Importance sampling and Sequential Monte Carlo

Exercise 1: Choice of proposal distribution

Let X be a random variable with probability density g with respect to the Lebesgue measure on $\mathbb{R}$. Let $\kappa_X : t \mapsto \log(\mathbb{E}[\mathrm{e}^{tX}])$. We want to estimate $\mathbb{P}(X \geq x)$ for $x \in \mathbb{R}$ using the proposal distribution $h_t : x \mapsto \mathrm{e}^{xt-\kappa_X(t)}g(x)$ for $t \in \mathbb{R}$.

1. Propose a naive Monte Carlo estimator of $\mathbb{P}(X \geq x)$ for $x \in \mathbb{R}$.
2. Show that
$$\mathbb{E}\left[\mathbb{1}_{Y\geq x}\mathrm{e}^{-2Yt+2\kappa_X(t)}\right] \leq \exp(-xt+\kappa_X(t))\,.$$
where Y has density h_t for $t \geq 0$.
3. Propose a choice t_x to select the proposal distribution h_t.
4. Apply this result when $X \sim \mathcal{N}(\mu, \sigma^2)$.
5. Apply this result when $X \sim \mathcal{P}(\lambda)$.

Exercise 2: Optimal kernel

We consider a linear ang Gaussian hidden Markov model given for $k \geq 0$ by

$$\begin{aligned} X_{k+1} &= \phi X_k + \sigma U_k\,,\\ Y_k &= X_k + \eta V_k\,, \end{aligned}$$

where $(U_k)_{k\geq 0}$ and $(V_k)_{k\geq 0}$ are independent standard Gaussian random variables independent of X_0. The distribution ν of X_0 is the stationnary distribution of the Markov Chain.

1. Write the joint probability density function of $(X_{0:n}, Y_{0:n})$.
2. Write the recursion defining the filtering distributions, i.e. the distributions of X_n given $Y_{0:n}$ for $n \geq 0$.
3. Propose a sequential Monte Carlo method to estimate the filtering distribution at time $n+1$ using weighted samples $\{(\xi_n^i, \omega_n^i)\}_{i=1}^N$ targetting the filtering distribution at time n. New particles are proposed using the prior kernel, i.e. the distribution of X_{n+1} given X_n.
4. The *optimal kernel* to propose new particles is defined as the distribution of X_{n+1} given (X_n, Y_{n+1}). Compute the optimal kernel and the weights $(\omega_{n+1}^i)_{i=1}^N$.
5. In other settings than linear and Gaussian HMM, the optimal kernel is usually not tractable. Propose an accept-reject mechanism to sample from the optimal kernel for general HMM.

Exercise 3: Smoothing distribution

Let $\{(X_k, Y_k)\}_{k\geq 0}$ be a HMM where $(X_k)_{k\geq 0}$ is a Markov chain with initial distribution ν and Markov transition density m. For all $k \geq 0$, the conditional distribution of Y_k given $X_{0:n}$ depends on X_k only and its probability density function is written $g(X_k, \cdot)$.

1. Prove that for all $0 \leq k \leq n-1$, the conditional distribution of X_k given X_{k+1} and $Y_{0:k}$ is proportional to $\phi_k(\cdot)m(\cdot, X_{k+1})$ where ϕ_k is the filtering distribution at time k. We write $b_k(X_{k+1}, \cdot)$ this distribution.
2. Prove that the joint density of $X_{0:n}$ given $Y_{0:n}$ can be written $x_{0:n} \mapsto \phi_n(x_n) \prod_{k=0}^{n-1} b_k(x_{k+1}, x_k)$.
3. Assume that at each time k, $\{(\xi_k^i, \omega_k^i)\}_{i=1}^N$ is a particle-based approximation of ϕ_k. Propose a particle-based approximation of $b_k(X_{k+1}, \cdot)$.
4. Deduce from the previous questions an algorithm to approximately sample from the joint distribution $X_{0:n}$ given $Y_{0:n}$.