Department of Mathematical Sciences
Division of Applied Mathematics
Stellenbosch University

2019 | MS763 | Applied Markov Processes

Problems for Chapter 2: Markov chains

Due: 3 September 2019

Theoretical

Q1. **(Birth-death process)** Consider the simple population model seen in class for which $P(i \to i+1) = \alpha$ (birth) and $P(i \to i-1) = \beta$ (death), where $i \in \{0, 1, 2, \ldots\}$ and $\alpha + \beta \leq 1$.

(a) Write down the matrix Π of transition probabilities.

(b) Write down the Chapman–Kolmogorov equation in components. [Hint: There are only two relevant equations.]

(c) Find the stationary distribution p^*, assuming $\beta > \alpha > 0$. [Hint: Guess the solution or truncate the system to a finite number of states.]

(d) Calculate the expected stationary population.

(e) Why do we need $\alpha < \beta$. What happens if $\alpha > \beta$?

Q2. Show that a Markov chain has the property that the future is independent of the past given the present. We also say that the past and future are conditionally independent given the present.

Q3. A discrete Markov chain with $|\mathcal{X}|$ states is called **bi-stochastic** or **doubly stochastic** if $\sum_i \Pi_{ij} = 1$ in addition to the usual normalisation property $\sum_j \Pi_{ij} = 1$. Show in this case that the stationary distribution is the uniform distribution $p_i^* = 1/|\mathcal{X}|$.

Q4. **(Random walk on graphs)** Show that the stationary distribution of the unbiased random walk on an undirected, connected graph is, as seen in class,

$$p_i^* = \frac{k_i}{2M}, \tag{1}$$

where k_i is the degree of the node i, obtained from the adjacency matrix A_{ij} by $k_i = \sum_j A_{ij}$, and $M = \frac{1}{2}\sum_i k_i$ is the total number of edges in the graph.

Q5. A Markov chain with transition matrix Π_{ij} is said to be **reversible** with respect to a distribution p_i if it satisfies the following condition:

$$p_i \Pi_{ij} = p_j \Pi_{ji}, \qquad i, j \in \mathcal{X}, \tag{2}$$

known as the **detailed balance condition**.

(a) Show that p_i is a stationary distribution of the Markov chain.

(b) Show that the transformed matrix

$$\hat{\Pi}_{ij} = (p_i)^{1/2} \Pi_{ij} (p_j)^{-1/2} \tag{3}$$

is symmetric.

(c) What can be said about the eigenvalues of Π?

Q6. **(Markov chain Monte Carlo)** Let $\pi(x)$ be a probability distribution with $\pi(x) > 0$ for all x. Show that the Metropolis algorithm, based on the following transition probability:

$$P(x \to x') = \min\left\{1, \frac{\pi(x')}{\pi(x)}\right\} \tag{4}$$

for going from x to x', defines a reversible Markov chain. What is its stationary distribution?

Numerical

Q7. **(Stationary distribution)** We have seen in class that the stationary distribution p^* of an ergodic Markov chain can be computed using three methods:

1. Solve $p^*\Pi = p^*$, that is, find the left eigenvector of Π of eigenvalue 1;
2. Simulate *many* independent Markov chains in parallel or one after the other and compute the histogram of their final state X_n for n large enough;
3. Simulate a *single* Markov chain $\{X_i\}_{i=1}^n$ with n time steps and compute the time-averaged occupation

$$\hat{p}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i,x} = \frac{\text{\# states with value } x}{n}. \tag{5}$$

Show for the two-state Markov chain seen in class (consider the symmetric or non-symmetric one) that the last two numerical methods agree with the analytical result of Method 1. Which method do you see as more effective and why?

Q8. **(Random walk on graphs)** Choose a large enough connected graph, say with 10 or more states, and write a program that simulates a trajectory of the unbiased random walk on that graph. The program should use the adjacency matrix of the graph as the internal representation of the graph and should output a trajectory $\{x_0, x_1, \ldots, x_n\}$ of length n starting at x_0. Use a long enough trajectory to estimate the stationary distribution of the random walk (use Method 3 of the previous question) and show that it agrees with the stationary distribution in Q4.

Q9. **(Markov chain Monte Carlo)** Let us revisit the estimation of π, as seen in CW1. Instead of dropping independent random points in the square $[-1,1]\times[-1,1]$, choose one point in the square – any point, e.g., $P_0 = (0,0)$ – and iterate the following steps to construct a Markov chain $P_1 \to P_2 \to \cdots \to P_L$ of L points:

Step 1: Choose a random displacement $\delta P = (\delta P_x, \delta P_y)$ according to *any* symmetric distribution;

Step 2: Set $P_{i+1} = P_i + \delta P$ if P_{i+1} stays in the square (accept move); otherwise, set $P_{i+1} = P_i$ (reject move).

For this algorithm,

(a) Show that the set of points $\{P_i\}_{i=1}^L$ is uniform, no matter what distribution is used for generating δP (!). [Hint: This is a Metropolis algorithm.]

(b) Construct from the Markov chain an estimator of π and show that it converges to the correct value. [Hint: Use a "good" distribution for δP.]

(c) Can we construct error bars for this estimator the way we have seen in class? Explain your answer.

Reading

- GS, Secs. 6.2 and 6.3: Classification of states and Markov chains (scanned pages on SunLearn)
- Wikipedia entry on Pseudorandom number generators.
- Wikipedia entry on Metropolis–Hastings algorithm. Also covered in Sec. 6.14 of GS.

Prize question

R100 for the best complete answer. Hand in your solution on a separate sheet.

We have seen that an ergodic Markov chain with transition matrix Π converges to a unique stationary distribution p^* starting from any initial distribution p_0. Presumably, some p_0's converge to p^* faster than others. Moreover, different Π with the same p^* might not converge to that distribution with the same speed. What properties of Π or p_0 determine the speed of convergence towards p^*?

Hugo Touchette Version: 16/08/2019 12:07