Probability and Computation: 24th Jan Problem session Solutions

Question 1. *Prove that if P is the transition matrix of a finite, irreducible, aperiodic Markov chain then there exists some time $K < \infty$ such that, for every $w, z \in \mathcal{I}$, if $t \geq K$ then $P^t_{w,z} > 0$.*

Solution: We use the following number-theoretic fact:

> Any set of non-negative integers which is closed under addition and has greatest common divisor 1 must contain all but finitely many of the non-negative integers.

For $x \in \mathcal{I}$, let $T(x) = \{t \geq 1 : P^t_{x,x} > 0\}$. Since the chain is aperiodic, the greatest common divisor of $T(x)$ is 1. The set $T(x)$ is closed under addition: if $s, t \in T(x)$, then

$$P^{s+t}_{x,x} = \sum_{y \in \mathcal{I}} P^s_{x,y} \cdot P^t_{y,x} \geq P^s_{x,x} \cdot P^t_{x,x} > 0,$$

and hence $s + t \in T(x)$. Therefore there exists a $t(x)$ such that $t \geq t(x)$ implies $t \in T(x)$.

By irreducibility we know that for any $y \in \mathcal{I}$ there exists $r = r(x, y)$ such that $P^r_{x,y} > 0$. Therefore, for $t \geq t(x) + r$, $P^t_{x,y} \geq P^{t-r}_{x,x} P^r_{x,y} > 0$. For $t \geq t'(x) := t(x) + \max_{y \in \mathcal{I}} r(x, y)$, we have $P^t_{x,y} > 0$ for all $y \in \mathcal{I}$. Finally, if $t \geq K := \max_{x \in \mathcal{I}} t'(x)$, then $P^t_{x,y} > 0$ for all $x, y \in \mathcal{I}$.

Question 2. *Prove that for any states x and y of a finite irreducible chain $\mathbf{E}_x\left[\tau_y^+\right] < \infty$.*

Solution: By irreducibility there exist $r > 0$ and $\varepsilon > 0$ such that for any states $z, w \in \Omega$, there exists a $j \leq r$ with $P^j_{z,w} > \varepsilon$.

Thus for any $z \in \mathcal{I}$,

$$\mathbf{P}[\, X_{t+1} \neq y, \dots, X_{t+r} \neq y \mid X_t = z \,] \leq 1 - \varepsilon.$$

Hence for $k \geq 0$,

$$\begin{aligned}
\mathbf{P}_x\left[\tau_y^+ > (k+1)r\right] &= \sum_{z \in \mathcal{I}} \mathbf{P}\left[\, X_{(k+1)r} \neq y, \dots, X_{kr+1} \neq y \mid X_{kr} = z \,\right] \mathbf{P}_x\left[\tau_y^+ > kr, X_{kr} = z\right] \\
&\leq (1 - \varepsilon) \sum_{z \in \mathcal{I}} \mathbf{P}_x\left[\tau_y^+ > kr, X_{kr} = z\right] \\
&= (1 - \varepsilon) \mathbf{P}_x\left[\tau_y^+ > kr\right] \\
&\leq (1 - \varepsilon)^{k+1}.
\end{aligned}$$

Since $\tau_y^+ \geq 0$ and $\mathbf{P}_x\left[\tau_y^+ > t\right]$ is a non-increasing function of t,

$$\mathbf{E}_x\left[\tau_y^+\right] = \sum_{t \geq 0} \mathbf{P}_x\left[\tau_y^+ > t\right] \leq \sum_{k \geq 0} r \mathbf{P}_x\left[\tau_y^+ > kr\right] \leq r \sum_{k \geq 0} (1 - \varepsilon)^k < \infty.$$

Question 3. *Prove the following Lemma from class: For any probability distributions μ and η on a countable state space Ω*

$$\|\mu - \eta\|_{tv} = \sup_{A \subset \Omega} |\mu(A) - \eta(A)|.$$

Solution: Let $\Omega^+ = \{\omega : \mu(\omega) \geq \eta(\omega)\}$ and $\Omega^- = \{\omega : \mu(\omega) < \eta(\omega)\}$. Then

$$\max_{A \subseteq \Omega} \mu(A) - \eta(A) = \mu(\Omega^+) - \eta(\Omega^+)$$

and

$$\max_{A \subseteq \Omega} \eta(A) - \mu(A) = \eta(\Omega^-) - \mu(\Omega^-).$$

Since $\Omega = \Omega^+ \cup \Omega^-$ and $\Omega^+ \cap \Omega^- = \emptyset$ we have

$$\mu(\Omega^+) + \mu(\Omega^-) = 1 \qquad \text{and} \qquad \eta(\Omega^+) + \eta(\Omega^-) = 1,$$

thus

$$\mu(\Omega^+) - \eta(\Omega^+) = \eta(\Omega^-) - \mu(\Omega^-).$$

Hence

$$\sup_{A \subset \Omega} |\mu(A) - \eta(A)| = |\mu(\Omega^+) - \eta(\Omega^+)| = |\mu(\Omega^-) - \eta(\Omega^-)|.$$

Combining the above yields

$$2 \sup_{A \subset \Omega} |\mu(A) - \eta(A)| = |\mu(\Omega^+) - \eta(\Omega^+)| + |\mu(\Omega^-) - \eta(\Omega^-)| = \sum_{\omega \in \Omega} |\mu(\omega) - \eta(\omega)| = 2 \left\| \mu - \eta \right\|_{tv} .$$
