\mathcal{R}_Y(t) > m)$.

4. Show that a simple random walk on an infinite tree in which every vertex has a degree at least three is transient.

5. Consider a simple random walk on a cycle $\mathbb{Z}_n = \{0, 1, \ldots n-1\}$ (where consecutive integers mod n are connected) started from 0. Find the distribution of the last vertex visited by this walk (i.e. for each $k > 0$ find a probability that all other vertices are visited before the vertex k).

6. (a) Let τ be a stopping time for a finite and irreducible Markov chain satisfying $\mathbb{E}[\tau] < \infty$ and $P_a(X_\tau = a) = 1$. Show that for all x

$$\mathbb{E}_a\left[\sum_{t=0}^{\tau-1} 1(X_t = x)\right] = \pi(x)\mathbb{E}_a[\tau]\,.$$

   (b) Consider a finite, irreducible and aperiodic Markov chain. Prove that for all x

$$\pi(x)\mathbb{E}_\pi[T_x] = \sum_{t=0}^{\infty}(P^t(x, x) - \pi(x)).$$

   Hint: Count the number of visits to x up until $T_x^m = \inf\{t \geq m : X_t = x\}$ in two different ways: using part (a) and also using the convergence to equilibrium theorem.

7. Show that for an irreducible Markov chain P on a finite state space S with stationary distribution π, the target time $t^a_{\odot} = \sum_{x\in S} \mathbb{E}_a[T_x]\,\pi(x)$ does not depend on $a \in S$. Deduce that $\max_{x,y\in S} \mathbb{E}_y[T_x] \leq 2\max_{x\in S} \mathbb{E}_\pi[T_x]$.

8. Show directly from the definition that positive recurrence is a class property.

9. If the size of the state space, S, of an irreducible Markov chain X_t is n show that

$$t_{\text{hit}} \leq t_{\text{cov}} \leq t_{\text{hit}} \left(1 + \frac{1}{2} + \ldots + \frac{1}{n-1}\right)$$

where $t_{\text{hit}} = \max_{x,y \in S} \mathbb{E}_x[T_y]$ where T_y is the hitting time of y and $t_{\text{cov}} = \max_{x \in S} \mathbb{E}_x[T_{\text{cov}}]$ where the cover time T_{cov} is the first time the Markov chain has visited all states [i.e. $T_{\text{cov}} = \inf\{t \geq 0 : \forall x \in S\, \exists s \leq t \text{ such that } X_s = x\}$]

1 Rate of convergence to stationarity (Mixing Times)

The following questions are about finding the rate of convergence to the stationary (invariant) distribution. All Markov chains mentioned in them are assumed to be finite (so positive recurrent), irreducible and aperiodic. To study the rate of convergence, we need to define a metric between distributions. The most widely used metric in this context is total variation distance, which for two distributions μ and ν on state space S evaluates what is the maximal disagreement they can have on any set and is formally defined by

$$\|\mu - \nu\|_{\text{TV}} = \max_{A \subset S} (\mu(A) - \nu(A)).$$

Further, a coupling of two probability measures μ, ν is a pair (X, Y) of random variables defined on the same probability space, such that $X \sim \mu$ and $Y \sim \nu$.

We also let

$$d_P(t) = \max_{x \in S} \|P^t(x, \cdot) - \pi\|_{\text{TV}}.$$

and we define the *ε-mixing time* of P to be

$$t_{\text{mix}}(\varepsilon) = \min\{t : d_P(t) \leq \varepsilon\}.$$

The *ε-mixing time* is supposed to quantify how long it takes to get ε close to the invariant distribtuion.

1. For μ and ν be two probability distributions on S check that

$$\begin{aligned}\|\mu - \nu\|_{\text{TV}} &= \sum_{x \in S} (\mu(x) - \nu(x))^+ = 1 - \sum_{x \in S} \mu(x) \wedge \nu(x) \\ &= \frac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)| = \frac{1}{2} \sup_{f : S \to [-1,1]} |\mu f - \nu f|\end{aligned}$$

where $a \wedge b = \min\{a, b\}$ denotes the minimum and $(a - b)^+ = \max\{a - b, 0\}$ and $\mu f = \sum_{x \in S} \mu(x) f(x)$. (Hint: consider a set $B = \{x \in S : \mu(x) \geq \nu(x)\}$.)

2. Let μ and ν be two probability distributions on S. Show that

$$\|\mu - \nu\|_{\text{TV}} = \inf\{\mathbb{P}(X \neq Y) : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\}$$

and that there is a coupling achieving the infimum above, which is called an optimal coupling. (Hint: Try to set X and Y to be equal as often as possible and show that this can be done with probability $p = \sum_{x \in S} \mu(x) \wedge \nu(x)$.)

3. Consider a Markov chain on the set of permutations of n cards, which from any permutation moves to a new permutation by picking a uniformly at random card (possibly the top card) and placing it on top of the deck while keeping the relative order of all other cards the same. Show that the invariant distribution for this chain is a uniform permutation. Show moreover that if P is the transition matrix of this chain, then for all $\varepsilon > 0$ there is a large enough constant $C > 0$ (which does not depend on n) such that $t_{\mathrm{mix}}(\varepsilon) \leq n \log n + Cn$. (Hint: Try running one chain from π and another one from the identity permutation and using the previous question).

4. Let P be the transition matrix of a finite reversible chain with invariant distribution π. Using the Cauchy-Schwarz inequality or otherwise prove that for all x, y and all t

$$\frac{P^{2t}(x,y)}{\pi(y)} \leq \sqrt{\frac{P^{2t}(x,x)}{\pi(x)} \cdot \frac{P^{2t}(y,y)}{\pi(y)}} \quad \text{and} \quad P^{2t+2}(x,x) \leq P^{2t}(x,x).$$

Show that if P is lazy $P^{t+1}(x,x) \leq P^t(x,x)$ for all $t \geq 0$ and deduce that $P^t(x,x) \geq \pi(x)$.

(Hint: For the lazy chain consider a suitable chain on the state space where you add a vertex for eah x, y such that $P(x,y) > 0$ and apply the previous part.)