CHAPTER 2

Markov chains

2.1. Examples

EXAMPLE 2.1.1 (Markov chain with two states). Consider a phone which can be in two states: "free"$= 0$ and "busy"$= 1$. The set of the states of the phone is

$$E = \{0, 1\}.$$

We assume that the phone can randomly change its state in time (which is assumed to be discrete) according to the following rules.

1. If at some time n the phone is free, then at time $n+1$ it becomes busy with probability p or it stays free with probability $1-p$.

2. If at some time n the phone is busy, then at time $n+1$ it becomes free with probability q or it stays busy with probability $1-q$.

Denote by X_n the state of the phone at time $n = 0, 1, \ldots$. Thus, $X_n : \Omega \to \{0, 1\}$ is a random variable and our assumptions can be written as follows:

$$p_{00} := \mathbb{P}[X_{n+1} = 0 | X_n = 0] = 1-p, \qquad p_{01} := \mathbb{P}[X_{n+1} = 1 | X_n = 0] = p,$$
$$p_{10} := \mathbb{P}[X_{n+1} = 0 | X_n = 1] = q, \qquad p_{11} := \mathbb{P}[X_{n+1} = 1 | X_n = 1] = 1-q.$$

We can write these probabilities in form of a *transition matrix*

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

Additionally, we will make the following assumption which is called the *Markov property*: Given that at some time n the phone is in state $i \in \{0, 1\}$, the behavior of the phone after time n does not depend on the way the phone reached state i in the past.

PROBLEM 2.1.2. Suppose that at time 0 the phone was free. What is the probability that the phone will be free at times $1, 2$ and then becomes busy at time 3?

SOLUTION. This probability can be computed as follows:

$$\mathbb{P}[X_1 = X_2 = 0, X_3 = 1] = p_{00} \cdot p_{00} \cdot p_{01} = (1-p)^2 p.$$

PROBLEM 2.1.3. Suppose that the phone was free at time 0. What is the probability that it will be busy at time 3?

SOLUTION. We have to compute $\mathbb{P}[X_3 = 1]$. We know the values $X_0 = 0$ and $X_3 = 1$, but the values of X_1 and X_2 may be arbitrary. We have the following possibilities:

(1) $X_0 = 0, X_1 = 0, X_2 = 0, X_3 = 1$. Probability: $(1-p) \cdot (1-p) \cdot p$.
(2) $X_0 = 0, X_1 = 0, X_2 = 1, X_3 = 1$. Probability: $(1-p) \cdot p \cdot (1-q)$.

(3) $X_0 = 0, X_1 = 1, X_2 = 0, X_3 = 1$. Probability: $p \cdot q \cdot p$.
(4) $X_0 = 0, X_1 = 1, X_2 = 1, X_3 = 1$. Probability: $p \cdot (1-q) \cdot (1-q)$.

The probability we look for is the sum of these 4 probabilities:

$$\mathbb{P}[X_3 = 1] = (1-p)^2 p + (1-p)(1-q)p + p^2 q + p(1-q)^2.$$

Example 2.1.4 (Gambler's ruin). At each unit of time a gambler plays a game in which he can either win 1€ (which happens with probability p) or he can loose 1€ (which happens with probability $1-p$). Let X_n be the capital of the gambler at time n. Let us agree that if at some time n the gambler has no money (meaning that $X_n = 0$), then he stops to play (meaning that $X_n = X_{n+1} = \ldots = 0$). We can view this process as a Markov chain on the state space $E = \{0, 1, 2, \ldots\}$ with transition matrix

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

2.2. Definition of Markov chains

Let us consider some system. Assume that the system can be in some states and that the system can change its state in time. The set of all states of the system will be denoted by E and called the *state space* of the Markov chain. We always assume that the state space E is a finite or countable set. Usually, we will denote the states so that $E = \{1, \ldots, N\}$, $E = \mathbb{N}$, or $E = \mathbb{Z}$.

Assume that if at some time the system is in state $i \in E$, then in the next moment of time it can switch to state $j \in E$ with probability p_{ij}. We will call p_{ij} the *transition probability* from state i to state j. Clearly, the transition probabilities should be such that

(1) $p_{ij} \geq 0$ for all $i, j \in E$.
(2) $\sum_{j \in E} p_{ij} = 1$ for all $i \in E$.

We will write the transition probabilities in form of a *transition matrix*

$$P = (p_{ij})_{i,j \in E}.$$

The rows and the columns of this matrix are indexed by the set E. The element in the i-th row and j-th column is the transition probability p_{ij}. The elements of the matrix P are non-negative and the sum of elements in any row is equal to 1. Such matrices are called *stochastic*.

Definition 2.2.1. A *Markov chain* with state space E and transition matrix P is a stochastic process $\{X_n : n \in \mathbb{N}_0\}$ taking values in E such that for every $n \in \mathbb{N}_0$ and every states $i_0, i_1, \ldots, i_{n-1}, i, j$ we have

$$\begin{aligned}\mathbb{P}[X_{n+1} = j | X_n = i] &= \mathbb{P}[X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i] \\ &= p_{ij},\end{aligned} \tag{2.2.1}$$

provided that $\mathbb{P}[X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i] \neq 0$ (which ensures that the conditional probabilities are well-defined).

Condition (2.2.1) is called the *Markov property*.

In the above definition it is not specified at which state the Markov chain starts at time 0. In fact, the initial state can be in general arbitrary and we call the probabilities

$$\alpha_i := \mathbb{P}[X_0 = i], \quad i \in E, \tag{2.2.2}$$

the *initial probabilities*. We will write the initial probabilities in form of a row vector $\alpha = (\alpha_i)_{i\in E}$. This vector should be such that $\alpha_i \geq 0$ for all $i \in E$ and $\sum_{i\in E} \alpha_i = 1$.

THEOREM 2.2.1. *For all $n \in \mathbb{N}_0$ and for all $i_0, \ldots, i_n \in E$ it holds that*

$$\mathbb{P}[X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n] = \alpha_0 p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n}. \tag{2.2.3}$$

PROOF. We use the induction over n. The induction basis is the case $n = 0$. We have $\mathbb{P}[X_0 = i_0] = \alpha_{i_0}$ by the definition of initial probabilities, see (2.2.2). Hence, Equation (2.2.3) holds for $n = 0$.

Induction assumption: Assume that (2.2.3) holds for some n. We prove that (2.2.3) holds with n replaced by $n + 1$. Consider the event $A = \{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\}$. By the induction assumption,

$$\mathbb{P}[A] = \alpha_{i_0} p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n}.$$

By the Markov property,

$$\mathbb{P}[X_{n+1} = i_{n+1} | A] = p_{i_n i_{n+1}}.$$

It follows that

$$\begin{aligned}
\mathbb{P}[X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n, X_{n+1} = i_{n+1}] &= \mathbb{P}[X_{n+1} = i_{n+1} | A] \cdot \mathbb{P}[A] \\
&= p_{i_n i_{n+1}} \cdot \alpha_{i_0} p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n} \\
&= \alpha_{i_0} p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n} p_{i_n i_{n+1}}.
\end{aligned}$$

This completes the induction. □

REMARK 2.2.2. If $\mathbb{P}[A] = 0$, then in the above proof we cannot use the Markov property. However, in the case $\mathbb{P}[A] = 0$ both sides of (2.2.3) are equal to 0 and (2.2.3) is trivially satisfied.

THEOREM 2.2.2. *For every $n \in \mathbb{N}$ and every state $i_n \in E$ we have*

$$\mathbb{P}[X_n = i_n] = \sum_{i_0, \ldots, i_{n-1} \in E} \alpha_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}.$$

PROOF. We have

$$\begin{aligned}
\mathbb{P}[X_n = i_n] &= \sum_{i_0, \ldots, i_{n-1} \in E} \mathbb{P}[X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n] \\
&= \sum_{i_0, \ldots, i_{n-1} \in E} \alpha_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n},
\end{aligned}$$

where the last step is by Theorem 2.2.1. □

2.3. n-step transition probabilities

NOTATION 2.3.1. If we want to indicate that the Markov chain starts at state $i \in E$ at time 0, we will write $\mathbb{P}_i$ instead of $\mathbb{P}$.

DEFINITION 2.3.2. The n-step transition probabilities of a Markov chain are defined as

$$p_{ij}^{(n)} := \mathbb{P}_i[X_n = j].$$

We will write these probabilities in form of the *n-step transition matrix* $P^{(n)} = (p_{ij}^{(n)})_{i,j\in E}$.

By Theorem 2.2.2 we have the formula

$$p_{ij}^{(n)} = \sum_{i_1,\dots,i_{n-1}\in E} p_{ii_1} p_{i_1 i_2} \dots p_{i_{n-1}j}.$$

The next theorem is crucial. It states that the n-step transition matrix $P^{(n)}$ can be computed as the n-th power of the transition matrix P.

THEOREM 2.3.1. *We have* $P^{(n)} = P^n = P \cdot \ldots \cdot P$.

PROOF. We use induction over n. For $n = 1$ we have $p_{ij}^{(1)} = p_{ij}$ and hence, $P^{(1)} = P$. Thus, the statement of the theorem is true for $n = 1$.

Let us now assume that we already proved that $P^{(n)} = P^n$ for some $n \in \mathbb{N}$. We compute $P^{(n+1)}$. By the formula of total probability, we have

$$p_{ij}^{(n+1)} = \mathbb{P}_i[X_{n+1} = j] = \sum_{k\in E} \mathbb{P}_i[X_n = k]\mathbb{P}[X_{n+1} = j | X_n = k] = \sum_{k\in E} p_{ik}^{(n)} p_{kj}.$$

On the right hand-side we have the scalar product of the i-th row of the matrix $P^{(n)}$ and the j-th column of the matrix P. By definition of the matrix multiplication, this scalar product is exactly the entry of the matrix product $P^{(n)}P$ which is located in the i-th row and j-th column. We thus have the equality of matrices

$$P^{(n+1)} = P^{(n)}P.$$

But now we can apply the induction assumption $P^{(n)} = P^n$ to obtain

$$P^{(n+1)} = P^{(n)}P = P^n \cdot P = P^{n+1}.$$

This completes the induction. □

In the next theorem we consider a Markov chain with initial distribution $\alpha = (\alpha_i)_{i\in E}$ and transition matrix P. Let $\alpha^{(n)} = (\alpha_j^{(n)})_{j\in E}$ be the distribution of the position of this chain at time n, that is

$$\alpha_j^{(n)} = \mathbb{P}[X_n = j].$$

We write both $\alpha^{(n)}$ and α as *row* vectors. The next theorem states that we can compute $\alpha^{(n)}$ by taking α and multiplying it by the n-step transition matrix $P^{(n)} = P^n$ from the right.

THEOREM 2.3.2. *We have*

$$\alpha^{(n)} = \alpha P^n.$$

PROOF. By the formula of the total probability

$$\alpha_j^{(n)} = \mathbb{P}[X_n = j] = \sum_{i \in E} \alpha_i \mathbb{P}_i[X_n = j] = \sum_{i \in E} \alpha_i p_{ij}^{(n)}.$$

On the right-hand side we have the scalar product of the row α with the j-th column of $P^{(n)} = P^n$. By definition of matrix multiplication, this means that $\alpha^{(n)} = \alpha P^n$. $\square$

2.4. Invariant measures

Consider a Markov chain on state space E with transition matrix P. Let $\lambda : E \to \mathbb{R}$ be a function. To every state $i \in E$ the function assigns some value which will be denoted by $\lambda_i := \lambda(i)$. Also, it will be convenient to write the function λ as a row vector $\lambda = (\lambda_i)_{i \in E}$.

DEFINITION 2.4.1. A function $\lambda : E \to \mathbb{R}$ is called a *measure* on E if $\lambda_i \geq 0$ for all $i \in E$.

DEFINITION 2.4.2. A function $\lambda : E \to \mathbb{R}$ is called a *probability measure* on E if $\lambda_i \geq 0$ for all $i \in E$ and

$$\sum_{i \in E} \lambda_i = 1.$$

DEFINITION 2.4.3. A measure λ is called *invariant* if $\lambda P = \lambda$. That is, for every state $j \in E$ it should hold that

$$\lambda_j = \sum_{i \in E} \lambda_i p_{ij}.$$

REMARK 2.4.4. If the initial distribution α of a Markov chain is invariant, that is $\alpha P = \alpha$, then for every $n \in \mathbb{N}$ we have $\alpha P^n = \alpha$ which means that at every time n the position of the Markov chain has the same distribution as at time 0:

$$X_0 \stackrel{d}{=} X_1 \stackrel{d}{=} X_2 \stackrel{d}{=} \ldots.$$

EXAMPLE 2.4.5. Let us compute the invariant distribution for the Markov chain from Example 2.1.1. The transition matrix is

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

The equation $\lambda P = \lambda$ for the invariant probability measure takes the following form:

$$(\lambda_0, \lambda_1) \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = (\lambda_0, \lambda_1).$$

Multiplying the matrices we obtain the following two equations:

$$\begin{aligned} \lambda_0(1-p) + \lambda_1 q &= \lambda_0, \\ \lambda_0 p + \lambda_1(1-q) &= \lambda_1. \end{aligned}$$

From the first equation we obtain that $\lambda_1 q = \lambda_0 p$. Solving the second equation we obtain the same relation which means that the second equation does not contain any information not contained in the first equation. However, since we are looking for invariant *probability* measures, we have an additional equation

$$\lambda_0 + \lambda_1 = 1.$$

Solving this equation together with $\lambda_1 q = \lambda_0 p$ we obtain the following result:

$$\lambda_0 = \frac{q}{p+q}, \quad \lambda_1 = \frac{p}{p+q}.$$

Problem 2.4.6. Consider the phone from Example 2.1.1. Let the phone be free at time 0. What is (approximately) the probability that it is free at time $n = 1000$?

Solution. The number $n = 1000$ is large. For this reason it seems plausible that the probability that the phone is free (busy) at time $n = 1000$ should be approximately the same as the probability that it is free (busy) at time $n + 1 = 1001$. Denoting the initial distribution by $\alpha = (1, 0)$ and the distribution of the position of the chain at time n by $\alpha^{(n)} = \alpha P^n$ we thus must have

$$\alpha^{(n)} \approx \alpha^{(n+1)} = \alpha P^{n+1} = \alpha P^n \cdot P = \alpha^{(n)} P.$$

Recall that the equation for the invariant probability measure has the same form $\lambda = \lambda P$. It follows that $\alpha^{(n)}$ must be approximately the invariant probability measure:

$$\alpha^{(n)} \approx \lambda.$$

For the probability that the phone is free (busy) at time $n = 1000$ we therefore obtain the approximations

$$p_{00}^{(n)} \approx \lambda_0 = \frac{q}{p+q}, \quad p_{01}^{(n)} \approx \lambda_1 = \frac{p}{p+q}.$$

Similar considerations apply to the case when the phone is busy at time 0 leading to the approximations

$$p_{10}^{(n)} \approx \lambda_0 = \frac{q}{p+q}, \quad p_{11}^{(n)} \approx \lambda_1 = \frac{p}{p+q}.$$

Note that $p_{00}^{(n)} \approx p_{10}^{(n)}$ and $p_{01}^{(n)} \approx p_{11}^{(n)}$ which can be interpreted by saying that the Markov chain almost forgets its initial state after many steps. For the n-step transition matrix we therefore may conjecture that

$$\lim_{n\to\infty} P^n = \lim_{n\to\infty} \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} \\ p_{10}^{(n)} & p_{11}^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_0 & \lambda_1 \\ \lambda_0 & \lambda_1 \end{pmatrix}.$$

The above considerations are not rigorous. We will show below that if a general Markov chain satisfies appropriate conditions, then

(1) The invariant probability measure λ exists and is unique.
(2) For every states $i, j \in E$ we have $\lim_{n\to\infty} p_{ij}^{(n)} = \lambda_j$.

Example 2.4.7 (Ehrenfest model). We consider a box which is divided into 2 parts. Consider N balls (molecules) which are located in this box and can move from one part to the other according to the following rules. Assume that at any moment of time one of the N balls is chosen at random (all balls having the same probability $1/N$ to be chosen). This ball moves to the other part. Then, the procedure is repeated. Let X_n be the number of balls at time n in Part 1. Then, X_n takes values in $E = \{0, 1, \dots, N\}$ which is our state space. The transition probabilities are given by

$$p_{0,1} = 1, \quad p_{N,N-1} = 1, \quad p_{i,i+1} = \frac{N-i}{N}, \quad p_{i,i-1} = \frac{i}{N}, \quad i = 1, \dots, N-1.$$

For the invariant probability measure we obtain the following system of equations

$$\lambda_0 = \frac{\lambda_1}{N}, \quad \lambda_N = \frac{\lambda_{N-1}}{N}, \quad \lambda_j = \frac{N-j+1}{N}\lambda_{j-1} + \frac{j+1}{N}\lambda_{j+1}, \quad j = 1, \ldots, N-1.$$

Additionally, we have the equation $\lambda_0 + \ldots + \lambda_N = 1$. This system of equations can be solved directly, but one can also guess the solution without doing computations. Namely, it seems plausible that after a large number of steps every ball will be with probability 1/2 in Part 1 and with probability 1/2 in Part 2. Hence, one can guess that the invariant probability measure is the binomial distribution with parameter 1/2:

$$\lambda_j = \frac{1}{2^N}\binom{N}{j}.$$

One can check that this is indeed the unique invariant probability measure for this Markov chain.

EXAMPLE 2.4.8. Let $X_0, X_1, \ldots$ be independent and identically distributed random variables with values $1, \ldots, N$ and corresponding probabilities

$$\mathbb{P}[X_n = i] = p_i, \quad p_1, \ldots, p_N \geq 0, \quad \sum_{i=1}^{N} p_i = 1.$$

Then, $X_0, X_1, \ldots$ is a Markov chain and the transition matrix is

$$P = \begin{pmatrix} p_1 & \ldots & p_N \\ \ldots & \ldots & \ldots \\ p_1 & \ldots & p_N \end{pmatrix}.$$

The invariant probability measure is given by $\lambda_1 = p_1, \ldots, \lambda_N = p_N$.

2.5. Class structure and irreducibility

Consider a Markov chain on a state space E with transition matrix P.

DEFINITION 2.5.1. We say that state $i \in E$ *leads* to state $j \in E$ if there exists $n \in \mathbb{N}_0$ such that $p_{ij}^{(n)} \neq 0$. We use the notation $i \rightsquigarrow j$.

REMARK 2.5.2. By convention, $p_{ii}^{(0)} = 1$ and hence, every state leads to itself: $i \rightsquigarrow i$.

THEOREM 2.5.1. *For two states $i, j \in E$ with $i \neq j$, the following statements are equivalent:*

(1) $i \rightsquigarrow j$.
(2) $\mathbb{P}_i[\exists n \in \mathbb{N} \colon X_n = j] \neq 0$.
(3) *There exist $n \in \mathbb{N}$ and states $i_1, \ldots, i_{n-1} \in E$ such that $p_{ii_1} \ldots p_{i_{n-1}j} > 0$.*

PROOF. We prove that Statements 1 and 2 are equivalent. We have the inequality

$$p_{ij}^{(n)} \leq \mathbb{P}_i[\exists n \in \mathbb{N} : X_n = j] \leq \sum_{n=1}^{\infty} \mathbb{P}_i[X_n = j] = \sum_{n=1}^{\infty} p_{ij}^{(n)}. \tag{2.5.1}$$

If Statement 1 holds, then for some $n \in \mathbb{N}$ we have $p_{ij}^{(n)} > 0$. Hence, by (2.5.1), we have $\mathbb{P}_i[\exists n \in \mathbb{N} : X_n = j] > 0$ and Statement 2 holds. If, conversely, Statement 2 holds, then

$\mathbb{P}_i[\exists n \in \mathbb{N} : X_n = j] > 0$. Hence, by (2.5.1), $\sum_{n=1}^{\infty} p_{ij}^{(n)} > 0$, which implies that at least one summand $p_{ij}^{(n)}$ must be strictly positive. This proves Statement 1.

We prove the equivalence of Statements 1 and 3. We have the formula

$$p_{ij}^{(n)} = \sum_{i_1,\ldots,i_{n-1}\in E} p_{ii_1} \ldots p_{i_{n-1}j}. \tag{2.5.2}$$

If Statement 1 holds, then for some $n \in \mathbb{N}$ we have $p_{ij}^{(n)} > 0$ which implies that at least one summand on the right-hand side of (2.5.2) must be strictly positive. This implies Statement 3. If, conversely, Statement 3 holds, then the sum on the right-hand side of (2.5.2) is positive which implies that $p_{ij}^{(n)} > 0$. Hence, Statement 1 holds. □

Definition 2.5.3. States $i, j \in E$ *communicate* if $i \rightsquigarrow j$ and $j \rightsquigarrow i$. Notation: $i \leftrightsquigarrow j$.

Theorem 2.5.2. *$i \leftrightsquigarrow j$ is an equivalence relation, namely*

1. $i \leftrightsquigarrow i$.
2. $i \leftrightsquigarrow j \Longleftrightarrow j \leftrightsquigarrow i$.
3. $i \leftrightsquigarrow j, j \leftrightsquigarrow k \Rightarrow i \leftrightsquigarrow k$.

Proof. Statements 1 and 2 follow from the definition. We prove statement 3. If $i \leftrightsquigarrow j$ and $j \leftrightsquigarrow k$, then, in particular, $i \rightsquigarrow j$ and $j \rightsquigarrow k$. By Theorem 2.5.1, Statement 3, we can find $r \in \mathbb{N}$, $s \in \mathbb{N}$ and states $u_1, \ldots, u_{r-1} \in E$ and $v_1, \ldots, v_{s-1} \in E$ such that $p_{iu_1}p_{u_1u_2}\ldots p_{u_{r-1}j} > 0$ and $p_{jv_1}p_{v_1v_2}\ldots p_{v_{s-1}k} > 0$. Multiplying both inequalities, we get

$$p_{iu_1}p_{u_1u_2}\ldots p_{u_{r-1}j}p_{jv_1}p_{v_1v_2}\ldots p_{v_{s-1}k} > 0.$$

By Theorem 2.5.1, Statement 3, we have $i \rightsquigarrow k$. In a similar way one shows that $k \rightsquigarrow i$. □

Definition 2.5.4. The *communication class* of state $i \in E$ is the set $\{j \in E : i \leftrightsquigarrow j\}$. This set consists of all states j which communicate to i.

Since communication of states is an equivalence relation, the state space E can be decomposed into a disjoint union of communication classes. Any two communication classes either coincide completely or are disjoint sets.

Definition 2.5.5. A Markov chain is *irreducible* if every two states communicate. Hence, an irreducible Markov chain consists of just one communication class.

Definition 2.5.6. A communication class C is *open* if there exist a state $i \in C$ and a state $k \notin C$ such that $i \rightsquigarrow k$. Otherwise, a communication class is called *closed*.

If a Markov chain once arrived in a closed communication class, it will stay in this class forever.

Exercise 2.5.7. Show that a communication class C is open if and only if there exist a state $i \in C$ and a state $k \notin C$ such that $p_{ik} > 0$.

Theorem 2.5.3. *If the state space E is a finite set, then there exists at least one closed communication class.*

PROOF. We use a proof by contradiction. Assume that there is no closed communication class. Hence, all communication classes are open. Take some state and let C_1 be the communication class of this state. Since C_1 is open, there is a path from C_1 to some other communication class $C_2 \neq C_1$. Since C_2 is open, we can go from C_2 to some other communication class $C_3 \neq C_3$, and so on. Note that in the sequence $C_1, C_2, C_3, \ldots$ all classes are different. Indeed, if for some $l < m$ we would have $C_l = C_m$ (a "cycle"), this would mean that there is a path starting from C_l, going to C_{l+1} and then to $C_m = C_l$. But this is a contradiction since then C_l and C_{l+1} should be a single communication class, and not two different classes, as in the construction. So, the classes $C_1, C_2, \ldots$ are different (in fact, disjoint) and each class contains at least one element. But this is a contradiction since E is a finite set. □

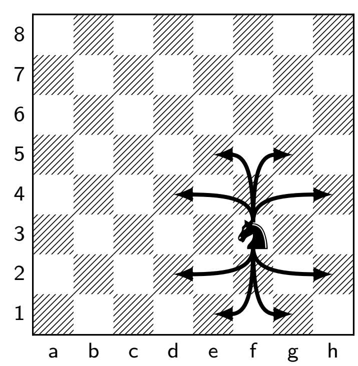
2.6. Aperiodicity

DEFINITION 2.6.1. The *period* of a state $i \in E$ is defined as

$$\gcd\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}.$$

Here, gcd states for the greatest common divisor. A state $i \in E$ is called *aperiodic* if its period is equal to 1. Otherwise, the state i is called *periodic*.

EXAMPLE 2.6.2. Consider a knight on a chessboard moving according to the usual chess rules in a random way. For concreteness, assume that at each moment of time all moves of the knight allowed by the chess rules are counted and then one of these moves is chosen, all moves being equiprobable.



This is a Markov chain on a state space consisting of 64 squares. Assume that at time 0 the knight is in square i. Since the knight changes the color of its square after every move, it cannot return to the original square in an odd number of steps. On the other hand, it can return to i in an even number of steps with non-zero probability (for example by going to some other square and then back, many times). So,

$$p_{ii}^{(2n+1)} = 0, \quad p_{ii}^{(2n)} > 0.$$

Hence, the period of any state in this Markov chain is 2.

EXAMPLE 2.6.3. Consider a Markov chain on a state space of two elements with transition matrix

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

We have

$$p_{ii}^{(2n+1)} = 0, \quad p_{ii}^{(2n)} = 1.$$

Hence, the period of any state in this Markov chain is 2.

Exercise 2.6.4. Show that in the Ehrenfest Markov chain (Example 2.4.7) every state is periodic with period 2.

Lemma 2.6.5. *Let $i \in E$ be any state. The following conditions are equivalent:*

(1) *State i is aperiodic.*
(2) *There is $N \in \mathbb{N}$ such that for every natural number $n > N$ we have $p_{ii}^{(n)} > 0$.*

Proof. If Statement 2 holds, then for some sufficiently large n we have $p_{ii}^{(n)} > 0$ and $p_{ii}^{(n+1)} > 0$. Since $\gcd(n, n+1) = 1$, the state i has period 1. Hence, Statement 1 holds.

Suppose, conversely, that Statement 1 holds. Then, we can find $n_1, \dots, n_r \in \mathbb{N}$ such that $\gcd\{n_1, \dots, n_r\} = 1$ and $p_{ii}^{(n_1)} > 0, \dots, p_{ii}^{(n_r)} > 0$. By a result from number theory, the condition $\gcd\{n_1, \dots, n_r\} = 1$ implies that there is $N \in \mathbb{N}$ such that we can represent any natural number $n > N$ in the form $n = l_1 n_1 + \ldots + l_r n_r$ for suitable $l_1, \dots, l_r \in \mathbb{N}$. We obtan that

$$p_{ii}^{(l_1 n_1 + \ldots + l_r n_r)} \geq (p_{ii}^{(n_1)})^{l_1} \cdot \ldots \cdot (p_{ii}^{(n_r)})^{l_r} > 0.$$

This proves Statement 2.

Lemma 2.6.6. *If state $i \in E$ is aperiodic and $i \leftrightsquigarrow j$, then j is also aperiodic.*

Remark 2.6.7. We can express this by saying that aperiodicity is a *class property*: If some state in a communication class is aperiodic, then all states in this communication class are aperiodic. Similarly, if some state in a communication class is periodic, then all states in this communication class must be periodic. We can thus divide all communication classes into two categories: the aperiodic communication classes (consisting of only aperiodic states) and the periodic communication classes (consisting only of periodic states).

Definition 2.6.8. An irreducible Markov chain is called aperiodic if some (and hence, all) states in this chain are aperiodic.

Proof of Lemma 2.6.6. From $i \leftrightsquigarrow j$ it follows that $i \rightsquigarrow j$ and $j \rightsquigarrow i$. Hence, we can find $r, s \in \mathbb{N}_0$ such that $p_{ji}^{(r)} > 0$ and $p_{ij}^{(s)} > 0$. Since the state i is aperiodic, by Lemma 2.6.5 we can find $N \in \mathbb{N}$ such that for all $n > N$, we have $p_{ii}^{(n)} > 0$ and hence,

$$p_{jj}^{(n+r+s)} \geq p_{ji}^{(r)} \cdot p_{ii}^{(n)} \cdot p_{ij}^{(s)} > 0.$$

It follows that $p_{jj}^{(k)} > 0$ for all $k := n + r + s > N + r + s$. By Lemma 2.6.5, this implies that j is aperiodic. $\square$