Chapter 2 Markov Chains

1 Introduction

The importance of Markov chains comes from the fact that it has a wide range of applications in physics, biology, economics and social sciences, and that there is a well-established theory that allows us to perform computations. Let us first examine an example before giving a formal definition of Markov chains.

Gambler's fortune Starting with a fortune k a gambler is determined to earn an extra $(n-k)$ playing the following game repeatedly: He tosses a coin and if it shows heads he wins 1 and he loses 1 if it is tails. The gambler can either manage to reach his objective of reaching n or lose everything. We are interested in the probability of either events happening. A convenient way of modelling the problems is as follows.

Let X_i be 1 with probability p and -1 with probability $1-p$ starting from of fortune k the amount of money (positive or negative) that the gambler's has after playing the game for n rounds is given by

$$S_n = k + X_1 + \cdots + X_n \, .$$

Note that his fortune (or debt) at time $n+1$ depends on his fortune S_n at time n that includes all the information necessary about his past performance. Indeed

$$\mathbf{P}\left(S_{n+1} = i+1 \mid S_n = i, S_{n-1} = i_{n-1}, \ldots, S_0 = k\right) = p = \mathbf{P}(S_{n+1} = i+1 \mid S_n = i) \, .$$

The process $(S_n)_{n\geq 0}$ is said to follow the Markov property. The process $(S_n)_{n\geq 0}$ is a *Markov chain.*

2 Markov property

Markov chains enable us to define a general framework to describe such dynamics. Let $\{X_0, X_1, \ldots\}$ be a sequence of discrete-random variables taking values in S, the state space.

Definition 1 *The process (sequence)* $X = \{X_0, X_1, \ldots\}$ *is a Markov chain if it satisfies*

$$\mathbf{P}(X_n = s \mid X_0 = x_0,\ X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = \mathbf{P}(X_n = s \mid X_{n-1} = x_{n-1}) \, ,$$

for all $n \geq 1$ *and* $s, x_0, \ldots, x_{n-1} \in S$.

It is said to be homogeneous if

$$\mathbf{P}(X_n = j \mid X_{n-1} = i) = \mathbf{P}(X_1 = j \mid X_0 = i) \, .$$

In this case we define the transition matrix $P = (p_{ij})_{i,j}$ *the matrix of transition probabilities*

$$p_{ij} = \mathbf{P}(X_n = j \mid X_{n-1} = i) = \mathbf{P}(X_1 = j \mid X_0 = i) \, .$$

The definition of a Markov chains can be restated as follows: *given the current state* X_n *any other information about the past is irrelevant for predicting the future state* X_{n+1} as it is already contained in X_n.

In the remainder of the chapter we will focus our attention on homogeneous Markov chains.

Example 1 *Gambler's ruin or random walk.* S_n *is a Markov chain with transition matrix*

$$p_{ij} = \begin{cases} p, & \text{if } j = i+1, \\ 1-p, & \text{if } j = i-1, \\ 0, & \text{otherwise} \end{cases}$$

Example 2 *Wright-Fisher model.*

We consider a fixed population of N individuals each having one copy of one of two types (alleles): A or a. As a first simplification let us assume that at time $n+1$ the model of the population is obtained by drawing with replacement from the population at time n. Let X_n be the number of A alleles at time n, then X_n is a Markov chain with transition probability

$$p_{ij} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j} .$$

To make the model more interesting we can add mutations, that is to say an A drawn from the n-th generation ends being an a in the next generation with probability u while the probability is v if the mutation is from a to A. Hence, using Bayes' rule, the probability that an A is produced by a given draw is

$$\rho_i = \frac{i}{N}(1-u) + \frac{N-i}{N} v ,$$

and the transition still has the binomial form

$$p_{ij} = \binom{N}{j} (\rho_i)^j \left(1 - \rho_i\right)^{N-j} .$$

Example 3 *Ehrenfest chain.*

Consider two cubical volumes of air connected by a small hole. We can picture this as two urns and N balls. We pick one ball at random from one urn and move it to the other urn. If X_n is the number of balls in first urn after the n-th draw it is clear that it forms a Markov chain with transition probability

$$p_{i(i+1)} = \frac{N-i}{N}, \quad p_{i(i-1)} = \frac{i}{N} .$$

For $N = 3$ the transition matrix is given by

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix} .$$

Proposition 1 *The transition matrix is a stochastic matrix, i.e.,*

1. *$p_{ij} \geq 0$, for all i, j*
2. *$\sum_j p_{ij} = 1$, for all i.*

Example 4 *Social mobility. Let X_n be a family's social class in the nth generation, which is assumed to be either lower, middle or upper. A simple version of social mobility we assume that changes of social status is described by the following transition matrix*

$$\begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} .$$

Suppose that a family starts in the middle class in generation 0. what is the probability that the generation one rises to the upper class and generation two falls to the lower class? What is the probability that it will be in the lower class at generation 2?

Define the n-step transition matrix $P(n) = (p_{ij}(n))_{ij}$

$$p_{ij}(n) = \mathbf{P}(X_{n+m} = j \mid X_m = i) = \mathbf{P}(X_n = j \mid X_0 = i) .$$

Example 5 *Bernoulli process. Let $S = \{0, 1, 2, \dots\}$ and define the Markov chain Y by $Y_0 = 0$ and*

$$\mathbf{P}(Y_{n+1} = s + 1 \mid Y_n = s) = p, \quad \mathbf{P}(Y_{n+1} = s \mid Y_n = s) = 1 - p$$

for $p \in (0, 1)$. you may think of Y_n as the number of heads thrown in n tosses of a biased coin. It is easy to see that

$$\mathbf{P}(Y_{n+m} = j \mid Y_m = i) = \binom{n}{j-i} p^{j-i}(1-p)^{n-j+i}, \quad 0 \le j - i \le n .$$

The transition matrix P of the Markov chain X and the probability function of the initial condition μ_0 entirely determine the Markov chain.

Theorem 1 (Chapman-Kolmogorov) *For all $i, j \in S$ and all integers m, n*

$$p_{ij}(m + n) = \sum_{k \in S} p_{ik}(n) p_{kj}(m)$$

i.e, for $P(n) = (p_{ij}(n))_{i,j \in S}$ we have

$$P(n + m) = P(n)\, P(m), \quad P(n) = P^n .$$

Example 6 *The weather chain. Let X_n be the weather on day n in London, which we assume either rainy or sunny. We can propose a very simple model of weather change with transition matrix*
Let $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$.

Example 7 *Virus mutation. Suppose that a virus can exist in N different strains and each generation either stays the same, or with probability α mutates to another strain, chosen at random. What is the probability that the strain in the nth generation is the same as the initial strain?*

Example 8 *Linear algebra. Let* $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$. *What is the probability of returning to 1 after n steps?*

To completely determine the evolution of the Markov chain, i.e. to give the distribution of X_n for any instant n, we also need to specify the initial condition or state from which the chain is started at time 0. In general we need to know the probability distribution $\mathbf{P}(X_n = i_0)$, $i_0 \in S$, given μ_0 and P. By Chapman-Komogorov

$$\begin{aligned} \mathbf{P}(X_n = j) &= \sum_{i_0 \in S} p_{i_0 j}(n) \mu_0(i_0) \\ &= \sum_{i_0 \in S} \cdots \sum_{i_{n-1} \in S} \mathbf{P}(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = j) \\ &= \sum_{i_0 \in S} \cdots \sum_{i_{n-1} \in S} \mu_0(i_0) p_{i_0, i_1} \dots p_{i_{n-1}, j} \\ &= (\mu_0 P^n)_j , \end{aligned}$$

where $\mu_0 = (\mu_0(i))_{i \in S}$.

Example 9 *The weather chain with transition matrix* $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$ *has different asymptotic behaviours depending on the choice of the parameters* α *and* β*. In fact,*

- *for* $\alpha+\beta \in (0,2)$*,* P^n *converges to* $P_\infty = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix}$*, i.e. the probabilities of being in state* 1 *and state* 2*, in the long run, are given by is* $\mathbf{P}(X_\infty = 1) = \frac{\beta}{\alpha+\beta}$ *and* $\mathbf{P}(X_\infty = 2) = \frac{\alpha}{\alpha+\beta}$ *respectively, regardless of the initial condition.*
- *For* $\alpha = \beta = 1$*, a rather different picture emerges as* $P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *and* $P^{2n+1} = P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$*. Unless* $\mu_0 = (1/2, 1/2)$*, there is no convergence. This is property is due to the periodicity of the chain.*
- *For* $\alpha = \beta = 0$*,* $P^n = P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *and the chain stays in its initial state which completely determines its future behaviour.*

In fact the convergence of the sequence P^n to some limit P_∞ when n goes to ∞ does not necessarily imply that the sequence X_n converges to a distribution that is independent of the initial condition.

To illustrate this let us return to the asymptotic behaviour of Markov chains. Given P and μ_0 we are interested in the asymptotic μ_∞ of the variable X_∞ which is given by $\mu_0 P_\infty$ where P_∞ is the limit of P_n when n goes to infinity. We start with an example.

Example 10 *We consider the gambler's ruin problem with f* $N = 4$ *and playing with a fair coin.*

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

We can easily check that, as n goes to infinity, that $P^n \to P_\infty = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3/4 & 0 & 0 & 0 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \\ 1/4 & 0 & 0 & 0 & 3/4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

In particular if the initial condition is $\mu_0 = [\mu_0(0), \mu_0(1), \mu_0(2), \mu_0(3), \mu_0(4)]$ *then*

$$\mu_\infty = \mu_0 P_\infty = [\mu_0(0) + \frac{3}{4}\mu_0(1) + \frac{1}{2}\mu_0(2) + \frac{1}{4}\mu_0(3), 0, 0, 0, \frac{1}{4}\mu_0(1) + \frac{1}{2}\mu_0(2) + \frac{3}{4}\mu_0(3) + \mu_0(4)].$$

3 Hitting times and absorption probabilities

We consider the random variable T^A corresponding to the time it takes the frog to enter a subset A of the state space S, i.e.

$$T^A = \inf\{n \geq 0,\ X_n \in A\}$$

The probability that starting from i the chain ever hits A is given by

$$h_i^A = \mathbf{P}(T^A < \infty \mid X_0 = i).$$

It is referred to as absorption probability in A of the Markov chain starting at i.

The mean time taken by X to reach A is given by

$$H_i^A = \mathbf{E}\left(T^A \mid X_0 = i\right) .$$

The quantities H_i^A and h_i^A can be calculated using the transition matrix P.

Example 11 *Let* $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. *Starting from state* 2*, what is the probability of absorption in* 4*? Howe long it takes until the chain is absorbed in* 1 *or* 4*?*

Theorem 2 *The vector of absorption probabilities* $h^A = (h_i^A, i \in S)$ *is the (non-negative minimal) solution of*

$$\begin{cases} h_i^A = 1, & \text{if } i \in A, \\ h_i^A = \sum_{j \in S} p_{ij} h_j^A, & \text{if } i \notin A. \end{cases}$$

The vector of hitting times $H^A = (H_i^A, i \in S)$ *is the (non-negative minimal) solution of*

$$\begin{cases} H_i^A = 0, & \text{if } i \in A, \\ H_i^A = 1 + \sum_{j \notin A} p_{ij} H_j^A, & \text{if } i \notin A. \end{cases}$$

Example 12 *Gambler's ruin revisited. A gambler has* 2 *pounds and needs to increase it to* 10 *pounds in a hurry. He can play a game with the following rules: a fair coin is tossed ; if the player bets on the side which actually turns up, he wins a sum equal to his stake and his stake is returned; otherwise he loses his stake. The gambler decides to use a strategy in which he stakes all his money if he has* 5 *pounds or less and otherwise stakes just enough to increase his capital, if he wins, to* 10 *pounds. Prove that the gambler will achieve his goal with probability* 1/5 *and that the expected number of tosses before he either achieves his aim or loses his capital is* 2*?*

4 Communicating classes

Definition 2 *We say that state* j *is accessible from* i *or* i *leads to* j*, if the chain, started in* i*, ever visits state* j*, with positive probability, i.e. if we have*

$$\mathbf{P}_i(X_n = j) = \mathbf{P}(X_n = j \mid X_0 = i) = p_{ij}(n) > 0$$

for some $n \geq 0$*.*

We can see this in the graphical representation as the existence of a path $i_0 = i, i_1 \in S, \ldots i_{n-1} \in S, i_n = j$ *such that* $p_{i_0 i_1} \ldots p_{i_{n-1} i_n} > 0$*.*

We say that the states i *and* j *communicate if* i *is accessible from* j *and* j *is accessible from state* i*.*

Let us now consider the relation communicates with $i \leftrightarrow j$ as a relation between states. It is not difficult to see that if $i \leftrightarrow j$ and $j \leftrightarrow k$ and that $i \leftrightarrow i$ so that the relation $\leftrightarrow$ is an equivalence relation on S which is, by the properties of equivalence relations, partitioned into equivalence classes which we refer to as *communicating* classes.

we explore the properties of the different states a chain can visit: certain states can be visited an infinite number of times these are known as recurrent (or persistent) others can only be visited finite number of times they are called transient. More precisely

Definition 3 *Given a Markov chain* $(X_n)_n$ *on the state space* S *and* $i \in S$

- *i is recurrent if $\mathbf{P}_i(X_n = i$ for infinitely many $n) = 1$, alternatively*

$$\sum_{n\geq 0} p_{ii}(n) = \infty\,.$$

- *i is transient if $\mathbf{P}_i(X_n = i$ for infinitely many $n) < 1$, alternatively*

$$\sum_{n\geq 0} p_{ii}(n) < \infty\,.$$

Assume that we have states i and j such that $i \to j$ and $j \nrightarrow i$. Then r_{ij} the probability of reaching j from i is strictly greater than 0. Letting r_i be the probability of returning to i then $1 - r_i \geq r_{ij}$ (one way of not returning to i is to go through j), i.e. $r_i \leq 1 - r_{ij} < 1$ which implies that i is transient.

Example 13 *For the Markov chain with transition matrix*

$$P = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Here $2 \to 1$ and $1 \nrightarrow 2$ so 2 is transient; $3 \to 4$ and $4 \nrightarrow 3$ so 3 is transient. The sets $\{1,5\}$ and $\{4,6,7\}$ are the remaining communicating classes

Definition 4 *A communicating class $C \subset S$ is said to be closed if for all $i \in C$ and $j \notin C$ $p_{ij} = 0$. Alternatively, for $i \in C$ and $j \in S$ such that $i \to j$ then $j \in C$.*

A set B is said to be irreducible if $i, j \in B$, we have $i \leftrightarrow j$.

In the previous example $\{1,5\}$ and $\{4,6,7\}$ are closed as well as $\{1,4,5,6,7\}$ and $\{1,2,3,4,5,6,7\}$. However, only $\{1\},\{2\},\{3\},\{5\},\{6\},\{1,5\},\{4,6,7\}$ are irreducible.

For closeness $\{1,4,5,6,7\}$ and $\{1,2,3,4,5,6,7\}$ are not of great interest as are $\{1\},\{2\},\{3\},\{5\},\{6\}$ for irreducibility.

Proposition 2 *All the elements of a closed irreducible set are recurrent.*

The only such ensembles are $\{1,5\}$ and $\{4,6,7\}$, so that the states $1,4,5,6,7$ are recurrent.

Example 14 *Ehrenfest chain.*

Consider two cubical volumes of air connected by a small hole. We can picture this as two urns and N balls. We pick one ball at random from one urn and move it to the other urn. If X_n is the number of balls in first urn after the n-th draw it is clear that it forms a Markov chain with transition probability

$$p_{i(i+1)} = \frac{N-i}{N}, \quad p_{i(i-1)} = \frac{i}{N}\,.$$

For $N = 3$ the transition matrix is given by

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

It is not difficult to check that we have, when n goes to infinity,

$$P^{2n} \to \begin{pmatrix} 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \end{pmatrix} \qquad P^{2n+1} \to \begin{pmatrix} 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix} .$$

This is not very surprising as the number of balls in one the urns changes from an even to an odd number and vice-versa at each step. Since the distribution of the states of the chain does not converge to a given distribution except for some particular initial conditions.

Definition 5 *We say that A state j of a Markov chain $(X_t)_t$ is periodic is there exists am integer Δ such that*

$$\mathbf{P}(X_{t+s} = j \mid X_t = j) = 0$$

unless s is divisible by Δ. The chain is periodic if all its states are periodic and is aperiodic if non of its states is periodic.

As a matter of example, for $\alpha, \beta \in (0, 1)$ the weather chain is aperiodic. For the gambler's ruin, the two extreme states are aperiodic, and all other states have period 2. For the Ehrenfest model all the states have period 2.

We now introduce an important class of Markov chains

Definition 6 (Ergodic Chains) *A finite state space, irreducible and aperiodic Markov chain is an ergodic chain.*

Remark 1 *If we do not assume that the chain has a finite state space, we will need to introduce the notions of positive recurrent chain: a state i is said to be positive recurrent if stating from i the average return time of the chain to i is finite otherwise its called null recurrent. For finite chain, it is not difficult to see that any recurrent state is positive recurrent.*

5 Stationary distribution

So far we looked at the long-run (asymptotic) behaviour of Markov chains by looking at hitting times and computing absorption probabilities In this section we are interested in the behaviour of the Markov chain after a large number of steps. In many random phenomena, although it is unlikely that the random system we are interested is going to settle in a particular state and not budge, under some conditions the distribution of X_n stops evolving. It reaches stationarity or equilibrium. Ergodic Markov chains have this desirable property.

Let us suppose the the initial state is random, i.e. we let $\mu_i(0) = \mathbf{P}(X_0 = i)$, and $\mu_i(n) = \mathbf{P}(X_n = i)$. The vector $\mu(n) = (\mu_i(n),\ i \in S)$ satisfies

$$\mu(n) = \mu(0)P(n), \quad \mu(m+n) = \mu(m)P(n) .$$

Repeatedly applying Bayes' formula we get

$$\mathbf{P}(X_0 = x_0, X_1 = i_1, \ldots, X_n = i_n) = \mu_0(i_0)p_{i_0 i_1} \ldots p_{i_{n-1} i_n} .$$

Definition 7 *The vector π is called a stationary or invariant distribution of the the chain X if $\pi = (\pi_i, i \in S)$ is such that*

- $\pi_i \geq 0$ *for all i and* $\sum_i \pi_i = 1$

- $\pi P = \pi$, *i.e.*, $\pi_j = \sum_i \pi_i p_{ij}$

In other words, if π is the stationary distribution then starting from an initial condition distributed according to π then this is going to be the distribution at all subsequent times.

Example 15 *Consider the chain with transition matrix* $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$. *solving* $\pi P = \pi$ *with* $\pi_1 + \pi_2 + \pi_3 = 1$ *we have* $\pi = (1/5, 2/5, 2/5)$.

The next result explains the idea of equilibrium.

Theorem 3 *If the state space* S *is finite, such that*

$$\lim_n p_{ij}(n) = \pi_j$$

then, $(\pi_j)_{j \in S}$ *is a probability distribution then it is an invariant distribution. Moreover*

$$\pi_j = \frac{1}{h_{j,j}}$$

where $h_{j,j}$ *the expected number of steps to return to* j *when starting in* j.

Remark 2 *It is not always the case that such limits exists as in the case where* $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

In example 8, we showed that for the chain in example 15 satisfies

$$p_{11}(n) = a + \left(\frac{1}{2}\right)^n (b \cos(n\pi/2) + c \sin(n\pi/2)) \ .$$

Hence using the above theorem and the derivation of the stationary distribution, we have $a = 1/5$, instead of working it out from $p_{11}(2)$ as we previously did.

Example 16 *The weather chain. Transition matrix* $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$. *It is not difficult to see that* $\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$. *Moreover, for* $i = 1, 2$

$$\left| p_{i1}(n) - \frac{\beta}{\alpha+\beta} \right| \leq |1 - \alpha - \beta|^n \ .$$

Theorem 4 *Let* $(X_n)_{n \geq 0}$ *be an ergodic Markov chain with transition matrix* P *then the chain has a unique stationary distribution pi and*

$$\lim_n p_{ij}(n) = \pi_j \ .$$

Example 17 (Simple queue) *A queue is a line where customer wait for service as follows: at each time step, exactly one of the following events occurs*

- *If the queue has fewer than* $N > 0$, *a new customer arrives with probability* λ
- *If the queue is not empty a customer departures with probability* μ
- *the queue is unchanged with the remaining probability*

6 Random walks on graphs

We denote by $G = (V, E)$ a graph with vertices (or nodes) in V and edges in E where the set of edges (or links) is in $V \times V$ such that $(i, j) \in E$ means that nodes i and j are connected.

Remark 3 *In what follows we focus on undirected[1] connected[2], non-bipartite[3] and simple[4] graphs.*

For each node $i \in V$ we say that j is a neighbour of i if $(i, j) \in E$ and we denote by d_i the number of neighbours of node i, a.k.a. the degree of node i.

A (simple) random walk on the graph G is the Mrkov chain $(X_n)_{n\geq 0}$ with state space V and transition matrix $p_{ij} = \frac{1}{d_i}$ if j is a neighbour of i that we denote by $i \sim j$ and 0 otherwise. The stationary distribution of this chain is given by

$$\pi_i = \frac{d_i}{2m}$$

where $2m = \sum_{i\in V} d_i$, and $m = |E|$ the number of edges in G. Let $h_{u,v}$ be the average number f steps for the walk to reach v from u. By theorem 3

$$h_{u,u} = \frac{1}{\pi_u} = \frac{2m}{d_u} \,.$$

As when deriving the absorption time, it is not difficult to show that

$$\frac{2m}{d_u} = h_{u,u} = \sum_{w\sim u} \frac{1}{d_u}(1 + h_{w,u})$$

which implies that for v a neighbour of u we have

$$h_{v,u} < \sum_{w\sim u}(1 + h_{w,u}) = 2m \,. \tag{1}$$

Definition 8 *We denote by $C_i(G)$ the times it takes the walk, stated in i to visit all the nodes in V at least once. the cover time of the graph G is given by*

$$C(G) = \max_{i\in V} C_i(G) \,.$$

We have already come across the notion of cover time when dealing with the coupon collector problem. In fact the duration of the coupon collection is equivalent to the cover time of the complete graph, the graph where all pairs of nodes are connected. In particular the cover time of the complete graph with N nodes is given by $C(K_N) \approx N \log N$ for N large.

Lemma 1

$$C(G) < 4|V||E| = 4mN < 2N^3 \,.$$

Proof We start from a given node v_0 in V and construct a spanning tree by means of breadth first exploration of the graph[5]. The spanning tree constructed will have $N - 1$ edges[6]. We now perform a tour of the nodes of the tree visiting each edge of the tree at most once, which can be done by means of depth first

[1] If $(i, j) \in E$ then so is (j, i).

[2] there is a path using edges of the graph between any pair of nodes i and j.

[3] A graph is bipartite if there two disjoint subsets of V V_1 and V_2 such that $V_1 \cup V_2 = V$ and $E \in (V_1 \times V_2) \cup (V_2 \times V_1)$, i.e. edges are always between nodes in V_1 and nodes in V_2

[4] The is a unique edge between any pair of connected nodes and there are no links between a node and itself.

[5] We include all the neighbours of v_0 as its children in the tree and then we go through each of the neighbours of v_0 and explore their neighbours until we have explored all the nodes in the graph which is bound to happen as the graph is connected.

[6] There are always $N - 1$ edges in a tree with N nodes.

exploration of the tree. Let us denote $v_0, v_1, \ldots, v_{2n-2}$ the sequence of nodes visited, there are obviously repetitions in this sequence. Now it is clear that it takes longer to make the tour than to cover the graph. Hence

$$C(G) \leq \sum_{i=0}^{2N-3} h_{v_i,v_{i+1}} .$$

By lemma 1 we have that

$$C(G) \leq 2(N-1)2m < 4mN .$$

Note that $m \leq \frac{N(N-1)}{2}$ which yields the bound $C(G) < 2N^3$.
Now we are going to apply the above result to the $s-t$ connectivity problem. The aim is to decide whether there exists a path in the graph between two nodes s and t. This can be done using standard breadth-first or depth first search[7]. Here we instead use our previous analysis of the cover time to devise a randomised algorithm to solve the s-t connectivity problem.

Pseudocode: $s-t$ **connectivity algorithm**

1. Start a random walk from s
2. If the walk reaches t within $4N^3$ steps, return that there is path.
3. Else, return that there is no path.

Theorem 5 *The above algorithm returns the correct answer with probability* $1/2$*, and it only errs when returning that there is no path from* s *to* t *when there is one.*

Proof If there is no path then the algorithm always returns the correct answer. If there is path, then it is possible that the walk does no visit n in the first $4N^3$ steps. The expected time to reach s from t, if a path exists, is bounded above by the cover time, which is at most $2N^3$. By Markov's inequality, the probability that the walk takes more than $4N^3$ steps to reach t from s is at most $1/2$. We can make the error smaller by repeating the algorithms

7 Continuous-time Markov chains

In this section we define the equivalent of Markov chains described so far when time is continuous rather then being discrete. We consider families of random variable X_t where $t \geq 0$.

The introduction of continuous Markov is a bit more intricate the in discrete time. However once the theory is set up many of the results for discrete time Markov chain can be easily stated in the continuous tine setting. Let us start with some terminology illustrated through an example.

For S a countable (state) space we say that a matrix Q is a rate matrix if

(i) $0 \leq -q_{ii} < \infty$, for all i,

(ii) $q_{ij} \geq 0$, for $i \neq j$,

(iii) $\sum_{j \in S} q_{ij} = 0$, for all i.

For convenience we denote $q_i = \sum_{j \neq i} q_{ij} = -q_{ii}$. As we will see this corresponds to the transition matrix of a continuous-time Markov chain. Indeed, let $Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$ corresponds to the following dynamic:

[7] The use of each of these standard technique requires $\Omega(N)$ space where as our algorithm only requires $O(\log n)$ bits. Such analysis is beyond the scope of this course.

In state 3 you take two independent exponential timers $T_1 \sim \text{Exp}(2)$ and $T_2 \sim \text{Exp}(1)$; if T_1 is smaller the chain jumps to 1, otherwise it jumps to 2.

Let us define $P(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k = e^{tQ}$. Note that $P'(t) = P(t)Q$ and $P''(t) = P(t)Q^2$

We consider the case of a following process on $\{0, \ldots, N\}$ with rate matrix

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \ldots & 0 \\ 0 & -\lambda & \lambda & \ldots & 0 \\ 0 & \ldots & \ddots & \ddots & \lambda \\ 0 & \ldots & 0 & -\lambda & \lambda \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}.$$

It is not difficult to see that $p_{ij}(t) = 0$ for $i > j$ and that

$$\begin{aligned} p'_{ii}(t) &= -\lambda p_{ii}(t), \quad p_{ii}(0) = 1 \quad , i < N \\ p'_{ij}(t) &= -\lambda p_{ij}(t) + \lambda p_{i,j-1}(t), \quad p_{ij}(0) = 0 \quad , i < j < N \\ p'_{iN}(t) &= \lambda p_{i,N-1}(t), \quad p_{iN}(0) = 0 \quad , i < N\,. \end{aligned}$$

Definition 9 *A process $X = \{X_t, t \geq 0\}$ is a continuous time Markov chain if*

$$\mathbf{P}(X_{t_n} = s \mid X_{t_1} = x_1,\, X_{t_2} = x_2, \ldots, X_{t_{n-1}} = x_{n-1}) = \mathbf{P}(X_{t_n} = s \mid X_{t_{n-1}} = x_{n-1})\,,$$

for all $n \geq 1$ and $s, x_1, \ldots, x_{n-1} \in S$ and $t_1 < t_2 < \cdots < t_n$.

We define the transition matrix $P(s,t) = (p_{ij}(s,t))_{i,j}$ the matrix of transition probabilities

$$p_{ij}(s,t) = \mathbf{P}(X_t = j \mid X_s = i)\,.$$

The chain is called homogeneous if $p_{ij}(s,t) = p_{ij}(0, t-s)$. In this case we focus on $P(t) = P(0,t)$.

Example 18 *In the previous example with $Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$, one can show that*

$$p_{11}(t) = \frac{3}{8} + \frac{1}{4} e^{-2t} + \frac{3}{8} e^{-4t}\,.$$

In the remainder of this section we focus our attention on homogeneous chains. Similar to discrete-time Markov chains we have, for $s \leq u \leq t$,

$$p_{ij}(s,t) = \sum_k p_{ij}(s,u) p_{ij}(u,t)$$

More generally we have the Chapman-Kolmogorov equation

$$P(s+t) = e^{(s+t)Q} = e^{sQ} e^{tQ} = P(s)P(t),\ s \leq u \leq t\,.$$

In fact, for a small time dt we can consider

$$\frac{P(t+dt) - P(t)}{dt} = P(t) \frac{P(dt) - I}{dt}\,,$$

where I is the identity matrix. When $dt \to 0$, letting $Q = \lim_{dt \to 0} \left(\frac{P(dt) - I}{dt} \right)$ we have

$$\frac{dP(t)}{dt} = P(t)Q\,,$$

the matrix Q is known as the transition rate or generator and

$$q_{ii} = \lim_{dt\to 0} \frac{p_{ii}(t,t+dt)-1}{dt}, \quad q_{ij} = \lim_{dt\to 0} \frac{p_{ij}(t,t+dt)}{dt},\ i \neq j\,.$$

It is not difficult to see that $\sum_{j\in S} q_{ij} = 0$. Moreover we have that from $\frac{dP(t)}{dt} = P(t)Q$, the forward equation given by

$$p'_{ij}(t) = \sum_{k\in S} p_{ik}(t) q_{kj}$$

and from $\frac{dP(t)}{dt} = QP(t)$ the backward equation given by

$$p'_{ij}(t) = \sum_{k\in S} q_{ik} p_{kj}(t)$$

Alternatively, since $P(0) = I$, we have

$$P(t) = e^{tQ} = I + tQ + \frac{t^2}{2!}Q^2 + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k\,.$$

8 Poisson process

The first important process we are interested in is the Poisson process on the integers with transition rate

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & -\lambda & \lambda & \dots \\ 0 & \dots & 0 & \ddots & \ddots \end{pmatrix}.$$

Using the ideas from the introduction with N going to infinity, we have that $p_{ij}(t) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}$ for $i \leq j$ and $p_{ij}(t) = 0$ if $j < i$.

Theorem 6 *First, $s > 0$, $N_{t+s} - N_t$ independent of N_u for all $u \leq t$.*

If $N_0 = 0$ then N_t has the Poisson distribution with parameter λt.

Let $T_1,\ T_2,\ \dots$ the instants where the Process N is incremented by 1 and let $X_i = T_{i+1} - T_i$, $i \geq 0$ with $T_0 = 0$.

The sequence X_i is i.i.d. with common distribution exponential with parameter λ and the T_n is distributed according to a Gamma distributions with parameters (n,λ), i.e.

$$f_{T_n}(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda e^{-\lambda t},\ t \geq 0\,.$$

9 Jump Markov chains

Given a continuous-time Markov chain with rate matrix Q we can define a discrete-time Markov chain known as the jump chain with transition matrix P with

$$p_{ij} = \begin{cases} \frac{q_{ij}}{q_i}, & \text{if } j \neq i,\ q_i \neq 0, \\ 0, & \text{if } j \neq i,\ q_i = 0, \end{cases} \qquad p_{ii} = \begin{cases} 0, & \text{if } q_i \neq 0, \\ 1, & \text{if } q_i = 0, \end{cases}$$

In fact the chain stays in state i for an exponential distribution of parameter $q_i > 0$ and then jumps to other states according to the transition matrix P. If $q_i = 0$ then once the chain reaches i is stays there forever.

Example 19 *In the previous example with* $Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$, *we have* $P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 2/3 & 1/3 & 0 \end{pmatrix}$

As a consequence of this, assuming that $q_i > 0$ for all $i \in S$ and letting h_i^A be the hitting probability of $A \subset S$ starting in state i and H_i^A the corresponding mean hitting time then

Theorem 7 *We have*

$$h_i^A = 1, \; i \in A; \quad \sum_j q_{ij} h_j^A = 0, \; i \notin A$$

and

$$H_i^A = 0, \; i \in A; \quad 1 + \sum_{j \notin A} q_{ij} H_j^A = 0, \; i \notin A$$

10 Stationary distribution

The previous analogy enables us to generalise all the notions related to discrete-time Markov chain to the continuous-time ones. In particular we can define the stationary distribution of the Markov chain as follows

Definition 10 *The vector* $\pi = (\pi_0, \pi_1, \pi_2, \dots)$ *is a stationary distribution of the chain if* $\pi_i \geq 0$, $\sum_i \pi_i = 1$ *and*

$$\pi = \pi P(t) \, .$$

The last property is equivalent to

$$\pi Q = 0 \, .$$

Theorem 8 *If the Markov has a stationary distribution* π *then*

$$\lim_{t \o \infty} p_{ij}(t) = \pi_j$$

We are going to focus our attention on a family of continuous-time Markov chains known as birth and death processes

Definition 11 *A birth and death process is a continuous-time Markov process taking values in* $\{0, 1, 2, \dots\}$ *with transition probabilities*

$$\mathbf{P}\left(X(t+h) = m + n \mid X(t) = n\right) = \begin{cases} \lambda_n h, & \text{if } m = 1, \\ \mu_n h, & \text{if } m = -1, \\ 0, & \text{otherwise} \end{cases}$$

Its generator is then given by

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \lambda_3 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} .$$

The birth rates are given by $\lambda_0, \lambda_1, \dots$ *and the death rates are* $\mu_0, \mu_1, \dots$

Remark 4 *The Poisson process is a special case of birth and death process where* $\mu_1 = \mu_2 = \dots = 0$ *and* $\lambda_0 = \lambda_1 = \dots = \lambda$ *the rate of the Poisson process.*

If $\sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n} < \infty$, then the stationary distribution of a birth and death process is given by

$$\pi_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \ldots \mu_n} \pi_0,\ n \geq 1$$

where

$$\pi_0 = \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \ldots \mu_n} \right)^{-1} .$$

Letting $\lambda_n = \lambda$ and $\mu_n = n\mu$ we have

$$\lim_{k \to \infty} \mathbf{P}(X(t) = k) = \frac{\rho^n}{n!} e^{-\rho}$$

The stationary distribution of $X(t)$ is a Poisson distribution with parameter $\rho = \frac{\lambda}{\mu}$.

11 Exercises

Exercise 1 Consider two candidates in an election in which candidate A gets a votes while candidate B gets b votes ($b < a$). Votes are counted in random order chosen uniformly at random over the permutation on the $a+b$ votes. Let $n = a+b$, S_k=number of votes by which candidate A is ahead after counting the first k votes (S_k can be negative), $S_n = a-b$, and $X_k = \frac{S_{n-k}}{n-k}, k = 0, ..., n-1$.

1. Derive the number of votes for candidates A and B after counting $n-k+1$ votes in terms of S_{n-k+1}.
2. Show that $\mathbb{E}(S_{n-1}|S_{n-k+1}) = \frac{S_{n-k+1}(n-k)}{n-k+1}$.
3. Show that $\mathbb{E}(X_k|X_0, ..., X_{k-1}) = X_{k-1}$.
4. Let $T = \inf\{k \geq 0 | X_k = 0\}$ if such T exists and $T = n-1$ otherwise.
   (a) Show that $\mathbb{E}(X_k) = \mathbb{E}(X_0), \forall k < n$
   (b) Show that $\mathbb{E}(X_T) = \frac{a-b}{a+b}$.
5. (a) Assume that A leads throughout the count (i.e. $T = n-1$). Show that $X_T = X_{n-1} = S_1 = 1$.
   (b) Assume that A does not lead throughout the count, i.e. $\exists k < n-1, X_k = 0$. Explain why $T = k < n-1$ and $X_T = 0$.
   (c) Let E_1 denote the probability of the event in 5a. Show that

$$\mathbb{P}(E_1) = \frac{a-b}{a+b}$$

Exercise 2 Assume that random variables U and V are chosen independently and uniformly out of the set {1,2,3,4,5}. From this, we derive random variables X =min(U, V) and Y =max(U, V).

1. Determine the joint law of (U, V).
2. Determine $\mathbb{E}[U|Y = n]$, for $n \in \{1, 2, 3, 4, 5\}$.
3. Derive $\mathbb{E}[U|Y]$.
4. Derive $\mathbb{E}[Y|U]$.
5. Derive also $\mathbb{E}[U|X]$ and $\mathbb{E}[X|U]$.

Exercise 3 We consider two random variables X and Y such that the joint distribution of the random vector (X, Y) is given by, for $1 \leq j \leq i \leq 5$,

$$p_{ij} = \mathbb{P}(X = i, Y = j) = \frac{1}{15}$$

1. (a) Show that $(p_{ij})_{i,j\in\{1,...,5\}}$ is a probability distribution.
   (b) Compute the marginal distributions of X and Y.
   (c) For $1 \leq j \leq i \leq 5$, compute the conditional distribution.

$$\mathbb{P}(X = i|Y = j)$$

2. (a) Compute $\mathbb{E}(X|Y = j), j = 1, ..., 5$. Show that $\mathbb{E}(X|Y) = \frac{Y+5}{2}$.
   (b) Compute $\mathbb{E}(Y|X = i), i = 1, ..., 5$, and $\mathbb{E}(Y|X)$.

(c) Using the previous two questions show that

$$\mathbb{E}(X) = \frac{1}{2}\mathbb{E}(Y) + \frac{5}{2}$$

$$\mathbb{E}(Y) = \frac{1}{2}\mathbb{E}(X) + \frac{1}{2}$$

(d) Using the previous question, compute $\mathbb{E}(X)$ and $\mathbb{E}(Y)$.

3. By analogy to the arguement above, derive $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ for X and Y having joint distribution

$$\mathbb{P}(X = i, Y = j) = \frac{2}{n(n+1)}$$

where n is a positive integer and $1 \leq j \leq i \leq n$.

Exercise 4 Let X be a Markov chain. Which of the following are Markov chains?

1. $(X_{m+r}, r \geq 0)$.
2. $(X_{2m}, m \geq 0)$.
3. The sequences of pairs $((X_n, X_{n+1}), n \geq 0)$.

Exercise 5 A die is rolled repeatedly. Which of the following are Markov chains? For those that are, supply the transition matrix.

1. The largest number X_n shown up to the nth roll.
2. The number N_n of sixes in n rolls.
3. At time r, the time C_r since the most recent six.
4. At time r, the time B_r until the next six.

Exercise 6 Let X be a Markov chain and let $n_r : r \geq 0$ be an unbounded increasing sequence of positive integers. Show that $Y_r = X_{n_r}$ constitutes a (possibly inhomogeneous) Markov chain. Find the transition matrix of Y when $n_r = 2r$ and X is simple random walk.

Exercise 7 *Virus mutation.* Suppose that a virus can exist in N different strains and each genereation either stays the same, or with probability α mutates to another strain, chosen at random. What is the probability that the strain in the nth generation is the same as the initial strain?

Exercise 8 *Markov's other chain.* Let $Y_1, Y_3, Y_5, \dots$ a sequence of iid r.v. such that, for $k = 0, 1, \dots$

$$\mathbb{P}(Y_{2k+1} = -1) = \mathbb{P}(Y_{2k+1} = +1) = 1/2\,.$$

Let $Y_{2k} = Y_{2k-1}Y_{2k+1}$, for $k = 1, 2 \dots$. It is not difficult to see that $Y_2, Y_4, \dots$ is a sequence of iid r.v. with the same distribution as $Y_1, Y_3, Y_5, \dots$.

Is $Y_1, Y_2, Y_3, \dots$ a Markov chain?

Exercise 9 A faulty digital video confrencing system shows a clustered error pattern. If a bit is received correctly, then the chance to receive the next bit correctly is 0.999. If a bit is received incorrectly, then the next bit is incorrect with probability 0.95.

1. Model the error pattern of this system using the discrete time Markov chain.
2. How many communicating classes does the Markov chain have? Is it irreducible?
3. In the long run, what is the fraction of correctly received bits and the fraction of incorrectly received bits?
4. After the system is repaired , it works properly for 99.9% of the time. A test sequence after repair shows that, when always starting with a correctly received bit, the next 10 bits are correctly received with probability 0.9999. What is the probability now that a correctly (and analogously incorrectly) received bit is followed by another correct (incorrect) bit?

Exercise 10 We consider $(X_n)_n$ be a Markov chain over $\{1, \ldots, N\}$ whose transition matrix is partitioned as

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

A being a k by k matrix, B a k by $N-k$ matrix, B a $N-k$ by k matrix and D a $N-k$ by $N-k$ matrix, $0 < k < N$.

We are interested in the Markov chain $(\tilde{X}_n)_n$ for which only visits to the states $1, \ldots, k$ are recorded. It is not difficult to show that its transition matrix is

$$\tilde{P} = A + B \sum_{n \geq 0} D^n C = A + B(I - D)^{-1} C \,.$$

Now consider the following problem

A businesswoman spends hers time in London between (1) her office in the City, (2) her mansion in West Hampstead , (3) the Pacha restaurant in Gloucester road , and (4) with her lover. She moves from one to another according to a transition matrix P, though her husband, who knows nothing of the lover believes her movements are governed by the transition matrix P_W:

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \end{pmatrix}, \quad P_W = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

1. The public only sees the businesswoman when she is in (1), (2) or (3); calculate the transition matrix $\tilde{P}$ which they believe controls her movements.
2. Each time she moves to a new location (except when she meets her lover), she phones her husband. Write down the transition matrix that governs the sequence of locations from which she phones, and calculate its invariant distribution.
3. The husband notes her calls and becomes suspicious as she is not enough at home. Justify the husband's suspicion.
4. Confronted the businesswoman swears hers fidelity and resolves to choose a new transition matrix

$$P^* = \begin{pmatrix} 1/4 & 1/4 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 & 0 \\ 0 & 3/8 & 1/8 & 1/2 \\ 1/5 & 1/10 & 1/10 & 3/5 \end{pmatrix}$$

Will this deal with the husband's suspicions? Explain your answer.

Exercise 11 Consider a computer that has two identical and independent processors. The time between failures has an exponential distribution. The mean value of this distribution is 1000 hours. The repair time for a damaged processor is exponentially distributed as well, with a mean value of 100 hours. We assume damaged processors can be repaired in parallel. There are clearly three states for this computer: (1) both processors work, (2) one processor is damaged and (3) both processors are damaged.

1. Make a continuous time Markov chain presentation of these states.
2. What is the generator matrix **Q** for this Markov chain? Give the relation between the state probability at time t and its derivative.

   Hint: The time between failures has an exponential distribution. This means that the failure rate is $\lambda = 0.001$ *per hour. Similarly, the repair rate will be* $\mu = 0.01$ *per hour.*
3. Calculate the steady state (π) of this process. Comment on the availability of the computer if (i) both processors are required to work, or (ii) at least one processor should work.

Exercise 12

1. Let us assume that $(X_t)_{t\geq 0}$ is a given continuous-time Markov chain on the state space S with rate matrix $Q = (q_{ij})_{i,j\in S}$. Show that if there exists a probability distribution $\pi = (\pi_i)_{i\in S}$ such that, for all $i, j \in S$, we have
$$\pi_i q_{ij} = \pi_j q_{ji} \ ,$$
   then π is the invariant distribution of $(X_t)_{t\geq 0}$. The Markov chain is said to be *reversible.*

   Hint: Use the fact that if Q is a rate matrix then $\sum_{j\in S} q_{ij} = 0$, for all $i \in S$.
2. We consider the dynamics of a Markovian single server queue in continuous time. Customers join the queue and are served on a first-in-first-out basis, i.e., according to the order in which they join the queue. We suppose that the time between two successive arrivals is exponentially distributed with parameter $\lambda > 0$ and that each customer requires a service time that is exponentially distributed with parameter $\mu > 0$.

   Let X_t, $t \geq 0$, be the random process that describes the number of customers waiting in the queue including the one being served.

   (a) Derive the transition matrix Q of $(X_t)_{t\geq 0}$.

   (b) Using question 3. a), show that, if $\rho = \frac{\lambda}{\mu} < 1$, the stationary (invariant) distribution of $(X_t)_{t\geq 0}$ is given by
$$\pi_i = (1-\rho)\rho^i \ .$$
   Comment on the *stability condition* $\rho < 1$.

   (c) Compute the average number of customers in the queue in the stationary regime.
3. Let us now assume that we have a post office with two cashiers. The two queues at the cashiers run (separately) in parallel. Each of these queues operates following the dynamics in 3) b) with the same parameters λ and μ.

(a) Derive the average number of customers in the post office.

(b) A clever employee suggests merging the two queues so that customers arrive at rate 2λ wait until one of the two cashiers is available and is then served. Describe the underlying continuous-time Markov chain and discuss its stability.

(c) Using similar arguments as in 3. b), show that the average number of customers in the post office in stationary regime is given by

$$\frac{2\mu\lambda}{(\mu+\lambda)(\mu-\lambda)}\,.$$

(d) Is the suggestion of the employee better than running the two queues separately? Justify your answer.