Homework 10

Exercise 1. [Metropolized independent sampling in a particular case]

Let $0 < \theta < 1$ and let us consider the following distribution π on $S = \{1, \ldots, N\}$:

$$\pi_i = \frac{1}{Z}\,\theta^{i-1}, \quad i = 1, \ldots, N$$

where Z is the normalization constant, whose computation is left to the reader.

a) Consider the base chain $\psi_{ij} = \frac{1}{N}$ for all $i, j \in S$ and derive the transition probabilities p_{ij} obtained with the Metropolis-Hastings algorithm.

b) Using the result of the course, derive an upper bound on $\|P_i^n - \pi\|_{\text{TV}}$. Compare the bounds obtained for $i = 1$ and $i = N$ (for large values of N).

c) Deduce an upper bound on the (order of magnitude of the) mixing time

$$T_\varepsilon = \inf\{n \geq 1 \,:\, \max_{i \in S} \|P_i^n - \pi\|_{\text{TV}} \leq \varepsilon\}$$

Exercise 2. On the state space $S = \{0, 1, 2\}$ and given $\beta > 0$, consider the following distribution:

$$\pi = \frac{1}{Z}\,\left(1, e^{-2\beta}, e^{-\beta}\right)$$

where the normalization constant $Z = 1 + e^{-2\beta} + e^{-\beta}$ is easy to compute in this case. For any given $\beta > 0$, we would like to sample from π, in order to obtain (by taking β large) an estimate of the global minimum of the function $f : S \to \mathbb{Z}$ defined as $f(0) = 0$, $f(1) = 2$ and $f(2) = 1$. Of course, in this situation, both finding the global minimum of f and sampling from the distribution π are trivial tasks, but the idea here is to get an idea of the performance (i.e. rate of convergence) of the Metropolis-Hastings algorithm in a simple case.

Consider the base chain on S with transition probabilities

$$\psi_{01} = \psi_{21} = 1 \quad \text{and} \quad \psi_{10} = \psi_{12} = \frac{1}{2}.$$

a) Compute the transition probabilities p_{ij} of the corresponding Metropolis chain.

b) Check that the detailed balance equation is satisfied.

c) Compute the eigenvalues $\lambda_0 \geq \lambda_1 \geq \lambda_2$ of P. (*Hint:* You already know that $\lambda_0 = 1$.)

d) Express the spectral gap γ as a function of β. How does it behave as β gets large?