Solutions 10

1. a) The transition probabilities are given by the following formula:

$$p(x,y) = \begin{cases} \dfrac{\pi(x'_u \mid x_1, \cdots, x_{u-1}, x_{u+1}, \cdots, x_d)}{d}, & \text{if } y = (x_1, \cdots, x_{u-1}, x'_u, x_{u+1}, \cdots, x_d) \\ & \text{for } u \in \{1, \cdots, d\} \text{ and } x_u \neq x'_u, \\ \dfrac{\sum_{u=1}^{d} \pi(x_u | x_1, \cdots, x_{u-1}, x_{u+1}, \cdots, x_d)}{d}, & \text{if } y = x, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

b) Let us now show that the detailed balance equation is satisfied.

First note from (1) that $p(x,y) = 0$ if and only if x and y differ in more than a single component. Hence, $p(x,y) = 0$ if and only if $p(y,x) = 0$ and the detailed balance equation is satisfied since both sides are zero.

Then, suppose that x and y differ in component u, i.e., $x = (x_1, \cdots, x_{u-1}, x_u, x_{u+1}, \cdots, x_d)$ and $y = (x_1, \cdots, x_{u-1}, x'_u, x_{u+1}, \cdots, x_d)$. Therefore, from (1) we have

$$\begin{aligned} \pi(x)p(x,y) &= \frac{\pi(x)\pi(x'_u \mid x_1, \cdots, x_{u-1}, x_{u+1}, \cdots, x_d)}{d} = \frac{g(x)g(y)}{Z \cdot \sum_{y_u \in S} g(x_1, \ldots, x_{u-1}, y_u, x_{u+1}, \ldots, x_d)} \\ &= \pi(y)p(y,x). \end{aligned}$$

As a result, this algorithm can be viewed as a Metropolis-Hastings algorithm where the base chain has the transition probabilities $p(x,y)$ defined in (1) and the acceptance probability $a(x,y) = 1$ for any $x, y \in S^d$. In words, every move is always accepted. Note that every base chain which satisfies the detailed balance equation induces a Metropolis Hastings algorithm in which every move is always accepted. To see this, consider the base chain ψ_{ij} s.t. $\pi_i \psi_{ij} = \pi_j \psi_{ji}$ for any i, j in some state space $\mathcal{S}$. Then, as the base chain is not necessarily symmetric, the acceptance probability is given by

$$a_{ij} = \min\left(1, \frac{\pi_j \psi_{ji}}{\pi_i \psi_{ij}}\right) = 1.$$

2. a) The transition probabilities are given by

$$\begin{array}{llll} p_{01} &= \psi_{01}\min\left(1, \frac{\psi_{10}\pi_1}{\psi_{01}\pi_0}\right) = \frac{e^{-2\beta}}{2} & p_{21} &= \psi_{21}\min\left(1, \frac{\psi_{12}\pi_1}{\psi_{21}\pi_2}\right) = \frac{e^{-\beta}}{2} \\ p_{10} &= \psi_{10}\min\left(1, \frac{\psi_{01}\pi_0}{\psi_{10}\pi_1}\right) = \frac{1}{2} & p_{12} &= \psi_{12}\min\left(1, \frac{\psi_{21}\pi_2}{\psi_{12}\pi_1}\right) = \frac{1}{2} \\ p_{02} &= p_{20} = p_{11} = 0 & p_{00} &= 1 - \frac{e^{-2\beta}}{2} \qquad p_{22} = 1 - \frac{e^{-\beta}}{2} \end{array} \tag{2}$$

b) Let us now check that the detailed balance equation is satisfied:

$$\begin{aligned} p_{01}\pi_0 &= \frac{1}{2}e^{-2\beta} = p_{10}\pi_1 \\ p_{02}\pi_0 &= 0 = p_{20}\pi_2 \\ p_{12}\pi_1 &= \frac{1}{2}e^{-2\beta} = p_{21}\pi_2. \end{aligned}$$

c) As usual, there are several methods to compute the eigenvalues. For example, one can find the three solutions λ_0, λ_1, and λ_2 to the equation

$$\det(P - \lambda I) = 0, \tag{3}$$

where I is the 3×3 identity matrix and P the matrix of the transition probabilites computed in (2).

Another (perhaps even simpler method) method is to solve the following system of equations:

$$\begin{cases} \lambda_0 = 1 \\ \lambda_0 + \lambda_1 + \lambda_2 = \mathrm{tr(P)} \\ \lambda_0 \cdot \lambda_1 \cdot \lambda_2 = \det(P) \end{cases},$$

as we know that the largest eigenvalue is 1, the sum of the eigenvalues equals the trace of P, and their product equals the determinant of P.

Consequently, we obtain

$$\begin{cases} \lambda_0 = 1 \\ \\ \lambda_1 = -\dfrac{e^{-2\beta}}{4} - \dfrac{e^{-\beta}}{4} + \dfrac{1}{2} + \dfrac{1}{4}\sqrt{e^{-4\beta} - 2e^{-3\beta} + e^{-2\beta} + 4} \\ \\ \lambda_2 = -\dfrac{e^{-2\beta}}{4} - \dfrac{e^{-\beta}}{4} + \dfrac{1}{2} - \dfrac{1}{4}\sqrt{e^{-4\beta} - 2e^{-3\beta} + e^{-2\beta} + 4} \end{cases},$$

d) The spectral gap is given by

$$\gamma = 1 - \lambda_1 = \frac{1}{2} + \frac{e^{-2\beta}}{4} + \frac{e^{-\beta}}{4} - \frac{1}{4}\sqrt{e^{-4\beta} - 2e^{-3\beta} + e^{-2\beta} + 4}. \tag{4}$$

Therefore, when β is large, we have

$$\gamma \approx \frac{1}{4}e^{-\beta}. \tag{5}$$

Remark. The value of β has to be tuned carefully and there is an inherent trade-off in its choice. If we pick β too large, then the spectral gap is small and the convergence to the global minimum occurs very slowly. On the other hand, if we pick β too small, we might not be able to visit all the states and, therefore, we might get stuck in the local minimum (=state 2).