UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Tuesday, 1 June, 2021 12:00 pm to 2:00 pm

PAPER 215

MIXING TIMES OF MARKOV CHAINS

Before you begin please read these instructions carefully

Candidates have TWO HOURS to complete the written examination.

Attempt no more than ***THREE*** *questions.*
There are ***FOUR*** *questions in total.*
The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Mixing times of Markov chains

Let Q be an irreducible and reversible transition matrix on the finite state space S with invariant distribution π. Let X be a Markov chain with matrix $P = (Q+I)/2$.

(a) Define the relaxation time t_{rel} of P. State without proof the spectral decomposition of P^t for every $t \in \mathbb{N}$.

(b) Prove that for all x

$$\sum_{k=0}^{\infty}(P^k(x,x) - \pi(x)) \leqslant \frac{e}{e-1}\sum_{k=0}^{\lceil t_{\mathrm{rel}}\rceil}(P^k(x,x) - \pi(x)).$$

(c) Let $t^{(2)}_{\mathrm{mix}}(x,\varepsilon)$ be the $\mathcal{L}_2$ ε-mixing time starting from x, i.e.

$$t^{(2)}_{\mathrm{mix}}(x,\varepsilon) = \min\left\{t \geqslant 0 : \left\|\frac{P^t(x,\cdot)}{\pi(\cdot)} - 1\right\|_{2,\pi} \leqslant \varepsilon\right\}.$$

Write $\tau_x = \inf\{t \geqslant 0 : X_t = x\}$ for the first hitting time of x. Prove that

$$t^{(2)}_{\mathrm{mix}}(x, 1/4) \leqslant 8\mathbb{E}_\pi[\tau_x]\,.$$

[You may use the identity $\pi(x)\mathbb{E}_\pi[\tau_x] = \sum_{k=0}^{\infty}(P^k(x,x) - \pi(x))$ without proof.]

(d) Using the identity $\pi(x)\mathbb{E}_\pi[\tau_x] = \sum_{k=0}^{\infty}(P^k(x,x) - \pi(x))$ or otherwise, show that there exists a universal constant C (independent of the chain) so that for all a

$$\mathbb{E}_a\left[\sum_{k=0}^{t^{(2)}_{\mathrm{mix}}(a,1/4)-1}\mathbf{1}(X_k = a)\right] \leqslant C \cdot \mathbb{E}_a\left[\sum_{k=0}^{\lceil t_{\mathrm{rel}}\rceil}\mathbf{1}(X_k = a)\right].$$

[Hint: A value of C that works is $9e/(e-1)$.]

2 Mixing times of Markov chains

(a) Define what it means for a family of Markov chains to exhibit pre-cutoff.

Let $X^{(n)}$ be a sequence of irreducible aperiodic Markov chains with relaxation times $t_{\mathrm{rel}}^{(n)}$ and 1/4-total variation mixing times $t_{\mathrm{mix}}^{(n)}$. Suppose that $t_{\mathrm{mix}}^{(n)} \to \infty$ as $n \to \infty$ and $t_{\mathrm{mix}}^{(n)}/t_{\mathrm{rel}}^{(n)}$ is bounded from above. Prove that there is no pre-cutoff.

[You may use results from the course relating the total variation mixing time and the relaxation time without proof.]

(b) The purpose of this part is to prove that the converse to the above is not true, i.e. if $t_{\mathrm{mix}}^{(n)}/t_{\mathrm{rel}}^{(n)}$ is unbounded, this does not imply cutoff.

Let P_n be a sequence of transition matrices with invariant distributions π_n and $t_{\mathrm{rel}}^{(n)}/t_{\mathrm{mix}}^{(n)} \to 0$ as $n \to \infty$ and with a cutoff. [If you wish, you may work with P_n being the transition matrix of lazy simple random walk on the hypercube $\{0,1\}^n$, for which recall that $t_{\mathrm{mix}}^{(n)}/(n \log n/2) \to 1$ as $n \to \infty$ and $t_{\mathrm{rel}}^{(n)} = n$.] Let $a_n = (t_{\mathrm{rel}}^{(n)} t_{\mathrm{mix}}^{(n)})^{-1/2}$ and define a new transition matrix for all x, y

$$\widetilde{P}_n(x,y) = (1-a_n)P_n(x,y) + a_n \pi_n(y).$$

(i) Show that
$$\|\widetilde{P}_n^t(x,\cdot) - \pi_n\|_{\mathrm{TV}} = (1-a_n)^t \cdot \|P_n^t(x,\cdot) - \pi_n\|_{\mathrm{TV}}.$$

(ii) Deduce that the family $(\widetilde{P}_n)$ does not exhibit pre-cutoff.

(iii) Let $\widetilde{t}_{\mathrm{rel}}^{(n)}$ and $\widetilde{t}_{\mathrm{mix}}^{(n)}$ be the relaxation time and 1/4-total variation mixing time respectively of $\widetilde{P}_n$. Show that

$$\frac{\widetilde{t}_{\mathrm{rel}}^{(n)}}{\widetilde{t}_{\mathrm{mix}}^{(n)}} \to 0 \quad \text{as } n \to \infty.$$

3 Mixing times of Markov chains

(a) Let P be an irreducible transition matrix on the finite set S and suppose it is reversible with respect to the invariant distribution π.

Suppose that $\widetilde{P}$ is another irreducible transition matrix on S reversible with respect to the invariant distribution $\widetilde{\pi}$. Let $E = \{(x,y) : P(x,y) > 0\}$ and $\widetilde{E} = \{(x,y) : \widetilde{P}(x,y) > 0\}$. For every (x,y) set $\mathcal{P}_{x,y}$ for the set of paths from x to y and let ν_{xy} be a probability measure on $\mathcal{P}_{x,y}$. Let

$$B = \max_{e \in E} \left(\frac{1}{Q(e)} \sum_{(x,y) \in \widetilde{E}} \widetilde{Q}(x,y) \sum_{\Gamma \in \mathcal{P}_{x,y} : e \in \Gamma} \nu_{xy}(\Gamma) |\Gamma| \right),$$

where $Q(x,y) = \pi(x)P(x,y)$ and $\widetilde{Q}(x,y) = \widetilde{\pi}(x)\widetilde{P}(x,y)$. Show that the spectral gaps γ and $\widetilde{\gamma}$ of P and $\widetilde{P}$ respectively satisfy

$$\widetilde{\gamma} \leqslant \left(\max_x \frac{\pi(x)}{\widetilde{\pi}(x)} \right) B\gamma,$$

[You may use results on the comparison of spectral gaps via comparison of Dirichlet forms without proof.]

(b) Let $G = (V, E)$ be a transitive graph on n vertices with vertex degree d and diameter Δ. Let γ be the spectral gap of simple random walk on G.

(i) For each $x, y \in V$ let $\mathcal{P}^*_{x,y}$ be the set of shortest paths from x to y and let ν_{xy} be the uniform measure on $\mathcal{P}^*_{x,y}$. For $e \in E$ and $x \in V$ we define

$$f(e) = \sum_{x,y} \sum_{\Gamma \in \mathcal{P}^*_{x,y}} \frac{\mathbf{1}(e \in \Gamma)}{|\mathcal{P}^*_{x,y}|} \quad \text{and} \quad \widetilde{f}(x) = \sum_{y \sim x} f(x,y).$$

Using transitivity show that $\widetilde{f}$ is a constant function and then show that for every $e \in E$

$$f(e) \leqslant 2n \cdot \Delta.$$

(ii) By comparing the transition matrix P of simple random walk on G to the matrix defined via $\widetilde{P}(x,y) = \pi(y)$ or othewise, prove that

$$\frac{1}{\gamma} \leqslant 2 \cdot d \cdot \Delta^2.$$

[You may use results from the course as long as they are stated clearly.]

4 Mixing times of Markov chains

Consider the following Markov chain on the hypercube $\{0,1\}^n$: for $x = (x_1, \ldots, x_n)$ and $x' \in \{(0, x_1, \ldots, x_{n-1}), (1, x_1, \ldots, x_{n-1})\}$

$$P(x, x') = \frac{1}{3}$$

while for $x' \in \{(0, x_3, \ldots, x_n, x_1), (1, x_3, \ldots, x_n, x_1)\}$ we have

$$P(x, x') = \frac{1}{6}.$$

In words, when the current state is $x = (x_1, \ldots, x_n)$, then first we either shift the vector to the right with probability $2/3$ or to the left with probability $1/3$. Then we refresh the bit in the first coordinate to 0 or 1 equally likely.

(a) Check that $\pi(x) = 1/2^n$ for all $x \in \{0,1\}^n$ is the invariant distribution.

(b) Show that this process exhibits total variation cutoff around time $3n$.

[Hint: Prove separately an upper and a lower bound on $t_{\mathrm{mix}}(\varepsilon)$. You may use the following facts about a random walk: if X is a random walk on $\mathbb{Z}$ with $P(i, i+1) = 2/3 = 1 - P(i, i-1)$ and $\tau_x = \inf\{t \geqslant 0 : X_t = x\}$ for $x \in \mathbb{Z}$, then for all $a, b \geqslant 0$

$$\mathbb{E}_0[\tau_a] = 3a \text{ and } \mathrm{Var}(\tau_a) = 24a,$$

$$\mathbb{E}_0[\tau_a \wedge \tau_{-b}] = 3a - 3(a+b) \cdot \frac{1 - 2^{-a}}{2^b - 2^{-a}} \quad \text{and}$$

$$\mathrm{Var}(\tau_a \wedge \tau_{-b}) \leqslant 24a + 3(a+b)\mathbb{E}_0[\tau_a \wedge \tau_{-b}] \cdot \frac{1 - 2^{-a}}{2^b - 2^{-a}}.]$$

END OF PAPER