University of
Cambridge

Mathematics Tripos

Part III

Mixing Times of Markov Chains

Example Sheet I

January 3, 2020

Solutions by
Joshua Snyder

1 Introduction

These are written solutions to Mixing Times of Markov Chains Example Sheet 1. Solutions are based on those handed out by Samuel Thomas and are not endorsed by the lecturer nor necessarily correct.

2 Questions

Question (Question 1)**.** Let P be the transition matrix of a Markov chain with values in E and let μ and ν be two probability distributions on E. Show that

$$\|\mu P - \nu P\|_{\mathrm{TV}} \leq \|\mu - \nu\|_{\mathrm{TV}} \ .$$

Deduce that $d(t) = \max_x \|P^t(x, \cdot) - \pi\|_{\mathrm{TV}}$ is decreasing as a function of t, where π is the invariant distribution.

Solution. Since P is a stochastic matrix, any eigenvalue λ of P satisfies $|\lambda| \leq 1$. To prove this suppose that $Pv = \lambda v$ and let $|v_k| = \max |v_1|, |v_2|, \ldots, |v_n|$; so $|v_k| > 0$. Then

$$|\lambda||v_k| = |\lambda v_k| = |(Pv)_k| = |\sum_1^n P_{k,j} v_j| \leq |v_k|$$

since P is stochastic. Therefore $|\lambda| \leq 1$ and so $\|P\| \leq 1$.
Then we can apply the inequality $\|(\mu - \nu)P\|_1 \leq \|P\|_1 \|\mu - \nu\|_1$.
For the last part, using the fact that $\pi P = P$ gives the result. □

Question (Question 2)**.** Let $\Sigma = \prod_{i=1}^n \Sigma_i$ where Σ_i are finite sets. For each i, let μ_i and ν_i be probability distributions on Σ_i and set $\mu = \prod_{i=1}^n \mu_i$ and $\nu = \prod_{i=1}^n \nu_i$. Show that

$$\|\mu - \nu\|_{\mathrm{TV}} \leq \sum_{i=1}^n \|\mu_i - \nu_i\|_{\mathrm{TV}}$$

Solution. For each $i = 1, \ldots, n$, let (X_i, Y_i) be an optimal coupling of (μ_i, ν_i). Let $X(X_1, \ldots, X_n)$ and $Y(Y_1, \ldots, Y_n)$. Then (X, Y) is a coupling of (μ, ν). From this we get that

$$\|\mu - \nu\|_{\mathrm{TV}} \leq \mathbb{P}(X \neq Y) \leq \sum_1^n \mathbb{P}(X_i \neq Y_i) = \sum_1^n \|\mu_i - \nu_i\|_{\mathrm{TV}}$$

□

Question. Let X and Y be Poisson random variables with parameters λ and μ respectively. Writing $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ for their laws, prove that

$$\|\mathcal{L}(X) - \mathcal{L}(Y)\|_{\mathrm{TV}} \leq |\lambda - \mu|$$

Solution 1. Take a coupling of (X, Y) by taking first $Y \sim Poiss(\mu)$ and $Z \sim Poiss(\lambda - \mu)$ independently and setting $X = Y + Z$. Then we have

$$\|\mathcal{L}(X) - \mathcal{L}(Y)\|_{\mathrm{TV}} \leq \mathbb{P}(X \neq Y) = \mathbb{P}(Z \neq 0) = 1 - e^{-(\lambda - \mu)} \leq |\lambda - \mu|$$

□

Solution 2. WLOG $\lambda \geq \mu$.

$$\begin{aligned}
\|\mathcal{L}(X) - \mathcal{L}(Y)\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{n\geq 1}\left|\frac{e^{-\lambda}\lambda^n}{n!} - \frac{e^{-\mu}\mu^n}{n!}\right| \\
&\leq \frac{1}{2}\sum_{n\geq 1}\frac{e^{-\lambda}}{n!}(\lambda^n - \mu^n) \\
&= \frac{1}{2}(\lambda-\mu)\sum_{n\geq 1}\frac{e^{-\lambda}}{n!}(\lambda^{n-1} + \lambda^{n-2}\mu + \cdots + \mu^{n-1}) \\
&\leq \frac{1}{2}(\lambda-\mu)\sum_{n\geq 1}\frac{e^{-\lambda}}{n!}(n\lambda^{n-1}) \\
&\leq \frac{1}{2}(\lambda-\mu)\left(1 + \sum_{n\geq 1}\frac{e^{-\lambda}\lambda^n}{n!}\right) = \frac{1}{2}\cdot 2\cdot(\lambda-\mu) = \lambda - \mu
\end{aligned}$$

□

> **Question** (Question 4)**.** Let Y be a random variable with values in $\mathbb{N}$ which satisfies
>
> $$\mathbb{P}(Y = j) \leq c, \text{ for all } j > 0 \text{ and } \mathbb{P}(Y = j) \text{ is decreasing in } j,$$
>
> where c is a positive constant. Let Z be an independent random variable with independent values in $\mathbb{N}$. Prove that
>
> $$\|\mathbb{P}(X + Y = \cdot) - \mathbb{P}(Y = \cdot)\|_{\mathrm{TV}} \leq c\mathbb{E}[Z]$$

Solution. Since the law of Y is decreasing, $\mathbb{P}(Y = j) \geq \mathbb{P}(Y + k = j)$ for all $k, j \geq 0$. Therefore

$$\begin{aligned}
\|\mathbb{P}(Y + k = \cdot) - \mathbb{P}(Y = \cdot)\|_{\mathrm{TV}} &= \sum_{j:\mathbb{P}(Y=j)\geq\mathbb{P}(Y+k=j)} (\mathbb{P}(Y = j) - \mathbb{P}(Y + k = j)) \\
&= \sum_j (\mathbb{P}(Y = j) - \mathbb{P}(Y + k = j) \\
&= 1 - \mathbb{P}(Y \geq k) = \mathbb{P}(Y \leq k) \leq ck
\end{aligned}$$

□

Remark. Using this definition of the total variation distance can be very helpful when we know that a probability distribution is monotone past a certain value, which a lot of distributions are.

> **Question** (Question 5)**.** Let X be a Markov Chain and let W and V be random variables taking values in $\mathbb{N}$ and suppose they are independent of X. Prove that
>
> $$\|\mathbb{P}(X_W = \cdot) - \mathbb{P}(X_V = \cdot)\|_{\mathrm{TV}} \leq \|\mathbb{P}(W = \cdot) - \mathbb{P}(V = \cdot)\|_{\mathrm{TV}}$$

Solution. Let (X_W, X_V) be the optimal coupling of $\mu = \mathbb{P}(W = \cdot)$ and $\nu = \mathbb{P}(V = \cdot)$. Then we have

$$\|\mathbb{P}(X_W = \cdot) - \mathbb{P}(X_V = \cdot)\|_{\mathrm{TV}} \leq \mathbb{P}(X_W \neq X_V) \leq \mathbb{P}(W \neq V)$$

$\square$

> **Question** (Question 6)**.** Let $G = (V, E)$ be a finite connected graph with maximal distance between any two vertices equal to D. Suppose that X is a lazy simple random walk on G. Prove that for all $\epsilon < \frac{1}{2}$ we have
>
> $$t_{\mathrm{mix}}(\epsilon) \geq \frac{D}{2}$$

Idea. Consider $\bar{d}$ instead of d for lower bounds.

Solution. If $t \leq \frac{D}{2}$ and $d(x, y) = D$ then $supp(P^t(x, \cdot)) \cap supp(P^t(y, \cdot)) = \emptyset$ i.e they never meet and therefore $\bar{d}(t) = 1$.
Therefore $d(t) \geq \frac{1}{2}\bar{d}(t) = \frac{1}{2}$ and the result follows. $\square$

> **Question** (Question 7)**.** Let X be a Markov chain in E with transition matrix P and invariant distribution π. Let $A \subset E$ be a subset with $\pi(A) \geq \frac{1}{8}$. Let $\tau_A = \inf t \geq 0 : X_t \in A$. Prove that there exists a positive constant c such that
>
> $$t_{\mathrm{mix}}\left(\frac{1}{4}\right) \geq c \max_x \mathbb{E}_x[\tau_A]$$

Idea. Haven't hit a set of size $\frac{1}{8}$, so you know you are lying in only $\frac{7}{8}$ and so you cannot be mixed. Take as the distinguishing set the bad $\frac{1}{8}$.

Solution. Let $t = t_{\mathrm{mix}}\left(\frac{1}{16}\right)$ then $d(t) \leq \frac{1}{16}$ and hence for all sets A we have $P^t(x, A) \geq \pi(A) - \frac{1}{16}$.
Let A be such that $\pi(A) = \frac{1}{8}$, then $P^t(X, A) \geq \frac{1}{16}$.
Hence $\tau(A) \lessapprox t \cdot Geom\left(\frac{1}{16}\right)$ since $\tau_A \leq \inf k \geq 0 : X_t \in A \lessapprox Geom\left(\frac{1}{16}\right)$ from the above.
Therefore $\mathbb{E}[\tau_A] \lessapprox t$. Using the submultiplicativity of $\bar{d}$ and the inequality $d(t) \leq 2\bar{d}(t)$, the claim is deduced. $\square$

> **Question** (Question 13)**.** Let P be a 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
>
> $$\frac{P^{2t}(x, y)}{\pi(y)} \leq \sqrt{\frac{P^{2t}(x, x)}{\pi(x)} \cdot \frac{P^{2t}(y, y)}{\pi(y)}} \text{ and } P^{2t+2}(x, x) \leq P^{2t}(x, x)$$

Solution 1.

$$\begin{aligned}
\left(\frac{P^{2t}(x,y)}{\pi(y)}\right)^2 &= \sum_z \frac{P^t(x,z)P^t(z,y)}{\pi(y)} \\
&= \sum_z \frac{P^t(x,z)P^t(z,y)}{\pi(z)} \\
&\leq \left(\sum_z \frac{P^{2t}(x,z)}{\pi(z)}\right)\left(\sum_z \frac{P^{2t}(y,z)}{\pi(z)}\right) \\
&\leq \left(\frac{P^{2t}(x,x)}{\pi(x)}\right)\left(\frac{P^{2t}(y,y)}{\pi(y)}\right)
\end{aligned}$$

which upon taking square roots of both sides completes the proof.
For the second part since $|\lambda_j| \leq 1$ we have that

$$\begin{aligned}
\frac{P^{2t}(x,x)}{\pi(x)} &= \sum_1^{|\Sigma|} f_j^2(x)\lambda_j^{2t} \\
&\geq \sum_1^{|\Sigma|} f_j^2(x)\lambda_j^{2t+2} \\
&= \frac{P^{2t+2}(x,x)}{\pi(x)}
\end{aligned}$$

□