Solutions 9

1. Because of the assumptions made, $a_{ij} > 0$ if $\psi_{ij} > 0$, so the chain with transition probabilities p_{ij} is also irreducible and aperiodic, therefore ergodic, as the state space S is finite. Let us check the detailed balance equation:

$$\pi_i \, p_{ij} = \pi_i \, \psi_{ij} \, a_{ij} = \frac{\pi_i \, \psi_{ij} \, \pi_j \, \psi_{ji}}{\pi_j \, \psi_{ji} + \pi_i \, \psi_{ij}}$$

which is clearly symmetric in i and j, and therefore equal to $\pi_j \, p_{ji}$.

2. Note that the base chain is clearly irreducible, but is not aperiodic. This is not an issue however, as for any distribution π which is not uniform on S (and such a distribution does not exist on $S = \mathbb{N}^*$), the Metropolis chain will have a self-loop and therefore be aperiodic. Besides, the matrix ψ satisfies the condition $\psi_{ij} > 0$ if and only if $\psi_{ji} > 0$.

a) The computation gives:

$$a_{1,2} = \min\left(1, \frac{\pi_2 \, \psi_{2,1}}{\pi_1 \, \psi_{1,2}}\right) = \frac{\pi_2}{2\pi_1} \quad \text{and} \quad a_{2,1} = \min\left(1, \frac{\pi_1 \, \psi_{1,2}}{\pi_2 \, \psi_{2,1}}\right) = 1$$

and for $i \geq 2$:

$$a_{i+1,i} = \min(1, \pi_i/\pi_{i+1}) = 1 \quad \text{and} \quad a_{i,i+1} = \min(1, \pi_{i+1}/\pi_i) = \pi_{i+1}/\pi_i$$

Correspondingly, we obtain ($i \geq 2$ below):

$$p_{1,1} = 1 - \frac{\pi_2}{2\pi_1}, \quad p_{1,2} = \frac{\pi_2}{2\pi_1}, \quad p_{i,i-1} = \frac{1}{2}, \quad p_{i,i} = \frac{1}{2}\,(1 - \pi_{i+1}/\pi_i) \quad \text{and} \quad p_{i,i+1} = \frac{1}{2}\,\pi_{i+1}/\pi_i$$

b) 1. We obtain in this case (for $i \geq 2$)

$$a_{i,i-1} = 1 \quad \text{and} \quad a_{i,i+1} = \pi_{i+1}/\pi_i = i^2/(i+1)^2, \quad \text{so} \ \lim_{i\to\infty} a_{i,i+1} = 1$$

2. We obtain in this case (for $i \geq 2$)

$$a_{i,i-1} = 1 \quad \text{and} \quad a_{i,i+1} = \pi_{i+1}/\pi_i = e^{-1} \quad \text{so} \ \lim_{i\to\infty} a_{i,i+1} = e^{-1}$$

3. We obtain in this case (for $i \geq 2$)

$$a_{i,i-1} = 1 \quad \text{and} \quad a_{i,i+1} = \pi_{i+1}/\pi_i = e^{-(i+1)^2 - i^2} = e^{-2i-1} \quad \text{so} \ \lim_{i\to\infty} a_{i,i+1} = 0$$

Observe also that in each case, the normalization constant C disappears and need therefore not be computed in order to run the Metropolis algorithm (but of course, in these cases, the computation of C is not a huge problem. . .).

3. The base chain must be irreducible and aperiodic, and such that $\psi_{ij} > 0$ if and only if $\psi_{ji} > 0$. There are of course many possible choice for ψ. The symmetric random walk on $\{0, \dots, n\}$

$$\psi_{ij} = \begin{cases} 1/2, & \text{if } |i-j| = 1 \\ 1/2, & \text{if } i = j = 0 \text{ or } i = j = n \\ 0, & \text{otherwise} \end{cases}$$

is a simple choice. The acceptance probabilities are then given by

$$a_{ij} = \begin{cases} \min\left(1, \frac{\pi_j}{\pi_i}\right), & \text{if } |i-j| = 1 \\ 0, & \text{otherwise} \end{cases}$$

Computing these exactly using the probability mass function of the binomial distribution, we get

$$a_{ij} = \begin{cases} \min\left(1, \frac{\binom{n}{i+1} p^{i+1}(1-p)^{n-i-1}}{\binom{n}{i} p^i (1-p)^{n-i}}\right) = \min\left(1, \frac{(n-i)p}{(i+1)(1-p)}\right), & \text{if } j = i+1 \\ \min\left(1, \frac{\binom{n}{i-1} p^{i-1}(1-p)^{n-i+1}}{\binom{n}{i} p^i (1-p)^{n-i}}\right) = \min\left(1, \frac{i(1-p)}{(n-i+1)p}\right), & \text{if } j = i-1 \\ 0, & \text{otherwise} \end{cases}$$

Again, we observe here that the factorials which enter into the binomial coefficients $\binom{n}{i}$ need not to be computed in order to run the Metropolis algorithm. Further computations show that the Metropolis chain favors moves towards the maximum of the distribution, which is located around np, over moves that go in the other direction: e.g., if $i = nq$, with $q < p$, then

$$a_{i,i+1} = 1 \quad \text{and} \quad a_{i,i-1} \simeq \frac{nq(1-p)}{n(1-q)p} = \frac{q(1-p)}{(1-q)p} < 1, \quad \text{as } q < p$$

4. First note that $Z = \frac{1-\theta^N}{1-\theta} \simeq \frac{1}{1-\theta}$ for large N.

a) The weights defined in class are given in this case by $w_i = \frac{\pi_i}{\psi_i} = \frac{N}{Z}\,\theta^{i-1}$, so that for $j \neq i$:

$$a_{ij} = \min\left(1, \frac{w_j}{w_i}\right) = \min\left(1, \theta^{j-i}\right) = \begin{cases} 1 & \text{if } j < i \\ \theta^{j-i} & \text{if } j > i \end{cases}$$

which leads to

$$p_{ij} = \begin{cases} \frac{1}{N} & \text{if } j < i \\ \frac{1}{N}\,\theta^{j-i} & \text{if } j > i \\ \frac{1}{N} + \frac{1}{N}\sum_{k=i+1}^{N}\left(1 - \theta^{k-i}\right) & \text{if } j = i \end{cases}$$

b) From the course, we know that

$$\|P_i^n - \pi\|^{\text{TV}} \leq \frac{\lambda_*^n}{2\sqrt{\pi_i}}$$

where

$$\lambda_* = 1 - \frac{1}{w_*} \quad \text{and} \quad w_* = \max_{i \in S} w_i = w_1 = \frac{N}{Z}$$

We conclude therefore that

$$\|P_i^n - \pi\|_{\text{TV}} \leq \frac{\sqrt{Z}}{2\sqrt{\theta^{i-1}}}\left(1 - \frac{Z}{N}\right)^n$$

For $i = 1$ and large N, this bound leads to:

$$\|P_1^n - \pi\|_{\mathrm{TV}} \leq \frac{1}{2\sqrt{1-\theta}} \exp\left(-\frac{n}{N(1-\theta)}\right)$$

while for $i = N$ and large N, this bound leads to:

$$\|P_N^n - \pi\|_{\mathrm{TV}} \leq \frac{1}{2\sqrt{(1-\theta)\,\theta^{N-1}}} \exp\left(-\frac{n}{N(1-\theta)}\right) = \frac{1}{2\sqrt{1-\theta}} \exp\left(\frac{N-1}{2}\log(1/\theta) - \frac{n}{N(1-\theta)}\right)$$

c) Because of the last estimate, in order for $\max_{i\in S} \|P_i^n - \pi\|_{\mathrm{TV}}$ to be smaller than ε, we need that $n \gg N^2$, which gives the desired upper bound on the mixing time. What can actually be shown in this case (but this was not asked) is the following: using the more precise estimate

$$\|P_i^n - \pi\|_{\mathrm{TV}} \leq \frac{1}{2}\sqrt{\sum_{k=1}^{N-1} \lambda_k^{2n} \left(\phi_i^{(k)}\right)^2}$$

we find that this quantity is small (uniformly in i) for $n \gg N$ already.