Problem: the i-SIR Markov chain

In all the problem, π, resp. $\tilde{\pi}$, are probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (where $\mathcal{B}(\mathbb{R})$ is the Borel sigma-field on $\mathbb{R}$) and we assume that these distributions have **strictly positive densities** with respect to the Lebesgue measure λ. For simplicity, we also denote by π, resp. $\tilde{\pi}$, their densities with respect to λ, that is, $\pi(\mathrm{d}x) = \pi(x)\lambda(\mathrm{d}x) = \pi(x)\mathrm{d}x$ and $\tilde{\pi}(\mathrm{d}x) = \tilde{\pi}(x)\lambda(\mathrm{d}x) = \tilde{\pi}(x)\mathrm{d}x$ where $\boxed{\pi(x) > 0 \text{ and } \tilde{\pi}(x) > 0 \text{ for any } x \in \mathbb{R}}$ and where we recall the abuse of notation $\lambda(\mathrm{d}x) = \mathrm{d}x$.

For integers $i \leqslant j$, the notation $[i : j]$ stands for $\{i, i+1, \ldots, j\}$. Define $\boxed{w(x) = \frac{\pi(x)}{\tilde{\pi}(x)}}$.

The i-SIR algorithm with d proposals (also called i-SIR(d)) consists in constructing a Markov chain $\{X_n : n \in \mathbb{N}\}$ starting with initial distribution μ in the following way:

- Draw $X_0 \sim \mu$ where μ is arbitrary

for $k \leftarrow 1$ **to** n **do**

- Set $Y_0 = X_{k-1}$ and draw independently d random variables $Y_i \sim \tilde{\pi}$ for $i = 1, \ldots, d$.
- Draw a random variable J taking values on $[0 : d]$ with probabilities: $\mathbb{P}(J = k) = \frac{w(Y_k)}{\sum_{\ell=0}^{d} w(Y_\ell)}$ for $k \in [0 : d]$.
- Set $X_k = Y_J$.

end

Algorithm 1: the i-SIR(d) algorithm

1. For any bounded measurable function f on $\mathbb{R}$, show that

$$\mathbb{E}_\mu[f(X_k)\mathbf{1}_{\{J=0\}} | X_{k-1}] = f(X_{k-1})\beta(X_{k-1})$$

where $\beta(x) = \int \cdots \int \frac{w(x)}{w(x) + \sum_{i=1}^{d} w(y_i)} \tilde{\pi}(\mathrm{d}y_1) \ldots \tilde{\pi}(\mathrm{d}y_d)$

2. Show that for any $\ell \in [1 : d]$,

$$\mathbb{E}_\mu[f(X_k)\mathbf{1}_{\{J=\ell\}} | X_{k-1}] = \int f(y_1) \left(\int \cdots \int \frac{1}{w(X_{k-1}) + \sum_{i=1}^{d} w(y_i)} \tilde{\pi}(\mathrm{d}y_2) \ldots \tilde{\pi}(\mathrm{d}y_d) \right) \pi(\mathrm{d}y_1)$$

3. Deduce that the Markov kernel P of the Markov chain $\{X_k : k \in \mathbb{N}\}$ writes

$$P(x, \mathrm{d}y) = \beta(x)\delta_x(\mathrm{d}y) + \gamma(x, y)\pi(\mathrm{d}y)$$

where $\gamma(x, y)$ should be expressed explicitely. Check that for any $x, y \in \mathbb{R}$, we have $\gamma(x, y) = \gamma(y, x)$ and $\gamma(x, y) > 0$.

4. Show that P is π-invariant.

5. Is π the unique invariant probability measure for P?

6. According to which theorem, we can obtain that for any measurable function f such that $\pi(|f|) < \infty$, we have $\mathbb{P}_\pi$-a.s.

$$\lim_{k \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \pi(f),$$

7. Let h be a bounded non-negative measurable function such that $\pi(h) = 0$ and $Ph(x) = h(x)$ for any $x \in \mathbb{R}$. Show that for any $x \in \mathbb{R}$, $h(x) = 0$.

8. Let f be a measurable function such that $\pi(|f|) < \infty$. Define

$$A = \left\{ \lim_{n \to \infty} \frac{\sum_{i=0}^{n-1} f(X_i)}{n} = \pi(f) \right\}$$

For any $x \in \mathbb{R}$, $h(x) = \mathbb{E}_x[\mathbf{1}_{A^c}] = \mathbb{P}_x(A^c)$. We admit that $Ph(x) = h(x)$ for any $x \in \mathbb{R}$. Show that $\pi(h) = 0$. Deduce from the previous questions, that the Law of Large Numbers actually holds for $\{X_n : n \in \mathbb{N}\}$ starting from any initial distribution, that is, for any probability measure ξ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we have $\mathbb{P}_\xi - a.s.$,

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n-1} f(X_i)}{n} = \pi(f)$$

9. We now assume that

   (A1) $\sup_{x\in\mathbb{R}} w(x) = \infty$

   Show that there exists a sequence of real numbers $(x_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \beta(x_n) = 1$.

10. Recall that a Markov kernel P is geometrically ergodic, if there exists a measurable non-negative function $V : \mathbb{R} \to \mathbb{R}^+$ and constants ρ such that $0 < \rho < 1$ satisfying: for any $n \in \mathbb{N}$ and any $x \in \mathbb{R}$,

$$\|P^n(x,\cdot) - \pi\|_{TV} \leqslant V(x)\rho^n. \tag{1}$$

    We will show by contradiction that under condition (A1) (of the previous question), the Markov kernel P cannot be geometrically ergodic. Indeed, using that $\pi(\{x\}) = 0$ for any singleton $\{x\}$, show that (1) implies that for any $x \in \mathbb{R}$ and any $n \in \mathbb{N}$, $2\beta^n(x) \leqslant V(x)\rho^n$.

11. Conclude.