Solutions 8

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. 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 \( w_i = \frac{\pi_i}{w_1} = \frac{N}{Z} \theta^{i-1} \), 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

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

b) From the course, we know that

\[
\|P^n_i - \pi\|_{TV} \leq \frac{\lambda_*^n}{2\sqrt{\pi_i}}
\]

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

\[
\|P^n_i - \pi\|_{TV} \leq \frac{\sqrt{Z}}{2\sqrt{\theta^{i-1}}} \left(1 - \frac{Z}{N}\right)^n
\]

For \( i = 1 \) and large \( N \), this bound leads to:

\[
\|P^n_1 - \pi\|_{TV} \leq \frac{1}{2\sqrt{1 - \theta}} \exp \left( -\frac{n}{N(1 - \theta)} \right)
\]

while for \( i = N \) and large \( N \), this bound leads to:

\[
\|P^n_N - \pi\|_{TV} \leq \frac{1}{2\sqrt{(1 - \theta) \theta^{N-1}}} \exp \left( -\frac{n}{N(1 - \theta)} \right) = \frac{1}{2\sqrt{1 - \theta}} \exp \left( \frac{N - 1}{2} \log(1/\theta) - \frac{n}{N(1 - \theta)} \right)
\]

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

\[
\|P^n_i - \pi\|_{TV} \leq \frac{1}{2} \sqrt{\sum_{k=1}^{N-1} \lambda_k^{2n} \left( \phi_i^{(k)} \right)^2}
\]
we find that this quantity is small (uniformly in \(i\)) for \(n \gg N\) already.

3. a) We have

\[
\mathbb{P}(X \neq Y) = 1 - \mathbb{P}(X = Y) = 1 - \sum_{i \in S} \xi_i = 1 - \sum_{i \in S} \min(\mu_i, \nu_i)
\]

\[
= \sum_{i \in S : \mu_i > \nu_i} (\mu_i - \nu_i) = \sum_{i \in S : \nu_i > \mu_i} (\nu_i - \mu_i)
\]

where the last two equalities hold because both \(\mu\) and \(\nu\) are distributions. Summing these last two equalities, we obtain

\[
2 \mathbb{P}(X \neq Y) = \sum_{i \in S} |\mu_i - \nu_i| = 2 \|\mu - \nu\|_{TV}
\]

b) Observe first that if \(\sum_{i \in S} \xi_i = 1\), then \(X = Y\) with probability one. When \(\sum_{i \in S} \xi_i < 1\), we obtain for \(i \in S:\)

\[
\mathbb{P}(X = i) = \mathbb{P}(X = i, Y = j) + \sum_{j \in S \setminus i} \mathbb{P}(X = i, Y = j) = \xi_i + \sum_{j \in S \setminus i} \frac{(\mu_i - \xi_i)(\nu_j - \xi_j)}{1 - \sum_{k \in S} \xi_k}
\]

\[
= \xi_i + \frac{\mu_i - \xi_i}{1 - \sum_{k \in S} \xi_k} \sum_{j \in S \setminus i} (\nu_j - \xi_j) = \xi_i + \frac{\mu_i - \xi_i}{1 - \sum_{k \in S} \xi_k} \left(1 - \nu_i - \sum_{j \in S} \xi_j + \xi_j\right)
\]

\[
= \xi_i + (\mu_i - \xi_i) - \frac{(\mu_i - \xi_i)(\nu_i - \xi_i)}{1 - \sum_{k \in S} \xi_k} = \mu_i
\]

as \((\mu_i - \xi_i)(\nu_i - \xi_i) = 0\) necessarily. A similar reasoning shows that \(\mathbb{P}(Y = j) = \nu_j\) for all \(j \in S\).

c) Fix \(i \in S\), let \(X_0 = i\) and \(Y_0 \sim \pi\) the stationary distribution, and fix also a time \(n\). By parts a) and b), we can find a coupling of \(X_n\) and \(Y_n\) such that \(d_i(n) = \|P^n_i - \pi\|_{TV} = \mathbb{P}(X_n \neq Y_n)\). We can now define a new coupling for \(X_{n+1}\) and \(Y_{n+1}\) in the following way:

- If \(X_n = Y_n\), then \(X_{n+1} = Y_{n+1}\);
- Else, let \(X\) and \(Y\) evolve independently according to \(P\).

Then

\[
d_i(n + 1) = \|P^{n+1}_i - \pi\|_{TV} \leq \mathbb{P}(X_{n+1} \neq Y_{n+1}) \leq \mathbb{P}(X_n \neq Y_n) = d_i(n)
\]

The first inequality holds by the coupling lemma, and the second inequality is by construction. Observe finally that \(d(n) = \max_{i \in S} d_i(n)\) is also non-increasing in \(n\) (being the maximum of non-increasing functions).