HOMEWORK 2

654 STOCHASTIC PROCESSES, PB

Note for the current and future homework: due to a large class size and time constraints, not all exercises will be graded.

A linear algebra example to clarify what's happening to prove uniqueness in section 1.5.4 (optional, skip if you are confident on your linear algebra). Let I be the 2×2 identity matrix and consider the matrix

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$

Compute the rank of $P - I$. Determine the subspaces

$$V = \{h \text{ column vector of size 2 such that } h = Ph\}.$$

$$W = \{\mu \text{ row vector of size 2 such that } \mu = \mu P\}.$$

If needed, here are some online resources on kernel (nullspace), row rank and column rank of a matrix:

- https://math.mit.edu/linearalgebra/ila0306.pdf
- https://math.stackexchange.com/questions/332908/looking-for-an-intuitive-explanation-why-the-row-rank-is-equal-to-the-column-ran
- https://ocw.mit.edu/courses/mathematics/18-701-algebra-i-fall-2010/study-materials/MIT18_701F10_rrk_crk.pdf

Basic properties of Markov Chains. Let $\mathcal{X}$ be a finite set. We assume that $\{X_t, t = 0, 1, 2, 3, ...\}$ satisfies the Markov property as in equation (1.1) in the book. Let P be the matrix of transition probabilities.

Note that there are several ways of proving following results, including (a) matrix calculus, (b) conditioning on the "right" event and then using the Markov property and the properties of conditional probabilities, or (c) invoking a random mapping representation $f : \mathcal{X} \times [0, 1]$ with an iid sequence of random variables $Z_1, Z_2, Z_3, ...$ uniform in $[0, 1]$ such that $\mathbf{P}(f(x, Z_t) = y) = P(x, y)$.

1. For brevity, in this question we use the notation $AB = A \cap B$ for any two events. Let A, B, C, D, E be events. Prove that

   $$\mathbf{P}(ABCDE) = \mathbf{P}(E|ABCD)\mathbf{P}(D|ABC)\mathbf{P}(C|AB)\mathbf{P}(B|A)\mathbf{P}(A)$$

   and state a generalization to n events $A_1, ..., A_n$.

2. For $x_0, ..., x_t \in \mathcal{X}$ and μ a distribution on $\mathcal{X}$, express

   $$\mathbf{P}_\mu(X_t = x_t, X_{t-1} = x_{t-1}, ..., X_0 = x_0)$$

   in terms of P and μ.

3. Let $g : \mathcal{X}^{k+1} \to \{0,1\}$ be a function. Show that
$$\mathbf{P}_\mu(g(X_0, ..., X_k) = 1) = \sum_{x \in \mathcal{X}} \mu(x)\mathbf{P}_x(g(X_0, ..., X_k) = 1).$$

4. Let $g : \mathcal{X}^{k+1} \to \{0,1\}$ be a function. Express the probability $\mathbf{P}_\mu(g(X_0, ..., X_k) = 1)$ as a sum over the set
$$\mathcal{P} = \{(x_0, ..., x_k) \in \mathcal{X}^{k+1} \quad \text{such that} \quad g(x_0, ..., x_k) = 1\}.$$

5. For the rest of the questions, let $g : \mathcal{X}^{k+1} \to \{0,1\}$ and $p : \mathcal{X}^{t+1} \to \{0,1\}$ be two functions (think of these two functions as indicator functions of events. g stands for "future" and p for "past"). For $x_t \in \mathcal{X}$, if the event $\{X_t = x_t, p(X_0, X_1, ..., X_t) = 1\}$ has positive probability, show that
$$\begin{aligned}&\mathbf{P}_\mu[g(X_t, X_{t+1}, ..., X_{t+k}) = 1 | X_t = x_t, p(X_0, X_1, ..., X_t) = 1]\\ &= \mathbf{P}_\mu[g(X_t, X_{t+1}, ..., X_{t+k}) = 1 | X_t = x_t]\\ &= \mathbf{P}_{x_t}[g(X_0, X_1, ..., X_k) = 1].\end{aligned}$$

6. Prove also that if $\{X_t = x_t\}$ has positive probability, then conditionally on $\{X_t = x_t\}$, the future is independent from the past in the sense that
$$\begin{aligned}&\mathbf{P}_\mu[g(X_t, X_{t+1}, ..., X_{t+k}) = 1, p(X_0, X_1, ..., X_t) = 1 | X_t = x_t]\\ &= \mathbf{P}_\mu[g(X_t, X_{t+1}, ..., X_{t+k}) = 1 | X_t = x_t]\mathbf{P}_\mu[p(X_0, X_1, ..., X_t) = 1 | X_t = x_t]\\ &= \mathbf{P}_{x_t}[g(X_0, X_1, ..., X_k) = 1]\mathbf{P}_\mu[p(X_0, X_1, ..., X_t) = 1 | X_t = x_t]\end{aligned}$$

7. Let $E = \{X_t = x_t, X_{t+1} = x_{t+1}, ..., X_{t+m} = x_{t+m}\}$. Prove that if E has positive probability, then conditionally on the event E, the future is independent from the past in the sense that
$$\begin{aligned}&\mathbf{P}_\mu[g(X_{t+m}, X_{t+m+1}, ..., X_{t+m+k}) = 1, p(X_0, X_1, ..., X_t) = 1 | E]\\ &= \mathbf{P}_\mu[g(X_{t+m}, X_{t+m+1}, ..., X_{t+m+k}) = 1 | E]\mathbf{P}_\mu[p(X_0, X_1, ..., X_t) = 1 | E]\\ &= \mathbf{P}_{x_{t+m}}[g(X_0, X_1, ..., X_k) = 1]\mathbf{P}_\mu[p(X_0, X_1, ..., X_t) = 1 | X_t = x_t]\end{aligned}$$

   Do the above equalities still hold if E is replaced by $\{X_t = x_t, X_{t+m} = x_{t+m}\}$?

Exercises from the book. Starting page 17, Exercises

- 1.2,
- 1.7,
- 1.8 (A Markov chain is said to be reversible if it satisfies the detailed balanced equations, cf. (1.29) in section 1.6 of the book),
- 1.9,
- 1.11.
- 1.6 (Optional) Here you must define the classes $C_1, ..., C_b$ that define a partition of $\mathcal{X}$, such that $P(x, y) > 0$ implies that $x \in C_i$ and $y \in C_{i+1}$ for some i,