UNIVERSITY OF COPENHAGEN

Lecture notes for Stochastic Processes 2011/2012

Tolver, Anders

Publication date:
2011

Document version
Peer reviewed version

Citation for published version (APA):
Tolver, A. (2011). *Lecture notes for Stochastic Processes 2011/2012*. Department of Natural Sciences, University of Copenhagen.

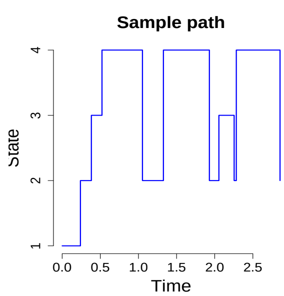
Lecture notes for Stochastic Processes 2011/2012

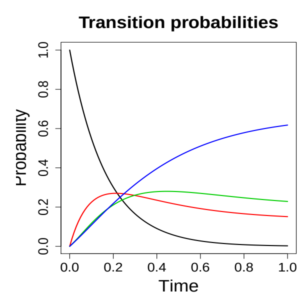
by Anders Tolver

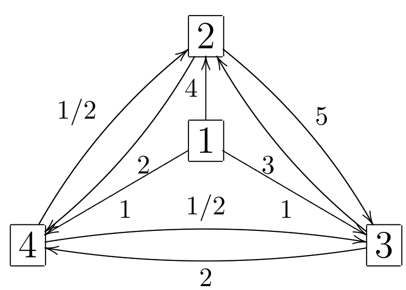
Intensity matrix Q

Invariant distribution $\overline{\pi}Q = 0$

Kolmogorov's diff. equation $P'(t) = QP(t)$







Transition diagram

Contents

Preface

The present material has been developed for the course *Stochastic Processes* at Department of Mathematical Sciences, University of Copenhagen during the teaching years 2010/2011 and 2011/2012. The topics covered are mainly Markov chains in discrete and continuous time on finite or countable state spaces.

The back bone of this work is the collection of exercises in Chapters 2 and 3. Hopefully, all of the theoretical results required to solve the exercises are contained in the first Section of Chapters 2 and 3. The manuscript was never intended to provide complete mathematical proofs of the main results since these may be found elsewhere in the literature. It is my intention to spend part of the lectures on (sketches of) proofs in order to illustrate how to work with Markov chains in a formally correct way. Some exercises in Chapter 4.2 are formulated as step-by-step instructions on how to construct formal proofs of selected theoretical results. It is definitely advisable (and probably necessary) to consult other textbooks on Markov chains in order to be able to solve *all* of the exercises given here. However, it is my strong belief that you should postpone the more theoretical exercises until you feel familiar with most of the exercises in Chapter 2 and 3. For further reading I can recommend selected Chapters of the books by Asmussen (2003), Brémaud (1999) and Lawler (2006). My own introduction to the topic was the lecture notes (in Danish) by Jacobsen & Keiding (1985).

A number of the exercises presented here are greatly inspired by examples in Ragner Nordberg's lecture notes on *Basic Life Insurance Mathematics* (Version: September 2002). The paragraphs presenting theoretical results on Markov chains are greatly inspired by various lecture notes by Jacobsen & Keiding (1985), M., by Nielsen, S. F., and by Jensen, S. T.

Part of this material has been used for the course *Stochastic Processes 2010/2011* at University of Copenhagen and I thank Massimiliano Tamborrino, Ketil Biering Tvermosegaard, and the students for many useful comments to this revised version.

Copenhagen, November 2011
Anders Tolver

Chapter 1

Introduction

1.1 About these lecture notes

The main purpose of developping the present manuscript has been to collect a number of problems providing an easy introduction to the most basic theory of Markov chains on finite or countable state spaces. The exposition differs from most textbooks on Markov chains in that the problems take up most of the space while only a limited number of pages are devoted to the presentation of the mathematical results. This reflects my personal point of view that you should learn the stuff by working with the problems.

The lecture notes are divided into two main parts: Chapter 2 deals with Markov chains in discrete time and Chapter 3 is about Markov chains in continuous time. Each chapter is subdivided into sections of which the first (i.e. Sections 2.0 and 3.0) contains a short summary of relevant definitions and theoretical results. The purpose of these sections is to give you an easy overview of the theory that you need to solve the problems. No proofs of the results are given in the lecture notes. To get the whole story behind the mathematical results you will need to consult other textbooks on Markov chains. During the lectures some time will be spend on mathematical proofs of selected results because this may help you building up your intuition about Markov chains. The problems in Chapter 4.2 explains how to construct mathematical proofs of certain theorems and might be used for some of the lectures.

The remaining sections of Chapter 2 and 3 present the problems divided into various subsections according to the size of the state space for the Markov chain. Whenever possible we have tried to put the exercises into a practical context if they deal with models that have reasonable interpretations in the real world. As the present course preceeds a course on life insurance a great number of the exercises are motivated by this particular application. However, other important examples from the vast area of applications of Markov chains have found their way to the present collection of problem, most notably from the field of queueing theory.

Clearly, the exercises vary in their difficulty and probably also in their relevance to a student just wanting to pass the course. We have made an effort to ensure that most exercises contains a mixture of (very) simple and more complicated questions. This is done because we know how frustrating it feels not to be able to get anywhere when trying to do the exercises at home. On the other hand this also implies that everyone should be able to prepare something before the exercise classes. If you show up at the exercise classes without having prepared any question for any exercise then the teaching assistant will not believe that you gave it a try. It is more likely that you will be seen as a lazy and unambitious student ;-) If I am mistaken on this point please let me know.

I want you to remember that you are supposed to do a written exam to pass this course. Therefore my general advice to you is to use one of the problems as offset anytime you work with the course. Use the questions in the problems to figure out what part of Chapters 2.0 or 3.0 that might be relevant to answer the question. Use the problems to find relevant pages from the slides used for the lectures. Do not expect it to work the other way around. It will probably take you a lot of time to read and understand the definitions and theorems of Chapters 2.0 or 3.0 and every little mathematical argument presented at the lectures. Even if you do manage to digest all the mathematics you will probably not find it straigh-forward to apply it to solve the problems. There is a huge discrepancy between reading (about) probability theory and being able to solve problems on probability theory. In my opinion this is the main reason why courses on Markov chains have a reputation for being very difficult.

Work with the problems if you want to do well at the exam!

Finally, we have decided to include a Chapter 5 that contains some mathematical tools that we think might be useful for solving the problems in these lecture notes.

1.2 Transition diagrams

We advocate for visualising the dynamics of a Markov chain whenever possible. This will be done using socalled *transition graph*s with *nodes* (or vertices) representing the states of the Markov chains and *edges* representing transitions.

For a discrete time Markov chain (at least on a finite state space) the dynamics of the chain is given by the *matrix of transitions probabilities*. On the graph the transition probabilities are given as labels to the arrow representing the individual transitions. Usually, we use the convention that an edge corresponding to a zero of the transition matrix need not be drawn on the graph. Remember that a discrete-time Markov chain need not jump to another new state at every time period. This is represented by a circular arrow on the transition diagram. As the transition probabilities for arrows pointing out from a state should always sum to one we will sometimes be a bit lazy and omit the arrows from a state pointing to itself putting our faith in the readers ability to add the remaining arrows to the transition diagram. To illustrate the points above complete and lazy examples of transition diagrams for the same three state discrete-time Markov chain are displayed below.

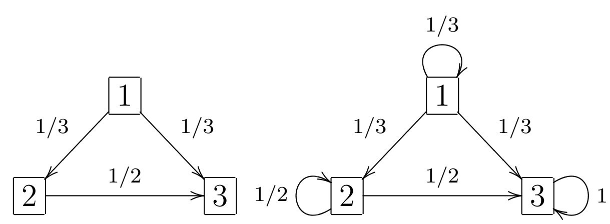


Figure 1.1: Two different versions of a transition diagram for the same discrete time Markov chain with three states.

For a continuous-time Markov chain the dynamics is given by the time spent in each state and the distribution of the jumps whenever they occur. For a finite state space Markov chain everything is summarized in the *transtion intensity matrix* with non-negative off diagonal entries and diagonals adjusted to make all rows sum to zero. The chain may be visualized by a transition diagram with nodes representing individual states and edges representing transitions. The correspondence between the transition intensity matrix and the transition diagram is obtained by labeling edges by the corresponding entry of the transition intensity matrix. In contrast to the discrete time case we (always!) omit edges of transition intensity zero. Further, there are no circular arrows from any state pointing to itself. An example of a transition diagram for a continuous-time Markov chain is given below.

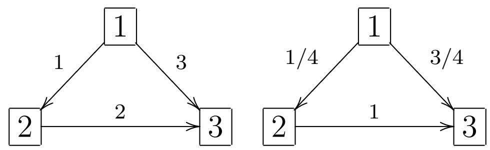


Figure 1.2: Transition diagram for a continuous time Markov chain with three states (left) and transition diagram for the corresponding discrete time Markov chain of jumps (right).

1.3 Overview of exercises

Below we have listed some important topics related to Markov chains and numbers of relevant problems dealing with the particular topic. Note that the list may incomplete, in particular if a problem deals with several topics.

Communication classes
Exercises: 2.2.1, 2.3.1, 2.3.7, 3.4.2

Transience or recurrence
Exercises: 2.2.1, 2.3.7

Null-recurrence or positive recurrence
Exercises: 2.4.2, 2.4.3, 2.4.4, 2.4.6, 3.2.4, 3.4.2, 3.4.4

Periodicity
Exercises: 2.3.2, 2.3.7

Absorption probabilities
Exercises: 2.3.7, 2.3.8, 3.2.4

Invariant distribution
Exercises: 2.2.2, 2.2.4, 2.3.3, 2.3.4, 2.3.7, 2.4.1, 3.2.1, 3.2.5, 3.3.4, 3.4.3

Recurrence (=return) time
Exercises: 2.1.2, 2.3.2, 2.4.6

Markov property
Exercises: 2.2.4, 2.3.6, 3.3.1

Kolmogorov's differential equation
Exerices: 3.2.2, 3.2.3, 3.3.3

Transition probabilities
Exercises: 2.1.4, 2.3.4, 3.1.3, 3.2.1

Poisson process
Exercises: 3.1.2, 3.2.4, 3.2.6, 3.4.1, 3.5.1, 3.5.2

Chapter 2

Discrete-time Markov chains

2.0 Results for discrete-time Markov chains

2.0.1 Definition of Markov chains

It all begins with a probability measure P. You may think of a probability measure, P, on a set Ω as a function assigning a number $P(A) \in [0,1]$ to subsets $A \subset \Omega$. If you are familiar with measure theory you may correctly insist that a probability measure only assigns a probability to subsets $A \in \mathcal{F}$ in a σ-algebra $\mathcal{F}$ on Ω but this point of view is not crucial for the story to come. Subsets of Ω are referred to as events. For two events A, B with $P(B) > 0$ we define the elementary conditional probability, $P(A|B)$, of A given B as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

A stochastic process in discrete-time is a family, $\{X(n)\}_{n \in \mathbb{N}_0}$, of random variables indexed by $\mathbb{N}_0$. The possible values, S, of $X(n)$ are referred to as the *state space* of the process. In this course we consider only stochastic processes with values in a finite or countable state space. The mathematician may then think of a random variable, X, on S as a measurable map

$$X : \quad (\Omega, \mathcal{F}) \to (S, \mathcal{P}(S))$$

where $\mathcal{P}(S)$ is the family of all subsets of S.

The distribution of a discrete-time stochastic process with at most countable state space, S, is characterized by the point probabilities

$$P(X(n) = i_n, X(n-1) = i_{n-1}, \ldots, X(0) = i_0)$$

for $i_n, i_{n-1}, \ldots, i_0 \in S$ and $n \in \mathbb{N}_0$. From the definition of elementary conditional probabilities it follows that

$$\begin{aligned}
& P(X(n) = i_n, \ldots, X(0) = i_0) \\
= \; & P(X(n) = i_n | X(n-1) = i_{n-1}, \ldots, X(0) = i_0) \\
\times \; & P(X(n-1) = i_{n-1} | X(n-2) = i_{n-2}, \ldots, X(0) = i_0) \\
\times \; & \ldots \\
\times \; & P(X(1) = i_1 | X(0) = i_0) \times P(X(0) = i_0).
\end{aligned}$$

This is a general identity that holds for any discrete-time stochastic process on a countable state space, but we are only going to consider the class of Markov chains.

A discrete-time Markov chain on a countable set, S, is a stochastic process satisfying the Markov property

$$
\begin{aligned}
& P(X(n) = i_n | X(n-1) = i_{n-1}, \dots, X(0) = i_0) \\
= \; & P(X(n) = i_n | X(n-1) = i_{n-1})
\end{aligned}
$$

for any $i_n, \dots, i_0 \in S$ and $n \in \mathbb{N}_0$. Introducing the notation

$$P_{i,j}(n-1) = P(X(n) = j | X(n-1) = i)$$

we immediately observe that for a Markov chain the formula for the point probabilities simplifies to

$$
\begin{aligned}
& P(X(n) = i_n, X(n-1) = i_{n-1}, \dots, X(0) = i_0) \\
= \; & P_{i_{n-1},i_n}(n-1) \cdot P_{i_{n-2},i_{n-1}}(n-2) \cdot \ldots \cdot P_{i_0,i_1}(0) \cdot P(X(0) = i_0).
\end{aligned}
$$

We shall make a final simplification by considering only time-homogeneous Markov chains for which the transition probabilities $P_{i,j}(n) = P_{i,j}$ do not depend on the time index $n \in \mathbb{N}_0$. For a discrete-time and time-homogeneous Markov chain on S we thus have that

$$P(X(n) = i_n, \dots, X(0) = i_0) = P_{i_{n-1},i_n} \cdot \ldots \cdot P_{i_0,i_1} \cdot \phi(i_0) \qquad (2.1)$$

where we use the notation

$$\phi(i_0) = P(X(0) = i_0)$$

for the initial distribution of $X(0)$.

Definition 1 (Homogeneous Markov chain in discrete-time) *A time-homogeneous Markov chain on a finite or countable set S is a family of random variables, $\{X(n)\}_{n \in \mathbb{N}_0}$, on a probability space $(\Omega, \mathcal{F}, P)$ such that*

$$P(X(n+1) = j | X(n) = i, X(n-1) = i_{n-1}, \dots, X(0) = i_0) = P_{i,j}$$

for $j, i, i_{n-1}, \dots, i_0 \in S$ and $n \in \mathbb{N}_0$. The distribution of the Markov chain is uniquely determined by the initial distribution and the transition probabilities

$$
\begin{aligned}
\phi(i) &= P(X(0) = i) \quad \leftarrow \textit{initial distribution} \\
P_{i,j} &= P(X(n+1) = j | X(n) = i) \quad \leftarrow \textit{transition probabilities.}
\end{aligned}
$$

□

From a practical point of view any probability vector $\overline{\phi} = (\phi(i))_{i \in S}$ and array of probabilities $P = (P_{i,j})_{i,j \in S}$ with $\sum_{j \in S} P_{i,j} = 1$ for all $i \in S$ defines the distribution of a time-homogeneous Markov chain on S through the identity (2.1). When the state space is finite we speak of the transition matrix $P = (P_{i,j})_{i,j \in S}$. As we will consider only time-homogeneous Markov chains we will throughout these lecture notes omit the phrase *time-homogeneous* by referring to the process simply as a Markov chain.

The dynamics of a discrete-time Markov chain with state space S is given by the matrix, P, of transition probabilities. A similar representation is given by a directed graph (*the transition diagram*) with nodes representing the individual states of the chain and directed edges labeled by the probability of possible transitions.

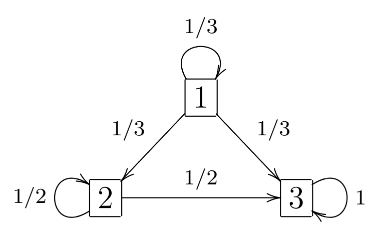


Figure 2.1: Transition diagram for a discrete-time Markov chain with three states.

The probability of any event involving the observations $X(0), \ldots, X(n)$ from a Markov chain may be obtained by splitting the event into disjoints sets of the form

$$(X(0) = i_0, \ldots, X(n) = i_n)$$

and summing up point probabilities of the form given by (2.1). For finite state space Markov chains the computation of the probability of certain events have simple representations in terms of matrix multiplication.

Theorem 1 (n-step transition probabilities) *For a Markov chain on a finite state space, $S = \{1, \ldots, N\}$, with transition probability matrix P and initial distribution $\overline{\phi} = (\phi(1), \ldots, \phi(N))$ (row vector) then the distribution of $X(n)$ is given by*

$$(P(X(n) = 1), \ldots, P(X(n) = N)) = \phi P^n.$$

□

2.0.2 Classification of states

For a discrete-time Markov chain on S with transition probabilities $P = (P_{i,j})_{i,j \in S}$ we say that there is a possible path from state i to state j if there is a sequence of states $i_0, i_1, \ldots, i_n \in S$ with $i = i_0, j = i_n$ and all $P_{i_{l-1},i_l} > 0$, $l = 1, \ldots, n$. We say that two states $i, j \in S$ communicate if there is a possible path from i to j and from j to i. We use the notation $i \leftrightarrow j$ when the two states i and j communicate. If we use the convention that $i \leftrightarrow i$ then the relation $\leftrightarrow$ partitions the state space, S, into disjoint *communication classes*. A Markov chain is said to be *irreducible* if there is only one communication class.

To understand and describe the dynamics and the long-run behaviour of a Markov chain we introduce the concepts of a *transient* and a *recurrent* communication class. For any state $i \in S$ we define the return time or the recurrence time to i by

$$T_i = \inf\{n > 0 | X(n) = i\}.$$

Definition 2 (Recurrence and transience) *For a discrete-time Markov chain on S we say that a state $i \in S$ is recurrent if and only if*

$$P(T_i < +\infty | X(0) = i) = 1.$$

If $P(T_i < +\infty | X(0) = i) < 1$ then i is said to be a transient state. □

The interpretation of recurrence is very important: if the Markov chain is started in a recurrent state i then with probability 1 it will eventually return to state i. On the contrary, if a Markov chain starts in a transient state i then the probability of returning to state i is strictly less than 1.

It is very important to emphasize the difference between a Markov chain being irreducible and being recurrent. For an irreducible Markov chain there is a positive probability of a transition between any two states $i \neq j$ while (as we shall see) for a recurrent Markov chain there will eventually be a transition between any two states $i \neq j$ with probability 1!

As stated in the following two results it is possible to extract a bit more information from the definition of recurrence.

Theorem 2 (Number of visits to state i) *For a discrete-time Markov chain on S with initial distribution $P(X(0) = i) = 1$ consider the total number of visits to state i*

$$N_i = \sum_{n=1}^{\infty} 1(X(n) = i).$$

If i is a recurrent state then $N_i = +\infty$ with probability 1. If i is a transient state then N_i follows a geometric distribution

$$P(N_i = k) = (1-q)^k q, \quad k \in \mathbb{N}_0$$

where $q = P(T_i = +\infty | X(0) = i)$. □

Theorem 3 (Recurrence is a class property) *All states in a communication class are either all recurrent or all transient.* □

The result above suggests that we first find the communication classes and then consider only one representative of each class for classification into either recurrent or transient states. Unfortunately, it is rarely possible to apply directly Definition 2 to verify that a state is recurrent. The following two results may be more useful for practical purposes.

Theorem 4 (Recurrence criterion 1) *For a discrete-time Markov chain with n-step transition probabilities $P^n = (P^n)_{i,j \in S}$ then state i is recurrent if and only if $\sum_{n=1}^{\infty} (P^n)_{i,i} = +\infty$.* □

In many cases it is not possible to get explicit formulaes for the n-step transition probability, $(P^n)_{i,i}$, allowing us to evaluate the sum from Theorem 4. However, it might be possible to bound $(P^n)_{i,i}$ from below (or above) by elements of a divergent (or convergent) series $\sum_{n=1}^{\infty} a_n$ allowing us to deduce that state i is recurrent (or transient).

The following result gives another way to demonstrate recurrence of a state without reference to n-step transition probabilities.

Theorem 5 (Recurrence criterion 2) *Let $\{X(n)\}_{n \geq 0}$ be an irreducible Markov chain on S with transition proabability $P = (P_{i,j})_{i,j \in S}$. Consider the system of equations*

$$\alpha(j) = \sum_{k \neq i} P_{j,k} \alpha(k), \quad j \in S, j \neq i, \tag{2.2}$$

where $i \in S$ is a fixed (but arbitrary) state.

The Markov chain is recurrent if and only if the only bounded solution to (2.2) is given by $\alpha(j) = 0, j \neq i$. □

Remark 1 *Let us try to explore a little further the result of Theorem 5. For fixed (but arbitrary) $i \in S$ define*

$$\alpha(j) = P(T_i = +\infty | X(0) = j) \quad \textit{where} \quad T_i = \inf\{n \geq 0 | X(n) = i\}.$$

Note that $\alpha(j)$ is the probability of never visiting state i given that the Markov chain starts in state j. Then it is straight-forward to show that $(\alpha(j))_{j \in S, j \neq i}$ is a bounded solution to (2.2). One deduces that if $\alpha(j) = 0$ is the only bounded solution then we must have for all $j \neq i$ that

$$P(T_i < \infty | X(0) = j) = 1.$$

In particular, state i is recurrent since we have by the Markov property that

$$\begin{aligned} P(T_i < \infty | X(0) = i) &= P_{i,i} + \sum_{j \neq i} P_{i,j} \cdot P(T_i < \infty | X(0) = j) \\ &= P_{i,i} + \sum_{j \neq i} P_{i,j} \cdot 1 = 1! \end{aligned}$$

The non-trivial part of Theorem 5 is show that if any non-zero solution to (2.2) exists then we must also have $P(T_i = +\infty | X(0) = j) > 0$ for at least one $j \neq i$ implying (trivally) that the irreducible Markov chain is not recurrent. □

For communication classes with only finitely many elements things are considerably easier as explained in the following result.

Theorem 6 (Closed communication classes) *A communication class, $\mathcal{C}$, is said to be closed if the submatrix of transition probabilities restricted to $\mathcal{C}$ has all row sums equal to* 1.

A finite communication class, $\mathcal{C}$, is recurrent if and only if it is closed. In general, closed communication classes with a countable number of states can be either recurrent or transient.

The restriction of a Markov chain to a closed communication class is an irreducible Markov chain. □

In order to discuss the long-run behavior of a Markov chain we need to introduce the period of a state. The period of a state $i \in S$ is the greatest common divisor of the length of all possible loops starting and ending in state i.

Definition 3 (Period of a Markov chain) *For a discrete-time Markov chain on S a loop of length n is a sequence of states $i_0, i_1, \ldots, i_n \in S$ with $i_0 = i_n$. We will speak of a possible loop if*

$$P_{i_0,i_1} \cdot P_{i_1,i_2} \cdot \ldots \cdot P_{i_{n-1},i_n} > 0.$$

Introduce

$$D_i = \{n \in \mathbb{N} | \text{ there exists a possible loop of length } n \text{ with } i_0 = i_n = i\}$$

and define the period of state i as the largest number dividing all numbers in the set D_i. All states of a communication class will have the same period and we shall use the phrase aperiodic about a class of period 1. An irreducible, aperiodic Markov chain is a Markov-chain with one communication class of period 1. □

2.0.3 Invariant distributions and absorption

We have previously considered the return time

$$T_i = \inf\{n > 0 | X(n) = i\}$$

to state i. For a transient state i there is (by definition!) a positive probability that the chain never returns to state i

$$P(T_i = +\infty | X(0) = i) > 0$$

hence trivally $E[T_i | X(0) = i] = +\infty$. If i belongs to a recurrent communication class then we know that

$$P(T_i < +\infty | X(0) = i) = 1$$

that is we are sure to get back to state i. We may, however, also consider the mean return time

$$E[T_i | X(0) = i]$$

which may or may not be finite.

Definition 4 (Positive recurrence and null-recurrence) *A recurrent state i is said to be positive recurrent if and only if the mean return time to state i is finite*

$$E[T_i|X(0) = i] < +\infty.$$

Otherwise the recurrent state is said to be null-recurrent. It can be shown that all states belonging to the same recurrent class are either positive recurrent or null-recurrent. □

The main results concerning the long-run behaviour of a discrete-time Markov chain are formulated below. You may think that the exposition here differs from the literature on Markov chains. However, we believe that the present formulation is suitable for students who should learn to apply the results for solving problems. Clearly, we would use another exposition if the intention was to present the mathematical theory underlying the results.

Theorem 7 *For an irreducible, aperiodic discrete-time Markov chain then for any state i and any initial distribution, $\overline{\phi}$, it holds that*

$$\lim_{n\to\infty} P(X(n) = i) = \tfrac{1}{E[T_i|X(0)=i]}.$$

If $E[T_i|X(0) = i] = +\infty$ the limit on the right hand side is defined to be 0.
□

Note that by choosing the initial distribution $P(X(0) = i) = 1$ we conclude from Theorem 7 that the n-step transition probabilities

$$(P^n)_{i,j} = P(X(n) = j|X(0) = i)$$

have a limit as $n \to \infty$. Using the identity

$$\begin{aligned}
& P(X(n+1) = j|X(0) = i) \\
= & \sum_{l\in S} P(X(n+1) = j, X(n) = l|X(0) = i) \\
= & \sum_{l\in S} P(X(n+1) = j|X(n) = l) \cdot P(X(n) = l|X(0) = i) \\
= & \sum_{l\in S} P(X(n) = l|X(0) = i) \cdot P_{l,j}
\end{aligned}$$

it follows by formal mathematical arguments (dominated convergence!) that

$$\begin{aligned}\pi(j) &= \lim_{n\to\infty} P(X(n+1)=j|X(0)=i) \\ &= \sum_{l\in S}\left\{\lim_{n\to\infty} P(X(n)=l|X(0)=i)\right\}\cdot P_{l,j} = \sum_{l\in S}\pi(l)P_{l,j}. \quad (2.3)\end{aligned}$$

A non-negative vector, $\overline{\pi} = (\pi(j))_{j\in S}$, solving the system of equations (2.3) for all $j \in S$ is called an invariant measure for the transition probabilities, $P = (P_{i,j})_{i,j\in S}$. If $\overline{\pi}$ is a probability (i.e. $\pi(j) \geq 0, \sum_{j\in S}\pi(j) = 1$) we will speak of an *invariant distribution* for P.

The following result states that an invariant probability exists exactly if the irreducible Markov chain is positive recurrent.

Theorem 8 *For an irreducible, recurrent Markov chain, $\{X(n)\}_{n\geq 0}$, there is a unique (up to multiplication!) invariant measure solving the equations*

$$\nu(j) = \sum_{i\in S}\nu(i)P_{i,j}, \quad j \in S. \quad (2.4)$$

The unique solution (up to mulitiplication) is given by

$$\nu(j) = \mathbb{E}\left[\sum_{n=0}^{T_i-1} 1(X(n)=j)|X(0)=i\right]$$

where $i \in S$ is any fixed state. The solution can be normalized into an invariant probability if and only if $\mathbb{E}[T_i|X(0)=i] < +\infty$. □

Corollary 1 *From Theorem 7 and 8 above we conclude that for an irreducible Markov chain there is an invariant distribution solving (2.4) if and only if the Markov chain is positive recurrent. The invariant distribution is given by the inverse of the mean return times*

$$\pi(i) = \frac{1}{\mathbb{E}[T_i|X(0)=i]}.$$

If the Markov chain is aperiodic it further holds that

$$\lim_{n\to\infty} P(X(n)=j) = \pi(j)$$

for any initial distribution $\overline{\phi} = (\phi(i))_{i \in S}$.

If we use the invariant distribution as initial distribution (i.e. $P(X(0) = j) = \pi(j)$*) then it holds that*

$$P(X(n) = j) = \pi(j)$$

for all $n \geq 0$*, and we will say that the Markov chain is stationary. For that reason we will also refer to* $\overline{\pi}$ *as the stationary distribution.* □

Remark 2 (Positive recurrence and null-recurrence) *It follows from Corollary 1 that a recurrent communication class is positive recurrent if and only if there exists an invariant distribution. If there does not exist an invariant distribution on a recurrent class then the class is null-recurrent.*

One can show that only communication classes with a countable number of states can be null-recurrent. All recurrent communication classes with a finite number of elements are positive recurrent. □

The results above are formulated for irreducible Markov chains and the main focus has been on aperiodic communication classes. We end this section by discussing what happens if we relax these two assumptions.

Corollary 2 (Irreducible, periodic Markov chains) *For an irreducible Markov chain with period* $d > 1$ *the limit* $\lim_{n\to\infty} P(X(n) = i)$ *does not exist for an arbitrary initial distribution. However, the average over a period of length* d *has a limit*

$$\pi(i) := \lim_{n\to\infty} \frac{P(X(n) = i) + P(X(n+1) = i) + \ldots + P(X(n+d-1) = i)}{d}.$$

If this limit is a probability distribution (i.e. if $\sum_{i\in S} \pi(i) = 1$*) then* $\overline{\pi} = (\pi(i))_{i\in S}$ *is a unique invariant distribution for the Markov chain.* □

If the Markov chain is not irreducible then one must apply Corollaries 1 and 2 to each communication class. For each of the positive recurrent classes there exists a unique invariant distribution with positive probabilities only for the states in the class. However, any convex combination of the invariant distribution for the positive recurrent subclasses constitutes an invariant probability on the entire state space of the Markov chain. In particular, the invariant probability distribution is no longer unique.

Recurrent classes are closed in the sense that once the Markov chain enters the class it stays there forever. On the contrary, a Markov chain may leave a transient class sooner or later by entering a recurrent class. If a Markov chain is not irreducible it may consist of several transient and recurrent classes. This naturally raises the following questions: if the Markov chain is started in a transient state i, how many times will it visit state i before it leaves the state for good, and what is the probability that it will end up in each of the recurrent classes?

Theorem 9 (Absorption probabilities - finite state space) *Consider a finite state Markov chain with transition matrix P. Suppose that the states are ordered such that P can be decomposed as a block matrix*

$$P = \left(\begin{array}{c|c} \tilde{P} & 0 \\ \hline S & Q \end{array} \right)$$

where $\tilde{P}$ is the transition matrix restricted to the recurrent states. Similarly, Q is the submatrix of P restricted to the transient states, and S describes transition probabilities from transient to recurrent states. The 0 block in the upper right part of P reflects the fact that transitions from recurrent to transients states are not possible.

The ij-th entry of the matrix $M = (I - Q)^{-1}$ describes the excepted number of visits to the transient state j before the Markov chain reaches one of the recurrent states under the assumption that the Markov chain starts in the transient state i (i.e. $P(X(0) = i) = 1$). Here, I denotes the identity matrix with zero off-diagonal and a diagonal of ones.

The ij-th entry of

$$A = (I - Q)^{-1}S$$

is the probability that j is the first recurrent state reached by the Markov chain when started in the transient state i (i.e. $P(X(0) = i) = 1$).

□

A common application of Theorem 9 is the case where all recurrent classes are sets with only one element. In this situation we have $\tilde{P} = I$ and the recurrent states are simply the absorbing states. The result then gives us the absorption probabilities for each of the absorbing states.

If the Markov chain contains more than one recurrent class then Theorem 9 may be used to compute the probability that the Markov chain will end its life in each of the recurrent classes. Note that in the long-run the total probability of being absorbed in a particular recurrent class will be redistributed on individual states according to the invariant distribution restricted to the relevant class.

Theorem 9 is stated in terms of matrix operations and are therefore restricted to Markov chains on a finite state space. For Markov chains on a countable state space absorption probabilities may be found by solving a countably infinite system of equations.

Theorem 10 (Absorption probabilities - general case) *For a Markov chain on S let $\mathcal{C}$ be a recurrent class. The probabilities*

$$\alpha(j) = P(X(n) \in C \text{ for some } n \geq 0 | X(0) = j), \quad j \in \mathcal{C}' \leftarrow \text{ transient states}$$

that the chain will ever visit class $\mathcal{C}$ (and stay there forever) then solves the system of equations

$$\alpha(j) = \sum_{l \in \mathcal{C}'} P_{j,l} \alpha(l) + \sum_{l \in \mathcal{C}} P_{j,l}. \tag{2.5}$$

The absorption probability $(\alpha(j))_{j \in \mathcal{C}'}$ is the smallest non-negative solution to (2.5). There is a unique bounded solution to (2.5) if and only if there is zero probability that the Markov chains stays in the transient states forever. □

2.1 Markov chains with two states

2.1.1 General two state Markov chain

We consider a Markov chain on $S = \{1, 2\}$ with transition diagram

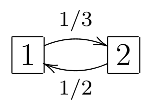


Figure 2.2: Transition diagram for Markov chain in Exercise 2.1.1 and 2.1.2.

Assume that $X(0) = 1$.

1. Write down the transition matrix P of the Markov chain.
2. Compute $P(X(1) = 1)$ and $P(X(1) = 2)$.
3. Find the distribution of $X(2)$ and $X(3)$.
4. Compute P^2 and P^3 and compare with the results of questions 2.-3.
5. Compute also P^{10}.

Assume in the following that the initial distribution of $X(0)$ is given by $\phi(1) = \phi(2) = 1/2$.

6. Compute the distribution of $X(1)$.
7. Let $\overline{\phi} = (\phi(1), \phi(2))$ and compute $\overline{\phi}P, \overline{\phi}P^2$, and $\overline{\phi}P^3$. What did you actually compute?
8. Find the distribution of $X(5)$.
9. Find the invariant probability vector $\overline{\pi} = (\pi_1, \pi_2)$ of the Markov chain by solving the matrix equation $\overline{\pi}P = \overline{\pi}$ that may be written out as
$$\pi_1 P_{11} + \pi_2 P_{21} = \pi_1 \quad \text{and} \quad \pi_1 P_{12} + \pi_2 P_{22} = \pi_2.$$
10. Compare the results of questions 5., 8., and 9.

2.1.2 Recurrence times

We consider again the Markov chain of Exercise 2.1.1 given by Figure 2.2. Assume that $X(0) = 1$ and define the *recurrence time* to state 1 by

$$T_1 = \inf\{n > 0 | X(n) = 1\}.$$

The purpose of this exercise is to study the distribution of the recurrence time and its relation to the invariant probability vector of the Markov chain.

1. Find $P(T_1 = 1)$.
2. Compute $P(T_1 = 2), P(T_1 = 3)$ and find the general expression for $P(T_1 = n)$, $n \geq 2$.
3. Find the mean recurrence time $\mu_1 = \mathbb{E}[T_1]$ to state 1.

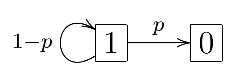
Assume that $X(0) = 2$ and define the *recurrence time* to state 2 by

$$T_2 = \inf\{n > 0 | X(n) = 2\}.$$

4. Compute $P(T_2 = n), n \geq n$ and the mean $\mu_2 = \mathbb{E}[T_2]$.
5. Find $1/\mu_1$ and $1/\mu_2$.
6. Compare the results of question 5. with the invariant probability vector $\overline{\pi}$ found in question 9. of exercise 2.1.1.

2.1.3 Two state absorbing Markov chain

We consider in this exercise the two-state Markov chain, $\{X(n)\}_{n \geq 0}$, given by the transition diagram



When discussing the Markov chain further we refer to the states through the following recoding: *1=alive,0=dead*.

1. Write down the transition matrix for the Markov chain.

2. Assuming that $X(0) = \textit{alive}$ find the probabilities $P(X(n) = \textit{alive})$ for $n \geq 1$.

Define the survival time, T, as the time of absorption in the state *dead*

$$T = \inf\{n > 0 | X(n) = \textit{dead}\}.$$

3. Argue that $P(T \leq n) = P(X(n) = \textit{dead})$.
4. Find the distribution of T, i.e. $P(T = n)$ for $n \geq 1$. What is the name of the distribution of T?
5. Compute the expected survival time $\mathbb{E}[T]$.

2.1.4 Transition probabilities for the two-state chain

Consider the general two-state Markov chain given by transition matrix

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

for $p, q \in [0, 1]$. The purpose of the exercise is to derive closed form expressions for the n step transition probabilities given by the matrix P^n.

1. Draw the transition diagram for the Markov chain.
2. Compute the characteristic polynomial for P given by

$$g(\lambda) = \det(P - \lambda I)$$

where I is the 2×2 identity matrix.

3. Argue that the equation $g(\lambda) = 0$ has two solutions

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_2 = 1 - p - q.$$

4. Find a (left) eigenvector $\overline{v} = (v_1, v_2)$ for P associated with the eigenvalue λ_2.
5. Show that $\overline{u} = (u_1, u_2) = (q, p)$ is an eigenvector for P with eigenvalue 1.

6. Verify that the matrix

$$O = \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix}$$

satisfies the matrix equation

$$OP = \underbrace{\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}}_{=D} O.$$

7. Find the inverse matrix O^{-1}.

8. Use that $P = O^{-1}DO$ to find a closed form expression for P^n and discuss the result.

2.2 Markov chains with three states

The general three-state Markov chain has transition matrix

$$P = \begin{pmatrix} P_{1,1} & P_{1,2} & P_{1,3} \\ P_{2,1} & P_{2,2} & P_{2,3} \\ P_{3,1} & P_{3,2} & P_{3,3} \end{pmatrix}$$

corresponding to the transition diagram on Figure 2.3

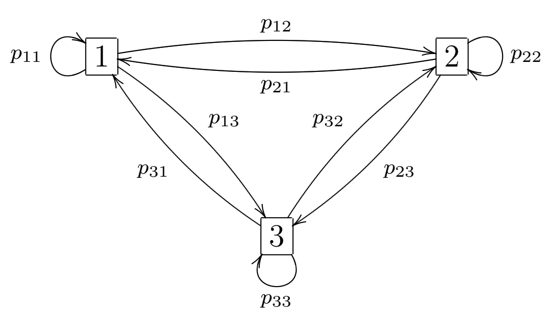


Figure 2.3: Transition diagram of a general three state Markov chain.

2.2.1 Classification of states

Consider a general three-state Markov chain as given by the transition diagram of Figure 2.3.

1. Argue that the chain is irreducible if $P_{i,j} > 0$, for all $i \neq j$.

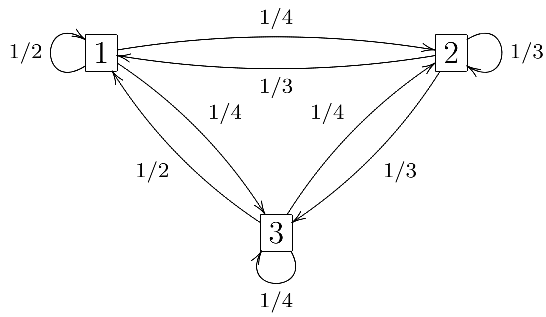
2. Give examples of irreducible three-state Markov chains for which $P_{i,j} = 0$ for at least one pair (i, j) of states.

3. Give examples of a three-state Markov chain with two communication classes.

4. Describe the relation between zero entries of the transition matrix P and the communication classes of the Markov chain. In each case determine if the communication classes are transient or recurrent.

2.2.2 General three state Markov chain

Consider a Markov chain given by the transition diagram



Assume that $X(0) = 1$ and let

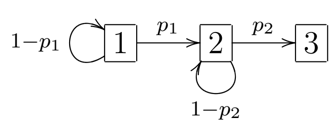
$$\tau = \inf\{n > 0 | X(n) \neq 1\}$$

be the time of the first jump away from state 1.

1. Find the transition matrix, P, of the chain.
2. Compute $P(X(1) = 1)$ and $P(X(2) = 1)$.
3. Use P^3, P^4, P^5 to find $P(X(n) = 1)$ for $n = 3, 4, 5$.
4. Find $P(\tau = 1), P(\tau = 2)$, and $P(\tau = 3)$. What is the name of the distribution of $\tau - 1$?
5. Write down the system of equations for the invariant distribution $\overline{\pi}$ and find $\overline{\pi}$.

2.2.3 The one-way Markov chain

We consider in this exercise a Markov chain given by transition diagram



Assume that $X(0) = 1$.

1. Find the probabilities $P(X(1) = j)$ for $j = 1, 2, 3$.
2. Find the probabilities $P(X(2) = j)$ for $j = 1, 2, 3$.

Denote by

$$\tau_1 := \inf\{n > 0 | X(n) = 2\}$$

the time of the jump between states 1 and 2. Similarly let

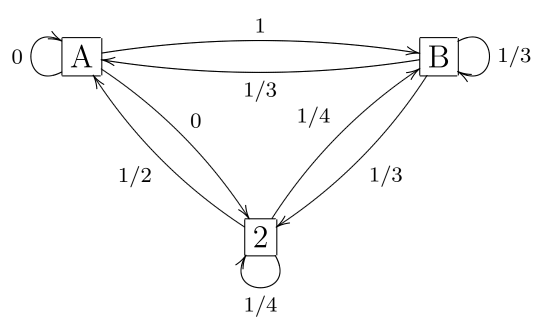
$$\tau_2 = \inf\{n > \tau_1 | X(n) = 3\}$$

between the time of the jump from state 2 to state 3.

3. Find the probabilities $P(\tau_1 = k)$ for $k \geq 1$. What is the name of the distribution of τ_1?
4. Find the probabilities $P(\tau_2 = k)$ for $k = 1, 2, 3$.
5. Try to find the general formula for $P(\tau_2 = k)$ for $k \geq 1$.
6. Assuming that $p_1 = p_2$ verify that $\tau_2 - 2$ follows a negative binomial distribution.

2.2.4 Markov property under aggregation of states

Consider a Markov chain given by the transition diagram



1. Find the transition matrix, P, for the Markov chain.
2. Write down the system of equations for the invariant distribution $\overline{\pi}$ and find $\overline{\pi}$.

Assume that the initial distribution is given by

$$P(X(0) = A) = P(X(0) = B) = P(X(0) = 2) = 1/3.$$

Use the transition matrix, P, of the chain to compute the following probabilities

3. $P(X(1) = A, X(0) = i)$ for $i \in \{A, B, 2\}$.
4. $P(X(1) = B, X(0) = i)$ for $i \in \{A, B, 2\}$.
5. $P(X(2) = 2, X(1) = A, X(0) = i)$ for $i \in \{A, B, 2\}$.
6. $P(X(2) = 2, X(1) = B, X(0) = i)$ for $i \in \{A, B, 2\}$.

Suppose that for some reason we are not able to distinguish between states A and B such that we only observe the process defined by

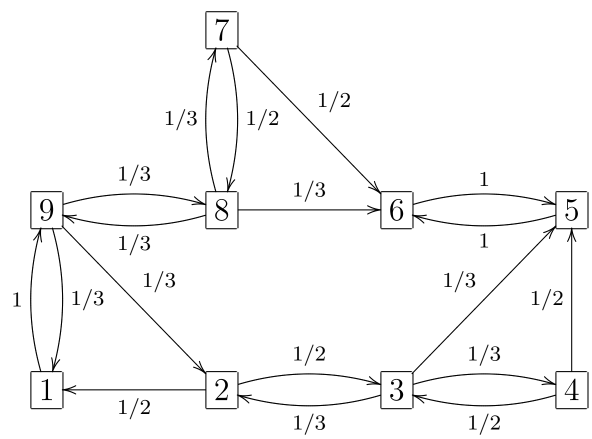
$$Y(n) = \begin{cases} 2 & , X(n) = 2 \\ 1 & , X(n) \in \{A, B\} \end{cases}$$

with state space $S = \{1, 2\}$

7. Use questions 3.-6. to compute $P(Y(2) = 2|Y(1) = 1, Y(0) = 1)$.
8. Use questions 3.-6. to compute $P(Y(2) = 2|Y(1) = 1, Y(0) = 2)$.
9. Argue that $\{Y(n)\}$ is not a Markov chain.
10. Show by an example that for certain choices of the transition probabilities for $\{X(n)\}_{n\geq 0}$ it holds that $\{Y(n)\}_{n\geq 0}$ is a Markov chain on $S = \{1, 2\}$.

2.3 Markov chains with finite state space

2.3.1 Find the communication classes



1. Argue that 7 and 8 belong to the same communication class.

2. Show that $P^2_{2,9} > 0$ and argue that 2 and 9 belong to the same communication class.

3. Find out if states 3 and 7 communicate.

4. Determine the communication class containing state 5.

5. Find all communication classes and determine if each class is recurrent or transient.

6. Is the chain irreducible?

The *loop trick* is a useful observation to speed up the process of determining the communication classes of a Markov chain. The basic observation is that if we can find a closed path of states

$$i_0 \to i_1 \to \ldots \to i_k \to i_0$$

such that all transition probabilities along the path are positive then all states in the path belong to the same communication class.

7. Use the loop trick to find the communication classes of the Markov chain.

8. Argue that there does exist an invariant probability vector, $\overline{\pi}$, for the chain and find it.

9. Suppose that we change the transition probabilities such that

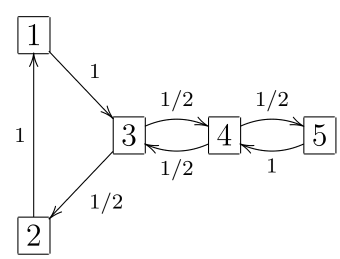
$$P_{7,6} = 0, P_{7,8} = 1, P_{2,3} = 0, P_{2,1} = 1, P_{8,7} = 1/2, P_{8,9} = 1/2, P_{8,6} = 0.$$

   Show that the modified version of the Markov chain has two recurrent subclasses.

10. Find an invariant probability vector, $\overline{\pi}$, for the Markov chain described in question 9. and discuss if $\overline{\pi}$ is uniquely determined.

2.3.2 A numeric example

Consider the Markov chain given by the transition diagram



and assume that $X(0) = 3$.

1. Write down the transition matrix, P, of the Markov chain.

2. Find $P(X(k) = 3)$ for $k = 1, 2, 3, 4$.

3. What is the period of all recurrent communication classes of the Markov chain?

4. Compute P^2, P^4, P^8, and P^{16}.

5. Argue that the Markov chain has an invariant distribution, $\overline{\pi}$, and find this.

6. Let $T_3 := \inf\{k \geq 1 | X(k) = 3\}$ be the first time the Markov chain visits state 3. Compute $P(T_3 = k)$, $k = 1, 2, 3, 4$ and try to find the entire distribution of T_3.

7. Compute the mean, $\mathbb{E}[T_3]$, of the return time to state 3 and compare with the invariant distribution $\overline{\pi}$ of question 5.

2.3.3 Two component repair system

Consider a technical device with two states *broken* and *functioning*. Suppose that every day there is a fixed probability p that the device breaks down. Every morning the state of the device is inspected and if it is broken it is replaced the following morning. Denote by $X(n)$ the state of the device on day n. Clearly, the process $\{X(n)\}_{n\geq 0}$ is a Markov chain.

1. Find the state space and the transition matrix P and draw the transition diagram.

2. Compute the invariant probability distribution $\overline{\pi}$ and find the long term fraction of time where the device is broken.

Consider now a system consisiting of two devices (working independently of each other) that can both take the values *broken* and *functioning*. Every day there is probability p_1 and p_2 of the individual devices breaking down. Every morning the system is inspected and the following morning the broken devices (if any) are replaced. The state of the system on the morning of day n can be described by a Markov chain with the four states

(broken,broken),(broken,funct.),(funct.,broken),(funct.,funct.)

Assume throughout the exercise that no device is broken on the morning of day $n = 0$. To ease notation we recode the state space as *0=broken,1=functioning*.

3. Find the possible transitions of the four state Markov chain and draw the transition diagram of the chain without transition probabilities.

4. Compute the distribution of $X(1)$ i.e. find $P(X(1) = (i, j)), i, j = 0, 1$.

5. Find the transition matrix of the Markov chain.

6. Let $\pi_{i,j} = \lim_{n\to\infty} P(X(n) = (i,j)), i,j = 0,1$, be the limiting distribution of $X(n)$. Show that
$$\pi_{0,0} = p_1 p_2 \pi_{1,1}.$$

7. Write down a similar equation as the one in question 6. for each of the probabilities $\pi_{1,0}, \pi_{0,1}, \pi_{1,1}$.

8. Show that the solution to the system of equations in question 6.-7. is given by
$$\pi_{i,j} = \frac{p_1^{1-i} p_2^{1-j}}{(1+p_1)(1+p_2)}, \quad i,j = 0,1.$$

9. Suppose that it is criticial to a production company that at least one of the individual devices is functioning since otherwise the production of the company ceases and all workers are sent home. What is the long run probability that the production must be stopped and how often does it happen (on average) that both devices break down and workers are sent home?

The company now changes its policy and decides no longer to replace a broken device as long as the other is still working.

10. Draw the transition diagram (with transition probabilities) corresponding to the new replacement strategy.

11. Write down an equation for the invariant probability $\pi_{1,1}$ and show that $\pi_{0,0} = (p_1 + p_2 - p_1 p_2)\pi_{1,1}$.

12. Write down a similar equation as the one in question 11. for each of the probabilities $\pi_{0,0}, \pi_{1,0}, \pi_{0,1}$.

13. Solve the system of equations in question 11.-12.

14. Answer question 9. for the Markov chain corresponding to the new replacement strategy of the company.

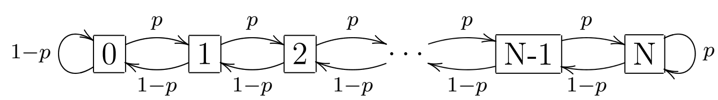
15. Compare the results of question 9. and 14. and discuss how large a fraction of the production is lost when the company only replaces a broken device if the other is broken too.

    [**Warning:** maybe the formulaes get too complicated to conclude anything for in general. In that case try to answer the question for different values of p_1 and p_2 for instance $p_1 = p_2 = p$.]

16. Assume that the maintenance costs for the two devices are proportional to $\frac{1}{p_1}+\frac{1}{p_2}$. Try to say something about what choices of p_1 and p_2 that are the best for the company under the boundary condition that $\frac{1}{p_1}+\frac{1}{p_2}=c$ (constant). The question should be discussed for each of the suggested replacement strategies.

2.3.4 Random walk reflected at two barriers

In this exercise we consider a Markov chain, $\{X(n)\}$, on the state space $\{0, 1, \ldots, N\}$ where only transitions between neighbouring states i and $i+1$ or i and $i-1$ are possible. When the Markov chain reaches the boundary 0 it stays there with probability $1-p$ and is otherwise reflected to state 1. At the upper boundary N the chain stays with probability p as is reflected to state $N-1$ with probability $1-p$. The transition diagram is given by



and we assume that $X(0)=0$.

1. Find the transition matrix, P, of the Markov chain.

2. Compute $P(X(1)=0)$ and $P(X(1)=1)$. What is the name of the distribution of $X(1)$?

3. Compute $P(X(2)=k)$, for $k=0,1,\ldots,N$.

4. Argue that there exists an invariant probability vector, $\overline{\pi}$, and write down the system of equations that should be satisfied by $\overline{\pi}=(\overline{\pi}_0, \overline{\pi}_1, \ldots, \overline{\pi}_N)$.

5. Argue that a vector of the form

$$\pi_i = c\left(\frac{p}{1-p}\right)^i, i=0,1,\ldots,N,$$

satisfies the system of equations from question 4. and find the constant c that turns $\overline{\pi}$ into a probability.

The purpose of the following questions is to find a simple expression for the n step transition matrix P^n for the case of $N = 2$ where the state space is $S = \{0, 1, 2\}$. For this particular case the transition matrix takes the form

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

It might be that you can guess the formula for P^n by looking at the expressions for P^2, P^3, and P^4. Another possibility is to follow the strategy outlined below.

6. Compute the characteristic polynomial $g(\lambda) = \det(P - \lambda I)$ of P.
7. Verify that $g(\lambda) = 0$ has three real valued solutions λ_1, λ_2, and λ_3.
8. For each of the eigenvalues $\lambda_k, k = 1, 2, 3$, above find an eigenvector, v_k, for P with eigenvalue λ_k.
9. Let O be the 3×3 matrix with rows $\overline{v}_1, \overline{v}_2, \overline{v}_3$ and verify that

$$OP = \underbrace{\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}}_{=D} O.$$

10. Show that $P^n = O^{-1}D^nO$ and try to get a closed form expression of $P(X(n) = k)$ for $k \geq 1$ under the initial condition that $P(X(0) = 0) = 1$.

    [**Warning:** maybe it is not worth spending too much time finding a closed form expression for O^{-1}.]

2.3.5 Yahtzee

Yathzee is a dice game. The object of the game is to score the most points by rolling five dice to make certain combinations. The dice can be rolled up to three times in a turn. After the first two rolls the player can save any dice that are needed to complete a combination and then re-roll the other dice. A Yahtzee is five-of-a-kind and holds the game's highest point value of 50.

The purpose of the present exercise is to compute the probability of ending up with a Yahtzee given that we use the strategy that maximizes the number-of-a-kind after each roll. To simplify the problem we consider initially in questions 1.-8. the probability of obtaining a Yathzee of five sixes. We deal with the general problem in questions 9.-17.

The problem may be put into the framework of Markov chains by defining a stochastic process as follows

- Let $X(0) = 0$, i.e. $P(X(0) = 0) = 1$.
- Roll five dice and let $X(1)$ denote the number of sixes.
- Define $X(n+1)$ recursively by the following rule.
  - If $X(n) = 5$ then $X(n+1) = 5$.
  - If $X(n) < 5$ then we let Y be the number of sixes after re-rolling the $5-X(n)$ dice and the value at time $n+1$ is given as $X(n+1) = X(n) + Y$.

1. Argue briefly that $\{X(n)\}$ is a Markov chain and write down the set, S, of possible states for the chain.
2. Find the distribution of $X(1)$ and explain which entries of the transition matrix P that correspond to the probabilities $P(X(1) = k), k \in S$.
3. Find the distribution of $X(2)$ given that $X(1) = 4$.
4. Find the distribution of $X(2)$ given that $X(1) = 3$.
5. Write down the entire transition matrix, P, of the Markov chain.
6. Compute P^2 and the probability $P(X(2) = 5)$.
7. Find the probability $P(X(3) = 5)$.
8. Use $P, P^2, P^3, \ldots$ to compute a numerical approximation to the expected number of rolls, $\mathbb{E}\tau_5$, where

$$\tau_5 = \inf\{n > 0 | X(n) = 5\}$$

   denotes the time before the Markov chain is absorbed in state 5.

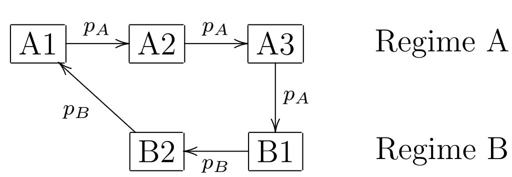
The solve the original problem posed above not restricting our selves to a Yathzee of sixes we need to modify the definition of the Markov chain above. More precisely after each roll we need to allow the player to switch from saving only dice with face six to dice with another number of eyes if this is more favorable. For example if we have two-of-a-kind after n rolls (i.e. $X(n) = 2$) and the next roll results in $5 - 2 = 3$ three dice of a different kind then we let $X(n + 1) = 3$ and only 2 dice are re-rolled.

The purpose of the following questions is to compute the transition matrix for the modified version, $\tilde{P}$, of the game. For all questions below compare the result to the relevant entry of the transition matrix for the original version of the game. Questions 13.-15. below are probably the most difficult.

9. Find the probabilities $P(X(1) = k), k = 3, 4, 5$.
10. Find the probabilities $P(X(2) = k|X(1) = j), k \in S$, for $j = 3, 4, 5$.
11. Find the probabilities $P(X(2) = k|X(1) = 2), k = 4, 5$.
12. Find the probabilities $P(X(2) = 5|X(1) = 1)$.
13. Find the probabilities $P(X(1) = k), k = 0, 1, 2$.
14. Find the probabilities $P(X(2) = k|X(1) = 2), k = 2, 3$.
15. Find the probabilities $P(X(2) = k|X(1) = 1), k = 1, 2, 3, 4$.
16. Write down the entire transition matrix $\tilde{P}$ and compute $\tilde{P}^2$ and $\tilde{P}^3$.
17. Find $P(X(2) = 5), P(X(3) = 5)$, and the mean $\mathbb{E}\tau_5$ and compare with the results in questions 6.-8.

2.3.6 Markov chain with two regimes

Consider the 5 state Markov chain with transition matrix



where $p_A, p_B \in (0,1)$.

We further define the stochastic process $\{Y(n)\}$ defined by

$$Y(n) = \begin{cases} A & , \quad X(n) = \text{A1,A2,A3} \\ B & , \quad X(n) = \text{B1,B2.} \end{cases}$$

In the following we will study the properties of the stochastic process $\{Y(n)\}$ on the state space $S = \{A, B\}$.

1. Find the conditional distribution of $Y(n+1)$ given that $Y(n) = B$ and $Y(n-1) = A$.

2. Argue that the conditional distribtuion of $Y(n+1)$ given that $Y(n) = Y(n-1) = B$ is different from the result of question 1.

Assume that we know that $P(Y(0) = A) = 1$. Clearly

$$P(Y(0) = A) = \underbrace{P(X(0) = A1)}_{=\phi_{A1}} + \underbrace{P(X(0) = A2)}_{=\phi_{A2}} + \underbrace{P(X(0) = A3)}_{=\phi_{A3}}$$

but if we only observe $\{Y(n)\}$ we do not know $\phi_{A1}, \phi_{A2}, \phi_{A3}$.

3. Let
$$\tau_B = \inf\{n > 0 | Y(n) = B\}$$
be the time of the first jump to state B. Express the probabilities $P(\tau_B = k)$ for $k = 1, 2, 3$ in terms of p_A, p_B and $\phi_{A1}, \phi_{A2}, \phi_{A3}$.

4. What should be the distribution of τ_B if $\{Y(n)\}$ was a Markov chain on $\{A, B\}$ with initial distribution $P(Y(0) = A) = 1$ and transition matrix
$$P = \begin{pmatrix} 1 - q_A & q_A \\ q_B & 1 - q_B \end{pmatrix}?$$

5. Use questions 1.-4. to discuss whether $\{Y(n)\}$ is a Markov chain on $S = \{A, B\}$.

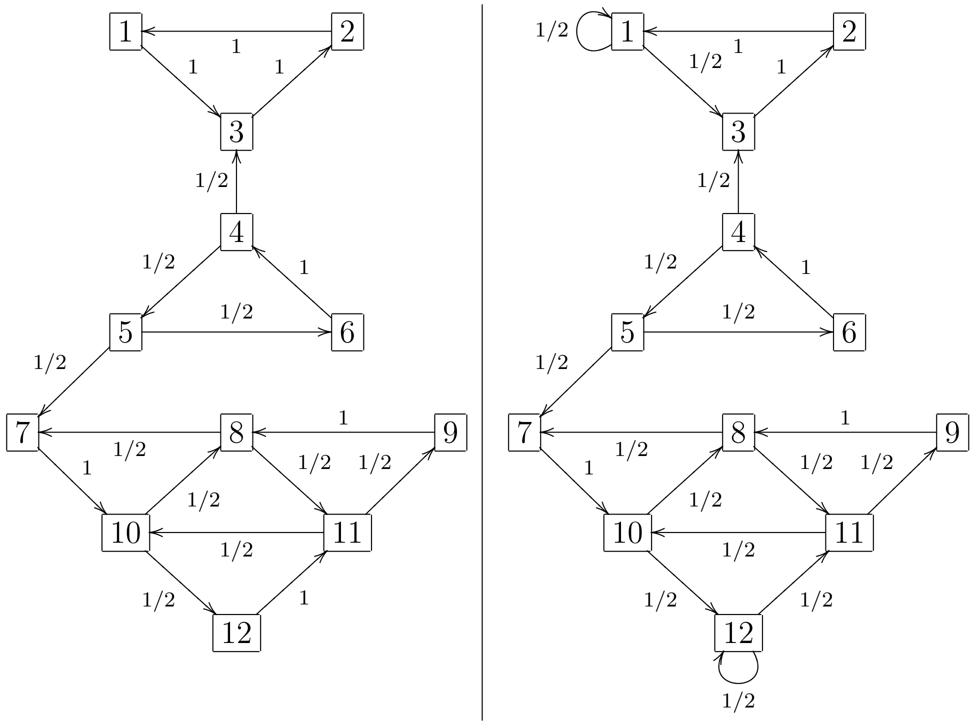


Figure 2.4: Transition diagram for exercise 2.3.7: left diagram should be used for questions 1.-11. and right diagram for the remaining questions 12.-16. of the exercise.

2.3.7 Periodicity of a Markov chain

1. Show that states 1-3 belong to the same communication class.

2. Show that states 10-12 belong to the same communication class.

3. Determine the communication class containing state 12.

4. Argue that states 1 and 6 do not belong to the same communication class.

5. Find all the disjoint communication classes in the partition of the state space. For each class determine whether the class is recurrent or transient.

6. Find the period of the Markov chain.

7. How would the communication classes be and what would be the period of the chain under the following changes: $P_{12,11} = 1/2 = P_{12,12}$

8. How would the communication classes be and what would be the period of the chain under the following changes: $P_{9,8} = 1/2 = P_{9,6}$

10. How would the communication classes be and what would be the period of the chain under the following changes: $P_{8,5} = P_{8,7} = P_{8,11} = 1/3$

11. Does there exist a unique invariant probability distribution for the original version of the chain of the left part of Figure 2.4?

For the rest of the exercise we make the following changes of the transition probabilities (-see also Figure 2.4):

$$P_{1,1} = P_{1,3} = 1/2 = P_{12,11} = P_{12,12}.$$

12. Argue that with the modified transition probabilities then all recurrent subclasses are aperiodic.

13. Find an invariant probability vector that is concentrated on each of the recurrent subclasses.

14. Assuming that the initial distribution of the chain is given by $P(X(0) = 1) = 1$ find the limiting distribution of $X(n)$ that is find

$$\lim_{n\to\infty} P(X(n) = i), \quad i = 1, \ldots, 12.$$

15. Find the limiting distribution of $X(n)$ for the initial distribution $P(X(0) = 12) = 1$.

16. Find the limiting distribution of $X(n)$ for the initial distribution $P(X(0) = 6) = 1$. [Hint: start by computing the probability that the first jump from state 4 to 3 occurs before the first jump between states 5 and 7.]

2.3.8 More about absorption probabilities

Consider a Markov chain on $S = \{0, 1, 2, 3, 4, 5, 6\}$ with transition probabilities

$$\begin{aligned} P_{0,0} &= 3/4, \quad P_{0,1} = 1/4 \\ P_{1,0} &= 1/2, \quad P_{1,1} = P_{1,2} = 1/4 \\ P_{j,0} &= P_{j,j-1} = P_{j,j} = P_{j,j+1} = 1/4, \quad j = 2, 3, 4, 5 \\ P_{6,0} &= 1/4, \quad P_{6,5} = 1/4, \quad P_{6,6} = 1/2 \end{aligned}$$

1. Is the Markov chain irreducible?

2. Is the Markov chain aperiodic?

3. What is the long-run probability of observing the sequence of states $4 \to 5 \to 0$?

4. For $X(0) = 1$ what is the probability of reaching state 6 before state 0?

5. For $X(0) = 3$ what is the expected number of steps until the chain is in state 3 again?

6. For $X(0) = 0$ what is the expected number of steps until the chain is in state 6?

2.4 Markov chains on countable state spaces

For Markov chains on a countable state space S things get a little more complicated by the fact that for a recurrent communication class it can take infinitely long time to get back to the starting point.

Concepts like perodicity as well as transience and recurrence of the communication classes carry over unchanged from the case of a finite state space. But unlike in the finite case it may happen that the chain is irreducible and recurrent yet still all n step transition probabilities tend to zero, i.e.

$$\lim_{n\to\infty} P^n_{i,j} = 0, i, j \in S. \tag{2.6}$$

An irreducible, recurrent Markov chain on a countable state space is said to be *null recurrent* if (2.6) holds. Otherwise we say that the chain is *positive recurrent*.

For an irreducible, positive recurrent Markov chain there exists a unique invariant probability distribution $\overline{\pi} = \{\pi(i)\}_{i \in S}$ such that for any $j \in S$

$$\sum_{i \in S} \pi(i) P_{i,j} = \pi(j).$$

For an irreducible, aperiodic, positive recurrent Markov chain the n step transition probabilities converge to the invariant probabilities

$$\lim_{n \to \infty} P^n_{i,j} = \pi(j)$$

and relates to the recurrence times

$$T_j = \inf\{n > 0 | X(n) = j\}$$

through the identity

$$\pi(j) = \tfrac{1}{\mathbb{E}[T_j | X(0)=j]}$$

where the expectation is computed under the assumption that the Markov chain starts in state j.

2.4.1 Queueing system

Markov chains are very popular as models for the number of customers in a queueing system. In this exercise we consider the socalled single server queue. Assume that no customers are present in the queue at time 0 i.e. $P(X(0) = 0) = 1$. In each time period (=step) there is probability $p \in (0, 1)$ that a new customer arrives and probability $q \in (0, 1)$ that the service of the customer at the service desk is completed. We denote by $X(n)$ the total number of customers in the queueing system at time n and note that this is a Markov chain on $\mathbb{N}_0$. The transition probabilities of the chain is given by the infinite transition matrix $P = (P_{i,j})_{i,j \geq 0}$.

1. Find $P_{0,1}$ and $P_{0,0}$.

2. Argue that $P_{1,1} = pq + (1 - p)(1 - q)$ and find $P_{1,0}, P_{1,2}$.

3. Use question 2. to find $P_{i,i-1}, P_{i,i}, P_{i,i+1}$ for $i > 1$ and draw the transition diagram of the Markov chain.

4. Find the communication classes.

5. For a vector $\overline{\pi} = (\pi_0, \pi_1, \pi_2, \ldots)$ to be an invariant distribution it must satisfy the system of equations $\pi_j = \sum_{i=0}^{\infty} \pi_i P_{i,j}, j \geq 0$. Write out the equation for $j = 0$ and deduce that $\pi_1 = \frac{p}{q(1-p)}\pi_0$.

6. Write down the equations for $\pi_j, j \geq 1$.

In the following questions 7.-10. we assume that $p = q$.

7. Show that for $p = q$ then $\pi_j = c_0 + c_1 \cdot j, j \geq 1$, ($c_0, c_1$ constants) solves the system of equations from question 6 (for $j \geq 2$). [One may show that any solution takes this form.]

8. Find a condition on the constants c_0, c_1 that ensures that the solution π_j from question 7. is bounded for $j \geq 1$.

9. Does there exist an invariant probability vector for the chain if $p = q$?

10. Discuss whether we have showed that the chain is positive recurrent, null recurrent, or transient for $p = q$?

For the remaining part of the exercise we consider the general case where $p \neq q$.

11. Argue that $\pi_j = c_0 + c_1 \cdot \left(\frac{p(1-q)}{q(1-p)}\right)^j, j \geq 1$, solve the system of equations from question 6. (for $j \geq 2$) for any choice of the constants c_0, c_1.

12. Use the equation from question 5. to express π_0 in terms of p, q, and the two constants c_0, c_1.

13. Determine when the chain is positive recurrent and find the invariant probability vector $\overline{\pi}$.

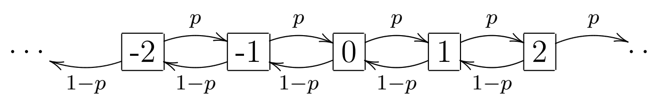
14. Compute the (long run) average number of customers in the system for the case where the chain is positive recurrent.

15. Give a complete description of when (i.e. for what conditions on p and q) the chain is transient, null recurrent, or positive recurrent.

    [**Warning:** this probably requires a little work!]

2.4.2 Random walk on $\mathbb{Z}$

Consider the random walk,$\{X(n)\}_{n\geq 0}$, on $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ given by the transition diagram



Assume that $X(0) = 0$.

1. What is the period of the chain?
2. Find the distribution of $X(1)$ and $X(2)$.
3. Compute $P(X(2) = 0), P(X(3) = 0)$, and $P(X(4) = 0)$.

Note that $(X(2k) = 0)$ if and only if there is exactly k upward jumps and k downward jumps among the first $2k$ jumps.

4. Argue that $P(X(2k) = 0) = \binom{2k}{k} p^k (1-p)^k$, $k \geq 1$.
5. Determine if $\sum_{n=1}^{\infty} P(X(n) = 0)$ is convergent and use this to decide if the random walk on $\mathbb{Z}$ is recurrent or transient.

2.4.3 Random walk on $\mathbb{Z}^2$

We now generalise Exercise 2.4.2 above and consider the symmetric random walk on the pairs of integers $\mathbb{Z} \times \mathbb{Z}$. More precisely, if the chain is in state (i, j) at time n then it jumps to either of the states

$$(i, j+1), (i, j-1), (i-1, j), (i+1, j)$$

with equal probabilities $(= 1/4)$ in step $n + 1$.

1. Draw (a part of) the transition diagram.
2. Argue that state $(0, 0)$ communicates with any other state and deduce that there is only one communication class.
3. Assuming that $P(X(0) = (0, 0)) = 1$ compute $P(X(n) = (0, 0))$ for $n = 1, 2, 3, 4$.

4. What is the period of the Markov chain.

5. Still assuming that $P(X(0) = (0,0)) = 1$ argue that $P(X(2n) = 0) = \sum_{k=0}^{n} \frac{(2n)!}{(n-k)!\cdot(n-k)!\cdot k!\cdot k!} 4^{-2n}$.

6. Use question 5. to determine if $\sum_{n=0}^{\infty} P(X(2n) = 0)$ is convergent and deduce if the random walk on $\mathbb{Z} \times \mathbb{Z}$ is recurrent or transient.

2.4.4 Random walk on $\mathbb{Z}^d$

It is a challenging exercise to determine if the extension of the random walk in the previous exercise to $\mathbb{Z}^d$ is recurrent or transient. The dynamics of the $d-$dimensional random walk is desribed by the fact that the process moves from state $(i_1, \ldots, i_d)$ to any of the $2d$ neighbouring states given by

$$(i_1, \ldots, i_{l-1}, i_l + j, i_{l+1}, \ldots, i_d)$$

where $j \in \{-1, 1\}$ with equal probability $(= 1/(2d))$.

1. Argue that the symmetric random walk on $\mathbb{Z}^d$ has period 2.

2. Assuming that the random walk starts in state $(0, \ldots, 0)$ at time 0 argue that

$$P(X(2n) = (0, \ldots, 0)) = \sum_{k_1,\ldots,k_d \in \mathbb{N}_0 : k_1 + \ldots + k_d = n} \frac{(2n)!}{(k_1! \cdot \ldots \cdot k_d!)^2} (2d)^{-2n}.$$

3. Show that $\sum_{n=1}^{\infty} P(X(n) = (0, \ldots, 0)) < \infty$ for $d > 2$.

4. Deduce from question 3. that the symmetric random walk on $\mathbb{Z}^d$ is transient for $d > 2$.

2.4.5 Brancing processes

In this exercise we consider a model for the number, $X(n)$, of individuals in a population at time n. During each time interval (generation) each individual (independently of each other) produces a number, Z, of offsprings described by a probability distribution on $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ with density $P(Z = k) = p_k$. Note that it is also possible for an individual to die without giving birth to any offspring if $Z = 0$. The total number of individuals born in $n-$th generation, $\{X(n)\}_{n\geq 0}$, is a Markov chain on $\mathbb{N}_0$

The important parameter for the large-time term behaviour of a branching process is the mean number of offsprings produced by an individual

$$\mu = \sum_{k=0}^{\infty} kp_k = \sum_{k=1}^{\infty} kp_k.$$

Not surprisingly one can prove that if $\mu > 1$ then the expected population size increases to infinity and that $\mu < 1$ implies that the population will eventually die out.

In this exercise we consider the (rather trivial) branching process with offspring distribution given by

$$p_1 = p, \quad p_0 = 1 - p,$$

with probability parameter $p \in (0,1)$. The interpretation is that each individual gives birth to one offspring with probability p while there is a probability of $1-p$ that no offspring is generated. Assume that we start out with a population of size $X(0) = N > 0$.

1. Find the probability $P(X(1) = N)$ and $P(X(1) = N - 1)$.
2. Argue that $X(1)$ follows a binomial distribution and find the integral parameter and the probability parameter.
3. Use 1.-2. to find the transition probabilities
$$P_{N,j} = P(X(n+1) = j | X(n) = N), j = 0, 1, \ldots, N.$$
4. Find an expression for the general transition probability $P_{i,j}, i, j \in \mathbb{N}_0$.
5. Compute $\mathbb{E}X(1)$ and give a (heuristic) argument that the population will eventually die out.
6. Show that $\sum_{n=1}^{\infty}(P^n)_{i,i} < +\infty$ for $i = 1, \ldots, n$, and deduce that $\lim_{n\to\infty} P(X(n) = 0) = 1$.

   [Hint: Find the communication classes and conclude that state i, $i = 1, \ldots, n$, (and its communication class) is transient.]

2.4.6 Positive recurrence and null-recurrence

The following exercise is greatly inspired by Exercise 2.2 in Lawler (2006). We consider a Markov chain, $\{X(n)\}_{n\geq 0}$, on $S = \{0, 1, 2, \ldots\}$ with transition probabilities

$$P_{0,i} = p_i > 0, i > 0, \quad P_{i+1,i} = 1, i \geq 0, \quad P_{i,j} = 0 \text{ otherwise}$$

where $(p_i)_{i\in S}$ is probability vector (i.e.$\sum_i p_i = 1$). Define the return time to state 0

$$T = \inf\{n > 0 | X(n) = 0\}.$$

1. Draw the transition diagram of the Markov chain.
2. Find the communication classes. Is the chain irreducible?
3. Compute
$$P(T = k|X(0) = 0), \quad k \geq 0$$
and argue that the Markov chain is recurrent.
4. What is the condition for the Markov chain to be null-recurrent or positive recurrent?
5. Find the invariant probability vector assuming that the Markov chain is positive recurrent.
6. Consider the time of the first visit to state 10
$$T_{10} = \inf\{n > 0 | X(n) = 10\}.$$
What is the expected return time to state 10
$$\mathbb{E}[T_{10}|X(0) = 10]$$
given that the Markov chain starts in state 10?

Chapter 3

Continuous-time Markov chains

3.0 Results for continuous-time Markov chains

3.0.1 Definition of Markov chains

A stochastic process in continuous time is a family, $\{X(t)\}_{t\geq 0}$, of random variables indexed by the positive real line $[0,\infty)$. The possible values of $\{X(t)\}_{t\geq 0}$, are referred to as the *state space*, S, of the process. On the course *Stochastic Processes* we shall only consider continuous-time processes on finite or countable state spaces. Further, we consider only processes with piecewise constant sample paths composed of sequences of the times, τ_n, of the jumps and the target states, $Y(n)$, of the following jump as visualized in Figure 3.1. The student who wants to dig deeper into the topic will at some point experience that there are Markov chains that may not be viewed as a sequence of waiting times and jumps.

Definition 5 (Homogeneous Markov chain in continuous-time) *A - continuous-time Markov chain on a finite or countable set, S, is a family of random variables $\{X(t)\}_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, P)$ such that*

$$\begin{aligned} & P(X(t_{n+1}) = j | X(t_n) = i, X(t_{n-1}) = i_{n-1}, \ldots, X(t_0) = i_0) \\ = \;& P(X(t_{n+1}) = j | X(t_n) = i) = P_{i,j}(t_{n+1} - t_n) \end{aligned}$$

for $j, i, i_{n-1}, \ldots, i_0 \in S$ and $t_{n+1} \geq t_n \geq \ldots \geq t_0 \geq 0$. The distribution of the Markov chain is determined by

$$\begin{aligned} \phi(i) &= P(X(0) = i) \quad \leftarrow \quad \textit{initial distribution} \\ P_{i,j}(t) &= P(X(t+s) = j | X(s) = i) \quad \leftarrow \quad \textit{transition probabilities} \end{aligned}$$

through the identity

$$\begin{aligned} & P(X(t_{n+1}) = j, X(t_n) = i, X(t_{n-1}) = i_{n-1}, \ldots, X(t_0) = i_0) \\ = \;& P_{i,j}(t_{n+1} - t_n) \cdot P_{i_{n-1},i}(t_n - t_{n-1}) \cdot \ldots \cdot P_{i_0,i_1}(t_1 - t_0) \cdot \phi(i_0) \quad (3.1) \end{aligned}$$

□

To characterize the distribution of a continuous-time Markov chain we therefore need a probability vector, $\overline{\phi} = (\phi(i))_{i\in S}$, of initial probabilities and a family of transition probabilities $(P_{i,j}(t))_{i,j\in S}$ for *any* $t \geq 0$. For any fixed $t \geq 0$ then $(P_{i,j}(t))_{i,j\in S}$ must be a transition probability as introduced for discrete-time Markov chains. However, the transition probabilities for different time arguments must fit together in accordance with the *Chapman-Kolmogorov equations* given below.

Definition 6 (Chapman-Kolmogorov equations) *The transition probabilities for a homogeneous continuous-time Markov chain satisfy the Chapman-Kolmogorov equations*

$$\forall 0 \leq s \leq t, i, j \in S : P_{i,j}(t+s) = \sum_{l \in S} P_{i,l}(t) \cdot P_{l,j}(s).$$

For finite state space, S, then $\{P_{i,j}(t)\}_{i,j \in S}$ may be treated as a matrix for any fixed $t \geq 0$ and the Chapman-Kolmogorov equations may be written as a matrix equation

$$P(t+s) = P(t) \cdot P(s).$$

□

As mentioned earlier we shall in this course consider only Markov chain with a pure jump structure. Intuitively, these are stochastic processes where a plot of the sample-path (=function)

$$t \to X(t)$$

is piecewise constant (-see Figure 3.1). In turns out that homogeneous continuous-time Markov jump processes are uniquely determined by its transition intensities $Q = (q_{i,j})_{i,j \in S}$ as described in Theorem 12. Conversely, any transition intensity, Q, satisfying the conditions in Definition 7 defines to a homogeneous continuous-time Markov jump process. Thus on this course continuous-time Markov chains are in a one-to-one correspondence with their transition intensity.

Definition 7 (Continuous-time MCs on Stochastic Processes) *Any - Markov chain may be defined from a vector of initial probabilities $\overline{\phi} = (\phi(i))_{i \in S}$ and a transition intensity matrix $Q = (q_{i,j})_{i,j \in S}$, with the following properties*

$$\begin{aligned} q_{i,j} &\geq 0 \quad i \neq j, i, j \in S \\ q_{i,i} &= -\sum_{j \neq i} q_{i,j}. \end{aligned}$$

Note that $q_{i,i} \leq 0$. We will use the notation $q_i = -q_{i,i}$ for the diagonal element with opposite sign. In the literature Q is often referred to as the infinitesimal generator. □

Definition 7 introduces in a rather non-informative way the ingredients needed to define a continuous-time Markov chain. The formal relation between the transition intensity of Definition 7 and the transition probabilities in Definition 5 is explained in Theorem 12. We shall now try to link the transition intensity, Q, and the initial distribution, $\overline{\phi}$, of Definition 7 to the dynamics of a continuous-time Markov chain.

For a continuous-time Markov jump process, $\{X(t)\}_{t\geq 0}$, with transition intensity, Q, we may consider the sequence of states

$$Y(0), Y(1), Y(2), \ldots$$

visited along the sample path while ignoring the waiting times between jumps. If $\{X(t)\}_{t\geq 0}$ can be absorbed there is a positive probability that only finitely many jumps $Y(0), Y(1), \ldots, Y(m)$ are observed. For technical reasons we then introduce an extra state Δ in the state space and let

$$Y(m+1) = Y(m+2) = \ldots = \Delta.$$

This construction ensures that any continuous-time Markov chain $\{X(t)\}_{t\geq 0}$ on S defines a discrete-time process $\{Y(n)\}_{n\geq 0}$ on a possibly extended state space $S \cup \Delta$.

Theorem 11 (Embedded Markov chain of jumps) *For a continuous-time Markov chain $\{X(t)\}_{t\geq 0}$ on S let*

$$A = \{i \in S | q_{i,i} = 0\}$$

be the subset of absorbing states. Let $\{Y(n)\}_{n\geq 0}$ be the sequence of states visited by $\{X(t)\}_{t\geq 0}$ with the convention that

$$Y(k) = \Delta, \quad k > m,$$

if $Y(m) \in A$ is an absorbing state. Then $\{Y(n)\}_{n\geq 0}$ is a discrete-time Markov chain on the extended state space $\overline{S} = S \cup \Delta$. The transition probabilities, $P = (P_{i,j})_{i\in\overline{S}}$, of the embedded Markov chain of jumps, $\{Y(n)\}_{n\geq 0}$, with state space $\overline{S} = S \cup \Delta$ are then given by

$$P_{i,j} = \begin{cases} -\frac{q_{i,j}}{q_{i,i}} = \frac{q_{i,j}}{q_i} & i \in S\backslash A, j \notin \{i, \Delta\} \\ 0 & i \in S\backslash A, j \in \{i, \Delta\} \\ 0 & i \in A, j \neq \Delta \\ 1 & i \in A, j = \Delta \\ 0 & i, j \neq \Delta \\ 1 & i, j = \Delta \end{cases}$$

If no absorbing states exist ($A = \emptyset$) then the embedded Markov chain of jumps has state space S and the transition probabilities are given by the first two lines above. □

We are now ready to describe the dynamics of a continuous-time Markov chain with transition intensity Q and initial distribution $\overline{\phi} = (\phi(i))_{i\in S}$. The initial value, $Y(0)$, of $X(0)$ is drawn (at random) according to the initial probability distribution $\overline{\phi} = (\phi(i))_{i\in S}$. The Markov chain then stays in the initial state for a random amount of time, τ_1, given by an exponential distribution with rate parameter $q_{Y(0)}$ depending on the initial state $Y(0)$. At the time of the jump the chain jumps to a new state, $Y(1)$, according to the transition probabilities from state $Y(0)$ corresponding to the embedded Markov chain of jumps in Theorem 11. The Markov chain then stays in state $Y(1)$ for an exponentially distributed amount of time, $\tau_2 - \tau_1$, with rate parameter $q_{Y(1)}$ and the following state, $Y(2)$, is drawn according to transition probabilities of the embedded jump process etc.

The construction described above is referred to as the *minimal construction* of a Markov jump process. There is a potential problem that there may be infinitely many jumps in finite time such that the random variable $\tau_\infty := \lim_{n\to\infty} \tau_n$ may have a positive probability

$$P(\tau_\infty < +\infty) > 0$$

of not defining a stochastic process $\{X(t)\}_{t\geq 0}$ for $t > \tau_\infty$. For technical we will then introduce an extra state Δ and let $X(t) = \Delta$ for $t \geq \tau_\infty$. We will discuss explosion a little further i Section 3.0.4.

In short, the sequence of states visited by a continuous-time Markov chain evolves as a (discrete-time) Markov chain with transition probabilities given by Theorem 11. The waiting time spent in state $i \in S$ follows an exponential distribution with rate parameter $q_i = \sum_{j\neq i} q_{i,j}$, i.e. the sum of the transition intensities corresponding to row i (or pointing away from state i).

3.0.2 Properties of the transition probabilities

On this course Markov chains are usually defined in terms of the transition intensity (or the infinitesimal generator), Q, from Definition 7. However, for

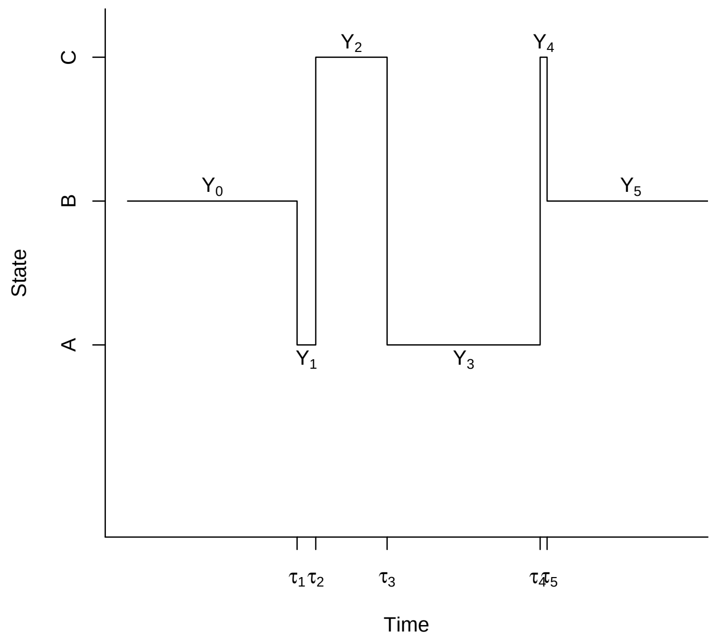


Figure 3.1: Sample path of a continuous-time Markov chain with three states. The sequence of states, $\{Y(n)\}_{n\geq 0}$, is a discrete-time Markov chain with transition probabilities given by Theorem 11. The waiting time $\tau_{n+1} - \tau_n$ between jump n and $n+1$ follows an exponential distribution with rate parameter $q_{Y(n)}$.

applications we are often more interested in the transtion probabilities

$$P(X(t+s) = j | X(s) = i) := P_{i,j}(t), \quad i, j \in S, t \geq 0$$

In this section we discuss the relation between the transition intensity, Q, of a continuous-time Markov chain and the transition probabilities $P(t) = (P_{i,j}(t))_{i,j \in S}$.

Theorem 12 (Infinitesimal generator of a Markov chain) *For a con- - tinuous-time Markov chain, $\{X(t)\}_{t\geq 0}$, the transition intensities may be* *obtained from transition probabilities* $P(t) = (P_{i,j}(t))_{i,j \in S}$ *as the limits*

$$\lim_{t \to 0+} \frac{P_{i,i}(t) - 1}{t} = q_{i,i} \tag{3.2}$$

$$\lim_{t \to 0+} \frac{P_{i,j}(t)}{t} = q_{i,j}, \quad i \neq j. \tag{3.3}$$

□

More generally, the transition probabilities and the transition intensities are related through the backward differential equations (or Kolmogorov's differential equation).

Theorem 13 (Backward differential equations) *For a continuous-time Markov chain,* $\{X(t)\}_{t \geq 0}$, *with transition intensity,* $Q = (q_{i,j})_{i,j \in S}$, *and transition probabilities* $\{P_{i,j}(t)\}_{i,j \in S}$ *it always holds that*

$$DP_{i,j}(t) = P'_{i,j}(t) = q_{i,i} P_{i,j}(t) + \sum_{k \neq i} q_{i,k} P_{k,j}(t) \tag{3.4}$$

□

Remark 3 (Backward integral equations) *An intermediate step in deriving the backward differential equations of Theorem 13 is the set of backward integral equations which may be of interest in itself*

$$P_{i,j}(t) = \delta_{i,j} \exp(q_{i,i} t) + \int_0^t \sum_{k \neq i} q_{i,k} \exp(q_{i,i}(t-s)) P_{k,j}(s) ds, \tag{3.5}$$

where $\delta_{i,j} = 0, i \neq j$, *and* $\delta_{i,i} = 1$. □

There also exist sets of forward differential and integral equations for continuous-time Markov chains.

Theorem 14 (Forward differential equations) *For a continuous-time Markov chain,* $\{X(t)\}_{t \geq 0}$, *with transition intensities,* $Q = (q_{i,j})_{i,j \in S}$, *and transition probabilities* $P(t) = (P_{i,j}(t))_{i,j \in S}$ *it holds that*

$$DP_{i,j}(t) = q_{j,j} P_{i,j}(t) + \sum_{l \neq j} P_{i,l}(t) q_{l,j}. \tag{3.6}$$

□

The proof of the forward differential equations is not trivial. It is often claimed in the literature that a sufficient condition for the forward differential equations to hold is that

$$\sum_{j\in S} p_{i,j}(t)(-q_{j,j}) < \infty,$$

and that the forward equations do not always hold. You are allowed to use the correct version of the result which says that the forward equations hold for any Markov jump process on a finite or countable state space.

Remark 4 *In Section 3.0.4 we discuss continuous-time Markov chains where explosion may occur. Explosion refers to the fact that there may be infinitely many jumps in finite time. One can show that explosion does not happen if the condition above is satisfied, i.e. if*

$$\sum_{j\in S} p_{i,j}(t)(-q_{j,j}) < \infty.$$

It is true that if explosion is not possible then the differential equations of Theorems 13 and 14 uniquely determines the transition probabilities $P(t) = (P_{i,j}(t))_{i,j\in S}$ *subject to the initial conditions* $P(0) = I$. *If explosion is possible then there is no unique solution to the differential equations. The minimal solution will give the transition probabilities corresponding to the process described by the minimal construction of a Markov jump process.* □

Theorem 15 (Transition probabilities for finite dimensional chains) *For a continuous-time Markov chain on a finite state space the backward differential equation may be expressed in matrix form as*

$$DP(t) = P'(t) = QP(t)$$

where $P(t) = (P_{i,j}(t))_{i,j\in S}$. *Using the boundary condition* $P(0) = I$ *it turns out that the transition probabilities may expressed in terms of exponential matrices as*

$$P(t) = \exp(Qt), \quad t \geq 0.$$

□

3.0.3 Invariant distributions and absorption

The state space of a continuous-time Markov chain is partitioned into communication classes. Two states i and j communicate if it is possible to move forth and back between them.

Definition 8 (Communication classes and irreducibility) *Two states $i, j \in S$ are said to communicate if there exists $s, t > 0$ such that*

$$P_{i,j}(s) > 0 \quad \textit{and} \quad P_{j,i}(t) > 0.$$

This definition partitions the state space, S, into (disjoint) communication classes. A continuous-time Markov chain is irreducible if there is only one communication class. For an irreducible Markov chain it holds that

$$\forall i, j \in S, \forall t > 0 \quad : \quad P_{i,j}(t) > 0.$$

□

From a practical point of view you are allowed to use that for a continuous-time Markov chain two states $i \neq j, i, j \in S$ communicate, if there exists a sequence of states $i_1, i_2, \ldots, i_n \in S$ containing state j such that

$$q_{i,i_1} \cdot q_{i_1,i_2} \cdot \ldots \cdot q_{i_{n-1},i_n} \cdot q_{i_n,i} > 0.$$

Since $P_{i,i}(t) > 0$ we always have that i communicate with itself.

Definition 9 (Recurrence and transience) *An irreducible continuous-time Markov chain is recurrent if and only if the embedded discrete-time process of jumps (-see Theorem 11) is recurrent. It is transient if and only if the embedded discrete-time process of jumps is transient. If the continuous-time Markov chain is not irreducible the definitions of recurrence and transience apply seperately to each communication class. Note that an absorbing state will always be transient.* □

As a consequence of Definition 9 to determine if a continuous-time Markov chain is recurrent or transient you should study the embedded discrete-time Markov chain of jumps and use the criterions for recurrence given in Definition 2, Theorem 4, and Theorem 5 of the Chapter 2.0 on discrete-time Markov chains.

Definition 10 (Invariant distribution) *A probability, $\overline{\pi} = (\pi(i))_{i \in S}$, is an invariant (or stationary) distribution for a continuous-time Markov chain if for all $t \geq 0$ and $j \in S$*

$$\pi(j) = \sum_{i \in S} \pi(i) P_{i,j}(t).$$

□

We state the following result concerning uniqueness of invariant distributions.

Theorem 16 (Uniqueness of invariant distribution) *For an irreducible continuous-time Markov chain the stationary distribution is unique if it exists.*

If for some $t_0 > 0$ there is a probability $\overline{\pi} = (\pi(i))_{i \in S}$ such that

$$\forall j \in S \quad : \quad \pi(j) = \sum_{i \in S} \pi(i) P_{i,j}(t_0)$$

then we may conclude that

1. $\forall i \in S \quad : \quad \pi(i) > 0$
2. *$P(t_0)$ is a transition probability, i.e.*

$$\forall i \in S \quad : \quad \sum_{j \in S} P_{i,j}(t_0) = 1$$

3. *$\overline{\pi}$ is an invariant distribution for the Markov chain, i.e.*

$$\forall t \geq 0, \forall j \in S \quad : \quad \pi(j) = \sum_{i \in S} \pi(i) P_{i,j}(t).$$

□

Since Markov chains are usually specified in terms of the transition intensities we can rarely apply Definition 10 directly to find the stationary distribution. The following result give a necessary condition for a stationary distribution expressed in terms of the transition intensity.

Theorem 17 (Necessary condition for a stationary distribution) *For a continuous-time Markov chain with transition intensity, Q, an invariant probability, $\overline{\pi} = (\pi(i))_{i \in S}$, must satisfy the system of equations*

$$\forall j \in S \quad : \quad \sum_{i \in S} \pi(i) q_{i,j} = 0 \tag{3.7}$$

or written in another form

$$\forall j \in S \quad : \quad \sum_{i \neq j} \pi(i) q_{i,j} = \pi(j)(-q_{j,j}) = \pi(j) q_j.$$

Thinking of $\overline{\pi}$ as a row vector and Q as a matrix the system of equations has a more compact formulation as

$$\overline{\pi} Q = 0.$$

□

From a practical point of view to find an invariant distribution for a continuous-time Markov chain it is advisable to start by solving the system of equations from Theorem 17. If a solution exists that is not zero in all coordinates there will always be infinitely many non-zero solutions since multiplication by a constant does not alter the system of equations (3.7). Therefore, an important step is to check the existence of a solution that can be normalized into a probability vector of non-negative coordinates with sum 1. It is very common that the coordinates of any non-zero solution sum to $+\infty$ so that no normalized solution may be found.

A probability distribution solving the system of equations from Theorem 17 will be a good candidate for an invariant distribution. However, it turns out that for Markov chains with an infinite state space additional conditions are required to ensure that we have indeed found an invariant probability. Note that the condition of the following Theorem 18 is trivially satisfied for Markov chains on a finite state space.

Theorem 18 (Sufficient condition for a stationary distribution) *If $\overline{\pi} = (\pi(i))_{i \in S}$ is a probability satisfying the condition*

$$\forall j \in S \quad : \quad \sum_{i \neq j} \pi(i) q_{i,j} = \pi(j)(-q_{j,j}) = \pi(j) q_j.$$

of Theorem 17 and furthermore

$$\sum_{j \in S} \pi(j)(-q_{j,j}) < \infty$$

then $\overline{\pi} = (\pi(i))_{i \in S}$ is a unique stationary distribution for an irreducible Markov chain. □

Remark 5 (Invariant distributions for non-irreducible chains) *If a continuous-time Markov chain is not irreducible (i.e. has more than one communication class) an invariant distribution might not be unique. However, the results of Theorems 17 and 18 are still necessary and sufficient conditions for any invariant distribution.* □

It might be a bit difficult to understand the role of the additional sufficient condition given in Theorem 18. Why don't we give a necessery *and* sufficient condition for a probability $\overline{\pi} = (\pi(i))_{i \in S}$ to be an invariant distribution for a continuous-time Markov chain? The main reason is that this can only be done with reference to the embedded discrete-time Markov chain of jumps.

Theorem 19 (Invariant distribution) *The continuous-time irreducible Markov chain $\{X(t)\}_{t \geq 0}$ has an invariant (or stationary) distribution if and only if the embedded discrete-time Markov chain of jumps is recurrent and there exists a probability vector $\overline{\pi} = (\pi(i))_{i \in S}$ such that (3.7) holds or written in a more compact notation such that $\overline{\pi} Q = 0$.* □

Theorem 20 (Limit results for transition probabilities) *For an irreducible Markov chain, $\{X(t)\}_{t \geq 0}$, with invariant distribution $\overline{\pi} = (\pi(i))_{i \in S}$ it holds for all $i, j \in S$ that*

$$\lim_{t \to \infty} P_{i,j}(t) = \pi(j).$$

Further, for any initial distribution $\overline{\phi} = (\phi(i))_{i \in S}$ and $j \in S$ it holds that

$$\lim_{t \to \infty} P(X(t) = j) = \pi(j).$$

If no invariant distribution exists then

$$\lim_{t \to \infty} P_{i,j}(t) = 0.$$

□

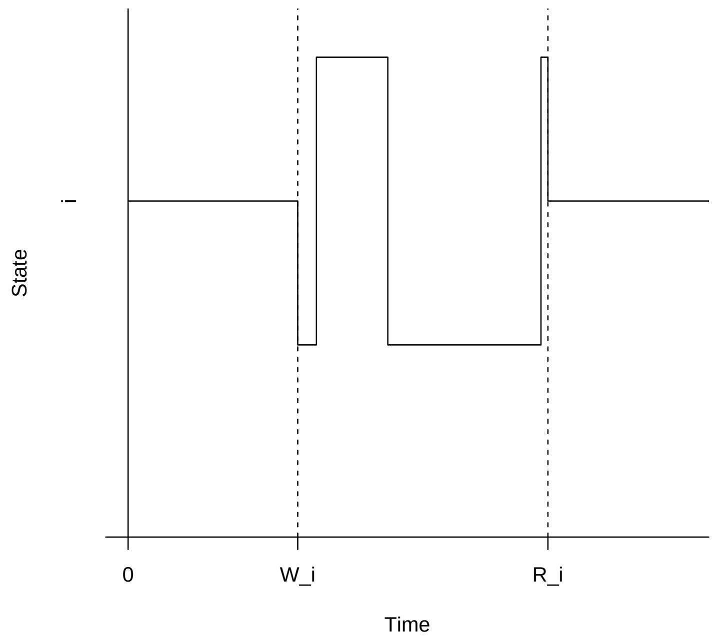


Figure 3.2: Interpretation of invariant distribution of continuous-time Markov chain. The invariant probability at state i is the average fraction of time spend in state i between two succesive jumps to state i.

For an irreducible continuous-time Markov chain there is a nice interpretation of the invariant distribution as the long-run fraction of time spend in individual states.

Theorem 21 (Interpretation of invariant distribution and positive recurrence) *For an irreducible continuous-time Markov chain $\{X(t)\}_{t\geq 0}$ define the escape time from state i*

$$W_i = \inf\{t \geq 0 | X(t) \neq i\}$$

and the return time to state i

$$R_i = \inf\{t > W_i | X(t) = i\}.$$

Then the invariant probability $\overline{\pi} = (\pi(i))_{i\in S}$ is given by

$$\pi(i) = \frac{\mathbb{E}[W_i|X(0)=i]}{\mathbb{E}[R_i|X(0)=i]} = \frac{1}{q_i\mathbb{E}[R_i|X(0)=i]}.$$

The result is also valid when all expectations $\mathbb{E}[R_i|X(0)=i] = +\infty$ if we take $\pi(i)=0$ to mean that no invariant distribution exists.

We say that a communication class is positive recurrent if $\mathbb{E}[R_i|X(0)=i] < \infty$ and note that this is equivalent to existence of an invariant distribution.
□

From Definition 9 we know that recurrence of a continuous-time Markov chain is (by definition!) equivalent to recurrence of the embedded Markov chain of jumps. This is not true when it comes to positive recurrence and existence of invariant distributions. For a discrete-time Markov chain an invariant distribution is a probability $\overline{\pi} = (\pi(i))_{i\in S}$ solving the system of equations

$$\forall j \in S \quad : \quad \pi(j) = \sum_{i\in S} \pi(i) P_{i,j}.$$

For continuous-time Markov chains a little more is required to verify the existence of an invariant distribution. Here we must both find a solution to the system of equations

$$\forall j \in S \qquad : \qquad \sum_{i\neq j} \pi(i) q_{j,i} = \pi(j)(-q_{j,j}) = \pi(j) q_j.$$

and verify that the Markov chain is recurrent! The invariant distribution for a continuous-time Markov chain and the invariant distribution for the embedded Markov chain of jumps are not identical (if they exist!).

Theorem 22 (Time-invariant vs. event-invariant distribution) *Consider a continous-time Markov chain with transition intensity Q and assume that the invariant distribution $\overline{\nu} = (\nu(i))_{i\in S}$ exists. Suppose that we have also verified the existence of an invariant distribution $\overline{\pi} = (\pi(i))_{i\in S}$ for the embedded Markov chain of jumps. Then the following relation holds*

$$\pi(i) \;=\; \frac{\nu(i) q_i}{\sum_{j\in S} \nu(j) q_j}, \quad i \in S. \tag{3.8}$$

□

An absorbing state is a state from where the Markov chain cannot escape. State $i \in S$ is absorbing if $q_{i,i} = 0$ or equivalently if $\sum_{j \neq i} q_{i,j} = 0$. A number of interesting questions are related to an absorbing state. First of all we might want to compute the probability that the Markov chain is eventually absorbed in state i. Secondly, one could be interested in the behavior of the Markov chain until absorption for example the average time spend in any other state $j \neq i$ before being caught in state i.

Theorem 23 (Time spent in state j before absorption) *For a continuous-time Markov chain the average number of periods (visits) spent in state j before reaching an absorbing state i (i.e. with $q_{i,i} = 0$) may be found by studying the transition probabilities of the embedded discrete-time Markov chain of jumps. For finite state space Markov chains this computation may be carried out using Theorem 9 while you may use Theorem 10 for Markov chains on countably infinite state spaces.*
If N_j is the mean number of visits to state j before absorption in state i then the average time spend in state j before absorption is given by $\frac{N_j}{q_j}$. □

3.0.4 Birth-and-death processes

In this section we discuss an important class of continuous-time Markov chains on a coutable state space. A *birth-and-death* process is a Markov chain on $S = \mathbb{N}_0$ that allows only jumps (upwards or downwards) of size one. Referring to our usual specification of Markov chains in terms of transition intensities this means that we assume that

$$q_{i,j} = 0, \quad i, j \in \mathbb{N}_0, |i - j| > 1$$

while the only non-zero intensities (except for the diagonal) are

$$\begin{aligned} q_{i,i+1} &= \beta_i, \quad i \in \mathbb{N}_0 \quad \leftarrow \text{ birth intensities} \\ q_{i,i-1} &= \delta_i, \quad i \in \mathbb{N} \quad \leftarrow \text{ death intensities.} \end{aligned}$$

The dynamics of a birth-and-death process is very simple. If the process is currently in state i then the waiting time to the next jump follows an exponential distribution with rate $\beta_i + \delta_i$ (i.e. mean $\frac{1}{\beta_i + \delta_i}$). At the time of the jump the process moves one step up with probability $\beta_i/(\beta_i + \delta_i)$ and one step down with probability $\delta_i/(\beta_i + \delta_i)$.

Consider the time, τ_n, of the $n-$th jump for a birth-and-death process. If the process is absorbed before the n-th jump then we let $\tau_n = +\infty$. Clearly, $(\tau_n)_{n\geq 0}$ is increasing

$$\tau_1 \leq \tau_2 \leq \ldots \leq \tau_n$$

and we may define the variable

$$\tau_\infty = \lim_{n\to\infty} \tau_n$$

with values in $[0, +\infty]$. From a mathematical point of view it is easy to give examples of birth intensities, $(\beta_i)_{i\in\mathbb{N}_0}$, and death intensities, $(\delta_i)_{i\in\mathbb{N}}$, such that

$$P(\tau_\infty = +\infty) < 1$$

or in other words such that there is a strictly positive probability of observing an infinite number of jumps in finite time. In this situation we will say that explosion is possible or that the transition intensities allow for explosion.

Example 1 (Pure birth process with explosion) *A Markov chain $\{X(t)\}_{t\geq 0}$ on $\mathbb{N}_0$ with transition intensities*

$$q_{i,i+1} = -q_{i,i} = \beta_i > 0, i \in \mathbb{N}_0 \quad \textit{and} \quad q_{i,j} = 0, \quad \textit{otherwise}$$

is called a pure birth process. Assuming that $P(X(0) = 1) = 1$ then we know that the $n-th$ jump will go from state n to $n+1$ with an average waiting time of $1/\beta_i$. The expected time of the $n-th$ jump will hence be

$$\mathbb{E}[\tau_n|X(0) = 1] = \sum_{i=1}^{n} 1/\beta_i$$

and by monotone convergence

$$\mathbb{E}[\tau_\infty|X(0) = 1] = \lim_{n\to\infty} \mathbb{E}[\tau_n|X(0) = 1] = \lim_{n\to\infty} \sum_{i=1}^{n} 1/\beta_i.$$

In particular, if $\sum_{i=1}^{\infty} 1/\beta_i < \infty$ then τ_∞ has finite mean and we conclude that

$$P(\tau_\infty = +\infty|X(0) = 1) = 0.$$

We conclude that for a pure birth process then $\sum_{i=1}^{\infty} 1/\beta_i < \infty$ implies that there will be infinitely many jumps in finite time (=explosion) with probability 1! □

Using the recurrence criterion given in Theorem 5 we get a simple characterization of recurrent birth-and-death processes.

Theorem 24 (Birth-and-death processes: recurrence) *A birth-and-death process is recurrent if and only if*

$$\sum_{i=1}^{\infty} \frac{\delta_i \cdot \ldots \cdot \delta_1}{\beta_i \cdot \ldots \cdot \beta_1} \quad = \quad \infty. \tag{3.9}$$

Equivalently, a birth-and-death process is transient if and only if

$$\sum_{i=1}^{\infty} \frac{\delta_i \cdot \ldots \cdot \delta_1}{\beta_i \cdot \ldots \cdot \beta_1} \quad < \quad \infty. \tag{3.10}$$

□

Theorem 25 (Birth-and-death processes: positive recurrence) *A birth-and-death process is positive recurrent if and only if*

$$\sum_{i=1}^{\infty} \frac{\beta_{i-1} \cdot \ldots \cdot \beta_0}{\delta_i \cdot \ldots \cdot \delta_1} < \infty \quad \textit{and} \quad \sum_{i=1}^{\infty} \frac{\delta_i \cdot \ldots \cdot \delta_1}{\beta_i \cdot \ldots \cdot \beta_1} = \infty \tag{3.11}$$

□

Remark 6 *We know from equation (3.7) of Theorem 17 that the invariant distribution of a continuous-time Markov chain must satisfy the system of equations*

$$\forall j \in S \quad : \quad \sum_{i \neq j} \pi(i) q_{j,i} = \pi(j)(-q_{j,j}) = \pi(j) q_j.$$

For a birth-and-death process the system of equations takes the form

$$\forall j \in \mathbb{N} \quad : \quad \pi(j-1)\beta_{j-1} + \pi(j+1)\delta_{j+1} = \pi(j)(\beta_j + \delta_j)$$

which turns out to have a solution that can be normalized into a probability vector provided that

$$\sum_{i=1}^{\infty} \frac{\beta_{i-1} \cdot \ldots \cdot \beta_0}{\delta_i \cdot \ldots \cdot \delta_1} < \infty.$$

You are reminded of Theorem 19 which tells us that positive recurrence of a continuous-time Markov chain requires both a solution to (3.7) and that the Markov chain is demonstrated to be recurrent. This is the reason that two conditions must be given in Theorem 25. □

From Example 1 we know that a pure birth process can have infinitely many jumps on a finite time interval (i.e. explosion may occur!). It is not possible to give a simple condition on the transition intensities for a continuous-time Markov chain that determines exactly when explosion is possible. For birth-and-death processes things are little easier as we have the following result.

Theorem 26 (Explosion for a birth-and-death processes) *For a birth-and-death process with intensities*

$$q_{i,i+1} = \beta_i, \quad q_{i+1,i} = \delta_{i+1}, \quad q_{i,i} = -(\delta_i + \beta_i), \quad q_{i,j} = 0 \text{ otherwise}, \quad i, j \in \mathbb{N}_0$$

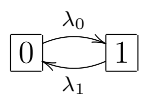
then explosion is possible if and only if

$$\sum_{i=1}^{\infty} \left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \ldots + \frac{\delta_i \cdot \ldots \cdot \delta_1}{\beta_i \cdot \ldots \cdot \beta_0} \right) < +\infty. \tag{3.12}$$

The inequality (3.12) is often referred to as Reuter's criterion for explosion.
□

3.1 Markov chains with two states

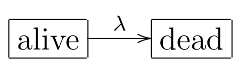
A continuous time Markov chain, $\{X(t)\}_{t\geq 0}$, with two states is given by the transition diagram



For any initial distribution and any choice of the jump intensities λ_0, λ_1 closed form expressions for the distribution of $X(t)$ can be given. There are essentially two types of two-state Markov chains: the absorbing chain and the recurrent chain.

3.1.1 Two-state absorbing Markov chain

In life insurance mathematics a two state Markov chain with one absorbing state is often used to model a single life with one cause of death. This corresponds to the following transition diagram where the states are labelled as *alive* or *dead*



Let $\{X(t)\}_{t\geq 0}$ be the Markov chain given by the diagram above, and assume that the person is *alive* at time $t = 0$.

1. What is (by definition) the distribution of the first (and only) jump time, τ_1?

2. Use the general formula for conditional probabilities

$$P(A|B) = P(A \cap B)/P(B)$$

   to compute $P(\tau_1 > s + t|\tau_1 > t)$ for $s, t > 0$. Give an interpretation of the result.

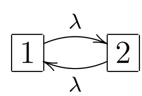
3. Assuming that this is a reasonable model for the life time of a Danish woman and that the mean life duration is 80 years what is then the probability of surpassing the age of 100 years given that one has already passed the age of 80 years?

4. Find the distribution of $X(t)$.

5. Consider n single lifs given by the absorping two-state Markov chain above. Let $N(t) = \sum_{i=1}^{n} 1(X_i(t) = \text{alive})$ be the number of individuals alive at time t (i.e. $X_i(t)$ is the state of i-th person at time t). Find $\mathbb{E}(N(t))$ and discuss what could be the distribution of $N(t)$.

3.1.2 Two-state Markov chain with equal intensities

Let $\{X(t)\}_{t\geq 0}$ be the Markov chain given by the transition diagram



and assume that the initial distribution is given by $P(X(0) = 1) = 1$. Note that we assume that the transition intensities are the same in both states.

1. Find the distribution of the jump times $\tau_1, \tau_2, \tau_3 \ldots,$ by referring to results from other exercises or from Chapter 5.

2. Let $N(t)$ denote the total number of jumps of the chain on the time interval $[0, t]$. Express the distribution of $X(t)$ (i.e. the probabilities $P(X(t) = 1)$ and $P(X(t) = 2)$) in terms of $N(t)$.

3. Find the distribution of $N(t)$ by referring to results from other exercises or from Chapter 5 and use this to obtain a formula for the distribution of $X(t)$.

4. Find the distribution of $X(t)$ under an arbitrary initial distribution given by $P(X(0) = 1) = p \in [0, 1]$. Does the distribution of $X(t)$ depend on t?

5. Show that for any initial distribution then the limits

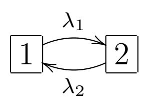
$$\nu_1 := \lim_{t\to\infty} P(X(t) = 1) \quad \text{and} \quad \nu_2 := \lim_{t\to\infty} P(X(t) = 2)$$

exist. Do ν_1 and ν_2 depend on the initial distribution?

6. Discuss what you can conclude from questions 4.-5.

3.1.3 Transition probabilities for a two-state chain

Consider the general two-state Markov chain, $\{X(t)\}_{t\geq 0}$, given by transition diagram



The corresponding transition matrix becomes

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}.$$

The general result says that for $i, j \in \{1, 2\}$ then the transition probabilities

$$P_{i,j}(s) := P(X(t+s) = j | X(t) = i), \quad t, s \geq 0,$$

are given by the entries of the *exponential matrix*

$$\exp(Qt) = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!},$$

where Qt is the matrix obtained by multiplying each entry of Q by t. The purpose of this exercise is to find closed form expressions for $P_{i,j}(s)$ for the general two-state Markov chain.

1. Suppose that we can find an invertible matrix U and a diagonal matrix

$$D = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix}$$

such that $Q = UDU^{-1}$. Argue that $(Qt)^n = U(D^n t^n)U^{-1}$ and deduce that

$$\exp(Qt) = U \begin{pmatrix} \exp(\delta_1 t) & 0 \\ 0 & \exp(\delta_2 t) \end{pmatrix} U^{-1}.$$

2. If U has entries u_{ij} and U^{-1} has entries $u_{ij}^{-1}, i,j \in \{1,2\}$ write down the formula for $P_{i,j}(t)$ which is given as the $ij-$th entry of $\exp(Qt)$ from question 1. above.

The last two questions 3.-4. show that it is always possible to obtain the representation $Q = UDU^{-1}$ given in question 1. above. This implies that for a two-state Markov chain then the transition probabilities

$$P(X(t+s) = j | X(t) = i)$$

are given as linear combinations of two exponential functions $\exp(\delta_i s), i = 1,2$.

3. Find expressions for δ_1, δ_2 (given as the eigenvalues of Q) by solving the equation

$$0 = \det(Q - \delta I) = \det \begin{pmatrix} -\lambda_1 - \delta & \lambda_2 \\ \lambda_2 & -\lambda_2 - \delta \end{pmatrix}.$$

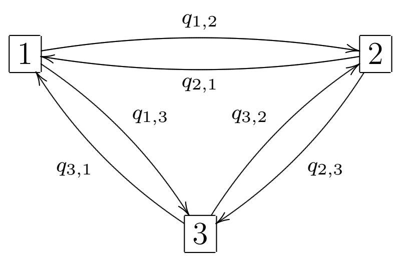
4. For each of the eigenvalues δ_1, δ_2 find the coordinates $u_{1j}, u_{2j}, j = 1,2$, of (right) eigenvectors for Q with eigenvalues δ_j, by solving the system of equations

$$Q \begin{pmatrix} u_{1j} \\ u_{2j} \end{pmatrix} = \rho_j \begin{pmatrix} u_{1j} \\ u_{2j} \end{pmatrix}$$

and verify that $Q = UDU^{-1}$.

3.2 Markov chains with three states

The general three-state Markov chain corresponds to the transition diagram



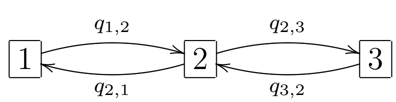
and the transition matrix

$$Q = \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix},$$

where the diagonal elements are given by $q_{ii} := -\sum_{j \neq i} q_{ij}$ such that all row sums equal zero.

3.2.1 Model for interest rates

In this exercise we consider the three-state Markov chain with $q_{1,3} = q_{3,1} = 0$ given by the transition diagram



Note that the model does not allow jumps between states 1 and 3. The model may for instance be used to describe an interest rate that may jump between three different levels but where direct jumps from lowest to highest level do not occur.

1. Write down the transition matrix, Q, for the Markov chain.

2. Find the limit distribution, $\pi(j) = \lim_{n\to\infty} P(X(t) = j)$, for the Markov chain.

3. Write down the matrix of transition probabilities, P, for the discrete-time Markov chain describing the jumps of the chain.

4. Find the invariant distribution for the discrete-time Markov chain given by P, and discuss when the probabilities of questions 2. and 4. coincide.

Assume in the following that $X(0) = 2$ and that all non-zero entries of Q are the same, i.e. $q_{1,2} = q_{2,1} = q_{2,3} = q_{3,2} = q$. Denote by $\tau_n, n \geq 1$, the time of the n-th jump of the Markov chain and let $N(t)$ be the number of jumps of the Markov chain on the interval $[0, t]$.

5. Argue that the distributions of τ_1 and $\tau_2 - \tau_1$ are exponential and find the rate parameters.

6. Using that τ_2 is the sum of the two independent random variables τ_1 and $\tau_2 - \tau_1$ show by applying the formula of Exercise 3.5.1 that τ_2 has density
$$g_2(t) = 2q(\exp(-qt) - \exp(-2qt)), t \geq 0.$$

7. Express the event $(X(t) = 2)$ in terms of events of the form $(N(t) = n)$.

8. Find a formula for the probabilities $P(X(t) = 1)$, $P(X(t) = 2)$, and $P(X(t) = 3)$ in terms of the (unknown) probabilities
$$p_n = P(N(t) = n), n \in \mathbb{N}_0.$$

Assume in the following that $X(0) = 1$ and that all non-zero entries of Q are the same, i.e. $q_{1,2} = q_{2,1} = q_{2,3} = q_{3,2} = q$.

9. Use the ideas from questions 5.-8. to express $P(X(t) = 2)$ in terms of p_n from question 8.

The computations above give us expressions for
$$P_{i,j}(s) = P(X(t+s) = j | X(t) = i)$$
for certain values of $i, j \in S = \{1, 2, 3\}$. Remember that in general the transition probability $P_{i,j}(s)$ is given as the ij-th entry of the exponential matrix $\exp(Qs)$.

10. For what values of $i, j \in S$ did we obtain expressions for $P_{i,j}(s)$ in terms of $P_n = P(N(t) = n)$ by the results of questions 7.-9?

11. Argue that $v_1 = (1, 1, 1)^T$, $v_2 = (1, 0, -1)^T$, and $v_3 = (1, -2, 1)^T$ are right eigenvectors for Q and find the corresponding eigenvalues λ_1, λ_2, and λ_3.

12. Find a matrix, O, and a diagonal matrix D such that

$$QO = OD.$$

13. Find the transition probabilites $P_{i,j}(s)$ by computing $\exp(Qs)$.

    [**Hint:** Argue first that $\exp(Qs) = O \exp(Ds) O^{-1}$.]

3.2.2 Model with two states of health and death

The present model may be used to analyse insurances with payments depending on the state of the insured. In this exercise we assume that the insured starts in state 0 (=*active0*). After a while the insured enters a more favorable state 1 (=*active1*)where she or he stays until death represented by state 2 (=*dead*). To put the model into a more practical setting we might label the states as *active0*, *active1*, and *dead*.

Assume that $X(0) = 0$ and denote by

$$T_1 = \inf\{t > 0 | X(t) = 1\}$$

the time of the jump to state 1. Further, let

$$P_{i,j}(t) = P(X(t+s) = j | X(s) = i), s, t \geq 0,$$

be the transition probabilities of the Markov chain.

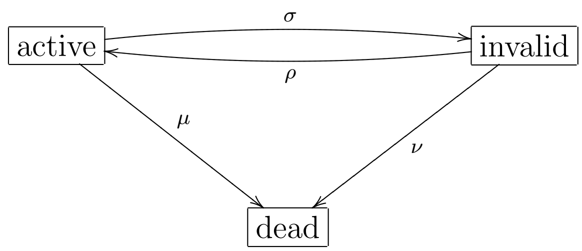
1. Find the matrix, Q, of transition intensities and explain for what $i, j \in S = \{0, 1, 2\}$ it holds that $P_{i,j}(t) = 0, t \geq 0$.

2. Write down the backward differential equation for $P_{0,0}(t)$ and determine
$$P(X(t) = 0).$$
3. What is $P(T_1 > t)$ and the expectation $\mathbb{E}[T_1]$ (Don't do the formal computations!).
4. Find $P_{2,2}(t)$.
5. Write down the backward differential equation for $P_{1,1}(t)$ and determine $P(X(t+s) = 1|X(s) = 1)$.
6. Find $P_{1,2}(t)$.
7. Argue that $P'_{0,2}(t) = -\lambda P_{0,2}(t) + \lambda P_{1,2}(t)$.
8. Define the function
$$h(t) = \exp(\lambda t) P_{0,2}(t)$$
and deduce from question 7. that
$$h'(t) = \lambda \exp(\lambda t) P_{1,2}(t).$$
9. Use the expression for $P_{1,2}(t)$ from question 6. and the boundary condition $h(0) = 0$ to solve the differential equation from question 8. to get a formula for $h(t)$.
10. Find a closed form expression for $P_{0,2}(t)$.

3.2.3 Model for disabilities, recoveries, and death

A model suitable for analysing insurances with payments depending on the state of health of the insured may be given by the three state Markov chain with transition matrix

Consider a portfolio for a person with initial state $X(0) = \textit{active}$ and denote by
$$\tau = \inf\{t > 0 | X(t) = \textit{dead}\}$$
the life length. To the insurance company it is important to know the distribution of τ. Further, if the payments depend on the state of the insured



(*active*/*invalid*) it is important to study the duration of the time spend in each of the states before absorption in the final state *dead*.

To simplify the notation below we relabel the states such that $0 = \textit{active}$, $1 = \textit{invalid}$, and $2 = \textit{dead}$. As usual we denote by

$$P_{i,j}(s) = P(X(t+s) = j | X(t) = i), \quad s, t \geq 0,$$

the transition probabilities of the Markov chain.

1. Write down the transition matrix, Q.

2. For what i, j does it hold that $P_{i,j}(s) = 0$?

3. Write down the forward differential equations for the transition probabilities $P_{i,j}(s)$ for $i = 0$ (=*active*).

Assume in the following questions 4.-9. that $\nu = \mu$.

4. Use question 3. and the fact that $P_{0,0}(t)+P_{0,1}(t)+P_{0,2}(t) = 1$ to obtain a simplified differential equation for $P_{0,2}(t)$ for $\nu = \mu$.

5. Find the distribution of the survival time, τ, for $\nu = \mu$.

   [**Hint:** First note that $P(\tau \leq t) = P_{0,2}(t)$ and then find (or guess!) the solution to the differential equation of question 4.]

6. Use question 3.+5. and that $P_{0,0}(t) + P_{0,1}(t) + P_{0,2}(t) = 1$ to obtain an equation for $P'_{0,1}(t)$ that involves $P_{0,1}(t)$ but no other transition probabilities $P_{i,j}(t)$. Solve the differential equation and find $P_{0,1}(t)$.

The total time spend in the active state (=0) may formally be expressed as

$$S_0 = \int_0^\infty 1(X(t) = 0)dt.$$

In a similar way we define

$$S_1 = \int_0^\infty 1(X(t) = 1)dt$$

that is the time spend in state 1 (=*invalid*). Note that we have the following formula

$$\mathbb{E}[S_i] = \mathbb{E}\left[\int_0^\infty 1(X(t) = i)dt\right] = \int_0^\infty P_{0,i}(t)dt$$

that may be useful for computing $\mathbb{E}[S_i]$ when the transition probabilities are known.

7. Use the results of questions 4.-6. to obtain an expression for $\mathbb{E}[S_i], i = 0, 1$, (-still assuming that $\mu = \nu$).

8. Compute $P(S_1 = 0)$.

9. Use question 5. and 8. to obtain an expression for $P(S_1 = 0, \tau \leq t)$ and compute the conditional probability

$$P(S_1 = 0|\tau \leq t)$$

   that a person dying before time t did not spend any time in the state 1 =*invalid*.

3.2.4 Model for single life with 2 causes of death

In the following questions let $P_{i,j}(t) = P(X(t+s) = j|X(s) = i)$ denote the transition probabilities and assume that $X(0) = alive$. For simplicity we recode the state space, S, such that: $0 = alive, 1 = dead\ 1, 2 = dead\ 2$.

1. Find the intensity matrix of the chain.

2. Determine the communication classes of the chain and argue for each class whether it is recurrent or transient.

3. For which $i, j \in S$ does it hold that $P_{i,j}(t) = 0$ or $P_{i,j}(t) = 1$?

4. Write down the backward equations for the non-constant transition probabilities.

5. Determine $P_{0,0}(t)$ (i.e. the probability of being alive at time t).

6. Find expressions for the remaining transition probabilities.

7. Assuming that $X(0) =$ *alive* find the probability that the process will eventually be absorbed in state *dead1*.

3.2.5 Model with one zero in the transition matrix

We consider a Markov chain $\{X(t)\}_{t\geq 0}$ with transition diagram given by Figure 3.3 and assume that $P(X(0) = 3) = 1$.

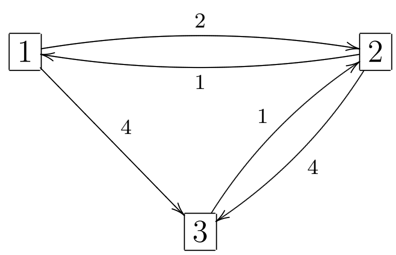


Figure 3.3: Transition diagram for Exercise 3.2.5

To solve questions 6.+7. you might find it useful to know that the equation

$$f'(t) = \alpha f(t) + \beta \exp(\gamma t) + \delta$$

has a solution of the form

$$f(t) = c_1 \cdot \exp(\gamma t) + c_2 \cdot \exp(\alpha t) + c_3$$

for $\gamma \neq \alpha$ and c_1, c_2, c_3 suitable constants.

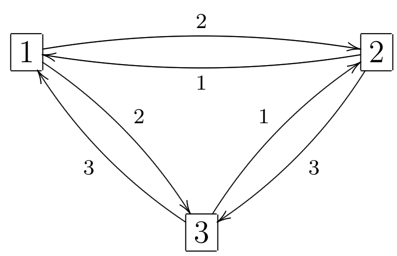
1. Find the infinitesimal generator, Q, (=intensity matrix) for the chain.

2. Find the transition probability matrix for the Markov chain of jumps.

3. Write down the system of equations for the invariant probability $\overline{\pi}$.

4. Compute $\overline{\pi}$.
5. Write down the forward differential equations for $P_{3,j}(t), j = 1, 2, 3$.
6. Use that $P_{3,1}(t) + P_{3,2}(t) + P_{3,3}(t) = 1$ to find $P_{3,3}(t)$.
7. Find $P_{3,1}(t)$ and $P_{3,2}(t)$.
8. Let $\tau_1 = \inf\{t > 0|X(t) = 1\}$ be the time of the first visit to state 1. Determine $\mathbb{E}\tau_1$.

3.2.6 A numerical example



1. Suppose that the chain starts in state 1 and let

$$\tau_1 = \inf\{t > 0|X(t) \neq 1\}$$

be the time of the first jump. What is the mean $\mathbb{E}\tau_1$ of τ_1?

2. Find the matrix, Q, of transition intensities and the transition matrix, P, for the jumps.
3. Argue (briefly) that $P(X(\tau_1) = 2) = 1/2$.
4. What is the distribution of the time between the first, τ_1, and the second, τ_2, jump if $X(\tau_1) = 2$.
5. What is the distribution of the time between the first and the second jump if $X(\tau_1) = 3$.
6. What is the distribution of τ_2?

7. Give an argument that the time, τ_n, of the n-th jump follows a $\Gamma-$distribution and find the parameters.

8. Find the equilibrium distribution $\overline{\pi}$ of the Markov chain.

The transition probabilities, $P(t)$, are given as the entries of the exponential matrix $\exp(Qt)$.

9. Find the characteristic polynomial $g(\lambda) = \det(Q - \lambda I)$ and show that g has one real root (= 0 of course!) and two complex roots (=$-6 \pm i$).

Remark We might continue to find a matrix O of linearly independent (column) eigenvectors for Q and compute the transition probabilities as

$$P(t) = O \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-6t}\exp(it) & 0 \\ 0 & 0 & e^{-6t}\exp(-it) \end{pmatrix} O^{-1}.$$

The eigenvectors will contain complex numbers but since we know that $P_{i,j}(t)$ are probabilities (in particular real numbers) all complex coefficients must cancel when we compute the matrix products. Consequently, since by definition

$$\exp(it) = \cos(t) + i \cdot \sin(t)$$

we can immediately conclude that all transition probabilities take the form

$$P_{i,j}(t) = a_{i,j} + b_{i,j}\exp(-6t) \cdot \cos(t) + c_{i,j}\exp(-6t) \cdot \sin(t),$$

for suitable *real* constants $a_{i,j}, b_{i,j}, c_{i,j}$. We can actually use the fact that $P_{i,j}(t) \to \pi_j$ for $t \to \infty$ to conclude that $a_{i,j} = \pi_j$. Further, we have that

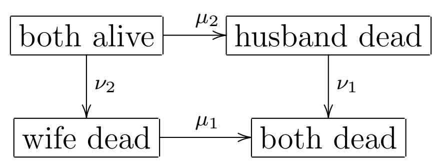
$$\begin{aligned} 1 &= P_{i,i}(0) = a_{i,i} + b_{i,i} + c_{i,i} = \pi_i + b_{i,i} + c_{i,i} \Rightarrow b_{i,i} = 1 - \pi_i - c_{i,i} \\ 0 &= P_{i,j}(0) = a_{i,j} + b_{i,j} + c_{i,j} = \pi_j + b_{i,j} + c_{i,j} \Rightarrow b_{i,j} = -\pi_j - c_{i,j}, \quad i \neq j, \end{aligned}$$

showing that only the constants $c_{i,j}$ need to be determined. Finally, using that $\sum_j P_{i,j}(t) = 1$ we get the additional constraint $\sum_j c_{i,j} = 0$ for any i. A system of equations for the remaining (6!) undetermined constants, $c_{i,j}$, may be obtained by the forward or backward differential equations for $P(t)$.

3.3 Markov chains with finite state space

3.3.1 Model for two lifes

Consider a Markov chain, $\{X(t)\}_{t\geq 0}$, with four states given by transition diagram



One may think of the model as a description of the two lifes of a married couple that wants to buy a combined life insurance and widow's pension policy.

In the following we assume that at both persons are alive at time $t = 0$.

1. Write down the transition matrix of the Markov chain and find the distribution of the first jump time, τ_1.

2. Find the probability that the husband dies before the wife.

3. Find the expected time before the last person dies.

4. Write down the backward differential equations for the transition probabilities needed to find the distribution of $X(t)$.

Consider now the stochastic process obtained by collapsing the states where one person of the couple is alive, i.e. define $\{\tilde{X}(t)\}_{t\geq 0}$ by

$$\tilde{X}(t) = \begin{cases} 0, & X(t) = \text{both alive} \\ 1, & X(t) \in \{\text{husband dead,wife dead}\} \\ 2, & X(t) = \text{both dead} \end{cases}$$

In general $\{\tilde{X}(t)\}_{t\geq 0}$ is not a Markov chain and we shall try to argue why.

5. Compute

$$P(X(3t) = \text{both dead}, X(2t) = \text{wife dead}, X(t) = \text{wife dead})$$

and

$$P(X(3t) = \text{both dead}, X(2t) = \text{husband dead}, X(t) = \text{husband dead})$$

using the formula

$$\begin{aligned} & P(X(3t) = k, X(2t) = j, X(t) = i) \\ = \; & P(X(t) = i) \cdot P(X(2t) = j | X(t) = i) \cdot P(X(3t) = k | X(2t) = j). \end{aligned}$$

6. Compute $P(\tilde{X}(3t) = 2, \tilde{X}(2t) = 1, \tilde{X}(t) = 1)$ by writing the event as a disjoint union of the sets in question 5.

7. Use a similar trick as in questions 5.-6. to compute

$$P(\tilde{X}(3t) = 2, \tilde{X}(2t) = 1, \tilde{X}(t) = 0).$$

8. Write the set $(\tilde{X}(2t) = 1, \tilde{X}(t) = 1)$ as a disjoint union of events involving $\{X(t)\}$ and use question 6. to compute

$$P(\tilde{X}(3t) = 2 | \tilde{X}(2t) = 1, \tilde{X}(t) = 1).$$

9. Use question 7. and the ideas of question 8. to compute

$$P(\tilde{X}(3t) = 2 | \tilde{X}(2t) = 1, \tilde{X}(t) = 0).$$

10. Argue that in general $\{\tilde{X}(t)\}_{t \geq 0}$ is not a Markov chain on $\{0, 1, 2\}$.

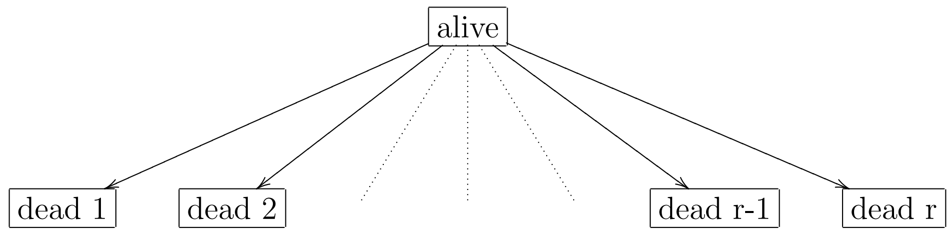
11. Under what restriction of the model parameters does it hold that $\{\tilde{X}(t)\}_{t \geq 0}$ is a Markov chain.

3.3.2 Single life with r causes of death

We consider in this exercise a stochastic model for a single life with r possible causes of death. We assume that $X(0) = \textit{alive}$ and that the waiting time, W_i, to death by cause i has density $f_i, i = 1, \ldots, r$.

Clearly, the life duration of the person is given by

$$W = \min(W_1, \ldots, W_r).$$



We assume that the waiting times $W_1, \ldots, W_r$ are independent. The description above leads to a stochastic process $\{X(t)\}_{t\geq 0}$ with state space $S = \{alive, dead\ 1, dead\ 2, \ldots, dead\ r\}$ through the definition

$$X(t) = \begin{cases} alive & , \quad t < W \\ dead\ i & , \quad t \geq W \text{ and } W_i = W = \min(W_1, \ldots, W_r) \end{cases}$$

The purpose of the exercise is to examine for what conditions on the densities f_i, for W_i that the process $\{X(t)\}_{t\geq 0}$ becomes a Markov chain.
Remember that the Markov property writes out as

$$P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}, \ldots, X(t_1) = i_1) = P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}),$$

for any choice of $0 \leq t_1 < \ldots < t_n$ and $i_1, \ldots, i_n \in S$.

1. Argue that when considering the Markov property for $\{X(t)\}_{t\geq}$ we only need to consider sets of the form above with all $i_1 = \ldots = i_n = alive$ or
$$i_1 = \ldots = i_j = alive, i_{j+1} = \ldots = i_n = dead\ i$$
for some $j = 1, \ldots, n-1$ and $i = 1, \ldots, r$.

2. Explain why the probabilitiy of an event of the form given in question 1. above may be expressed as
$$P(t_j < W_i \leq t_{j+1}, W_i \leq W = \min(W_1, \ldots, W_r)),$$
for some $j = 1, \ldots, n$ (for $j = n$ interpret t_{n+1} as $+\infty$).

3. Use the independence of $W_1, \ldots, W_r$ to argue that
$$P(t_j < W_i \leq t_{j+1}, W_i = W = \min(W_1, \ldots, W_r)$$
$$= \int_{t_j}^{t_{j+1}} f_i(w) \prod_{j=1:j\neq i}^{r} \left[\int_w^\infty f_j(w_j) dw_j \right] dw.$$

We assume in the following that the waiting times W_i are exponentially distributed with rate $\lambda_i > 0, i = 1, \ldots, r$, i.e. that

$$f_i(t) = \lambda_i \exp(-\lambda_i t), \quad t \geq 0.$$

4. Use questions 1.-3. to show that for $0 \leq t_1 < \ldots < t_n, n \in \mathbb{N}$ then

$$\begin{aligned} & P(X(t_n) = \ldots = X(t_{j+1}) = \textit{dead } i, X(t_j) = \ldots = X(t_1) = \textit{alive}) \\ = \; & \frac{\lambda_i}{\lambda} \left((\exp(-t_j \lambda) - \exp(-t_{j+1} \lambda) \right), \end{aligned}$$

   for $j = 1, \ldots, n$ (with $t_{n+1} = +\infty$) where we let $\lambda = \sum_{j=1}^{r} \lambda_j$.

5. Use question 4. to compute $P(X(t) = \textit{dead } i), i = 1, \ldots, r$.

6. Use question 4. to compute $P(X(t) = \text{alive})$. What is the probability $P(\tau_1 > t)$ that the first (and only) jump, τ_2, occurs after time t.

7. Based on the result in question 4. compute the conditional probability

$$P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}, \ldots, X(t_1) = i_1)$$

   for the three cases

   - $i_1 = \ldots = i_n = \textit{alive}$
   - $i_1 = \ldots = i_{n-1} = \textit{alive}, i_n = \textit{dead } i$
   - $i_1 = \ldots = i_j = \textit{alive}, i_{j+1} = \ldots = i_{n-1} = i_n = \textit{dead } i, j \leq n-2$

   and show the result only depends on $i_1, \ldots, i_{n-2}$. This demonstrates that $\{X(t)\}_{t \geq 0}$ is a Markov chain.

8. Find the transition diagram for the Markov chain in the usual representation in terms of the matrix of transition intensities.

Remark The exercise gives an alternative interpretation of the time dynamics for a (finite state) Markov chain. Suppose that we are at state $i \in S$ at time t. For each state $j \neq i$ that can be reached from i (i.e. $q_{ij} > 0$) generate independent exponentially distributed random variables (alarm clocks!) W_j with rate q_{ij}. The time when the first bell rings, $W = \min\{W_j\}$, determines the time and the target of the next jump. It can be shown that the waiting time to the jump is exponentially distributed with rate $q_i = \sum_{j \in S} q_{ij}$ and that the probability of the jump being directed to state j is given by $q_{ij}/q_i, j \neq i$.

3.3.3 Forward differential equations for four state chain

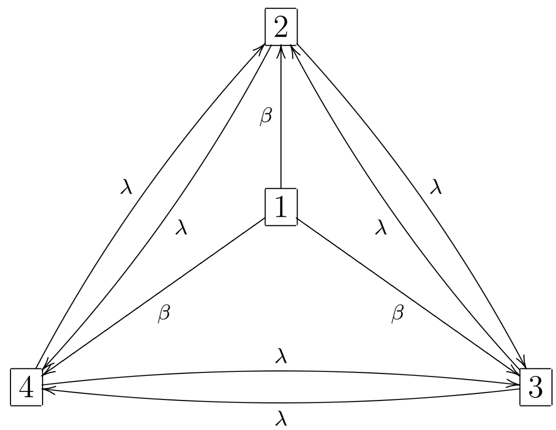


Figure 3.4: Transition diagram of four state Markov chain of Exercise 3.3.3.

Consider the Markov chain on Figure 3.4 with state space $S = \{1, 2, 3, 4\}$, where the initial distribution is given by $P(X(0) = 1)1$.

1. Find the intensity matrix Q of the chain.

2. Write down the system of equations for the invariant probability, $\overline{\pi}$, of the chain and find $\overline{\pi}$.

3. Find the transition matrix P for the jumps of the chain.

4. Find the distribution of the first jump time and use this to find an expression for $P_{1,1}(t)$.

5. Write down the forward differential equation, $P'(t) = P(t)Q$, for the transition probability $P_{1,2}(t) = P(X(t) = 2|X(0) = 1)$.

6. Solve the differential equation from question 5. by using the result from question 4. and that by symmetry we must have $P_{1,2}(t) = P_{1,3}(t) = P_{1,4}(t)$. Try also to give an even simpler derivation of $P_{1,2}(t)$ refering only to symmetry but without using the differential equation.

7. Write down the forward differential equation for $P_{2,2}(t)$.

8. Using that $P_{2,1}(t) = 0$ and $\sum_{j\in S} P_{2,j}(t) = 1$ show that the equation from question 7. has a solution of the form

$$P_{2,2}(t) = c_1 + c_2 \exp(-3\lambda t)$$

and determine the constants c_1, c_2.

9. Find the remaining transition probabilities.

[**Hint:** For an easy solution to this question start by listing transition probabilities that are zero and transition probabilities that must be the same due to symmetry. You will probably also find it useful to remember that the rows of $P(t)$ sum to one.]

3.3.4 Time to absorption

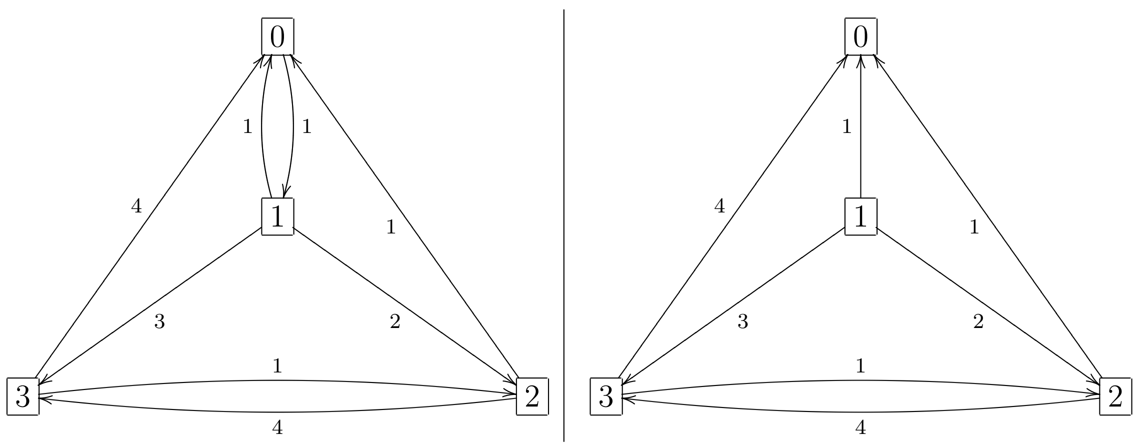


Figure 3.5: Transition diagram for Exercise 3.3.4: left diagram should be used for questions 1.-4. and right diagram for the remaining questions 5.-13.

Suppose that the chain starts in state 1, i.e. $P(X(0) = 1)$.

1. Find the intensity matrix, Q, of the chain.

2. Write down the system of equations for the invariant probability $\overline{\pi}$ of the chain and find $\overline{\pi}$.

3. What is the distribution and the mean of the time to the first jump.

4. Find the transition matrix, P, for the jumps of the chain.

In the rest of the exercise we exclude the possibility that the chain can jump from state 0 to state 1. This situation corresponds to the transition diagram on the right of Figure 3.5.

5. Write down the transition matrix for the jumps of the modified version of the chain. By convention for an absorbing state i let us make the convention that $P_{i,i} = 1$.

6. Considering only the Markov chain of jumps compute (using a computer) the expected number of times the chain will visit state 2 before absorption in state 0. Answer the same question for state 3.

7. Write down the matrix, Q, of transition intensities for the modified version of the continuous-time Markov chain.

8. Verify that $v_1 = (1,1,1,1)^T$ and $v_2 = (0,7/3,2,1)^T$ are (right) eigenvectors for Q with eigenvalues $\lambda_1 = 0$ and $\lambda_2 = -3$.

9. Verify that $v_3 = (0,1,0,0)^T$ and $v_4 = (0,1,-2,1)^T$ are (right) eigenvectors for Q and find the corresponding eigenvalues λ_3 and λ_4.

10. Let O be the 4×4 matrix with columns given by v_i (i.e. $O = (v_1, v_2, v_3, v_4)$). Use the fact that the transition probabilities, $P(t) = (P_{i,j}(t))_{i,j\in S}$, are given by the exponential matrix

$$\exp(Qt) = O \begin{pmatrix} \exp(\lambda_1 t) & 0 & 0 & 0 \\ 0 & \exp(\lambda_2 t) & 0 & 0 \\ 0 & 0 & \exp(\lambda_3 t) & 0 \\ 0 & 0 & 0 & \exp(\lambda_4 t) \end{pmatrix} O^{-1}$$

to compute the probability $P_{1,0}(t)$.

Hint: You can use without proof that

$$O^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3/4 & 0 & 1/4 & 1/2 \\ 1 & 1 & -1/3 & -5/3 \\ -1/4 & 0 & -1/4 & 1/2 \end{pmatrix}.$$

11. Define the time of the first visit to state 0

$$T = \inf\{t > 0 | X(t) = 0\}$$

and argue that $P(T \leq t) = P(X(t) = 0)$.

12. Use (without proof) that the expectation of the nonnegative random variable T may be expressed as

$$\mathbb{E}[T] = \int_0^\infty P(T > t)dt$$

to compute the expected time to absorption in state 0 when the chain is started at state 1 (i.e. $P(X(0) = 1) = 1$).

13. Use the following heuristic argument to compute the expected time to absorption in state 0 : First compute the expected number of the time periods where the chain visits states $1, 2,$ and 3. Then multiply the expected number of visits in each state with the average waiting time before the chain jump to another state. This gives you the expected time spend in each state.

[**Hint:** You already computed many of the necessary quantities in previous questions.]

3.4 Markov chains on countable state spaces

3.4.1 Pure death process with constant intensity

A birth-and-death process is a continuous time Markov chain on $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ that moves only in jumps of size one. The process may describe the size of a population and a jump between states i and $i+1$ is interpreted as a *birth* whereas jumps from i to $i-1$ corresponds to a *death*.

Denoting by $q_{i,j}$ the transition intensities from state i to state j the birth-and.death process has the following structure

$$q_{ij} = \begin{cases} \beta_i & , \quad j = i+1, i \geq 0 \\ \delta_i & , \quad j = i-1, i \geq 1 \\ 0 & , \text{otherwise} \end{cases}$$

for suitable nonnegative birth- and death intensities $\beta_i, \delta_i \geq 0$. Many important stochastic processes belong to the class of birth-and-death processes and may be obtained by imposing various restrictions on the birth- and death intensities.

Assume that the Markov chain $\{X(t)\}_{t\geq 0}$ is a birth- and death process with initial distribution $\rho_i = P(X(0) = i)$.

1. Draw a part of the transition diagram for the Markov chain under the assumption that all $\beta_i, \delta_i > 0$.

2. Find under the assumption of question 1. the transition probabilities for the corresponding discrete time Markov chain of jumps for $\{X(t)\}_{t\geq 0}$.

3. What choice of initial distribution and birth- and death intensities implies that $\{X(t)\}_{t\geq 0}$ is a Poisson process?

The pure death process is characterized by all the birth intensities, β_i, being equal to zero. In the following we consider a pure death process with $\delta_k > 0, k \geq 1$, and initial distribution $P(X(0) = k) = 1$ for some $k \geq 2$.

4. What is the distribution of the first jump time

$$\tau_1 = \inf\{t > 0 | X(t) \neq k\}?$$

5. Find $P(X(t) = k)$.

6. Assuming that all death intensities are the same, $\delta_i = \delta > 0$, what is then the distribution of the time, τ_2, of the second jump of the chain?

7. Under the assumption of question 6. one may argue that $\{X(t)\}_{t\geq 0}$ behaves like a modified Poisson process $\{N(t)\}_{t\geq 0}$ with downward jumps of intensity $\delta > 0$ until the time of the $k-$th jump. Use this to compute $P(X(t) = j), j = 1, 2, \ldots, k-1$.

8. Find $P(X(t) = 0)$ under the assumption of question 6.

3.4.2 Linear birth-and-death process

The linear birth-and-death process is a continuous-time Markov process on $\mathbb{N}_0$ with birth intensities $\beta_i = i\beta$ and death intensities $\delta_i = i\delta$. It may be thought of as a model for a population where at any time an individual dies with intensity $\delta > 0$ and gives rise to a birth with intensity $\beta > 0$.

1. Find the communicating classes of the linear birth-and-death process.

2. Assume that $P(X(0) = 1) = 1$ and let $\tau_1 = \inf\{t > 0 | X(t) \neq 1\}$ be the time of the first jump. Find the probability $P(\tau_1 > 1)$ and the distribution, $P(X(\tau_1) = i), i \geq 0$, of the chain observed just after the first jump.

3. Let $T = \inf\{t > \tau_1 | X(t) = 1)$ be the time of the first return to state 1. Use the result of question 2. to get an upper bound for the probability $P(T < +\infty | X(0) = 1)$. Discuss what you can conclude from this observation.

4. Still assuming that $P(X(0) = 1) = 1$ argue that

$$P(X(1) = 0) > \tfrac{\delta}{\delta+\beta}(1 - \exp(-(\delta + \beta))).$$

For the rest of the exercise we modify the birth intensities such that $\beta_i = i\beta + \lambda$ for some $\beta, \lambda > 0$. The resulting model has a very nice interpretation as a linear birth-and-death process with immigration intensity λ.

5. Argue that the linear birth-and-death process with immigration is irreducible.

6. Let $\{Y(n)\}_{n\geq 0}$ be the discrete-time Markov chain of jumps. Find the transition probabilities of $\{Y(n)\}_{n\geq 0}$.

The following questions seek to clarify for what values of the parameters that a linear birth-and-death process is transient, null-recurrent and positive recurrent. A the written exam you should directly apply the results in Chapter 3.0.4 to answer questions 8., 11. and 12. below. Questions 7., 9. and 10. are only relevant for those of you who want to understand better how to arrive at the main results in Chapter 3.0.4.

7. Write down the system of equations in Theorem 5 of Chapter 3.0 where you use $i = 0$ as *fixed state*.

It is rather technical to write down a complete solution to questions 8. and 11.-12. below covering all choices of the parameters β, δ, λ. To make things a bit more easy try to consider first the cases $\beta > \delta$ and $\beta < \delta$.

8. Use the system of equations from question 7. (or some other argument) to determine for what choices of $\beta, \delta, \lambda > 0$ that the linear birth-and-death process with immigration is recurrent or transient.

   [**Hint:** Use the system of equations from question 7. to deduce that

   $$\alpha(j+1) = \alpha(1)\left\{1 + \sum_{k=1}^{j} \frac{k\delta \cdot \ldots \cdot \delta_1}{(\lambda + k\beta) \cdot \ldots \cdot 1\beta}\right\}, \quad j \geq 1.$$

   A simpler approach is just to apply a suitable result in Chapter 3.0.4.]

9. Show that an invariant probability vector $\pi = (\pi_i)_{i \in \mathbb{N}_0}$ for the linear birth-and-death process with immigration must satisfy the following system of equations

   $$\begin{aligned} 0 &= \delta\pi_1 - \lambda\pi_0 \\ 0 &= ((i-1)\beta + \lambda)\pi_{i-1} + (i+1)\delta\pi_{i+1} - (i\beta + i\delta + \lambda)\pi_i, i \geq 1. \end{aligned}$$

10. Verify that the vector $\nu = (\nu_i)_{i \in \mathbb{N}_0}$ where

    $$\nu_i = \nu_0 \cdot \prod_{k=1}^{i} \tfrac{(k-1)\beta + \lambda}{k\delta}, i \geq 1,$$

    solves the system of equations from question 9.

11. Determine for what values of $\beta, \delta, \lambda > 0$ that the solution , ν, of question 10. can be normalized into a probability vector π.

12. For what choice of the parameters β, δ, λ is the birth-and-death process with immigration null-recurrent?

    [**Hint:** If we already know that the chain is recurrent then the chain is positive recurrent if and only if there exists a probability vector solving the system of equations from question 10.]

Answer the following two questions 13.-14. for the three cases $\lambda = \beta = 2\delta$, $\lambda = \beta = \delta$, and $\lambda = \beta = \delta/2$.

13. Find $\lim_{t\to\infty} P(X(t) = i)$ for $i \geq 0$ assuming that $P(X(0) = 1) = 1$.

14. What is the long run average population size?

3.4.3 Queueing systems

There is an entire branch of applied probability that deals with mathematical modeling of queueing systems. In this exercise we show by an example how continuous-time Markov chains may be used to model the number of customers in a queueing system. Throughout the exercise we assume that new customers arrive to the system according to a Poisson process with intensity $\beta > 0$ independently of the state of the system.

We consider initially the single server queue where customers are served according to the first-come-first-served queueing discipline. Upon arrival at the service desk the service time distribution is assumed to be exponential with rate $\delta > 0$ no matter how many customers are waiting in line. One can show (but you are not supposed to do so!) that under the given assumptions then the number, $\{X(t)\}_{t\geq 0}$, of customers present in the system constitutes a continuous-time Markov chain on $\mathbb{N}_0$ with transition intensities

$$q_{i,j} = \begin{cases} \beta & , \quad j = i+1, i \geq 0 \\ \delta & , \quad j = i-1, i \geq 1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

1. Argue that the chain is a birth-and-death process.
2. Write down the system of equations that must be satisfied for an invariant probability vector $\pi = (\pi_i)_{i\in\mathbb{N}_0}$. Find the invariant distribution, π, of the chain for the case where $\beta < \delta$.
3. Assuming that $\beta < \delta$ compute the (long run) average number of customers in the queue.
4. What is the distribution of the waiting time before arrival to the service desk if 4 customers are waiting in front of you when you arrive to the queueing system? (You are not expected to do any computations here!)

We now assume that the customers are served in their order of arrival by two servers with exponentially distributed service time distributions of (possibly different) rates $\delta_1, \delta_2 > 0$. If a customer arrives at an empty system she or he is by default served at service desk number 1. With some effort one can show that the system may be regarded as a continuous-time Markov chain on the state space

$$S = \{0, 1:0, 0:1, 2, 3, 4, \ldots\}.$$

This needs a little more explanation: state 0 means that no customers are present, state 0 : 1 means that one customer is being served at service desk 2 while service desk 1 is vacant. Similarly, state 1 : 0 represents the situation where service desk 1 is occupied and desk 2 is vacant. States $2, 3, 4, \ldots$ refer to situations where at least two customers are present of which two are currently being served at service desks 1 and 2.

The transition diagram (without transition intensities) is given on Figure 3.6.

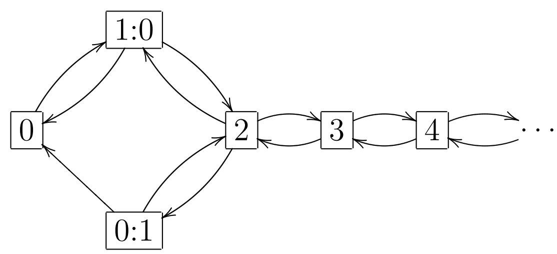


Figure 3.6: Transition diagram for the Markov chain considered in questions 5.-14. of Exercise 3.4.3.

5. What is the intensity $q_{0:1,0}$ of a jump from state 0 : 1 to 0?

6. What is the intensity $q_{1:0,0}$ of a jump from state 1 : 0 to 0?

7. Find the intensities $q_{0:1,2}, q_{1:0,2}, q_{0,0:1}$.

8. Argue that the intensity of a jump from state 3 to 2 equals $\delta_1 + \delta_2$.

9. Draw the transition diagram of the Markov chain with all intensities.

10. Argue very carefully that an invariant probability vector $\pi = (\pi_i)_{i \in S}$ must satisfy the system of equations

$$\begin{aligned}
0 &= \delta_1 \pi_{1:0} + \delta_2 \pi_{0:1} - \beta \pi_0 \\
0 &= \beta \pi_0 + \delta_2 \pi_2 - (\beta + \delta_1) \pi_{1:0} \\
0 &= \delta_1 \pi_2 - (\beta + \delta_2) \pi_{0:1} \\
0 &= \beta \pi_{1:0} + \beta \pi_{0:1} + (\delta_1 + \delta_2) \pi_3 - (\delta_1 + \delta_2 + \beta) \pi_2 \\
0 &= (\delta_1 + \delta_2) \pi_{i+1} + \beta \pi_{i-1} - (\beta + \delta_1 + \delta_2) \pi_i, \quad i \geq 3.
\end{aligned}$$

11. Show that for any constant c there is a vector, $\pi = (\pi_i)_{i \in S}$, with

$$\pi_i \;=\; c\left(\tfrac{\beta}{\delta_1+\delta_2}\right)^{i-2}, \quad i \geq 2$$

that solves the system of equations from 10 and derive expressions for $\pi_{0:1}$, $\pi_{1:0}$, and π_0.

12. For what values of the parameters $\beta, \delta_1, \delta_2$ can the solution in question 11. be normalized into an invariant probability vector? (You don't need to find a closed form expression for c to answer this question!)

13. Consider the case where $\delta_1 = \delta_2 = \delta$ and $\beta = \delta/2$. Find the invariant probability vector from questions 11.-12. and compute the (long run) average number of customers for the two-server queue.

    [Hint: You can use (or verify) that $c = 3/40$.]

14. Still assuming that $\delta_1 = \delta_2 = \delta$ and $\beta = \delta/2$ discuss how much the (long run) average queue length decreased by the introduction of the second server compared to the single server system (question 1.-4.).

We finally consider the case where arriving customers physically lines up in two different queues. Upon arrival a customer enters the shortest of the two lines. If there are the same number of customers in each queue any customer by default enters the queue nearest to the entrance of the building (let us call this queue number 1). If at any time the difference between the length of two queues is two the last customer in the longest queue will instantly switch to the last position in the shorter queue. The purpose of the following questions is to study the differences between the two-line queueing disciplin and the one line first-come-first-served disciplin considered in questions 5.-14.

It is possible to show that the joint number of customers, $\{(X_1(t), X_2(t))\}_{t \geq 0}$, in the two queues is a continuous-time Markov chain on $\mathbb{N}_0 \times \mathbb{N}_0$.

15. Technically speaking the state space of the chain is much smaller than $\mathbb{N}_0 \times \mathbb{N}_0$ because a large number of the states will never be visited by the chain. What is the trimmed version, S, of the state space that represents the truly possible states of the queueing system?

16. Draw the transition diagram (with transition intensities) of the Markov chain that displays only the trimmed state space, S, from question 15. You probably need to be careful to get all the transition intensities right in particular for jumps between states $(i, i+1)$ and (i, i) or between states $(i+1, i)$ and (i, i).

17. Argue that an invariant probability vector $\pi = (\pi_{i,j})_{(i,j)\in S}$ must satisfy the following system of equations

$$\begin{aligned}
0 &= \delta_2\pi_{0,1} + \delta_1\pi_{1,0} - \beta\pi_{0,0} \\
0 &= \beta\pi_{0,0} + \delta_2\pi_{1,1} - (\delta_1 + \beta)\pi_{1,0} \\
0 &= \delta_1\pi_{1,1} - (\delta_2 + \beta)\pi_{0,1} \\
0 &= \beta(\pi_{i-1,i} + \pi_{i,i-1}) + (\delta_1 + \delta_2)(\pi_{i+1,i} + \pi_{i,i+1}) - (\beta + \delta_1 + \delta_2)\pi_{i,i}, \quad i \geq 1, \\
0 &= \beta\pi_{i,i} + \delta_2\pi_{i+1,i+1} - (\beta + \delta_1 + \delta_2)\pi_{i+1,i}, \quad i \geq 1, \\
0 &= \delta_1\pi_{i+1,i+1} - (\beta + \delta_1 + \delta_2)\pi_{i,i+1}, \quad i \geq 1.
\end{aligned}$$

18. Verify that for any constant c there is a vector, $\pi = (\pi_i)_{i\in S}$, with

$$\pi_{i,i} = c\left(\tfrac{\beta^2}{(\delta_1+\delta_2)^2}\right)^i, \quad i \geq 1,$$

that solves the system of equations from question 17 and derive expressions for the remaining coordinates of π.

[Hint: Start by plugging in to the last equation of question 17. to get an expression for $\pi_{i,i+1}$ and do not try to find the constant c.]

19. For what values of the parameters $\beta, \delta_1, \delta_2$ can the solution of 18. be normalized into an invariant probability vector? (You don't need to find a closed for expression for c to answer this question!)

20. Consider the case where $\delta_1 = \delta_2 = \delta$ and $\beta = \delta/2$. Find the invariant probability vector from questions 17.-18. and compute the (long run) average number of customers present in the queueing system.

21. Are there any reason to prefer one of the two suggested two-server queueing disciplines to the other from the customers point-of-view? To answer the question you may find it useful to include a discussion of your results from questions 13. and 20.

22. The total service capacity (per time unit) of a queueing system with two servers is given by the sum $\delta_1 + \delta_2$. Which of the queueing systems with two servers exploit the service capacity in the must efficient way? (Don't do any computations!)

23. Try to do some numerical computations to examine if there are any differences betweeen the two suggested two-server queueing systems when $\delta_1 \neq \delta_2$. Look at the problemer from the customers point-of-view.

Comments: The results of this exercise do not carry over to real life queueing systems for several reasons of which we shall mention a few: the unrealistic assumption of exponentially (=memoryless) distributed service times and intervals between arriving customers, the assumption of customers arriving at the same rate at all times, and the independence of the service distributions on both time and on the number of customers already present in the queueing system.

It is trival that queues build up if the (average) service capacity is lower than the average rate of arriving customers. Another important lecture you may learn by digging further into the field of queueing theory is that even for a sufficient average service capacity queues are caused by variation in interarrival times and service times. The generel message is that inducing more variation deteriorates the performance of a queueing system.

3.4.4 Positive recurrence and null-recurrence

We consider a Markov chain, $\{X(t)\}_{t\geq 0}$, on $S = \{0, 1, 2, \ldots\}$ with transition intensities

$$q_{0,n} = p_n > 0, q_{n,n-1} = \delta_n > 0, n > 0, \quad q_{i,j} = 0 \text{ for any other } i \neq j$$

where $\sum_n p_n = 1$.

1. Find the transition probabilities for the embedded Markov chain of jumps.

2. Argue that $\{X(t)\}_{t\geq 0}$ is recurrent.

So far we have demonstrated that the Markov chain $\{X(t)\}_{t\geq 0}$ and the embedded Markov chain of jumps are always recurrent no matter the values of $p_n > 0$ and $\delta_n > 0$. The purpose of the following is to show that all four combinations of positive recurrence and null-recurrence for $\{X(t)\}_{t\geq 0}$ and the embedded jump chain may occur.

NN $\{X(t)\}_{t\geq 0}$ and the embedded jump chain are null-recurrent.

NP $\{X(t)\}_{t\geq 0}$ is null-recurrent and the embedded jump chain is postive recurrent.

PN $\{X(t)\}_{t\geq 0}$ is positive recurrent and the embedded jump chain is null-recurrent.

PP $\{X(t)\}_{t\geq 0}$ and the embedded jump chain are positive recurrent.

Find out how the four cases listed above correspond to the four sets of parameters described in questions 3.-6. below.

3. $p_n = (1-p)p^{n-1}, 0 < p < 1$, and $\delta_n = \delta > 0$
4. $p_n = c/n^2$ and $\delta_n = \delta > 0$
5. $p_n = (1-p)p^{n-1}$ and $\delta_n = (1-p/2)^{-1}(p/2)^n$ where $0 < p < 1$
6. $p_n = c/n^2$ and $\delta_n = n(n+1)$

   [**Hint:** We did already study the embedded Markov chain of jumps in Exercise 2.4.6.]

3.4.5 More examples of birth-and-death processes

We consider in this exercise four different birth-and-death processes. The purpose of this exercise is to get some experience using the results stated in Section 3.0.4.

1. Show that the birth-and-death process with intensities

$$q_{i,i+1} = \beta_i = i+1, \quad q_{i+1,i} = \delta_{i+1} = 1, \quad i \geq 0$$

   is transient.

2. Show that the birth-and-death process with intensities

$$q_{i,i+1} = \beta_i = i+1, \quad q_{i+1,i} = \delta_{i+1} = i+1, \quad i \geq 0$$

   is null-recurrent.

3. Show that the birth-and-death process with intensities

$$q_{i,i+1} = \beta_i = i + 1, \quad q_{i+1,i} = \delta_{i+1} = i + 3, \quad i \geq 0$$

is positive recurrent.

4. Show that for $q < p < cq$ then the birth-and-death process with intensities

$$q_{i,i+1} = \beta_i = c^i p, \quad q_{i+1,i} = \delta_{i+1} = c^{i+1} q, \quad i \geq 0$$

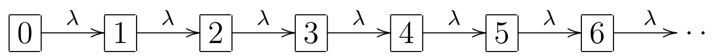
is transient *and* there exists a probability vector $\overline{\pi} = (\pi(i))_{i \in S}$ solving the system

$$\sum_{i \in S} \pi(i) q_{i,j} = 0, \quad j \in S.$$

[Hint: Use Section 5.6 on linear recurrence equations.]

Remark Question 4. shows that there exists a probability vector satisfying the necessary condition of Theorem 17 for an invariant distribution. However, since the Markov chain is transient the invariant distribution does not exist. One can show using Reuter's criterion from Theorem 26 that explosion may occur for the birth-and-death process given in Question 4.

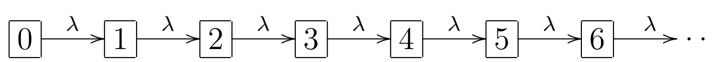
3.5 The Poisson process



Let $\{X(t)\}$ be the stochastic process given by the diagram above. The process takes values in the state space $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, it starts at state zero, $X(0) = 0$, and it moves in upward jumps of size 1. A stochastic process with this property is commonly refered to as a *counting process*. The *Poisson proces* is the counting process with independent identically exponentially distributed waiting times between jumps. The rate, λ, of the exponential distribution is called the intensity of the Poisson process.

3.5.1 Basic properties of the Poisson process

The *Poisson process* with intensity λ is the continuous-time and time-homogeneous Markov chain on $\mathbb{N}_0$, given by the following transition diagram



In particular the times between jumps are independent and exponentially distributed with density function

$$f(s) = \lambda \exp(-\lambda s), \quad s \geq 0.$$

Further, we denote by $\tau_1, \tau_2, \ldots$ the jump times of $\{X(t)\}$.

1. Compute $\mathbb{E}[\tau_1], P(X(t) = 0)$, and $P(X(t) \geq 1)$.

2. For $0 < s < t$ compute $P(X(s) = 0, X(t) = 0)$

For non-negative independent random variables V with density g and W with density h then the density of the sum $Y = V + W$ has density given by

$$k(y) := h * g(y) := \int_0^y h(y - v) g(v) dv, \quad y \geq 0.$$

3. Find the distribution (=density) of τ_2 by using that τ_2 is the sum of two independent exponential distributions with rate parameter λ.

4. Compute $P(X(t) \geq 2)$.

5. Verify by induction that the time, τ_n, of the jump to state n follows a distribution with density

$$f_n(s) = \frac{\lambda(\lambda s)^{n-1}}{(n-1)!} \exp(-\lambda s), \quad s \geq 0.$$

6. Compute $P(X(t) \geq n)$ and $P(X(t) = n)$.

   [**Hint:** Use without proof that

$$P(\tau_n \leq t) = 1 - \sum_{k=0}^{n-1} \tfrac{(\lambda t)^k}{k!} \exp(-\lambda t)$$

   or you may even try to prove the formula by induction.]

7. What is the name of the distribution of τ_n and $X(t)$?

3.5.2 Advanced exercise involving the Poisson process

In this exercise we consider a Poisson process $\{X(t)\}$ with intensity λ. You may use that from Exercise 3.5.1 we know the distribution of $X(t)$ i.e. the probabilities $P(X(t) = n), n \in \mathbb{N}$.

The purpose of this exercise is to study further the times of the jumps of the Poisson process which we will denote by $\tau_1, \tau_2, \tau_3, \ldots$. From the previous Exercise 3.5.1 we know the distribution of τ_n. In this exercise we consider what can be said about the distribution (=location) of the n first jumps times given that we know that $X(1) = n$ i.e. that exactly n jumps occured on the time interval $[0, 1]$.

For simplicity we consider only the distribution of τ_1 by asking the following question: given that we know that exactly one jump happened before time 1 (i.e. $X(1) = 1$) when was the must likely time on $[0, 1]$ for the jump, τ_1, to happen? Clearly, the conditional distribution of the first jump time, τ_1, given that $X(1) = 1$ is a distribution on the interval $[0, 1]$. The purpose of the following questions 1.-8. is to compute $P(a < \tau_1 \leq b | X(1) = 1)$ for $0 \leq a \leq b \leq 1$.

1. Try to argue, for instance on a suitable figure, that

$$(\tau_1 \leq b, X(1) = 1) = (X(b) = 1, X(1) = 1), \quad \text{for } 0 \leq b \leq 1.$$

2. Find the probability that $P(X(a) = 1)$, $a \geq 0$.

3. Explain how it follows from the Markov property (and the stationarity) of the Poisson process that for $s, t \geq 0$ and $i, j \in \mathbb{N}_0$ then

$$P(X(t+s) = i + j | X(s) = i) = \frac{(\lambda t)^j}{j!} \exp(-\lambda t).$$

4. Find the probability that $P(X(b) = 1, X(1) = 1)$, for $0 \leq b \leq 1$.

5. Compute $P(\tau_1 \leq b | X(1) = 1)$ using questions 1.-4.

6. Argue that for $0 \leq a \leq b \leq 1$ then

$$(a < \tau_1 \leq b, X(1) = 1) = (X(a) = 0, X(b) = 1, X(1) = 1).$$

7. Write $P(a < \tau_1 \leq b, X(1) = 1)$ as a product of three probabilities that are known from questions 1.-6. above.

8. Compute the conditional probability $P(a < \tau_1 \leq b | X(1) = 1)$.

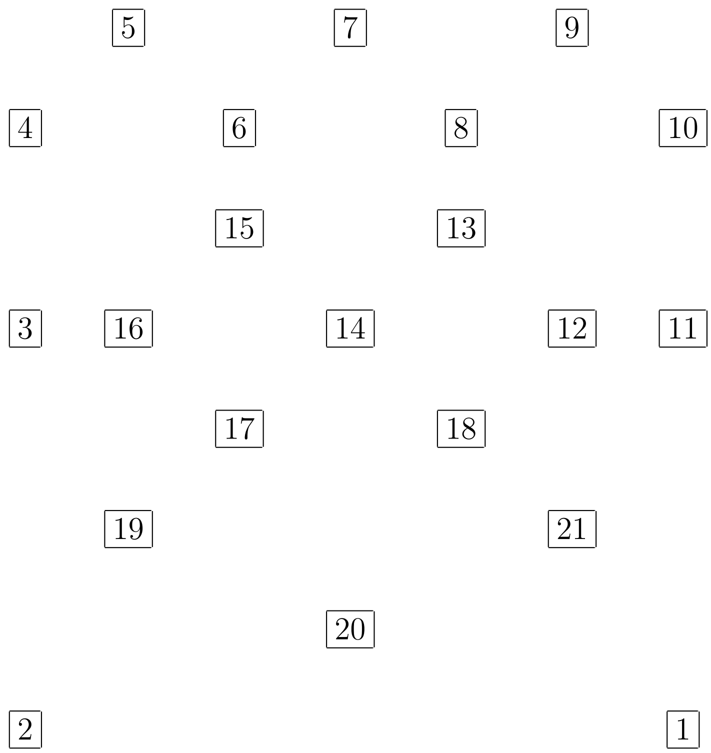
Remark The result shows that if exactly one jump of a Poisson process occurs on the interval $[0, 1]$ then the (conditional) distribution of the jump follows a uniform distribution on $[0, 1]$. The result generalises to the case where we consider the conditional distribution of the n first jumps given that exactly n jumps occured on the interval $[0, t]$. The location of the n jumps will behave as if they had been uniformly scattered over the interval $[0, t]$ independently of each other. For that reason the event times of a Poisson process is often said to describe a completely random pattern of points.

Chapter 4

Additional exercises

4.1 Assignments

4.1.1 Assignment 1 from 2010/2011



Consider a Markov chain on $S = \{1, \ldots, 21\}$ with transition matrix given by

$$\begin{aligned}
1 &= p_{1,2} = p_{2,1} = p_{13,14} = p_{18,14} = p_{15,16} = p_{16,3} = p_{17,16} \\
1/2 &= p_{5,5} = p_{5,7} = p_{9,7} = p_{9,9} = p_{12,13} = p_{12,18}, \quad 1/3 = p_{7,5} = p_{7,7} = p_{7,9} \\
p &= p_{3,2} = p_{4,5} = p_{10,9} = p_{11,1} = p_{14,20} \\
p/2 &= p_{6,5} = p_{6,7} = p_{8,7} = p_{8,9} \\
1-p &= p_{3,4} = p_{4,6} = p_{6,8} = p_{8,10} = p_{10,11} = p_{11,12} \\
(1-p)/2 &= p_{14,15} = p_{14,17} \\
1/2 &= p_{19,19} = p_{19,20} = p_{20,19} = p_{20,21} = p_{21,21} = p_{21,20}
\end{aligned}$$

1. Use this page to draw the transition diagram for the Markov chain. You are strongly encouraged to check your answer to this question with other groups as you may otherwise go wrong on the remaining questions.

2. Simulate and make a plot of two sample paths from the Markov chain with initial distribution $P(X(0) = 21) = 1$. Your plot must be handed in as part of your solution.

3. Compute the distribution of $X(5)$ when the initial distribution is given by $P(X(0) = 20) = P(X(0) = 21) = 1/2$.

4. Find the communication classes and determine for each class if they are recurrent or transient.

5. Find an invariant probability for each (if any!) aperiodic and recurrent class.

6. Does there exist an invariant probability vector, $\overline{\pi} = (\pi_1, \ldots, \pi_{21})$, for the Markov chain with $\pi_1 = 0.5$?

7. Does there exist an invariant probability vector, $\overline{\pi} = (\pi_1, \ldots, \pi_{21})$, for the Markov chain with $\pi_7 = 0.15$ and $\pi_{20} = 0.2$?

In the last part of the exercise we shall study the limiting probabilities

$$\lim_{n\to\infty} P(X(n) = j)$$

under the assumption that $P(X(0) = i) = 1$ for some $i \in S$. You are expected to argue both that the limit exists (or does not exist) and for the value of the limit.

8. Find
$$\lim_{n\to\infty} P(X(n) = j), \quad j \in S,$$
when the initial state i belong to any *recurrent* communication class.

9. Find
$$\lim_{n\to\infty} P(X(n) = 20),$$
for initial distribution given by $P(X(0) = 14) = 1$.

[**Hint:** you may start by considering the distribution of the time
$$\tau_{20} = \inf\{n > 0 | X(n) = 20\}$$
of the first jump to state 20.]

The following question 10. is optional and not a requirement for passing the assignment.

10. Give a complete characterisation of the limiting probabilities

$$\lim_{n\to\infty} P(X(n) = j), \quad j \in S,$$

for $P(X(0) = i)$ where i is any state in a transient communication class.

4.1.2 Assignment 2 from 2010/2011

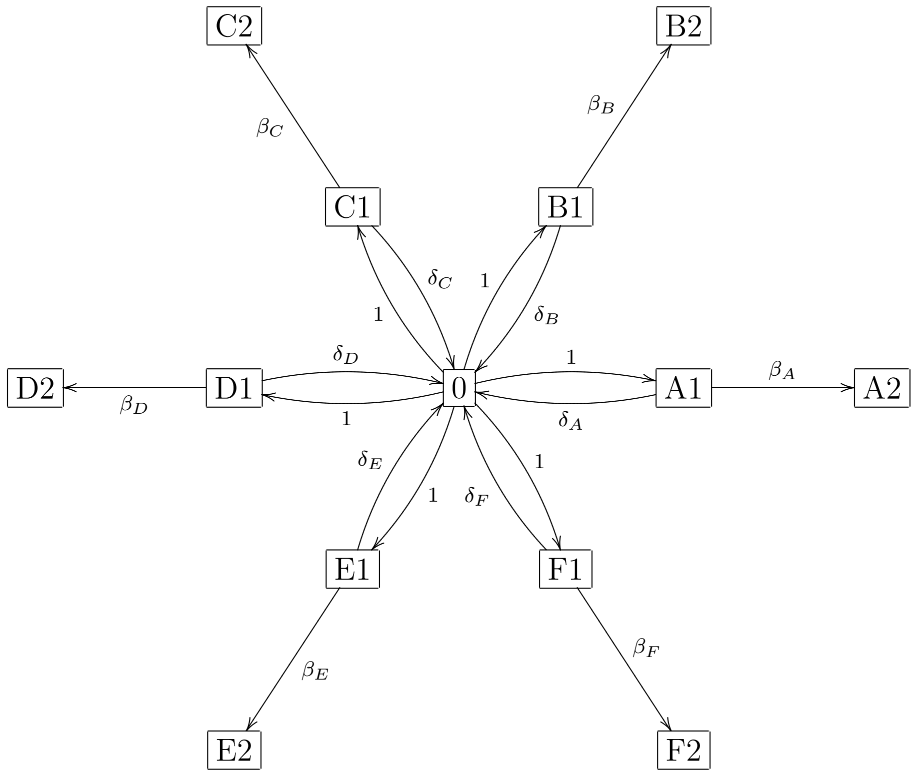


Figure 4.1: Transition diagram of Markov chain in questions 1.-7.

We consider in questions 1.-7. the continuous-time Markov chain on

$$S = \{0, A1, B1, C1, D1, E1, F1, A2, B2, C2, D2, E2, F2\}$$

with transition diagram given by Figure 4.1 where we assume that all $\beta_i, \delta_i > 0$, $i \in \mathcal{A} = \{A, B, C, D, E, F\}$. In questions 1.-4. we consider only the discrete-time Markov chain of jumps.

1. Write down the transition matrix, P, for the jumps of the chain.

2. Write down the submatrix, P_0, of transition probabilities corresponding to the transient states.

3. Compute the first row of $(I-P_0)^{-1}$ where I is the identity matrix.

   [**Hint:** Express the first row of $(I-P_0)^{-1}$ as

$$(I-P_0)^{-1} = \begin{pmatrix} x_0 & x_A & x_B & \dots & x_F \\ & & ? & & \end{pmatrix}$$

   and write out the entries of the first row of the matrix product

$$(I-P_0)^{-1}(I-P_0)$$

   in terms of $x_0, x_A, \dots, x_F$. Now use that the first row equals $(1, 0, \dots, 0)$ to get the (very simple!) formula for $x_0, x_A, \dots, x_F$. Another possibility is to use a computer to find $(I-P_0)^{-1}$.]

4. Assuming that $P(X(0)=0)=1$ find the expected number of times the continuous-time Markov chain jumps to state $D1$ before absorption. Answer the same question for state $B1$.

We now consider the continuous-time Markov chain. We assume throughout the exercise that $P(X(0)=0)=1$ and denote by

$$P_{0,j}(t) := P(X(t)=j), j \in S,$$

the transition probabilities from state 0.

5. Find the expected amount of time spend in state $D1$ before absorption. Answer the same question for state $B1$.

6. Find the absorption probabilities $\lim_{t\to\infty} P(X(t)=j)$ for all of the absorbing states, $j \in S$.

7. Write down the forward differential equations for $P_{0,A2}(t)$, $P_{0,A1}(t)$, and $P_{0,0}(t)$.

For the remaining questions 8.-14. we consider the continuous-time Markov chain on the countable state space

$$S = \{0\} \cup \{(i,j) | i = A, B, C, E, D, F, j \in \mathbb{N}\} = \{0\} \cup \mathcal{A} \times \mathbb{N}.$$

We assume that the transition intensities are given by

$$\begin{aligned} q_{(i,j),(i,j+1)} &= j \cdot \beta_i, \quad j \in \mathbb{N}, i \in \mathcal{A}, \\ q_{(i,j),(i,j-1)} &= j \cdot \delta_i, \quad j \geq 2, i \in \mathcal{A}, \\ q_{(i,1),0} &= \delta_i, \quad i \in \mathcal{A}, \\ q_{0,(i,1)} &= 1, \quad i \in \mathcal{A}. \end{aligned}$$

8. Draw a transition diagram for the Markov chain on the countable state S by entending the diagram on Figure 4.1 in a suitable way. Argue that the chain is irreducible.

The dynamics of the chain can be interpreted in the following way. Given that $X(0) = 0$ then (after an exponentially distributed time) there will be a jump to one of the states

$$(A, 1), (B, 1), (C, 1), (D, 1), (E, 1), (F, 1).$$

Then the Markov chain evolves as a birth-and-death process with intensities β_i, δ_i (where $i \in \mathcal{A}$) along the relevant *branch* of the transition diagram. The process is restarted any time the Markov chain returns to state 0 (-but we are not sure that this will ever happen!).

9. Try to give a heuristic argument that explosion does not occur for instance by giving a lower bound on the expected time of the n-th jump of the Markov chain.

In the following questions we examine for what choices of the parameters that the chain is recurrent. Remember that recurrence for a continuous-time Markov chain is defined in terms of the discrete-time Markov chain of jumps. Denote by τ_n the time of the n-th jump of the Markov chain and consider the Markov chain, $\{Y(n)\}_{n\geq 0} = \{X(\tau_n)\}_{n\geq 0}$, of jumps. We further consider the return time to state 0

$$T = \inf\{n \geq 0 | Y(n) = X(\tau_n) = 0\}.$$

10. Find the transition probabilities for the Markov chain of jumps.

11. Let $\alpha(0) = 1$ and define

$$\alpha(i, j) = P(T < +\infty | Y(0) = (i, j)), (i, j) \in \mathcal{A} \times \mathbb{N}.$$

Write down the system of equations for $\alpha(A, j), j \geq 2$ as given on page 49 of the textbook by Gregory F. Lawler.

12. Use without proof that any solution to the system of equations from question 11. can be written on the form

$$\alpha(A, j) = \begin{cases} c_A + d_A \left(\frac{\delta_A}{\beta_A}\right)^j & , j \geq 1, \delta_A \neq \beta_A \\ c_A + d_A \cdot j & , j \geq 1, \delta_A = \beta_A \end{cases}$$

and find a solution $\alpha(A, j), j \geq 1$, with $0 \leq \alpha(A, j) \leq 1, j \geq 1$, that also satisfies the equation for $\alpha(A, 1)$ given by

$$\alpha(A, 1) = \tfrac{\delta_A}{\beta_A + \delta_A}\alpha(0) + \tfrac{\beta_A}{\beta_A + \delta_A}\alpha(A, 2).$$

Note that the answer depends on the value of the parameters β_A, δ_A.

13. Use the ideas from questions 11.-12. to determine for what values of the parameters that the chain is recurrent or transient.

    [**Hint:** Combine the results from questions 11.-12. and the transience criterion on page 50 of the textbook by Gregory F. Lawler.]

In the final question we examine for what values of the parameters that the chain is positive recurrent.

14. Under what conditions does there exist an invariant distribution, $\overline{\pi}$, and what is the formula for $\overline{\pi}$ in case it exists?

    [**Some help:** Start by writing down the equations for $\pi(A, j)$, $j \geq 2$. Then show that for certain values of the parameters the vector

    $$\pi(A, j) = k_A \cdot \tfrac{1}{j} \left(\beta_A / \delta_A\right)^j, j \geq 1,$$

    solves the system of equations for $\pi(A, j), j \geq 2$. Next, write down the equation for $\pi(A, 1)$ to get an expression for the constant k_A. Repeat the exercise for the remaining $i \in \mathcal{A}$ and end up by adjusting $\pi(0)$ such that you get a probability vector. Then try to see how much you can say about the cases not covered by the solution strategy outlined above.]

4.1.3 Assignment 1 from 2011/2012

To be handed out during the course.

4.1.4 Assignment 2 from 2011/2012

To be handed out during the course.

4.2 Mathematical proofs of selected results

4.2.1 Strong Markov property

To be included next year

4.2.2 Recurrence criterion 1

In this exercise we outline a strategy for proving the recurrence criterion given in Theorem 4 of Section 2.0. The complete proof may be found on page 35-36 in Jacobsen & Keiding (1985). We consider a discrete-time Markov chain $\{X(n)\}_{n\geq 0}$ on S and define the time of the first visit to state j

$$T_j = \inf\{n > 0 | X(n) = j\}.$$

We further introduce the probability

$$f_{ij}^{(n)} = P(X(n) = j, X(n-1), \ldots, X(1) \neq j | X(0) = i)$$

that the first visit to state j happens at time n given that $P(X(0) = i) = 1$. The quantity

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

then describes the probability of ever reaching state j if $P(X(0) = i) = 1$ and you are reminded that state i is recurrent if and only if $f_{ii} = 1$.

1. By splitting the event $(X(n) = j)$ according to the time of the *first* visit to state j show that

$$(P^n)_{i,j} = \sum_{m=1}^{n} (P^{n-m})_{j,j} f_{ij}^{(m)}.$$

2. Summing the expression of question 1. over $n = 1, \ldots, N$ verify the following upper bound

$$\sum_{n=1}^{N} (P^n)_{i,j} \leq \sum_{m=1}^{N} f_{ij}^{(m)} \cdot \sum_{k=0}^{N} (P^k)_{j,j}$$

for $N > 0$.

3. Show that for any $M < N$ then we have a lower bound
$$\sum_{n=1}^{N}(P^n)_{i,j} \geq \sum_{m=1}^{M} f_{ij}^{(m)} \cdot \sum_{k=0}^{N-M}(P^k)_{j,j}.$$

4. Keeping $M < N$ fixed divide the inequalities of 2. and 3. by $\sum_{n=0}^{N}(P^{nn})_{jj}$ to show that
$$\sum_{m=1}^{M} f_{ij}^{(m)} \cdot \left(1 - \frac{\sum_{n=N-M+1}^{N}(P^n)_{j,j}}{\sum_{n=0}^{N}(P^n)_{j,j}}\right) \leq \frac{\sum_{n=1}^{N}(P^n)_{i,j}}{1+\sum_{n=1}^{N}(P^n)_{j,j}} \leq \sum_{m=1}^{N} f_{ij}^{(m)}.$$

5. Let $N \to \infty$ in the expression of question 4. to get that for any $M > 0$ then
$$\begin{aligned} \sum_{m=1}^{M} f_{ij}^{(m)} &\leq \liminf_{N\to\infty} \frac{\sum_{n=1}^{N}(P^n)_{i,j}}{1+\sum_{n=1}^{N}(P^n)_{j,j}} \\ &\leq \limsup_{N\to\infty} \frac{\sum_{n=1}^{N}(P^n)_{i,j}}{1+\sum_{n=1}^{N}(P^n)_{j,j}} \leq \sum_{m=1}^{\infty} f_{ij}^{(k)} = f_{ij}. \end{aligned}$$

6. Finally, let $M \to \infty$ to get that
$$f_{ij} = \frac{\sum_{n=1}^{\infty}(P^n)_{i,j}}{1+\sum_{n=1}^{\infty}(P^n)_{j,j}}.$$

7. Consider the case $i = j$ to get the result in Theorem 4 of Section 2.0.

4.2.3 Number of visits to state j

In this exercise we give a proof of Theorem 2 from Section 2.0. We study the total number of visits to state j defined as
$$N_j = \sum_{n=1}^{\infty} 1(X(n) = j).$$

More precisely, we show that
$$P(N_j = n | X(0) = j) = f_{jj}^{n}(1 - f_{jj}), \quad n \geq 0, \tag{4.1}$$

where we refer to Exercise 4.2.2 for an explanation of the notation.

1. Split the event $(N_j \geq n+1)$ into the time of the first visit to state j and use the Markov property to obtain

$$\begin{aligned} & P(N_j \geq n+1|X(0)=i) \\ = & \sum_{k=1}^{\infty} P(N_j \geq n|X(0)=j)P(T_j = k|X(0)=i) \\ = & P(N_j \geq n|X(0)=j) f_{ij}, \quad n \geq 1. \end{aligned}$$

2. Use question 1., the identity

$$P(N_j = n|X(0)=j) = P(N_j \geq n|X(0)=j) - P(N_j \geq n+1|X(0)=j),$$

and the initial condition $P(N_j \geq 1|X(0)=j) = f_{jj}$ to verify that

$$P(N_j = n|X(0)=j) = f_{jj}^n(1-f_{jj}), \quad n \geq 1.$$

If the following questions we discuss the implication of (4.1) a little further. Remember that (by definition!) state j is recurrent if and only if $f_{jj} = 1$.

3. Argue that $P(N_j = 0|X(0)=j) = 1 - f_{jj}$.

4. Argue that if state j is transient then the numbers of visits to state j follows a geometric distribution. Write down an expression for the expected number, $\mathbb{E}[N_j|X(0)=j]$, of visits to state j.

5. Argue that if state j is recurrent then

$$P(N_j = +\infty|X(0)=j) = 1.$$

6. Give a heuristic argument that for any initial state $i \neq j$ it holds that

$$\begin{aligned} P(N_j = 0|X(0)=i) &= 1 - f_{ij} \\ P(N_j = n|X(0)=i) &= f_{jj}^{n-1}(1-f_{jj})f_{ij} \quad , n \in \mathbb{N} \\ P(N_j = +\infty|X(0)=i) &= f_{ij} \cdot 1(f_{jj} = 1). \end{aligned}$$

Comment From questions 6. and 7. of Exercise 4.2.2 it follows easily that for a transient state j then

$$\sum_{n=0}^{\infty} (P^n)_{jj} = \frac{1}{1-f_{jj}} \text{ and } \sum_{n=0}^{\infty} (P^n)_{ij} = \frac{f_{ij}}{1-f_{jj}}.$$

Since it is trivial to see that

$$\mathbb{E}[N_j|X(0)=i] = \sum_{n=0}^{\infty}(P^n)_{ij}$$

this gives us an expression for the expected number of visits to state j. The present exercise, however, gives a complete description of the distribution of N_j providing us with an expression for the density

$$P(N_j = n|X(0) = i).$$

4.2.4 Recurrence is a class property

In this exercise we give a proof of Theorem 3 from Section 2.0.

1. Use Theorem 4 of Section 2.0 to show that if i is recurrent and if $P^n_{i,j} > 0$ for some $j \in S$ and $n > 0$ then there is an $m > 0$ such that $P^m_{j,i} > 0$. In particular, from a recurrent state i it is only possible to reach state j if if i and j belong to the same communication class.

2. Use question 1. and Theorem 4 of Section 2.0 to show that if i is recurrent and if $P^n_{i,j} > 0$ for some $n > 0$ then j is also recurrent.

3. Deduce from question 2. that states in the same communication class are either all recurrent or all transient.

4.2.5 Recurrence criterion 2

To be included next year

4.2.6 Invariant distribution - finite state space

In this exercise we discuss various steps in a formal proof for the existence of an invariant distribution for a discrete-time Markov chain with N states and transition matrix P.

1. Argue that 1 is always a (right) eigenvalue of P.

2. Show that 1 is always a (left) eigenvalue of P.

3. Argue that if $\overline{v} = (v(i))_{i\in S}$ is a (left) eigenvector for P with eigenvalue 1 and if $v(i) \geq 0, i \in S$, then there exists an invariant distribution, $\overline{\pi} = (\pi(i))_{i\in S}$, for the Markov chain.

4. Show that if there are $N-1$ distinct (right) eigenvalues with $|\cdot| < 1$ then $(P^n)_{i,j}$ converges to a number $a(j)$ that is independent of i.

5. Show that if the assumptions of questions 3. and 4. hold then we must have that $(P^n)_{i,j} \to \pi(j)$ as $n \to \infty$.

Comment The Perron-Frobenius theorem may be used to verify the existence of a unique version of the eigenvector in question 3. and a slightly weaker condition than the one given in queston 4. above. The proof is valid in the case where all entries of P^m are strictly positive for $m > 0$ sufficiently large. The last condition holds for irreducible and aperiodic Markov chains.

4.2.7 Invariant distribution

In this exercise we sketch the proof of Theorem 8 and Corollary 1 in Section 2.0. We consider an irreducible, recurrent Markov chain on S and we shall discuss the existence of solutions to the system of equations

$$\nu(j) = \sum_{i\in S} \nu(i) P_{i,j}, \quad j \in S. \tag{4.2}$$

For any $i \in S$ we introduce the return time to any fixed state i

$$T_i = \inf\{n > 0 | X(n) = i\}$$

and the expected number of visits to state j before first visit to state i

$$\nu^{(i)}(j) \;=\; \mathbb{E}\left[\sum_{n=0}^{T_i-1} 1(X(n) = j) | X(0) = i\right].$$

1. Argue that $\nu^{(i)}(i) = 1$ and by recurrence of state i then that

$$\begin{aligned}\nu^{(i)}(j) &= \mathbb{E}\left[\sum_{n=0}^{T_i-1} 1(X(n) = j) | X(0) = i\right] \\ &= \mathbb{E}\left[\sum_{n=1}^{T_i} 1(X(n) = j) | X(0) = i\right].\end{aligned}$$

2. Verify that

$$\begin{aligned} & \mathbb{E}\left[\sum_{n=1}^{T_i} 1(X(n)=j)|X(0)=i\right] \\ = & \mathbb{E}\left[\sum_{n=1}^{\infty} 1(X(n)=j, X(n-1),\ldots,X(1)\neq i|X(0)=i\right]. \end{aligned}$$

3. Use the Markov property to show that for $n \geq 2$ then

$$\begin{aligned} & P(X(n)=j, X(n-1),\ldots,X(1)\neq i|X(0)=i) \\ = & \sum_{l\neq i} P(X(n)=j, X(n-1)=l, X(n-2),\ldots,X(1)\neq i|X(0)=i) \\ = & \sum_{l\neq i} P_{l,j}\cdot P(X(n-1)=l, X(n-2),\ldots,X(1)\neq i|X(0)=i) \\ = & \sum_{l\neq i} P_{l,j}\cdot \mathbb{E}[1(X(n-1)=l, X(n-2),\ldots,X(1)\neq i)|X(0)=i]. \end{aligned}$$

4. Use questions 1.-3. to show that

$$\begin{aligned} & \nu^{(i)}(j) \\ = & P(X(1)=j|X(0)=i) \\ + & \sum_{n=2}^{\infty} P(X(n)=j, X(n-1),\ldots,X(1)\neq i|X(0)=i) \\ = & P_{i,j} + \sum_{l\neq i} P_{l,j}\cdot \sum_{n=2}^{\infty} P(X(n-1)=l, X(n-2),\ldots,X(1)\neq i|X(0)=i) \\ = & P_{i,j} + \sum_{l\neq i} P_{l,j}\cdot \sum_{n=2}^{\infty} \mathbb{E}[1(X(n-1)=l, X(n-2),\ldots,X(1)\neq i)|X(0)=i] \\ = & P_{i,j} + \sum_{l\neq i} P_{l,j}\cdot \mathbb{E}[\sum_{n=1}^{T_i} 1(X(n)=l)|X(0)=i] \\ = & \sum_{l\in S} P_{l,j}\cdot \mathbb{E}[\sum_{n=1}^{T_i} 1(X(n)=l)|X(0)=i] \\ = & \sum_{l\in S} \nu^{(i)}(l) P_{l,j}. \end{aligned}$$

In questions 1.-4. we have demonstrated that for any $i \in S$ then the vector $(\nu^{(i)}(j))_{j \in S}$ solves (4.2). Trivially, $\nu^{(i)}(j) \geq 0$ and we shall discuss when the total mass $\sum_{j \in S} \nu^{(i)}(j)$ is finite.

5. Show that $\sum_{j \in S} \nu^{(i)}(j) = \mathbb{E}[T_i | X(0) = i]$ and argue that $(\nu^i(j))_{j \in S}$ may be normalized into an invariant distribution (=probability) exactly if state i is positive recurrent.

We have now showed the existence of an invariant distribution for any irreducible, positive recurrent Markov chain in discrete time. An almost complete proof of the uniqueness part of Theorem 8 in Section 2.0 may be constructed along the lines given in the following questions 6.-10.

6. Show that for any solution to (4.2) it holds for $m \geq 1$ and any $l \in S$ that
$$\nu(j) = \sum_{i \in S} \nu(i)(P^m)_{i,j} \geq \nu(l)(P^m)_{l,j}.$$
Deduce that for any non-negative solution (different from zero!) we have that $\nu(j) > 0$ for all $j \in S$.

7. Let $\overline{\nu} = (\nu(j))_{j \in S}$ be any non-zero solution to (4.2). Argue from question 6. that we may assume that $\nu(i) = 1$ where i is any fixed state $i \in S$.

8. Use (without proof!) that for any solution to (4.2) with $\nu(i) = 1$ it holds that for all $j \in S$
$$\nu(j) \geq \nu^{(i)}(j).$$

9. Use question 1., 7. and 8. to argue that the vector $\overline{\mu} = (\mu(j))_{j \in S}$ defined by $\mu(j) = \nu(j) - \nu^{(i)}(j)$ is a non-negative solution to (4.2)

10. Use question 6. and 9. to deduce that for all $j \in S$ then
$$\mu(j) = 0$$
and conclude that $(\nu^{(i)}(j))_{j \in S}$ is the unique solution to (4.2) with i-th coordinate equal to 1.

Note that once we have showed the uniqueness (up to multiplication!) of solutions to (4.2) then it follows from question 5. that if for some $i_0 \in S$

$\mathbb{E}[T_{i_0}|X(0) = i_0] < +\infty$ for some $i_0 \in S$ then for *any* $i \in S$ it holds that $\mathbb{E}[T_i|X(0) = i] < +\infty$. In particular, the states in a recurrent are either all positive recurrent or all null-recurrent. This was postulated in Remark 2.

Since the solution to (4.2) is unique (up to multiplication) we further conclude that there is a unique solution, $\overline{\pi} = (\pi(j))_{j\in S}$, to (4.2)) with $\sum_{j\in S} \pi(j) = 1$ and that the solution may be represented as

$$\pi(j) = \frac{\nu^{(i)}(j)}{\mathbb{E}[T_i|X(0) = i]} = \frac{\mathbb{E}[\sum_{n=0}^{T_i-1} 1(X(n) = j)|X(0) = i]}{\mathbb{E}[T_i|X(0) = i]}$$

for any fixed $i \in S$. Choosing $j = i$ above we conclude that

$$\pi(j) = \frac{1}{\mathbb{E}[T_i|X(0) = i]}$$

hence the invariant probability mass in state j equals the inverse mean return time to state j.

4.2.8 Absorption probabilities - finite state space

To be included next year!

4.2.9 Absorption probabilities - countable state space

To be included next year!

4.2.10 Backward equations

To be included next year!

4.2.11 Uniqueness of invariant distribution (cont)

In this exercise we consider the proof of Theorem 16 in Section 3.0. Let $\{X(t)\}_{t\geq 0}$ be an irreducible continuous-time Markov chain on S and let $\overline{\pi} = (\pi(i))_{i\in S}$ be a probability. Assume that for some $t_0 > 0$ then

$$\forall j \in S \quad : \quad \pi(j) = \sum_{i\in S} \pi(i) P_{i,j}(t_0). \tag{4.3}$$

1. Use that $\pi(i_0) > 0$ and that for irreducible Markov chain then $P_{i_0,j}(t_0) > 0$ to verify from (4.3) that

$$\forall j \in S \quad : \quad \pi(j) > 0.$$

2. Sum (4.3) over $j \in S$ to obtain

$$\sum_{j \in S} \pi(j) = \sum_{j \in S} \sum_{i \in S} \pi(i) P_{i,j}(t_0) = \sum_{i \in S} \pi(i) \sum_{j \in S} P_{i,j}(t_0) \leq \sum_{i \in S} \pi(i)$$

   and conclude that

$$\forall i \in S \quad : \quad \sum_{j \in S} P_{i,j}(t_0) = 1.$$

3. Argue that $\overline{\pi}$ is the unique invariant distribution for the discrete-time Markov chain on S with transition probabilities $P(t_0) = \{P_{i,j}(t_0)\}_{i,j \in S}$.

4. For an arbitrary $t > 0$ find n such that $t < nt_0$. Use that $P(nt_0) = (P(t_0))^n$ to deduce that $P(t)$ is a transition matrix, i.e. that

$$\forall i \in S \quad : \quad \sum_{j \in S} P_{i,j}(t) = 1.$$

5. Verify that

$$\overline{\pi} P(t) P(t_0) = \overline{\pi} P(t_0) P(t) = \overline{\pi} P(t)$$

   and show that $\overline{\pi} P(t)$ is an invariant distribution for $P(t_0)$. Use question 3. to conclude that $\overline{\pi} = \overline{\pi} P(t)$.

Chapter 5

Auxiliary results

This short chapter contains some mathematical results that might be useful to solve the exercises on Markov chains from Chapters 2 and 3.

5.1 Elementary conditional probabilities

For two events (=sets) A, B with $P(B) > 0$ the conditional probability of $A|B$ is defined by the formula

$$P(A|B) = \tfrac{P(A \cap B)}{P(B)}.$$

When working with Markov chains the events will often be expressed by random variables for example as $A = (X(2) = j)$ and $B = (X(1) = i)$. One may show that for three sets A, B, C with $P(B \cap C) > 0$ then

$$P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C).$$

For Markov chains with three set given as $A = (X(2) = k)$, $B = (X(1) = j)$, and $C = P(X(0) = i)$ this may be written out as

$$\begin{aligned} & P(X(2) = k, X(1) = j, X(0) = i) \\ = \;& P(X(2) = k|X(1) = j, X(0) = i) \cdot P(X(1) = j|X(0) = i) \cdot P(X(0) = i) \\ = \;& P(X(2) = k|X(1) = j)P(X(1) = j|X(0) = i)P(X(0) = i) \\ = \;& P_{jk}P_{ij}\phi_i, \end{aligned}$$

where $\overline{\phi} = (\phi_i)$ is the initial distribution and $P = (P_{ij})$ the matrix of transition probabilities for the Markov chain. Note that only the second equality above explicitly makes use of the fact that $\{X(n)\}$ is a Markov chain.

5.2 Some important probability distributions

5.2.1 The binomial distribution

The binomial distribution with integral parameter n and probability parameter p has support on the set $\{0, 1, \ldots, n\}$ and the density is given by

$$p_j = \binom{n}{j} p^j (1-p)^{n-j}, \quad j = 0, 1, \ldots, n.$$

The binomial distribution has mean np and variance $np(1-p)$.

The binomial distribution describes the distribution of the number of successes in n independent replications of an experiment with two possible outcomes (success/failure) with probability of success equal to p.

5.2.2 The Poisson distribution

The Poisson distribution with parameter λ has support on the set $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and the density is given by

$$p_j = \frac{\lambda^j}{j!} e^{-\lambda}, \quad j \geq 0.$$

The Poisson distribution has mean λ and variance λ.

5.2.3 The geometric distribution

The geometric distribution with probability parameter p has support on the set $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and the density is given by

$$p_j = (1-p)^j p, \quad j = 0, 1, 2, \ldots.$$

The geometric distribution has mean $\frac{1-p}{p}$ and variance $\frac{1-p}{p^2}$.

The geometric distribution describes the number of failures before the first success in a sequence of experiments with two possible outcomes (success/failure) with probability of success equal to p.

5.2.4 The negative binomial distribution

The negative binomial distribution with integral parameter r and probability parameter p has support on the set $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and the density is given by

$$p_j = \binom{r+j-1}{j} p^r (1-p)^j, \quad j \geq 0.$$

The negative binomial distribution has mean $\frac{r(1-p)}{p}$ and variance $\frac{r(1-p)}{p^2}$.

The negative binomial disribution with probability parameter p and integer-valued integral parameter $r \in \mathbb{N}$ describes the distribution of the sum of r independent geometrically distributed random variables with probability parameter p.

5.2.5 The exponential distribution

The exponential distribution with rate parameter $\mu > 0$ is a continuous distribution on $[0, \infty)$ with density

$$f(t) = \mu \exp(-\mu t), \quad t > 0,$$

and cumulative distribution function

$$F(t) = \int_0^t f(s)ds = 1 - \exp(-\mu t), \quad t > 0.$$

The exponential distribution has mean $1/\mu$ and variance $1/\mu^2$.

For a continuous time Markov chain the distribution of the waiting time to the next jump follows an exponential distribution.

5.2.6 The gamma distribution

The gamma distribution with shape parameter λ and rate parameter $\mu > 0$ is a continuous distribution on $[0, \infty)$ with density

$$f(t) = \frac{t^{\lambda-1}\mu^{\lambda}}{\Gamma(\lambda)} \exp(-\mu t), \quad t > 0,$$

and cumulative distribution function

$$F(t) = \int_0^t f(s)ds, \quad t > 0.$$

The normalising constant in the density for the gamma distribution is given by the gamma integral

$$\Gamma(\lambda) = \int_0^\infty s^{\lambda-1} \exp(-s)ds$$

and for integer-valued shape parameter λ it holds that $\Gamma(\lambda) = (\lambda - 1)!$ The gamma distribution has mean λ/μ and variance λ/μ^2.

The gamma distribution with rate parameter μ and integer-valued shape parameter $\lambda \in \mathbb{N}$ is the distribution of the sum of λ independent exponentially distributed random variables with rate parameter μ.

5.3 Useful formulaes for sums and series

In many of the exercises you are asked to compute the mean of the invariant distribution for Markov chains on a finite or countable state space, S. If the invariant probability vector is given as $\overline{\pi} = (\pi_i)$ then the mean is given as

$$\mu = \sum_{i \in S} i\pi_i.$$

For other exercises you have an unnormalized version $\overline{\nu} = (\nu_i)$ of an invariant vector and you need to find out if $\sum_{i \in S} \nu_i < \infty$ such that you can define the invariant probability as

$$\pi_j = \frac{\nu_j}{\sum_{i \in S} \nu_i}.$$

Some of the frequently occuring sums or series in this connection are

$$\begin{aligned}
\sum_{i=0}^{N} \alpha\beta^i &= \alpha\tfrac{1-\beta^{N+1}}{1-\beta}, \quad \alpha \in \mathbb{R}, \beta \neq 1 \\
\sum_{i=0}^{\infty} \alpha\beta^i &= \alpha\tfrac{1}{1-\beta}, \quad \alpha \in \mathbb{R}, |\beta| < 1 \\
\sum_{i=0}^{\infty} \alpha i\beta^i &= \alpha\tfrac{\beta}{(1-\beta)^2}, \quad \alpha \in \mathbb{R}, |\beta| < 1 \\
\sum_{i=0}^{\infty} \alpha\tfrac{\beta^i}{i!} &= \alpha\exp(\beta), \quad \alpha, \beta \in \mathbb{R} \\
\sum_{i=0}^{\infty} \alpha i\tfrac{\beta^i}{i!} &= \alpha\beta\exp(\beta), \quad \alpha, \beta \in \mathbb{R}.
\end{aligned}$$

5.4 Some results for matrices

5.4.1 Determinants of a square matrix

For a 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

the determinant is defined as $\det(A) = a_{11}a_{22} - a_{12}a_{21}$. For a higher order square matrix A of dimension k the determinant may be defined recursively as

$$\det A = \sum_{j=1}^{k} (-1)^{1+j} a_{1j} \det A^*_{1j}, \quad \leftarrow \text{expansion by first row}$$

where

$$A^*_{1j} = \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2j-1} & a_{2j+1} & \dots & a_{2k-1} & a_{2k} \\ a_{31} & a_{32} & \dots & a_{3j-1} & a_{3j+1} & \dots & a_{3k-1} & a_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kj-1} & a_{kj+1} & \dots & a_{kk-1} & a_{kk} \end{pmatrix}$$

is the $(k-1) \times (k-1)$ matrix obtained by removing from A all entries from row 1 or column j.

For a 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

the definition leads to the following formula for the determinant

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}.$$

5.4.2 Diagonalisation of matrices

Let A be a $k \times k$ matrix. An eigenvalue for A is a (real or complex) number λ such that there exists a nonzero eigenvector v with

$$Av = \lambda v.$$

The eigenvalues of A are exactly the zeroes of the characteristic polynomial

$$g(\lambda) = \det(A - \lambda I).$$

If $\lambda_1, \dots, \lambda_k$ are the k roots of the characteristic polynomial P and $v_1, \dots, v_k$ are corresponding eigenvectors then

$$A \underbrace{\begin{pmatrix} v_1 & \dots & v_k \end{pmatrix}}_{:=O} = \underbrace{\begin{pmatrix} v_1 & \dots v_k \end{pmatrix}}_{:=O} \underbrace{\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{pmatrix}}_{:=D}.$$

If O is invertible then we get the useful identity

$$A = ODO^{-1}.$$

Note that above we consider socalled right eigenvectors. Similarly one may consider left eigenvectors defined as row vectors $v \neq 0$ solving the equation

$$vA = \lambda v.$$

For some of the exercises in Chapters 2 and 3 we consider right eigenvectors.

5.4.3 Exponential matrices

For any $k \times k$ matrix A consider the matrices obtained by raising A to higher powers A^n. It turns out that the finite sums

$$\sum_{n=0}^{N} \frac{A^n}{n!} = I + \frac{A}{1!} + \frac{A^2}{2!} + \ldots + \frac{A^N}{N!}$$

converge as $N \to \infty$ (entry-by-entry). This allows us to define the exponential matrix $\exp(A)$ as the limit

$$\exp(A) := \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Note that for convenience we use the notation A^0 for the identity matrix.

It is very important to note that the exponential matrix does not satisfy the same rules as the usual exponential function. In particular, except for very special cases it holds that

$$\exp(A + B) \neq \exp(A) \cdot \exp(B).$$

Closed form expressions for exponential matrices are rarely available. One important exception is the case where we can find an invertible matrix O such that

$$O^{-1}AO = D = \begin{pmatrix} d_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 & 0 & \ldots & 0 \\ 0 & 0 & d_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & d_k \end{pmatrix}$$

is a diagonal matrix. Using that $A = ODO^{-1}$ direct computations show that

$$\exp(A) = O^{-1}\begin{pmatrix} \exp(d_1) & 0 & 0 & \dots & 0 \\ 0 & \exp(d_2) & 0 & \dots & 0 \\ 0 & 0 & \exp(d_3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \exp(d_k) \end{pmatrix} O.$$

5.5 First order differential equations

For a continuous-time Markov chain on S with transition intensities $Q = (q_{i,j})_{i,j\in S}$, the transition probabilties

$$P_{i,j}(t) = P(X(s+t) = j|X(s) = i), i, j \in S,$$

always satisfy the backward differential equations

$$P'_{i,j}(t) = \sum_{l\in S} q_{i,l} \cdot P_{l,j}(t), j \in S$$

with the boundary condition that $P_{i,j}(0) = 0, i \neq j$, and $P_{i,i}(0) = 1$.
The solution has an explicit solution given as

$$P(t) = \exp(Qt), t \geq 0,$$

when the state space, S, is finite but computation of the exponential matrix may be infeasible. For Markov chains on countable state spaces no closed form formula for the transition probabilities exist. Sometimes we can get nice explicit formulas for some of the transition probabilities, $P_{i,j}(t)$, by solving some of the backward or forward differential equations. Remember that the forward differential equations take the form

$$P'_{i,j}(t) = \sum_{l\in S} P_{i,l}(t) \cdot q_{l,j}, j \in S.$$

To solve the differential equations you might find it useful to know that

$$\begin{aligned} f'(t) = \alpha \exp(\beta t) \quad &\Rightarrow \quad f(t) = \tfrac{\alpha}{\beta}\exp(\beta t) + c \\ f'(t) = \beta f(t) \quad &\Rightarrow \quad f(t) = c\exp(\beta t) \\ f'(t) = \alpha f(t) + \beta\exp(\gamma t) + \delta \quad &\Rightarrow \quad f(t) = \tfrac{\beta}{\gamma-\alpha}\exp(\gamma t) - \tfrac{\delta}{\alpha} + c\exp(\alpha t) \end{aligned}$$

where c is a constant. Note that the last expression is only valid for $\gamma \neq \alpha$.

5.6 Second order linear recurrence equations

Many of the results in Sections 2.0 and 3.0 are stated in terms of the solution to a system of equations. For Markov chains allowing only jumps of size one the system of equations will often take the following form

$$az_{i-1} + bz_{i+1} \quad = \quad z_i, \quad l < i < u, \tag{5.1}$$

where l or u can be $-\infty$ or $+\infty$. It is clear that if we know z_j, z_{j+1} for some time index j (and if $a, b \neq 0$) then we may recursively determine the values of z_i for the remaining indicies i. In mathematical terms one can formally show that the solution to (5.1) is a vector space of dimension 2 and we shall below describe two linearly independent solutions.

We express the solution in terms of the roots

$$\alpha_1 = \frac{1+\sqrt{1-4ab}}{2b} \quad , \quad \alpha_2 = \frac{1-\sqrt{1-4ab}}{2b}$$

to the characteristic equation for (5.1), which is given as

$$\alpha = a + b\alpha^2.$$

We give the solution for the two cases depending on whether there are two distinct roots.

($\alpha_1 \neq \alpha_2$) Any solution to (5.1) can be written on the form

$$z_i = c_1\alpha_1^i + c_2\alpha_2^i, \quad l \leq i \leq u$$

($\alpha_1 = \alpha_2$) Any solution to (5.1) can be written on the form

$$z_i = c_1\left(\frac{1}{2b}\right)^i + c_2 i\left(\frac{1}{2b}\right)^i, \quad l \leq i \leq u$$

The constants c_1, c_2 can be found from boundary conditions imposed by the relevant model.

5.7 R stuff for dealing with Markov chains

5.7.1 Matrices

The following code defines a 4×4 matrix P of transition probabilities for a discrete time Markov chain.

```
> P<-matrix(nrow=4,ncol=4)
> P[1,]<-c(0,4/6,1/6,1/6)
> P[2,]<-c(0,0,5/7,2/7)
> P[3,]<-c(0,3/5,0,2/5)
> P[4,]<-c(0,1/2,1/2,0)
> P
     [,1]      [,2]      [,3]      [,4]
[1,]    0 0.6666667 0.1666667 0.1666667
[2,]    0 0.0000000 0.7142857 0.2857143
[3,]    0 0.6000000 0.0000000 0.4000000
[4,]    0 0.5000000 0.5000000 0.0000000
```

To compute the n step probabilities given by P^n you need to know how to do matrix multiplication. Below we demonstrate how to compute P^2, P^4, and P^8.

```
> P2<-P%*%P
> P2
     [,1]      [,2]      [,3]      [,4]
[1,]    0 0.1833333 0.5595238 0.2571429
[2,]    0 0.5714286 0.1428571 0.2857143
[3,]    0 0.2000000 0.6285714 0.1714286
[4,]    0 0.3000000 0.3571429 0.3428571
> P4<-P2%*%P2
> P4
     [,1]      [,2]      [,3]      [,4]
[1,]    0 0.2938095 0.4697279 0.2364626
[2,]    0 0.4408163 0.2734694 0.2857143
[3,]    0 0.2914286 0.4848980 0.2236735
[4,]    0 0.3457143 0.3897959 0.2644898
> P8<-P4%*%P4
> P8
```

```
     [,1]      [,2]      [,3]      [,4]
[1,]    0 0.3481567 0.4002902 0.2515532
[2,]    0 0.3727913 0.3645248 0.2626839
[3,]    0 0.3471067 0.4020098 0.2508835
[4,]    0 0.3574321 0.3866506 0.2559174
```

The invariant distribution $\overline{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)$ for the Markov chain must solve the equation $\overline{\pi}P = \overline{\pi}$. In other words the invariant probability is a normalized left eigenvector of P associated with eigenvalue 1. Below we demonstrate how to find left eigenvectors and extract a normalized version of the eigenvector with eigenvalue 1.

```
> lefteigen<-eigen(t(P))
> lefteigen
$values
[1]  1.000000 -0.646385 -0.353615  0.000000

$vectors
          [,1]       [,2]       [,3]       [,4]
[1,] 0.0000000  0.0000000  0.0000000  0.7467709
[2,] 0.6133511 -0.5457060 -0.3294501  0.1633561
[3,] 0.6571619  0.7988318 -0.4822656 -0.4278375
[4,] 0.4381080 -0.2531258  0.8117158 -0.4822895

> normInv<-lefteigen$vectors[,1]/sum(lefteigen$vectors[,1])
> normInv
[1] 0.0000000 0.3589744 0.3846154 0.2564103
```

Note that (in accordance with theory) the colums of P^n approaches the invariant probability vector computed above as $n \to \infty$.

In Section 5.7.2 below we define a transition matrix Q of a continuous time Markov chain on four states. An invariant distribution $\overline{\pi}$ for the this chain must satisfy the matrix equation $\overline{\pi}Q = 0$ as well as the condition $\sum_{i \in S} \pi_i = 1$. One way to compute the invariant distribution in R is to define the matrix $\overline{Q}$ obtained by adding to Q a column of ones and then solve the equation $\pi\overline{Q} = (0, 0, 0, 0, 1)$. The code below works to find the invariant distribution in any case where only one recurrent class of states exist such that $\overline{\pi}$ is unique.

```
> Q1<-cbind(Q,1)
```

```
> Q1
     [,1] [,2] [,3] [,4] [,5]
[1,]   -6  4.0  1.0    1    1
[2,]    0 -7.0  5.0    2    1
[3,]    0  3.0 -5.0    2    1
[4,]    0  0.5  0.5   -1    1
> lm.fit(t(Q1),c(0,0,0,0,1))$coefficients
           x1            x2            x3            x4
-3.552509e-18  1.333333e-01  2.000000e-01  6.666667e-01
> round(lm.fit(t(Q1),c(0,0,0,0,1))$coefficients,digits=6)
      x1       x2       x3       x4
0.000000 0.133333 0.200000 0.666667
```

Note that the last line of code rounds the solution down to 6 significant digits showing that the invariant probability is given by $\overline{\pi} = (0, 2/15, 1/5, 2/3)$.

5.7.2 Computing exponential matrices

For a continuous time Markov chain on a finite state space the transition probabilities

$$P_{ij}(t) = P(X(t+s) = j | X(s) = i)$$

is given as the entries of the matrix $\exp(Qt)$, where Q is the intensity matrix of the Markov chain. The `MatrixExp` function of the `msm`-package may be used to compute exponential matrices in R. Below we demonstrate how to compute the transition probabilities of the four state Markov chain with transition intensity matrix

$$Q = \begin{pmatrix} -6 & 4 & 1 & 1 \\ 0 & -7 & 5 & 2 \\ 0 & 3 & -5 & 2 \\ 0 & 0.5 & 0.5 & -1 \end{pmatrix}.$$

Note that before running the following code on your computer you must install the `msm` package. Initially we define the intensity matrix Q.

```
> library(msm)
> Q<-matrix(nrow=4,ncol=4)
> Q[1,]<-c(-6,4,1,1)
> Q[2,]<-c(0,-7,5,2)
```

```
> Q[3,]<-c(0,3,-5,2)
> Q[4,]<-c(0,0.5,0.5,-1)
> Q
     [,1] [,2] [,3] [,4]
[1,]   -6  4.0  1.0    1
[2,]    0 -7.0  5.0    2
[3,]    0  3.0 -5.0    2
[4,]    0  0.5  0.5   -1
```

Then we compute the matrix of transition probabilities at time $t = 0.1$.

```
> P_1<-MatrixExp(Q,0.1)
> P_1
          [,1]       [,2]       [,3]      [,4]
[1,] 0.5488116 0.22501400 0.11738870 0.1087857
[2,] 0.0000000 0.54095713 0.28625502 0.1727879
[3,] 0.0000000 0.17307769 0.65413446 0.1727879
[4,] 0.0000000 0.03988527 0.04650866 0.9136061
```

You will often need to find the transition probabilities at several values of the time argument for instance if you want to visualize the development of the transtion probabilities over time. Below we compute the transition probabilities at all time points between 0 and 1 at a density of 0.01. The result is stored as a three dimensional array and we demonstrate how to plot the function $t \to P_{12}(t)$.

```
> timearg<-seq(0,1,by=0.01)
> res<-lapply(timearg,function(t){MatrixExp(Q,t)})
> trprob<-array(unlist(res),dim=c(dim(res[[1]]),length(res)))
> plot(timearg,trprob[1,2,],lwd=2,col="blue",type='l'
,xlab="Time",ylab="Probability")
```

5.7.3 Simulation of Markov chains

A simple way to simulate the sample path of a Markov chain is to stick to the description of the dynamics given by the transition diagram. The following code defines a function that can simulate sample paths for both discrete and continous time Markov chains.

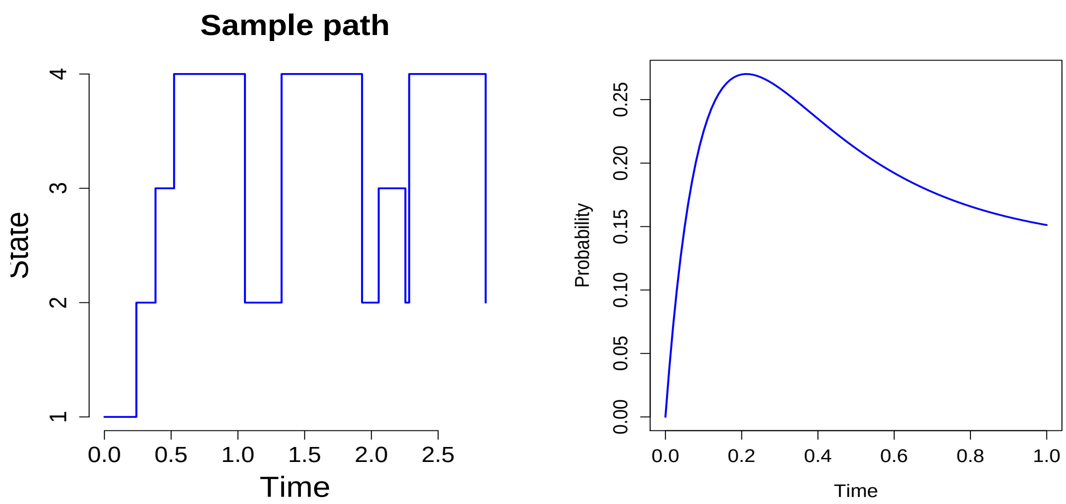


Figure 5.1: Simulated sample path of Markov chain (left) and transiton probability, P_{12} (right).

```
simMC<-function(tr,nJump=10,phi0=NULL){
nStates<-dim(tr)[1]
cont<-(sum(tr)==0)
if(cont){pJump<-matrix(data=0,nrow=nStates,ncol=nStates)
for(i in 1:nStates){pJump[i,-i]<-(-tr[i,-i]/tr[i,i])}}
else{pJump<-tr}
if(is.null(phi0)){phi0<-c(1,rep(0,nStates-1))}
states<-rep(0,nStates+1)
jumps<-rep(0,nStates+1)
states[1]<-sample(nStates,1,prob=phi0)
tmax<-0
for(i in 1:nJump){
if(cont){jumps[i+1]<-(-rexp(1)/tr[states[i],states[i]])+max(jumps)}
else{jumps[i+1]<-i}
states[i+1]<-sample(nStates,1,prob=pJump[states[i],])
}
return(list(y=states,t=jumps))
}
```

The following code simulates and displays the sample path up to step 25 for the discrete time Markov chain with transition matrix P of Section 5.7.1 and initial distribution $\overline{\phi} = (1/4, 1/4, 1/4, 1/4)$.

```
> mcDisc<-simMC(P,nJump=25,phi0=c(0.25,0.25,0.25,0.25))
> mcDisc$y
 [1] 3 2 3 4 2 4 3 2 3 2 4 3 2 3 2 3 4 3 2 4 2 3 4 3 2 3
```

Below we simulate and plot the sample path for the first 10 jumps of the continuous time Markov chain with intensity matrix Q of Section 5.7.2 for initial distribution given by $\overline{\phi} = (1, 0, 0, 0)$.

```
> mcCont<-simMC(Q,nJump=10,phi0=c(1,0,0,0))
> plot(mcCont$t,mcCont$y,type='s',xlab="Time",ylab="State"
,axes=F,col="blue",lwd=2)
> axis(side=1)
> axis(side=2,at=1:4)
```

The function `simMC` does not apply for simulation of Markov chains on countable state spaces. However, for the most common examples discussed in this collection of exercises it should be easy (or at least possible) to write simple functions for simulation of the sample paths based on the transition diagram of the chain.

The Poisson process is a continuous time Markov chain on $\mathbb{N}_0$ moving only in jumps of size 1 hence everything simplifies a great deal as we only need to simulate the jump times. The waiting times between jumps are independent and identically distributed exponential variables with rate λ (-the intensity of the Poisson process). Below we show how to simulate the first 50 jump times of a Poisson process with intensity 1 and plot the resulting sample path.

```
wait<-rexp(50,rate=1)
t<-cumsum(c(0,wait))
plot(0:50,t,type='s',lwd=2,col="blue",xlab="Time"
,ylab="Number of jumps")
```

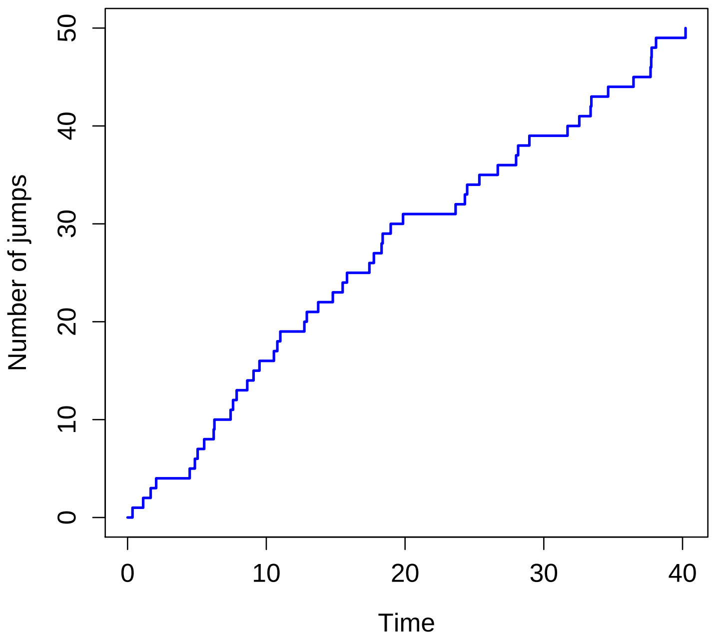


Figure 5.2: Simulated sample path of a Poisson process with intensity 1.

Bibliography

Asmussen, S. (2003), *Applied probability and queues*, Vol. 51 of *Applications of Mathematics (New York)*, second edn, Springer-Verlag, New York. Stochastic Modelling and Applied Probability.

Brémaud, P. (1999), *Markov chains*, Vol. 31 of *Texts in Applied Mathematics*, Springer-Verlag, New York. Gibbs fields, Monte Carlo simulation, and queues.

Jacobsen, M. & Keiding, N. (1985), *Markovkæder*, Lecture notes, University of Copenhagen.

Lawler, G. F. (2006), *Introduction to stochastic processes*, second edn, Chapman & Hall/CRC, Boca Raton, FL.