Solutions to Midterm Exam

Exercise 1. (11 points) Let $f : \mathbb{N} \rightarrow \mathbb{R}$ be a function such that $f(0) = 1$ and $0 < f(i) < 1$ for $i \geq 1$. Consider the Markov chain with state space $S = \mathbb{N}$ and transition probabilities

$$p_{ij} = \begin{cases} f(i), & j = i + 1, \ i \geq 0 \\ 1 - f(i), & j = i - 1, \ i \geq 1. \end{cases}$$

a) Is the chain irreducible? Justify your answer using the Chapman-Kolmogorov equation.

Solution: Let $i, j$ two natural integers satisfying $i < j$. According to Chapman-Kolmogorov equation we have

$$p_{ij}(j - i) \geq \prod_{k=i}^{j-1} p_{k,k+1} = \prod_{k=i}^{j-1} f(k) > 0 \quad (f(k) > 0 \ \forall k \in \mathbb{N}) ,$$

and

$$p_{ji}(j - i) \geq \prod_{k=i+1}^{j} p_{k,k-1} = \prod_{k=i+1}^{j} (1 - f(k)) > 0 \quad (1 - f(k) > 0 \ \forall k \in \mathbb{N} \ \setminus \{0\}) .$$

It follows that the chain is irreducible.

b) Consider two initial probability distributions $\pi_{j}^{(0)} = \mu_{j}$ and $\pi_{j}^{(0)} = \nu_{j}$ where $\mu_{2k} = 0$, $\nu_{2k} \neq 0$ and $\mu_{2k+1} \neq 0$, $\nu_{2k+1} = 0$, $k \in \mathbb{N}$. We set $\mu^{(n)} = \mu P^{n}$, $\nu^{(n)} = \nu P^{n}$. Prove that for any time $n \geq 1

$$\|\mu_{i}^{(n)} - \nu_{i}^{(n)}\|_{TV} = 1$$

Solution: Let $\pi$ be any probability distribution on the natural integers. For any positive integer $i$ we have

$$(\pi P)_{i} = f(i - 1)\pi_{i-1} + (1 - f(i + 1))\pi_{i+1}$$

and $(\pi P)_{0} = (1 - f(1))\pi_{1}$. If $\pi_{i}$ is zero for all $i \in \mathbb{N}$ even then $(\pi P)_{i}$ is zero for all $i \in \mathbb{N}$ odd, while if $\pi_{i}$ is zero for all $i \in \mathbb{N}$ odd then $(\pi P)_{i}$ is zero for all $i \in \mathbb{N}$ even.

We can now apply this observation recursively to $\mu^{(n)}$ and $\nu^{(n)}$. Starting from $\mu_{2k} = 0$ and $\nu_{2k+1} = 0$ for all $k \in \mathbb{N}$, it comes

- $n$ even: $\mu_{2k}^{(n)} = 0$ and $\nu_{2k+1}^{(n)} = 0 \ \forall k \in \mathbb{N}$;
- $n$ odd: $\mu_{2k+1}^{(n)} = 0$ and $\nu_{2k}^{(n)} = 0 \ \forall k \in \mathbb{N}$.
For \( n \) even, the total variation distance between \( \mu^{(n)} \) and \( \nu^{(n)} \) reads

\[
\|\mu^{(n)} - \nu^{(n)}\|_{TV} = \frac{1}{2} \sum_{k=0}^{\infty} |\mu_k^{(n)} - \nu_k^{(n)}|
\]

\[
= \frac{1}{2} \sum_{p=0}^{\infty} |\mu_{2p}^{(n)} - \nu_{2p}^{(n)}| + \frac{1}{2} \sum_{p=0}^{\infty} |\mu_{2p+1}^{(n)} - \nu_{2p+1}^{(n)}|
\]

\[
= \frac{1}{2} \sum_{p=0}^{\infty} \nu_{2p}^{(n)} + \frac{1}{2} \sum_{p=0}^{\infty} \mu_{2p}^{(n)} \quad (\mu_{2p}^{(n)} = \nu_{2p}^{(n)} = 0 \forall p \in \mathbb{N})
\]

\[
= \frac{1}{2} \sum_{k=0}^{\infty} \nu_k^{(n)} + \frac{1}{2} \sum_{k=0}^{\infty} \mu_k^{(n)} \quad (\mu_{2p}^{(n)} = \nu_{2p}^{(n)} = 0 & \mu_{2p+1}^{(n)}, \nu_{2p}^{(n)} \geq 0 \forall p \in \mathbb{N})
\]

\[
= 1
\]

A similar chain of thought leads to \( \|\mu^{(n)} - \nu^{(n)}\|_{TV} = 1 \) for \( n \) odd.

c) We recall the “ratio test” for the convergence of a series \( \sum_{j \in \mathbb{N}} a_j \) of positive terms \( a_j > 0 \). Let \( \lim_{j \to +\infty} \frac{a_{j+1}}{a_j} = L \). If \( L < 1 \) the series converges; if \( L > 1 \) the series diverges; and if \( L = 1 \) the test is inconclusive. Prove that the stationary distribution exists if

\[
\lim_{j \to +\infty} \frac{f(j)}{1 - f(j + 1)} < 1
\]

and does not exist if

\[
\lim_{j \to +\infty} \frac{f(j)}{1 - f(j + 1)} > 1
\]

Does the stationary distribution exist for the function \( f(0) = 1, f(i) = \frac{1}{2}, i \geq 1 \)?

**Solution:** Assume a stationary distribution \( \pi^* \) exists. Then, detailed balance must hold because the transition matrix is tridiagonal. We obtain for any \( k \in \mathbb{N} \)

\[
p_{k,k+1} \cdot \pi^* = p_{k+1,k} \cdot \pi^* \quad \text{i.e.} \quad \pi^* = \pi^* \frac{f(k)}{1 - f(k + 1)} \pi^*. 
\]

Hence, for all \( k \in \mathbb{N}, \pi^*_k = \pi^*_0 \prod_{\ell=0}^{k-1} \frac{f(\ell)}{1 - f(\ell + 1)} \). \( \pi^*_0 > 0 \), otherwise the \( \pi^*_k \)'s do not sum to one. Finally, the same argument

\[
1 = \sum_{k=0}^{+\infty} \pi^*_k = \pi_0^* \sum_{k=0}^{+\infty} \prod_{\ell=0}^{k-1} \frac{f(\ell)}{1 - f(\ell + 1)}
\]

shows that the series \( \sum_{j \in \mathbb{N}} a_j \) of positive terms \( a_j = \prod_{\ell=0}^{j-1} \frac{f(\ell)}{1 - f(\ell + 1)} \) converges if and only if a stationary distribution exists.

Note that \( \frac{a_{j+1}}{a_j} = \frac{f(j)}{1 - f(j + 1)} \). Assuming the limit \( \lim_{j \to +\infty} \frac{f(j)}{1 - f(j + 1)} \) exists, the ratio test says

1. \( \lim_{j \to +\infty} \frac{f(j)}{1 - f(j + 1)} < 1 \) then the stationary distribution exists, is unique, and

\[
\pi^*_0 = \left( \sum_{k=0}^{+\infty} \prod_{\ell=0}^{k-1} \frac{f(\ell)}{1 - f(\ell + 1)} \right)^{-1}, \quad \pi^*_k = \pi^*_0 \prod_{\ell=0}^{k-1} \frac{f(\ell)}{1 - f(\ell + 1)} \forall k \geq 1.
\]
2. If \( \lim_{j \to +\infty} \frac{f(j)}{1-f(j+1)} > 1 \) then the chain does not admit a stationary distribution.

For the function \( f(0) = 1, f(i) = \frac{1}{2}, i \geq 1 \), we are not in one of the two cases above, as
\[
\lim_{j \to +\infty} \frac{f(j)}{1-f(j+1)} = 1.
\]
Still, the stationary distribution does not exist because the above-mentioned series \( \sum_{j \in \mathbb{N}} a_j \) diverges \( (a_j = 1 \text{ for } j \geq 1) \).

**Remark:** The relation \( \pi^*_k = \frac{f(k)}{1-f(k+1)} \pi^*_k \) for any \( k \in \mathbb{N} \) can also be obtained from the equation \( \pi^* = \pi^* P \) by recurrence. First, \( (\pi^* P)_0 = \pi^*_0 \) gives \( \pi^*_1 = \frac{1}{1-f(1)} \pi^*_0 = \frac{f(0)}{1-f(1)} \pi^*_0 \). Then, for \( k \geq 1 \), we have

\[
(\pi^* P)_k = \pi^*_k \iff f(k-1)\pi^*_{k-1} + (1-f(k+1))\pi^*_{k+1} = \pi^*_k \iff \pi^*_{k+1} = \frac{\pi^*_k - f(k-1)\pi^*_{k-1}}{1-f(k+1)}.
\]

If \( \pi^*_k = \frac{f(k-1)}{1-f(k)} \pi^*_{k-1} \) then \( \pi^*_{k+1} = \frac{\pi^*_k - f(k-1)\pi^*_{k-1}}{1-f(k+1)} = \frac{f(k)}{1-f(k+1)} \pi^*_k \) follows by recurrence.

**d)** In each case (1) and (2) in the above question: Is the chain positive recurrent? Is the chain ergodic?

**Solution:** In the case (1), the chain is irreducible and the stationary distribution exists and is unique. Therefore, it is positive-recurrent. The chain is periodic of period 2: it is not ergodic.

In the case (2), the chain is irreducible and the stationary distribution does not exist. Hence the chain is neither positive-recurrent nor ergodic.

**e)** Let now \( 0 < p, q < 1 \) and consider the case where \( f(0) = 1, f(2k) = p \) and \( f(2k-1) = q \) for every \( k \geq 1 \). For what values of \( p \) and \( q \) does the chain admit a stationary distribution?

**Solution:** Note that the sequence \( \{\frac{f(j)}{1-f(j+1)}\}_{j \in \mathbb{N}} \) alternates between two values (one for \( j \) even, one for \( j \) odd), so that it does not converge and the result of question c) cannot be applied.

However, we have seen in question c) that a stationary distribution exists if, and only if, the series \( \sum_{j \in \mathbb{N}} a_j \) of positive terms \( a_j = \prod_{\ell=0}^{j-1} \frac{f(\ell)}{1-f(\ell+1)} \) converges. For \( j = 2k \) even we have

\[
a_j = \prod_{i=0}^{2k-1} \frac{f(i)}{1-f(i+1)} = \prod_{\ell=0}^{k-1} \frac{f(2\ell)}{1-f(2\ell+1)} \frac{f(2\ell+1)}{1-f(2\ell+2)} = \prod_{\ell=0}^{k-1} \frac{p}{1-q} \frac{q}{1-p} = \left(\frac{pq}{(1-q)(1-p)}\right)^k
\]

For \( j = 2k + 1 \) odd, it follows directly that \( a_j = \left(\frac{pq}{(1-q)(1-p)}\right)^k \frac{p}{1-q} \). The series \( \sum_{j \in \mathbb{N}} a_j \) converges if, and only if, the common ratio \( \frac{pq}{(1-q)(1-p)} \) is strictly less than one. For \( 0 < p, q < 1 \), this condition
simplifies to

\[
\frac{pq}{(1-q)(1-p)} < 1 \iff \frac{p}{1-p} < \frac{1-q}{1-(1-q)}
\]
\[
\iff p < 1 - q
\]
\[
\iff p + q < 1 ,
\]
where in the second line we used that the mapping \( x \mapsto \frac{x}{1-x} \) is strictly increasing on \([0,1]\).

**Exercise 2.** (9 points) Let \( N \) be an odd number greater than or equal to 3 and let \( S = \{0, \ldots, N-1\} \). Let then \((X_n, n \geq 0)\) be a Markov chain with state space \( S \) and \( N \times N \) transition matrix

\[
P = \begin{pmatrix}
0 & a & 0 & \ldots & 0 & b \\
b & 0 & a & 0 & \ldots & 0 \\
0 & b & 0 & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & b & 0 & a \\
a & 0 & \ldots & 0 & b & 0
\end{pmatrix}
\]

with \( a, b > 0 \) and \( a + b = 1 \).

**a)** Explain why the chain is ergodic for all such values of \( a \) and \( b \).

**Solution:** \( \forall (i,j) \in S^2 \) such that \( i < j \) we have

\[
p_{ij}(j-i) \geq \prod_{k=i}^{j-1} p_{k,k+1} = a^{j-i} > 0 ,
\]

\[
p_{ji}(j-i) \geq \prod_{k=i+1}^{j} p_{k,k-1} = b^{j-i} > 0 .
\]

The chain is thus irreducible. As it is irreducible and finite, it is also positive-recurrent. Finally, the chain is aperiodic because \( p_{00}(2) \geq p_{01} \cdot p_{10} = ab > 0, p_{00}(N) \geq \prod_{k=0}^{N-2} p_{k,k+1} \cdot p_{N-1,0} = a^N > 0 \) and \( N \) is odd. The chain is ergodic.

**b)** Compute its unique stationary and limiting distribution \( \pi \).

**Solution:** The transition matrix is clearly doubly-stochastic. It follows that the (unique) stationary distribution of the chain is the uniform distribution: \( \pi_i = \frac{1}{N} \ \forall i \in S \). \( \pi \) is also a limiting distribution because the chain is ergodic.

**c)** For what values of \( a \) and \( b \) does detailed balance hold?

**Solution:** \( \pi \) is the uniform distribution. For the detailed balance to hold we need

\[
\forall (i,j) \in S^2 : p_{ij}\pi_i = p_{ji}\pi_j \iff \forall (i,j) \in S^2 : p_{ij} = p_{ji} \iff a = b .
\]

The only solution for \( a, b > 0 \) such that \( a + b = 1 \) is \( a = b = \frac{1}{2} \).
d) Among the following three matrices, which are transition matrices?

\[ P^T, \quad P^T P, \quad PP^T \]

(where \( P^T \) stands for the transpose of \( P \)).

**Solution:** \( P \) is doubly-stochastic, hence \( P^T \) is a valid transition matrix (each row sums to one, the entries are positive) and doubly-stochastic (each column sums to one).

The matrix \( P^T P \) has clearly non-negative entries. The sum of its entries along the \( i^{th} \) row, \( 1 \leq i \leq N \), is

\[
\sum_{j=1}^{N} (P^T P)_{i,j} = \sum_{j=1}^{N} \sum_{k=1}^{N} P_{i,k} P_{k,j} = \sum_{k=1}^{N} P_{k,i} \sum_{j=1}^{N} P_{k,j} = \sum_{k=1}^{N} P_{k,i} = 1.
\]

\( P^T P \) is thus a transition matrix. Finally, to prove \( P^T P \) is a transition matrix, we only relied on the fact that \( P \) is a doubly-stochastic transition matrix. As \( P^T \) is a doubly-stochastic transition matrix, a similar argument shows that \( P P^T \) is a transition matrix too. And more generally, it holds that if \( P, Q \) are both transition matrices, then the product \( PQ \) is also a transition matrix.

Let now \( Q = P^T P \). Even though the chain is not necessarily reversible for all values of \( a \) and \( b \), it can be shown here that for all \( i \in \mathcal{S} \) and \( n \geq 1 \),

\[
\|P_i^n - \pi\|_{TV} \leq \frac{1}{2\sqrt{\pi_i}} \left(\lambda_* (Q)\right)^{n/2}
\]

where \( \lambda_0 (Q) \geq \lambda_1 (Q) \geq \ldots \geq \lambda_{N-1} (Q) \) are the (real) eigenvalues of the \( N \times N \) matrix \( Q \), and

\[
\lambda_* (Q) = \max_{k \in \{1, \ldots, N-1\}} |\lambda_k (Q)|
\]

**e)** In the case \( \lambda_3 = 3 \), compute the eigenvalues of \( Q \) and deduce the value of \( \lambda_* (Q) \).

**Solution:** For \( N = 3 \)

\[
Q = P^T P = \begin{pmatrix} 0 & b & a \\ a & 0 & b \\ b & a & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & a & b \\ a & 0 & b \\ a & b & 0 \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ab & ab \\ ab & a^2 + b^2 & ab \\ ab & ab & a^2 + b^2 \end{pmatrix}.
\]

\( Q \) is a transition matrix: \( u_0 = (1 \quad 1 \quad 1)^T \) is an eigenvector associated to the eigenvalue \( \lambda_0 (Q) = 1 \).

One observes also that both \( u_1 = (1 \quad -1 \quad 0)^T \) and \( u_2 = (1 \quad 0 \quad -1)^T \) are eigenvectors associated to the eigenvalue

\[
a^2 + b^2 - ab = (a + b)^2 - 3ab = 1 - 3ab < 1.
\]

The triplet \((u_0, u_1, u_2)\) forms a basis of \( \mathbb{R}^3 \): \( \lambda_0 (Q) = 1 \) and \( \lambda_1 (Q) = \lambda_2 (Q) = 1 - 3ab. \) Finally, \( \lambda_* (Q) = |1 - 3ab| = 1 - 3ab \) (remember that \( 1 - 3ab = a^2 + b^2 - ab = (a - b)^2 + ab > 0 \)).

\( \lambda_1 (Q) \) and \( \lambda_2 (Q) \) could also be deduced by finding the roots of the degree-two polynomial

\[
X^2 - (\lambda_1 (Q) + \lambda_2 (Q))X + \lambda_1 (Q)\lambda_2 (Q)
\]

with \( \lambda_1 (Q)\lambda_2 (Q) = \frac{\det (Q)}{\lambda_0 (Q)} = \det (Q) \) and \( \lambda_1 (Q) + \lambda_2 (Q) = \text{Tr} (Q) - \lambda_0 (Q) = \text{Tr} (Q) - 1. \)
f) Again in the case $N = 3$, deduce from (3) an upper bound on the mixing time

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

where $\varepsilon > 0$.

**Solution:** For $n \geq 1$ and $i \in \{0, 1, 2\}$

$$\|P_i^n - \pi\|_{TV} \leq \frac{1}{2\sqrt{\pi_i}} (\lambda_*(Q))^{n/2} = \frac{\sqrt{3}}{2} (1 - 3ab)^{n/2} \leq \frac{\sqrt{3}}{2} \exp\left(-\frac{3ab}{2}n\right)$$

Hence $\max_{i \in S} \|P_i^n - \pi\|_{TV} \leq \varepsilon$ for $n \geq \frac{2}{3ab} \ln \left(\frac{\sqrt{3}}{2\varepsilon}\right)$, so $T_\varepsilon \leq \frac{2}{3ab} \ln \left(\frac{\sqrt{3}}{2\varepsilon}\right)$. 