IMPERIAL

MATH70031

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2024

This paper is also taken for the relevant examination for the
Associateship of the Royal College of Science

Markov Processes

Date: Monday, May 13, 2024

Time: 14:00 – 16:30 (BST)

Time Allowed: 2.5 hours

This paper has 5 Questions.

Please Answer All Questions in 1 Answer Booklet

Candidates should start their solutions to each question on a new sheet of paper.

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Allow margins for marking.

Credit will be given for all questions attempted.

Each question carries equal weight.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO

1. (a) Give an example of a state space $\mathcal{X}$, an $\mathcal{X}$-valued Markov process $X = (X_n)_{n=0}^{\infty}$, and a bounded measurable function $f : \mathcal{B}_b(\mathcal{X})$ such that $Z = (Z_n)_{n=0}^{\infty}$ given by $Z_n = f(X_n)$ is not a Markov process (give a proof for the last assertion). (7 marks)

   (b) Given a transition probability P on a state-space $\mathcal{X}$, precisely define what it means for a probability measure $\mu \in \mathcal{P}(\mathcal{X})$ to be P-invariant. (4 marks)

   (c) Let $A = (A_n)_{n=0}^{\infty}$ be an $\mathbb{R}$-valued stochastic process which satisfies, for each bounded measurable $g \in \mathcal{B}_b(\mathbb{R})$, and any $i < j < k$, $\mathbb{E}\Big[g(A_k)|A_i, A_j\Big] = \mathbb{E}\Big[g(A_k)|A_j\Big]$.

   Prove that for any bounded measurable $f, h \in \mathcal{B}_b(\mathbb{R})$ and $i < j < k$, one has

$$\mathbb{E}\Big[f(A_i)h(A_k)|A_j\Big] = \mathbb{E}[f(A_i)|A_j]\mathbb{E}[h(A_k)|A_j] \ .$$

(9 marks)

(Total: 20 marks)

2. Let $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration generated by a $\mathcal{X}$-valued time homogenous Markov process $X = (X_n)_{n=0}^{\infty}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

   (a) Define what it means for a $\mathbb{N} \sqcup \{\infty\}$-valued random variable T to be an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time. (3 marks)

   (b) Suppose that $\{T_j\}_{j=0}^{\infty}$ are a family of $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping times, show that $\sup_j T_j$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time. (4 marks)

   (c) Given a $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time T, define the corresponding stopped sigma-algebra $\mathcal{F}_T$. (3 marks)

   (d) Prove, using the Markov property for X (**not** the strong Markov property) that for any $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time T with $\{T = \infty\} = \emptyset$, any $f \in \mathcal{B}_b(\mathcal{X})$, and any $n \in \mathbb{N}$, one has

$$\mathbb{E}\Big(f(X_{n+T})|\mathcal{F}_T\Big) = \int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y) \ ,$$

   where P^n is the n-step transition function for X. (10 marks)

(Total: 20 marks)

3. In all parts of this question, we specialize to looking at time homogenous Markov processes on a countable state space $\mathcal{X}$ and stochastic matrices P on $\mathcal{X}$.

(a) **True or False**: For each question, state whether the statement is true or false. If it is true, give a proof and if it is false give a counterexample with justification.

(i) Suppose P is irreducible, then P^2 is also irreducible. (4 marks)

(ii) Suppose P is reducible and $\mathcal{X}$ is finite, then P must have more than one invariant probability measure. (4 marks)

(b) Suppose we have distinct $i, j \in \mathcal{X}$. with j accessible from i, and with i not accessible from j. Prove that i is transient. (7 marks)

(c) Find the communication classes of the given stochastic matrix P, along with their partial ordering.

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \end{pmatrix}$$

(5 marks)

(Total: 20 marks)

4. (a) Let P be a stochastic matrix for the state space $\mathbb{Z}$. State what it means for a probability measure π on $\mathbb{Z}$ to satisfy *detailed balance* with respect to P. (4 marks)

(b) Let $Y = (Y_n)_{n=0}^{\infty}$ be a sequence of i.i.d $\mathcal{N}(0, 1)$ (i.e. standard Gaussian) random variables, and define an $\mathbb{R}$-valued Markov process $X = (X_n)_{n=0}^{\infty}$ by setting $X_0 = 0$ and for $n \in \mathbb{N}$,

$$X_{n+1} = Y_n + \sin(X_n/2 + Y_n) \ .$$

(i) Prove that there exists a invariant probability measure for X. (8 marks)

(ii) Prove that an invariant probability measure for X must be unique. (8 marks)

(Total: 20 marks)

Mastery Question

5. In the problems below, P is a transition probability on $\mathcal{X}$. Let θ be the shift map

$$a = (a_0, a_1, a_2, \dots) \mapsto \theta(a) = (a_1, a_2, \dots) \ .$$

We write $\mathcal{I} \subset \mathcal{B}\Big(\mathcal{X}^{\mathbb{N}}\Big)$ for the sigma algebra of invariant sets. Note that we work with one-sided sequences here.

Let $\pi \in \mathcal{P}(\mathcal{X})$ be a P-invariant probability measure, and X be the $\mathcal{X}$-valued Markov process transition P with initial distribution π. We write $\mathbb{P}_\pi$ for the law of X. Suppose that for every bounded measurable $f \in \mathcal{B}_b(\mathcal{X})$, one has, $\mathbb{P}_\pi$-almost surely,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \int_{\mathcal{X}} f(x)\, \pi(\mathrm{d}x) \ .$$

(a) Let $A \in \mathcal{I}$, and define

$$B = \{x \in \mathcal{X} : \mathbb{P}_x(A) = 1\} \ .$$

Prove that, $T\mathbf{1}_B = \mathbf{1}_B$ where T is the transition operator associated to P. **Hint:** Write $\mathbf{1}_B$ in terms of an expectation of $\mathbf{1}_A$. (10 marks)

(b) Prove that π as above is ergodic. (10 marks)

(Total: 20 marks)

1. (a) We set $\mathcal{X} = \{1,2,3,4\}$, and define a time-homogenous Markov process $X = (X_n)_{n=0}^{\infty}$ with $\mathbb{P}(X_0 = 1) = \mathbb{P}(X_0 = 3) = \frac{1}{2}$ and using the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Note that we can obtain this Markov process through the following dynamical system:

$$X_{n+1} = \begin{cases} 2 & \text{if } X_n = 1 \\ 1 & \text{if } X_n = 2 \\ 4 & \text{if } X_n = 3 \\ 3 & \text{if } X_n = 4 \; . \end{cases}$$

We define $f : \{1,2,3,4\} \to \{1,2,3\}$ by setting $f(1) = 1$, $f(3) = 3$ and $f(2) = f(4) = 2$. (3 marks for example that works)

We then have

$$\mathbb{P}(Z_2 = 1 | Z_1 = 2) = \mathbb{P}(Z_2 = 1) = \mathbb{P}(X_2 = 1) = \mathbb{P}(X_0 = 1) = \frac{1}{2} \; .$$

In the computation above, we used that $\{Z_1 = 1\} = \{X_1 = 2 \text{ or } 4\}$ has probability 1. On the other hand, we have

$$\mathbb{P}(Z_2 = 1 | Z_1 = 2, Z_0 = 1) = \mathbb{P}(Z_2 = 1 | Z_0 = 1) = 1 \; .$$

Above we used that $\{Z_0 = 1\} \subset \{Z_1 = 2\}$ in the first equality and in the second we observe that $\{Z_0 = 1\} = \{X_0 = 1\} = \{X_2 = 1\} = \{Z_2 = 1\}$. (4 marks for finishing justification).

(b) We have that μ is P-invariant if, for every Borel subset $A \in \mathcal{B}(\mathcal{X})$,

$$\mu(A) = \int_{\mathcal{X}} P(x, A) \; \mu(\mathrm{d}x) \; .$$

(4 marks)

(c) We have that

$$\begin{aligned} \mathbb{E}\Big[f(A_i)h(A_k)|A_j\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[f(A_i)h(A_k)|A_i, A_j\Big]\Big|A_j\Big] \\ &= \mathbb{E}\Big[f(A_i)\mathbb{E}\Big[h(A_k)|A_i, A_j\Big]\Big|A_j\Big] \\ &= \mathbb{E}\Big[f(A_i)\mathbb{E}\Big[h(A_k)|A_j\Big]\Big|A_j\Big] \\ &= \mathbb{E}\Big[h(A_k)|A_j\Big]\mathbb{E}\Big[f(A_i)|A_j\Big] \; . \end{aligned}$$

In the first equality we used the tower property (3 marks), in the second we factored out $f(A_i)$ since it is $\sigma(A_j, A_k)$ measurable (2 marks), then in the third equality we used our assumption that $\mathbb{E}\Big[h(A_k)|A_i, A_j\Big] = \mathbb{E}\Big[h(A_k)|A_j\Big]$ (2 marks), and then finally we factored out $\mathbb{E}\Big[h(A_k)|A_j\Big]$ since it is $\sigma(A_j)$-measurable. (2 marks)

(Total: 4 marks)

2. (a) A $\mathbb{N} \sqcup \{\infty\}$-valued random variable T on $(\Omega, \mathcal{F}, \mathbb{P})$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time if for each $n \in \mathbb{N}$, $\{T \le n\} \in \mathcal{F}_n$. (3 marks)

   (b) We have that $\{\sup_j T_j \le n\} = \bigcap_{j=0}^{n} \{T_j \le n\} \in \mathcal{F}_n$, so $\sup_j T_j$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time. (4 marks)

   (c) The stopped sigma-algebra $\mathcal{F}_T$ is defined by setting $\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \le n\} \in \mathcal{F}_n \text{ for all } n \in \mathbb{N}\}$ where $\mathcal{F}_\infty$ is the σ-algebra generated by $\bigcup_{n=0}^{\infty} \mathcal{F}_n$. (3 marks)

   (d) It suffices to prove that, for any $A \in \mathcal{F}_T$,

$$\mathbb{E}\Big(f(X_{n+T}) 1_A\Big) = \mathbb{E}\Big[\int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y) 1_A\Big].$$

   (2 marks)

   We have that $\Omega = \bigcup_{n=0}^{\infty} \{T = n\}$ (2 marks for decomposing based on value of T), so we can write

$$\begin{aligned}
\mathbb{E}\Big(f(X_{n+T}) 1_A\Big) &= \sum_{j=0}^{\infty} \mathbb{E}\Big(f(X_{n+j}) 1_{A \cap \{T=j\}}\Big) \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[\mathbb{E}\Big(f(X_{n+j}) 1_{A \cap \{T=j\}}\Big) | \mathcal{F}_j\Big)\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[1_{A \cap \{T=j\}} \mathbb{E}\Big(f(X_{n+j}) 1_A\Big) | \mathcal{F}_j\Big)\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[1_{A \cap \{T=j\}} \int_{\mathcal{X}} f(y) P^n(X_j, \mathrm{d}y)\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[1_{A \cap \{T=j\}} \int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y)\Big] \\
&= \mathbb{E}\Big[\int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y) 1_A\Big]. \qquad \text{(6 marks)}
\end{aligned}$$

(Total: 0 marks)

3. (a) (i) This is false (1 mark). For example, consider the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

   (2 marks for counter-example) Then $P^2 = I$, so P^2 is reducible since $\{1\}$ and $\{2\}$ are communication classes. (1 mark for justification)

   (ii) This is false (1 mark). For example, consider the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then P is reducible and has a unique invariant probability measure (2 marks for counter-example) - note that for any measure μ with $\mu(1) = \alpha$, $\mu(2) = \beta$, we have $(\mu P)(1) = 0$ and $(\mu P)(2) = \alpha + \beta$, one must have $\alpha = 0$, $\beta = 1$. (1 mark for justification)

(b) Suppose that j is accessible from i, and that i is not accessible from j. Then $\{n \in \mathbb{N}_{>0} : P^n_{i,j} > 0\}$ is non-empty, let m be the minimum element of this set. It follows that there exits $i_1, \dots, i_m = j$ such that

$$\mathbb{P}_i(X_1 = i_1, \dots, X_m = i_m = j) > 0 \ .$$

Since we chose m to be the minimum of the set above, it also follows that $i_1, \dots, i_m \neq i$ (3 marks for using accessibility with a minimal path) Since i is not accesible from j, it also follows that

$$\mathbb{P}_i\Big(\{X_m = j\} \setminus \{X_n \neq i \text{ for all } n \geq m\}\Big) = 0 \ .$$

(2 marks for using inaccessibility correctly) It follows that

$$\mathbb{P}_i(T_i = \infty) \geq \mathbb{P}_i(X_k \neq i \text{ for } 1 \leq k < m \text{ and } X_m = j) \geq \mathbb{P}_i(X_1 = i_1, \dots, X_m = i_m = j) > 0 \ .$$

(2 marks for rest of proof)

(c) The communication classes of P are $\{1, 2, 6, 7\}$ and $\{3\}$ and $\{4, 5\}$ (3 marks), and the partial ordering is $\{3\} \prec \{1, 2, 6, 7\}$ and $\{3\} \prec \{4, 5\}$. (2 marks)

(Total: 0 marks)

4. (a) For every $i, j \in \mathcal{X}$, $\pi(i)P_{i,j} = \pi(j)P_{j,i}$. (4 marks)

(b) (i) We appeal to the Krylov-Bogoliubov argument. (2 marks for strategy) Since, for any fixed y, the function $x \mapsto \sin(x/2 + y)$ is continuous, it follows the associated transition operator T is Feller. (2 marks for Feller) For tightness, we note that for any fixed $x \in \mathbb{R}$ and any $n \in \mathbb{N}_{>0}$

$$\begin{aligned}\mathbb{E}_x[X_n^2] &= \mathbb{E}_x[\Big(Y_n + \sin(X_{n-1}/2 + Y_n)\Big)^2] \\ &\leq 2\mathbb{E}_x[Y_n^2] + 2\mathbb{E}_x[\sin(X_{n-1}/2 + Y_n)^2] \leq 4\end{aligned}$$

Above we used that $(a + b)^2 \leq 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$. Since Y_n is a standard Gaussian, it follows that $\mathbb{E}_x[Y_n^2] = 1$. It then follows that $\mathbb{E}_x[X_n^2] \leq 4$, so by Chebyshev inequality, or any $K > 0$,

$$P^n(x, \mathbb{R} \setminus [-K, K]) = \mathbb{E}_x[\mathbf{1}\{X_n^2 > K^2\}] \leq \frac{1}{K^2}\mathbb{E}_x[X_n^2] \leq \frac{4}{K^2} \ .$$

Since we the right hand side goes to 0 as $K \uparrow \infty$, uniformly in n, it follows that the measures $\Big(P^n(x, \bullet)\Big)_{n=1}^{\infty}$ are tight. (4 marks for tightness)

(ii) We argue via deterministic contraction. (3 marks for strategy) Note that for any fixed value of y,

$$\Big(y + \sin(x/2 + y)\Big) - \Big(y + \sin(x'/2 + y)\Big) = \sin(x/2 + y) - \sin(x'/2 + y) \leq \frac{1}{2}|x - x'| \ ,$$

where the last inequality follows from the mean value theorem and the fact that $\left|\frac{d}{dx}\sin(x/2 + y)\right| \leq \frac{1}{2}$. Since the bound above holds for any fixed y, it follows that we have the we have the same bound when we integrate over y. (5 marks)

(Total: 0 marks)

5. (a) By definition, we have (2 marks)

$$\mathbb{P}(x, A) = \mathbb{E}_x(\mathbf{1}_A(X))$$

It suffices to show $h(x) = \mathbb{P}(x, A)$ is invariant since $B = \{x \in \mathcal{X} : h(x) = 1\}$. It follows, by using the Markov property in the third equality below (4 marks) and invariance (4 marks) of A in the fourth equality below, that

$$(Th)(x) = \mathbb{E}_x(h(X_1)) = \mathbb{E}_x\Big[\mathbb{E}_{X_1}(\mathbf{1}_A)\Big] = \mathbb{E}_x[\mathbf{1}_A \circ \theta] = \mathbb{E}_x[\mathbf{1}_A] = h(x);.$$

(b) Let A and B be as in the part above. By assumption, $\mathbb{P}_\pi$-almost surely,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_B(X_j) = \pi(B) .$$

By bounded convergence theorem for conditional expectations, we have, for π-almost every X_0,

$$\lim_{n\to\infty} \mathbb{E}\Big[\frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_B(X_j)|X_0\Big] = \pi(B) .$$

(3 marks) On the other hand, for each $j \geq 0$, we have

$$\mathbb{E}[\mathbf{1}_B(X_j)|X_0] = \Big(T^n \mathbf{1}_B\Big)(X_0) = \mathbf{1}_B(X_0) .$$

It follows that, for π-almost every x,

$$\mathbf{1}_B(x) = \pi(B) .$$

This is only possible if $\pi(B) = 0$ or $\pi(B) = 1$. (3 marks) On the other hand, we have

$$\mathbb{P}_\pi[A] = \mathbb{E}_\pi[\mathbf{1}_A] = \int_{\mathcal{X}} \mathbb{E}_x[\mathbf{1}_A]\pi(dx) = \int_{\mathcal{X}} \mathbf{1}_B(x)\pi(dx) = \pi(B) .$$

It follows that $\mathbb{P}_\pi(A) \in \{0, 1\}$, so π is ergodic. (4 marks for finishing proof)

(Total: 0 marks)

Module: MATH960031/MATH970031
Setter: Chandra
Checker: Krasovsky
Editor: Pal
External: Abban
Date: April 10, 2024
Version: Final version

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May – June 2024

MATH960031/MATH970031 Markov Processes

The following information must be completed:

Is the paper suitable for resitting students from previous years:

Category A marks: available for basic, routine material (excluding any mastery question) (40 percent = 32/80 for 4 questions):
1b) (4 marks); 2a) (3 marks); 2b) (3 marks); 2c) (3 marks); 3b) (7 marks); 3c) (5 marks); 4a) (4 marks)

Category B marks: Further 25 percent of marks (20/ 80 for 4 questions) for demonstration of a sound knowledge of a good part of the material and the solution of straightforward problems and examples with reasonable accuracy (excluding mastery question):
1a) (7 marks); 1c) (9 marks); 2d) (10 marks);

Category C marks: the next 15 percent of the marks (= 12/80 for 4 questions) for parts of questions at the high 2:1 or 1st class level (excluding mastery question):
3a) (8 marks);

Category D marks: Most challenging 20 percent (16/80 marks for 4 questions) of the paper (excluding mastery question):
4b) (16 marks);

Signatures are required for the final version:

Setter's signature	Checker's signature	Editor's signature
....	....	....

BSc, MSc and MSci EXAMINATIONS (MATHEMATICS)

May – June 2024

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science.

Markov Processes

Date: ??

Time: ??

Time Allowed: 2 Hours for MATH96 paper; 2.5 Hours for MATH97 papers

This paper has *4 Questions (MATH96 version); 5 Questions (MATH97 versions)*.

Statistical tables will not be provided.

- Credit will be given for all questions attempted.
- Each question carries equal weight.

1. (a) Give an example of a state space $\mathcal{X}$, an $\mathcal{X}$-valued Markov process $X = (X_n)_{n=0}^{\infty}$, and a bounded measurable function $f : \mathcal{B}_b(\mathcal{X})$ such that $Z = (Z_n)_{n=0}^{\infty}$ given by $Z_n = f(X_n)$ is not a Markov process (give a proof for the last assertion). (7 marks)

   (b) Given a transition probability P on a state-space $\mathcal{X}$, precisely define what it means for a probability measure $\mu \in \mathcal{P}(\mathcal{X})$ to be P-invariant. (4 marks)

   (c) Let $A = (A_n)_{n=0}^{\infty}$ be an $\mathbb{R}$-valued stochastic process which satisfies, for each bounded measurable $g \in \mathcal{B}_b(\mathbb{R})$, and any $i < j < k$, $\mathbb{E}\Big[g(A_k)|A_i, A_j\Big] = \mathbb{E}\Big[g(A_k)|A_j\Big]$.

   Prove that for any bounded measurable $f, h \in \mathcal{B}_b(\mathbb{R})$ and $i < j < k$, one has

$$\mathbb{E}\Big[f(A_i)h(A_k)|A_j\Big] = \mathbb{E}[f(A_i)|A_j]\mathbb{E}[h(A_k)|A_j] \; .$$

(9 marks)

(Total: 20 marks)

2. Let $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration generated by a $\mathcal{X}$-valued time homogenous Markov process $X = (X_n)_{n=0}^{\infty}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

   (a) Define what it means for a $\mathbb{N} \sqcup \{\infty\}$-valued random variable T to be an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time. (3 marks)

   (b) Suppose that $\{T_j\}_{j=0}^{\infty}$ are a family of $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping times, show that $\sup_j T_j$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time. (4 marks)

   (c) Given a $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time T, define the corresponding stopped sigma-algebra $\mathcal{F}_T$. (3 marks)

   (d) Prove, using the Markov property for X (**not** the strong Markov property) that for any $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time T with $\{T = \infty\} = \emptyset$, any $f \in \mathcal{B}_b(\mathcal{X})$, and any $n \in \mathbb{N}$, one has

$$\mathbb{E}\Big(f(X_{n+T})|\mathcal{F}_T\Big) = \int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y) \; ,$$

   where P^n is the n-step transition function for X. (10 marks)

(Total: 20 marks)

3. In all parts of this question, we specialize to looking at time homogenous Markov processes on a countable state space $\mathcal{X}$ and stochastic matrices P on $\mathcal{X}$.

(a) **True or False**: For each question, state whether the statement is true or false. If it is true, give a proof and if it is false give a counterexample with justification.

(i) Suppose P is irreducible, then P^2 is also irreducible. (4 marks)

(ii) Suppose P is reducible and $\mathcal{X}$ is finite, then P must have more than one invariant probability measure. (4 marks)

(b) Suppose we have distinct $i, j \in \mathcal{X}$. with j accessible from i, and with i not accessible from j. Prove that i is transient. (7 marks)

(c) Find the communication classes of the given stochastic matrix P, along with their partial ordering.

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \end{pmatrix}$$

(5 marks)

(Total: 20 marks)

4. (a) Let P be a stochastic matrix for the state space $\mathbb{Z}$. State what it means for a probability measure π on $\mathbb{Z}$ to satisfy *detailed balance* with respect to P. (4 marks)

(b) Let $Y = (Y_n)_{n=0}^{\infty}$ be a sequence of i.i.d $\mathcal{N}(0, 1)$ (i.e. standard Gaussian) random variables, and define an $\mathbb{R}$-valued Markov process $X = (X_n)_{n=0}^{\infty}$ by setting $X_0 = 0$ and for $n \in \mathbb{N}$,

$$X_{n+1} = Y_n + \sin(X_n/2 + Y_n) \ .$$

(i) Prove that there exists a invariant probability measure for X. (8 marks)

(ii) Prove that an invariant probability measure for X must be unique. (8 marks)

(Total: 20 marks)

Mastery Question

5. In the problems below, P is a transition probability on $\mathcal{X}$. Let θ be the shift map

$$a = (a_0, a_1, a_2, \dots) \mapsto \theta(a) = (a_1, a_2, \dots)\ .$$

We write $\mathcal{I} \subset \mathcal{B}\big(\mathcal{X}^{\mathbb{N}}\big)$ for the sigma algebra of invariant sets. Note that we work with one-sided sequences here.

Let $\pi \in \mathcal{P}(\mathcal{X})$ be a P-invariant probability measure, and X be the $\mathcal{X}$-valued Markov process transition P with initial distribution π. We write $\mathbb{P}_\pi$ for the law of X. Suppose that for every bounded measurable $f \in \mathcal{B}_b(\mathcal{X})$, one has, $\mathbb{P}_\pi$-almost surely,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \int_{\mathcal{X}} f(x)\, \pi(\mathrm{d}x)\ .$$

(a) Let $A \in \mathcal{I}$, and define

$$B = \{x \in \mathcal{X} : \mathbb{P}_x(A) = 1\}\ .$$

Prove that, $T\mathbf{1}_B = \mathbf{1}_B$ where T is the transition operator associated to P. **Hint:** Write $\mathbf{1}_B$ in terms of an expectation of $\mathbf{1}_A$. (10 marks)

(b) Prove that π as above is ergodic. (10 marks)

(Total: 20 marks)

1. (a) We set $\mathcal{X} = \{1, 2, 3, 4\}$, and define a time-homogenous Markov process $X = (X_n)_{n=0}^{\infty}$ with $\mathbb{P}(X_0 = 1) = \mathbb{P}(X_0 = 3) = \frac{1}{2}$ and using the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Note that we can obtain this Markov process through the following dynamical system:

$$X_{n+1} = \begin{cases} 2 & \text{if } X_n = 1 \\ 1 & \text{if } X_n = 2 \\ 4 & \text{if } X_n = 3 \\ 3 & \text{if } X_n = 4 \, . \end{cases}$$

We define $f : \{1, 2, 3, 4\} \to \{1, 2, 3\}$ by setting $f(1) = 1$, $f(3) = 3$ and $f(2) = f(4) = 2$. (3 marks for example that works)

We then have

$$\mathbb{P}(Z_2 = 1 | Z_1 = 2) = \mathbb{P}(Z_2 = 1) = \mathbb{P}(X_2 = 1) = \mathbb{P}(X_0 = 1) = \frac{1}{2} \, .$$

In the computation above, we used that $\{Z_1 = 1\} = \{X_1 = 2 \text{ or } 4\}$ has probability 1. On the other hand, we have

$$\mathbb{P}(Z_2 = 1 | Z_1 = 2, Z_0 = 1) = \mathbb{P}(Z_2 = 1 | Z_0 = 1) = 1 \, .$$

Above we used that $\{Z_0 = 1\} \subset \{Z_1 = 2\}$ in the first equality and in the second we observe that $\{Z_0 = 1\} = \{X_0 = 1\} = \{X_2 = 1\} = \{Z_2 = 1\}$. (4 marks for finishing justification).

(b) We have that μ is P-invariant if, for every Borel subset $A \in \mathcal{B}(\mathcal{X})$,

$$\mu(A) = \int_{\mathcal{X}} P(x, A) \, \mu(\mathrm{d}x) \, .$$

(4 marks)

(c) We have that

$$\begin{aligned} \mathbb{E}\Big[f(A_i)h(A_k) | A_j\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[f(A_i)h(A_k) | A_i, A_j\Big] \Big| A_j\Big] \\ &= \mathbb{E}\Big[f(A_i)\mathbb{E}\Big[h(A_k) | A_i, A_j\Big] \Big| A_j\Big] \\ &= \mathbb{E}\Big[f(A_i)\mathbb{E}\Big[h(A_k) | A_j\Big] \Big| A_j\Big] \\ &= \mathbb{E}\Big[h(A_k) | A_j\Big]\mathbb{E}\Big[f(A_i) | A_j\Big] \, . \end{aligned}$$

In the first equality we used the tower property (3 marks), in the second we factored out $f(A_i)$ since it is $\sigma(A_j, A_k)$ measurable (2 marks), then in the third equality we used our assumption that $\mathbb{E}\Big[h(A_k) | A_i, A_j\Big] = \mathbb{E}\Big[h(A_k) | A_j\Big]$ (2 marks), and then finally we factored out $\mathbb{E}\Big[h(A_k) | A_j\Big]$ since it is $\sigma(A_j)$-measurable. (2 marks)

(Total: 4 marks)

2. (a) A $\mathbb{N} \sqcup \{\infty\}$-valued random variable T on $(\Omega, \mathcal{F}, \mathbb{P})$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time if for each $n \in \mathbb{N}$, $\{T \leq n\} \in \mathcal{F}_n$. (3 marks)

   (b) We have that $\{\sup_j T_j \leq n\} = \bigcap_{j=0}^{n} \{T_j \leq n\} \in \mathcal{F}_n$, so $\sup_j T_j$ is an $(\mathcal{F}_n)_{n=0}^{\infty}$-stopping time. (4 marks)

   (c) The stopped sigma-algebra $\mathcal{F}_T$ is defined by setting $\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n \in \mathbb{N}\}$ where $\mathcal{F}_\infty$ is the σ-algebra generated by $\bigcup_{n=0}^{\infty} \mathcal{F}_n$. (3 marks)

   (d) It suffices to prove that, for any $A \in \mathcal{F}_T$,

$$\mathbb{E}\Big(f(X_{n+T})1_A\Big) = \mathbb{E}\Big[\int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y) 1_A\Big].$$

(2 marks)

We have that $\Omega = \bigcup_{n=0}^{\infty} \{T = n\}$ (2 marks for decomposing based on value of T), so we can write

$$\begin{aligned}
\mathbb{E}\Big(f(X_{n+T})1_A\Big) &= \sum_{j=0}^{\infty} \mathbb{E}\Big(f(X_{n+j}) 1_{A \cap \{T=j\}}\Big) \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[\mathbb{E}\Big(f(X_{n+j}) 1_{A \cap \{T=j\}}\Big) | \mathcal{F}_j\Big)\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[1_{A \cap \{T=j\}} \mathbb{E}\Big(f(X_{n+j}) 1_A\Big) | \mathcal{F}_j\Big)\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[1_{A \cap \{T=j\}} \int_{\mathcal{X}} f(y) P^n(X_j, \mathrm{d}y)\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[1_{A \cap \{T=j\}} \int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y)\Big] \\
&= \mathbb{E}\Big[\int_{\mathcal{X}} f(y) P^n(X_T, \mathrm{d}y) 1_A\Big]. \qquad \text{(6 marks)}
\end{aligned}$$

(Total: 0 marks)

3. (a) (i) This is false (1 mark). For example, consider the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(2 marks for counter-example) Then $P^2 = I$, so P^2 is reducible since $\{1\}$ and $\{2\}$ are communication classes. (1 mark for justification)

   (ii) This is false (1 mark). For example, consider the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then P is reducible and has a unique invariant probability measure (2 marks for counter-example) - note that for any measure μ with $\mu(1) = \alpha$, $\mu(2) = \beta$, we have $(\mu P)(1) = 0$ and $(\mu P)(2) = \alpha + \beta$, one must have $\alpha = 0$, $\beta = 1$. (1 mark for justification)

(b) Suppose that j is accessible from i, and that i is not accessible from j. Then $\{n \in \mathbb{N}_{>0} : P^n_{i,j} > 0\}$ is non-empty, let m be the minimum element of this set. It follows that there exits $i_1, \ldots, i_m = j$ such that

$$\mathbb{P}_i(X_1 = i_1, \ldots, X_m = i_m = j) > 0 \ .$$

Since we chose m to be the minimum of the set above, it also follows that $i_1, \ldots, i_m \neq i$ (3 marks for using accessibility with a minimal path) Since i is not accesible from i, it also follows that

$$\mathbb{P}_i\Big(\{X_m = j\} \setminus \{X_n \neq i \text{ for all } n \geq m\}\Big) = 0 \ .$$

(2 marks for using inaccessibility correctly) It follows that

$$\mathbb{P}_i(T_i = \infty) \geq \mathbb{P}_i(X_k \neq i \text{ for } 1 \leq k < m \text{ and } X_m = j) \geq \mathbb{P}_i(X_1 = i_1, \ldots, X_m = i_m = j) > 0 \ .$$

(2 marks for rest of proof)

(c) The communication classes of P are $\{1, 2, 6, 7\}$ and $\{3\}$ and $\{4, 5\}$ (3 marks), and the partial ordering is $\{3\} \prec \{1, 2, 6, 7\}$ and $\{3\} \prec \{4, 5\}$. (2 marks)

(Total: 0 marks)

4. (a) For every $i, j \in \mathcal{X}$, $\pi(i)P_{i,j} = \pi(j)P_{j,i}$. (4 marks)

(b) (i) We appeal to the Krylov-Bogoliubov argument. (2 marks for strategy) Since, for any fixed y, the function $x \mapsto \sin(x/2 + y)$ is continuous, it follows the associated transition operator T is Feller. (2 marks for Feller) For tightness, we note that for any fixed $x \in \mathbb{R}$ and any $n \in \mathbb{N}_{>0}$

$$\begin{aligned}\mathbb{E}_x[X_n^2] &= \mathbb{E}_x[\Big(Y_n + \sin(X_{n-1}/2 + Y_n)\Big)^2] \\ &\leq 2\mathbb{E}_x[Y_n^2] + 2\mathbb{E}_x[\sin(X_{n-1}/2 + Y_n)^2] \leq 4\end{aligned}$$

Above we used that $(a + b)^2 \leq 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$. Since Y_n is a standard Gaussian, it follows that $\mathbb{E}_x[Y_n^2] = 1$. It then follows that $\mathbb{E}_x[X_n^2] \leq 4$, so by Chebyshev inequality, or any $K > 0$,

$$P^n(x, \mathbb{R} \setminus [-K, K]) = \mathbb{E}_x[\mathbf{1}\{X_n^2 > K^2\}] \leq \frac{1}{K^2}\mathbb{E}_x[X_n^2] \leq \frac{4}{K^2} \ .$$

Since we the right hand side goes to 0 as $K \uparrow \infty$, uniformly in n, it follows that the measures $\Big(P^n(x, \bullet)\Big)_{n=1}^{\infty}$ are tight. (4 marks for tightness)

(ii) We argue via deterministic contraction. (3 marks for strategy) Note that for any fixed value of y,

$$\Big(y + \sin(x/2 + y)\Big) - \Big(y + \sin(x'/2 + y)\Big) = \sin(x/2 + y) - \sin(x'/2 + y) \leq \frac{1}{2}|x - x'| \ ,$$

where the last inequality follows from the mean value theorem and the fact that $\left|\frac{d}{dx}\sin(x/2 + y)\right| \leq \frac{1}{2}$. Since the bound above holds for any fixed y, it follows that we have the we have the same bound when we integrate over y. (5 marks)

(Total: 0 marks)

5. (a) By definition, we have (2 marks)

$$\mathbf{1}_B(x) = \mathbb{E}_x(\mathbf{1}_A(X))$$

It follows, by using the Markov property in the third equality below (4 marks) and invariance (4 marks) of A in the fourth equality below, that

$$(T\mathbf{1}_B)(x) = \mathbb{E}_x(\mathbf{1}_B(X_1)) = \mathbb{E}_x\Big[\mathbb{E}_{X_1}(\mathbf{1}_A)\Big] = \mathbb{E}_x[\mathbf{1}_A \circ \theta] = \mathbb{E}_x[\mathbf{1}_A] = \mathbf{1}_B(x) \ .$$

(b) Let A and B be as in the part above. By assumption, $\mathbb{P}_\pi$-almost surely,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_B(X_j) = \pi(B) \ .$$

By bounded convergence theorem for conditional expectations, we have, for π-almost every X_0,

$$\lim_{n\to\infty} \mathbb{E}\Big[\frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_B(X_j)|X_0\Big] = \pi(B) \ .$$

(3 marks) On the other hand, for each $j \geq 0$, we have

$$\mathbb{E}[\mathbf{1}_B(X_j)|X_0] = \Big(T^n \mathbf{1}_B\Big)(X_0) = \mathbf{1}_B(X_0) \ .$$

It follows that, for π-almost every x,

$$\mathbf{1}_B(x) = \pi(B) \ .$$

This is only possible if $\pi(B) = 0$ or $\pi(B) = 1$. (3 marks) On the other hand, we have

$$\mathbb{P}_\pi[A] = \mathbb{E}_\pi[\mathbf{1}_A] = \int_{\mathcal{X}} \mathbb{E}_x[\mathbf{1}_A]\pi(dx) = \int_{\mathcal{X}} \mathbf{1}_B(x)\pi(dx) = \pi(B) \ .$$

It follows that $\mathbb{P}_\pi(A) \in \{0, 1\}$, so π is ergodic. (4 marks for finishing proof)

(Total: 0 marks)

MATH60031 Markov Processes

Question	Marker's comment
1	Most students did well.
2	Most students did well, but some students used the strong markov property in part 2)d) even when asked not to. Note that 2)d) was in both the lectures and lecture notes.
3	Some students tried to use Perron Frobenius for 3)a)ii), but this doesn't work in this direction (the statement is indeed false) Additionally, some students tried to do 3)b) using a verbal explanation rather then explicit computations
4	Most students did well, although some were a bit sloppy with proving tightness in 4)b)i)

MATH70031 Markov Processes

Question	Marker's comment
1	For the most part, students did well on this question
2	Most students did well, but some students used the strong markov property in part 2)d) even when asked not to. Note that 2)d) was in both the lectures and lecture notes.nbsp;
3	Some students tried to use Perron Frobenius for 3)a)ii), but this doesn't work in this direction (the statement is indeed false)Additionally, some students tried to do 3)b) using a verbal explanation rather then explicit computations
4	Most students did well, although some were a bit sloppy with proving tightness in 4)b)i)
5	Students had a lot of difficulty with this problem, they showed some familiarity with the definitions but weren't able to use them effectively to argue proofs.nbsp;