Bayesian Modelling – Problem Sheet 5
Part B
(Solutions)

Problem 1

1. Let P be a transition matrix on $\mathcal{Y}$, $S = \{\mu \in \mathbb{R}^m_{\geq 0}$ such that $\sum_{i=1}^m \mu_i = 1\}$ and f be the mapping $\mu \mapsto P^T\mu$. Note that S is closed and bounded (and therefore compact) while f is continuous on S. To use Brouwer's fixed point theorem it remains to show that $f(S) = S$. Clearly, for any $\mu \in S$ all the components of $f(\mu)$ are non-negative because $\min_{i,j\in\mathcal{Y}} p_{ij} \geq 0$ and $\min_{i\in\mathcal{Y}} \mu_i \geq 0$. Let $f_i(\mu)$ be the i-th component of $f(\mu)$ and $\mathbf{1}_m = (1, \ldots, 1)$. Then,

$$\sum_{i=1}^m f_i(\mu) = \mathbf{1}_m^T(P^T\mu) = (P\mathbf{1}_m)^T\mu = \mathbf{1}_m^T\mu = 1$$

so that $f(S) = S$. Therefore, by Brouwer's fixed point theorem, there exists a $\mu \in S$ such that

$$\mu = f(\mu) = P^T\mu \Leftrightarrow \mu^T = \mu^T P.$$

2. Let $\mu = (\mu_1, \mu_2, 1 - \mu_1 - \mu_2)$. Then,

$$P^T\mu = \mu \Leftrightarrow (P^T - \mathbf{1}_3)\mu = 0$$

with

$$P^T - \mathbf{1}_3 = \begin{pmatrix} 0 & 1/3 & 0 \\ 0 & -2/3 & 0 \\ 0 & 1/3 & 0 \end{pmatrix}.$$

Therefore, $(\mu_1, 0, 1-\mu_1)$ is an invariant distribution of P for any $\mu_1 \in [0,1]$ so that P has infinity many invariant distributions.

Using the result in part 1 and in Theorem 7.2, if P is irreducible and aperiodic then P must have a unique invariant distribution. Since this is not the case we deduce that P is not an irreducible and aperiodic transition matrix.

3. It is easily checked that the system of equations $P^T\mu = \mu$ has a unique solution at $\mu = (1/3, 1/3, 1/3)$. However, $\mu_1 p_{12} = 1/3$ while $\mu_2 P_{21} = 0$ so that the detailed balance condition is not satisfied.

4. a) We have $P^t = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ if t is even and $P^t = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ if t is odd. Hence P is irreducible but periodic.

   b) It is easily checked that the system of equations $P^T\mu = \mu$ has a unique solution at $\mu = (1/2, 1/2)$.

c) Let $i = j = 1$ for instance. Then, from part 4.a),

$$1 = \limsup_{t\to+\infty} p_{11}^{(t)} > \liminf_{t\to+\infty} p_{11}^{(t)} = 0$$

and therefore $\lim_{t\to+\infty} p_{11}^{(t)}$ does not exist. A similar argument shows that, for any $(i, j) \in \mathcal{Y}$, we have

$$\limsup_{t\to+\infty} p_{ij}^{(t)} \neq \liminf_{t\to+\infty} p_{ij}^{(t)}.$$

Problem 2

1. Let $(\tilde{y}, y) \in \mathcal{Y}^2$. Then,

$$\begin{aligned}\alpha(y,\tilde{y}) = \min\left\{1, \frac{\mu(\tilde{y})q_i(y|\tilde{y})}{\mu(y)q_i(\tilde{y}|y)}\right\} &= \min\left\{1, \mathbb{1}_{\{y^{(-i)}\}}\big(\tilde{y}^{(-i)}\big)\frac{\mu(\tilde{y})\mu^{(i)}(y^{(i)}|\,\tilde{y}^{(-i)})}{\mu(y)\mu^{(i)}(\tilde{y}^{(i)}|\,y^{(-i)})}\right\}\\ &= \min\left\{1, \mathbb{1}_{\{y^{(-i)}\}}\big(\tilde{y}^{(-i)}\big)\frac{\mu(\tilde{y})\mu^{(i)}(y^{(i)}|\,y^{(-i)})}{\mu(y)\mu^{(i)}(\tilde{y}^{(i)}|\,y^{(-i)})}\right\}\end{aligned}$$

where (with obvious notation)

$$\mu^{(i)}(y^{(i)}|\,y^{(-i)}) = \frac{\mu(y)}{\mu(y^{(-i)})}, \quad \mu^{(i)}(\tilde{y}^{(i)}|\,y^{(-i)}) = \frac{\mu(\tilde{y}^{(i)},\, y^{(-i)})}{\mu(y^{(-i)})}$$

so that

$$\frac{\mu^{(i)}(y^{(i)}|\,y^{(-i)})}{\mu^{(i)}(\tilde{y}^{(i)}|\,y^{(-i)})} = \frac{\mu(y)}{\mu(\tilde{y}^{(i)},\, y^{(-i)})}.$$

Therefore

$$\begin{aligned}\alpha(y,\tilde{y}) &= \min\left\{1, \mathbb{1}_{\{y^{(-i)}\}}\big(\tilde{y}^{(-i)}\big)\frac{\mu(\tilde{y})}{\mu(y)}\frac{\mu(y)}{\mu(\tilde{y}^{(i)},\, y^{(-i)})}\right\}\\ &= \min\left\{1, \mathbb{1}_{\{y^{(-i)}\}}\big(\tilde{y}^{(-i)}\big)\frac{\mu(\tilde{y})}{\mu(y)}\frac{\mu(y)}{\mu(\tilde{y})}\right\}\\ &= \min\left\{1, \mathbb{1}_{\{y^{(-i)}\}}\big(\tilde{y}^{(-i)}\big)\right\}\\ &= \mathbb{1}_{\{y^{(-i)}\}}\big(\tilde{y}^{(-i)}\big)\end{aligned}$$

and thus

$$\begin{aligned}\mathbb{P}\big(Y_t = \tilde{Y}_t\big) = \mathbb{E}\big[\mathbb{P}\big(Y_t = \tilde{Y}_t|\tilde{Y}_t,\, Y_{t-1}\big)\big] &= \mathbb{E}\big[\alpha(Y_{t-1}, \tilde{Y}_t)\big]\\ &= \mathbb{E}\Big[\mathbb{E}\big[\alpha(Y_{t-1}, \tilde{Y}_t)|\, Y_{t-1}\big]\Big]\\ &= \mathbb{E}\big[\mathbb{P}\big(Y_{t-1}^{(-i)} = \tilde{Y}_t^{(-i)}|Y_{t-1}\big)\big]\\ &= \mathbb{E}[1]\\ &= 1.\end{aligned}$$

2. Let $\tilde{y} \in \mathcal{Y}$. Then,

$$
\begin{aligned}
\int_{\mathcal{Y}} k(\tilde{y}|y)\mu(y)\mathrm{d}y &= \int_{\mathcal{Y}} \Big(\int_{\mathcal{Y}} k_2(\tilde{y}|y')k_1(y'|y)\mathrm{d}y' \Big) \mu(y)\mathrm{d}y \\
&= \int_{\mathcal{Y}} k_2(\tilde{y}|y') \Big(\int_{\mathcal{Y}} k_1(y'|y)\mu(y)\mathrm{d}y \Big) \mathrm{d}y' \\
&= \int_{\mathcal{Y}} k_2(\tilde{y}|y')\mu(y')\mathrm{d}y' \\
&= \mu(\tilde{y})
\end{aligned}
$$

where the first equality uses the definition of $k(\tilde{y}|y)$, the second one uses Fubini's theorem, the third one the fact that $k_1(\tilde{y}|y)$ has μ as invariant distribution and the last one the fact that $k_2(\tilde{y}|y)$ has μ as invariant distribution.

3. From part 1, for any $i \in \{1, \dots, d\}$ the transition kernel $q_i(\tilde{y}|y)$ has μ has invariant distribution. Indeed, when $\mathbb{P}(Y_t = \tilde{Y}_t) = 1$ the proposal distribution and the Metropolis-Hastings kernel P^{MH} coincide, and by construction this latter has μ as invariant distribution.

   Next, let $\bar{q}_1(\tilde{y}|y) = q_1(\tilde{y}|y)$ and

$$
\bar{q}_i(\tilde{y}|y) = \int_{\mathcal{Y}} q_i(\tilde{y}|y')\bar{q}_{i-1}(y'|y)\mathrm{d}y', \quad i = 2, \dots d.
$$

   Then, using the result in part 2, $\bar{q}_d(\tilde{y}|y)$ has μ as invariant distribution.

   To show that $\bar{q}_d(\tilde{y}|y)$ is the Gibbs kernel remark that we can generate a random draw from $\bar{q}_d(\tilde{y}|y)\mathrm{d}\tilde{y}$ using the following algorithm

   Set $Z_0 = y$

   For $i = 1, ..., d$

   $Z_i \sim q_i(z_i|Z_{i-1})\mathrm{d}z_i$

   End for

   return $\tilde{Y} = Z_d$

   or equivalently, since q_i updates only component $Z_i^{(i)}$ of Z_i (and using obvious notation)

   Set $\tilde{Y}_0 = y$

   For $i = 1, ..., d$

   $Z^{(i)} \sim \mu^{(i)}(z^{(i)}|\tilde{Y}_{i-1}^{(-i)})\mathrm{d}z^{(i)}$

   $\tilde{Y}_i = (\tilde{Y}_i^{(1:(i-1))}, z^{(i)}, \tilde{Y}_i^{(i+1:d)})$

   End for

   return $\tilde{Y} = \tilde{Y}_d$

   which is the Gibbs sampler (Algorithm (A3) in the lecture notes).