Assignment 2

What you are allowed to do:

- Discuss the problems with your classmates.
- Ask the teacher for help (some points are hard, it is ok not to know how to proceed).
- Use a computer for linear operations (e.g. matrix multiplication, matrix vector etc), provided you know how to do it manually.

What you are bound to do:

- Write down the solutions yourself, while being alone.

What you are not allowed to do:

- Copy the written solutions from a classmate.
- Use AI to output the solutions.

Problems

Exercise 0.1. Let P be the Markov operator of an irreducible Markov Chain on a finite state space S. Let m be its invariant measure, and $P^\dagger$ the adjoint operator of P in $L^2(m)$.

a. Check that if P, Q are Markov operators, then PQ and $\alpha P + (1-\alpha)Q$ are Markov operators. In particular $PP^\dagger$ and P^2 are Markov operators.
b. Find the invariant measure of $PP^\dagger$ and P^2.
c. Give an example where P is irreducible, but P^2 and $PP^\dagger$ are not irreducible.
d. Prove that, if P is irreducible and aperiodic, then P^2 is irreducible and aperiodic.
e. Prove that $PP^\dagger$ is aperiodic and each communicating class is closed.
f. Assume that P is irreducible and aperiodic. Let X_t and Y_t be Markov Chains with Markov operators P^2 and $PP^\dagger$ respectively, prove that X_t converges to its "equilibrium" faster than Y_t in the following sense. Let μ_t^x (respectively ν_t^x) be the law of X_t when $X_0 = x$ (respectively Y_t when $Y_0 = x$). Prove that

$$\sup_x \lim_{t\to\infty} \frac{1}{t} \log \|\mu_t^x - m\|_{TV} \le \sup_x \lim_{t\to\infty} \frac{1}{t} \log \|\nu_t^x - m\|_{TV}$$

Solution

a. Let $R = PQ$, thus $r_{x,z} = \sum_y p_{x,y} q_{y,z}$. And exchanging summations $r_{x,z} \ge 0$ and $\sum_z r_{x,z} = \sum_y p_{x,y} \sum_z q_{y,z} = \sum_y p_{x,y} = 1$. As for the convex combination, the proof is immediate by linearity.
b. The invariant measure of all these operators is m. In general, if P, Q have the same invariant measure, then $m((PQ)f) = m(P(Qf)) = m(Qf) = m(f)$. So m is invariant for PQ.
c. We can take a unique cycle $S = \mathbb{Z}_{2k}$ with an even number of points, P corresponds to moving clockwise with probability 1, $p_{x,x+1} = 1$ (sums are understood $\pmod{2}k$). Then P is irreducible since there is a

path joining any two points. $PP^\dagger$ corresponds to making one step clockwise and one counter-clockwise, that is $PP^\dagger = I$, the walk just does not move. P^2 corresponds to making two steps clockwise $p^{(2)}_{x,x+2} = 1$. So there is not path joining points with different parity.

d. For an irreducible aperiodic chain on a finite state space, there exists T large enough so that $p^{(t)}_{x,y} > 0$ for all $t > T$ and $x, y \in S$. In particular the same holds true, changing T to $T' = \lceil T/2 \rceil$, for the entries of P^2.
e. Let $(q_{x,y})$ denote the entries of $Q = PP^\dagger$. $q_{x,z} > 0$ iff there exists $y \in S$ such that $p_{x,y}, p_{z,y} > 0$. In particular $q_{x,z} > 0$ iff $q_{z,x} > 0$ and $q_{x,x} > 0$ for all $x \in S$ (since each point x has at least one *successor* y_x such that $p_{x,y_x} > 0$). Therefore each Q-communicating class is closed.
f. By the exponential convergence theorem, we need to check that $|\lambda_1(P^2)| \leq |\lambda_1(PP^\dagger)|$, where $\lambda_1(Q)$ is the largest (by modulus) eigenvalue of an irreducible Markov operator Q, except $\lambda_0 = 1$; or equivalently, the largest eigenvalue of Q restricted to functions f with $m(f) = 0$, where m is the invariant measure of Q.

For P^2 we have $\lambda_1(P^2) = \lambda_1(P)^2 = \overline{\lambda_1(P^\dagger)}^2$. Take f_1 an associated (complex) eigenfunction (with $m(f_1) = 0$) of $P^\dagger$, so that $P^\dagger f_1 = \overline{\lambda_1(P^\dagger)} f_1$. Then, since $PP^\dagger$ is symmetric, $\lambda_1(PP^\dagger) > 0$ and thus denoting $\langle \cdot, \cdot \rangle$ the Hermitian product in $L^2(m)$

$$|\lambda_1(PP^\dagger)| = \sup_{f\,:\,m(f)=0} \frac{\langle f, PP^\dagger f\rangle}{m(|f|^2)} = \sup_{f\,:\,m(f)=0} \frac{\langle P^\dagger f, P^\dagger f\rangle}{m(|f|^2)} \geq \frac{\langle P^\dagger f_1, P^\dagger f_1\rangle}{m(|f_1|^2)} = \overline{\lambda_1(P^\dagger)}\lambda_1(P^\dagger) = |\lambda_1(P^2)|$$

Exercise 0.2. Consider an irreducible Markov Chain on a finite state space S with transition probabilities $p^0_{x,y}$. Let $c(x,y) \in \mathbb{R}$ be defined for all edges (x,y) such that $p_{x,y} > 0$. Fix $\beta > 0$ and define the tilted transition probabilities

$$p^\beta_{x,y} := \frac{e^{-\beta c(x,y)}}{Z^\beta_x} p^0_{x,y}$$

a. Determine the value of Z^β_x and check that the Markov Chain with transition probabilities $p^\beta_{x,y}$ is still irreducible.
b. Explain why we can assume that $\min_{y\,:\,p^0_{x,y}>0} c(x,y) = 0$ and verify that in this case

$$\inf_y p^0_{x,y} \leq Z^\beta_x \leq 1$$

c. Let m^β be the invariant probability measure associated to the transition probabilities p^β. Characterize the limit $\lim_{\beta\to\infty} m^\beta$. *Hint: For the sake of simplicity, assume that the values* $\sum_{e\in\theta} c(e)$ *are distinct for distinct* $\theta \in \Theta$*, where* Θ *is the space of rooted spanning in-arborescences.*
d. Give an explicit answer in the case $c(x,y) = (V(y) - V(x))^+$, where $V : S \to \mathbb{R}$ is a given injective function, and for $a \in \mathbb{R}$, $a^+ = \max(a, 0)$. *Hint: For the sake of simplicity, assume that* $p^0_{x,y} > 0$ *whenever* $p^0_{y,x} > 0$.

Solution

a. Since $p^\beta_{x,\cdot}$ is a probability, we have

$$Z^\beta_x := \sum_y e^{-\beta c(x,y)} p^0_{x,y}$$

The Markov chain is still irreducible since $p^\beta_{x,y} > 0$ iff $p^0_{x,y} > 0$.

b. If we change $c(x,y)$ to $c(x,y) - a(x)$, then $p^\beta_{x,y}$ does not change, regardless of $a(x)$. Therefore, to make the computation simpler, we can take $a(x) = \min_y c(x,y)$. In other words, we just assume $\min_y c(x,y) = 0$. In particular

$$\inf_y p^0_{x,y} \le \max_y e^{-\beta c(x,y)} p^0_{x,y} \le Z^\beta_x \le \sum_y e^{-0} p^0_{x,y} = 1$$

c. We know that

$$m^\beta_x = c^\beta \sum_{\theta \in \Theta_x} w^\beta(\theta), \qquad w^\beta(\theta) := \prod_{e \in \theta} p^\beta_e$$

where $c^\beta > 0$ is the suitable normalization constant (which is just the same sum over $\theta \in \Theta$). Using the definition of $p^\beta_{x,y}$

$$m^\beta_x = c^\beta \sum_{\theta \in \Theta_x} \frac{e^{-\beta \sum_{e\in\theta} c(e)}}{\prod_{y \neq x} Z^\beta_y} w^0(\theta)$$

where the product in the denominator runs over all the $y \in S$ from which there is an outgoing edge, namely all the $y \in S$ except the root x. For $Z^\beta := \prod_x Z^\beta_x$ and $c(\theta) := \sum_{e\in\theta} c(e)$:

$$m^\beta_x = \frac{c^\beta}{Z^\beta} \sum_{\theta \in \Theta_x} Z^\beta_x e^{-\beta c(\theta)} w^0(\theta)$$

We know that for an irreducible chain Θ_x is non-empty for any $x \in S$. And simply estimating the sum with the largest summand (as above)

$$Z^\beta_x \max_{\theta \in \Theta_x} e^{-\beta c(\theta)} w^0(\theta) \le m^\beta_x \frac{Z^\beta}{c^\beta} \le |\Theta_x| Z^\beta_x \max_{\theta \in \Theta_x} e^{-\beta c(\theta)} w^0(\theta)$$

and thus using the bound in point b.

$$\min_y p^0_{x,y} \min_{\theta \in \Theta_x} w^0(\theta) e^{-\beta \min_{\theta\in\Theta_x} c(\theta)} \le m^\beta_x \frac{Z^\beta}{c^\beta} \le |\Theta_x| \max_{\theta \in \Theta_x} w^0(\theta) e^{-\beta \min_{\theta\in\Theta_x} c(\theta)}$$

This implies that m^β_x concentrates (exponentially fast as $\beta \to \infty$) on the point x that carries the **minimal spanning arborescence** for the cost $c(x,y)$, that is on the root of the tree θ that minimizes $c(\theta)$. Indeed, for such an x and $y \neq x$, $m^\beta_y / m^\beta_x \to 0$ as $\beta \to \infty$.

d. We want to show that in the case $c(x,y) = (V(y) - V(x))^+$ the root of the minimal spanning arborescence is the minimizer of $V(\cdot)$ (which is unique since we assumed V injective). One may give a combinatorial proof in this case using $c(x,y) - c(y,x) = V(y) - V(x)$. Instead, let us give a *Markov Chain* proof, making good use of the fact that the minimal spanning arborescence does not depend on the values of $p^0_{x,y}$, but just on the edges on which this probability is non-vanishing.

First assume that the invariant measure m^0 of $(p^0_{x,y})$ is reversible, then $m^\beta_x := c^\beta e^{-\beta V(x)} Z^\beta_x m^0_x$ is reversible for $(p^\beta_{x,y})$ since

$$m^\beta_x p^\beta_{x,y} = Z^\beta_x e^{-\beta V(x)} \frac{e^{-\beta (V(y)-V(x))^+}}{Z^\beta_x} m^0_x p^0_{x,y} = e^{-\beta \max(V(x),V(y))} m^0_x p^0_{x,y} = e^{-\beta \max(V(x),V(y))} m^0_y p^0_{y,x} = m^\beta_y p^\beta_{y,x}$$

Thus, since $V(x) + (V(y) - V(x))^+ = \max(V(y), V(x))$, reasoning as above

$$m_x^\beta = c^\beta m_x^0 \sum_y e^{-\beta \max(V(y),V(x))} p_{x,y}^0$$

concentrates on the points x minimizing

$$x \mapsto \min_{y\,:\,p_{x,y}^0>0} \max(V(y), V(x))$$

In particular, if $p_{x,x}^0 > 0$, it concentrates on the minimizer of $V(x)$.

If m^0 is not reversible, we can change $p_{x,y}^0$ to $q_{x,y}^0 := p_{x,y}^0 + p_{y,x}^0 m_y^0 / m_x^0$ (for which m^0 is reversible) without changing the minimal spanning arborescence (which only depends on the $c(x,y)$), so the answer does not change even when m^0 is not reversible.

Exercise 0.3. Consider an irreducible Markov Chain on a finite state space S with transition probabilities $(p_{x,y})$. Recall that we consider S as a graph where the edges are the $e = (x,y)$ such that $p_{x,y} > 0$. The weight of a rooted spanning in-forest F is the product

$$w(F) := \prod_{e \in F} p_e$$

For $A \subset S$ nonempty, let $\mathfrak{F}_A$ be the set of rooted spanning in-forests, whose set of roots is exactly A. For $x \in S$ and $r \in A$, define $\mathfrak{F}_A(x \to r)$ as the set of forests $F \in \mathfrak{F}_A$ such that x is in an in-arborescence rooted at r (in other words, following the unique path in F outgoing from x, one reaches $r \in A$).

Let now $r \in A \subset S$. Since the chain is irreducible and finite, $\mathbb{P}(\tau_A < \infty) = 1$, and let $h(x) \equiv h_{r,A}(x) := \mathbb{P}_x(X_{\tau_A} = r)$. Prove that

$$h(x) = \frac{\sum_{F \in \mathfrak{F}_A(x \to r)} w(F)}{\sum_{F \in \mathfrak{F}_A} w(F)} \tag{1}$$

Solution

Since the Markov Chain is irreducible and has an invariant measure (being the state finite), h is the unique solution to the problem for the unknown u

$$(I - \mathbf{1}_{A^c} P) u = \mathbf{1}_r$$

Let $g(x)$ be the right hand side of Equation 1, we show that g also shows this equation, so $g = h$.
Clearly $g(x) = \mathbf{1}_r(x)$ for $x \in A$, so we only need to check that $g(x) = (Pg)(x)$ for $x \notin A$. The denominator $\sum_{F \in \mathfrak{F}_A} w(F)$ simplifies on both sides of this equation so it remains to verify

$$\sum_{G \in \mathfrak{F}_A(x \to r)} w(G) \overset{?}{=} \sum_y p_{x,y} \sum_{F \in \mathfrak{F}_A(y \to r)} w(F), \qquad x \notin A$$

Since $\sum_y p_{x,y} = 1$ can we write this as

$$\sum_{z \neq x} \sum_{G \in \mathfrak{F}_A(x \to r)} p_{x,z} w(G) \overset{?}{=} \sum_{y \neq x} \sum_{F \in \mathfrak{F}_A(y \to r)} p_{x,y} w(F), \qquad x \notin A$$

We can match terms on both sides, building a bijection between indices (z, G) (with $x \to r$ in G) on the left and (y, F) (with $y \to r$ in F) on the right as follows.

- If $z \to r$ in G, then we take $y = z$ and $G = F$, in other words these terms are on both sides.
- If $z \nrightarrow r$, then (x, z) is not in G. We match such a (z, G) with the element (y, F) in the r.h.s., where y is the unique vertex such that $(x, y) \in G$; and F is the forest obtained from G removing (x, y) and adding (x, z).

This is a bijection between summands in the l.h.s. and r.h.s. and clearly $p_{x,z}w(G) = p_{x,y}w(F)$.

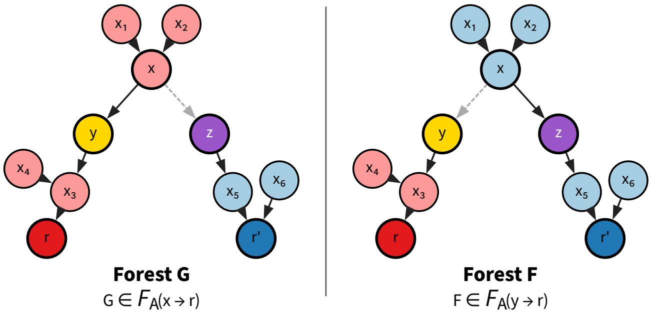

Figure 1: A pair (z, G) with $G \in \mathfrak{F}_A(x \to r)$ and $z \nrightarrow r$, is mapped bijectively to a (y, F) with $F \in \mathfrak{F}_A(y \to r)$, just switching the successor of x.