Repetition

The defintion fix three things: a sequence of **random variables** (the chain), a **state space** (in which the random variables takes values), and the **rules for transition** (the transition matrix).

Definition 2.1 ***Markov chain***
Let P be a $k \times k$-matrix with elements $\{ P_{i,j} \ : i,j = 1,...k \}$ A random process $(X_0, X_1, ...)$ with finite state space $S = \{s_1, s_2, ..., s_k\}$ is said to be a (homogenous) Markov chain with transition matrix P if for all n, all $i, j \in \{1, ..., k\}$ and all $i_0, ..., i_{n-1} \in \{1, ..., k\}$ we have

$$\mathbb{P}(X_{n+1} = s_j | X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = s_j | X_n = i) = P_{i,j}$$

The Markov property "The future depends on the past through the present."

$$\mathbb{P}(X_{n+1} = s_j | X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = s_j | X_n = i)$$

Representation I: transition matrix We can represent the Markov chain by a matrix containing the transition probabilities, or ...

Representation II: transition graph ... we can represent the Markov chain with a transition graph where a positive transition probability is represented by an arrow.

Time homogeneity The property that the transition probabilities doesn't change over time.

About simulation of Markov chains

Simulating a Markov chain is about simulating the sequence of variables $X_0, X_1, X_2, ...$ For each variable we know the distribution to simulate from.

The distribition for X_0 is often stated explicitly. The distribution for X_n, $n \geq 1$ is given by the rows of the transition matrix, that is

$$\mathbb{P}(X_n = j | X_{n-1} = i) = P_{i,j}$$

When writing simulation programs this is about using $U[0, 1]$ random numbers to get the correct distribution in every step.

A thorough treatment of this is given in the coursebook, chapter 3.

Important properties

Two important properties of Markov chains is irreducibility and aperiodicity. The final goal of our work with Markov Chain theory is the property of having stationary distributions. It is very important when using Markov chains as tools for simulation.

The theoretical property is the existence of the stationary distribution. The practical utility is that the distribution of X_n approaches the stationary distribution as n grows larger.

Both irreducibility and aperiodicity are properties such that if they are fulfilled by a finite state Markov chain there exists a stationary distribution. Since we are studying Markov chains in connection with computer algorithms we always have a finite state space due to a finite computer memeory.

Irreducible Markov chains

Let us consider Markov chains on a small state space $S = \{s_1, s_2, s_3, s_4, s_5\}$.

Some examples ...

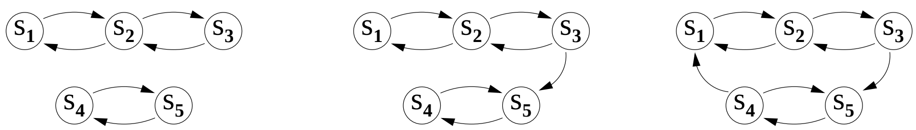

Figure 2.1: Examples of three Markov chains the one to the right is irreducible while the other two are not.

Irreducibility is the property that **regardless the present state we can reach any other state in finite time** . Mathematically it is expressed as ...

$$\forall i, j \in S, \exists m < \infty : \mathbb{P}(X_{n+m} = j | X_n = i) > 0$$

... and is easily (depending on size, of course) seen in the transition graph representation of the Markov chain.

Example 4.1 see page 27 in the course book.

Aperiodic Markov chains

Consider the following three small Markov chains, here represented by their transition graphs. An arrow means positive transition probability, no arrow means zero transition probability.

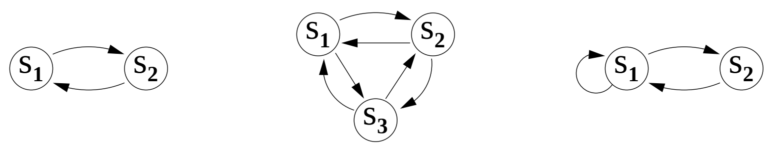

Figure 2.2: Examples of three Markov chains of which the left one has period 2 while the other two both are aperiodic.

Periodicity has to do with what period the occurrance of a state has. If a state s_i have period is 2 the chain can be in s_i every second time, that is on even or odd times depending on where we start, but not both. If a state has period 1 we say that it is aperiodic.

When considering periodicity we always look at the set of **possible** times we can be in a certain state, it is more general than any realization (outcome of a simulation) of the chain.

The next two theorems gives us properties of the n-step transition matrices if the chain is aperiodic.

Theorem 4.1
Suppose that we have an aperiodic Markov chain $(X_0, X_1, ...)$ with state space $S = \{s_1, ..., s_k\}$ and transition matrix P. Then there extists an $N < \infty$ such that

$$(P^n)_{i,i} > 0$$

for all $i \in \{1, ..., k\}$ and all $n \geq N$.

Interpretation/Consequences:
Give us a property regarding n-step return probabilities of aperiodic finite state Markov Chains.

Idea of proof:
We fix an arbitrary state $s_i \in S$ and let $A_i = \{n : (P^n)_{i,i} > 0\}$, that is, n such that if $X_0 = s_i$ then with positive probability we can have $X_n = s_i$. This set is closed under addition, that is,

$$a, b \in A_i \quad \Rightarrow \quad a + b \in A_i$$

and then uses a theorem from number theory to get the result for s_i. Since $s_i \in S$ was arbitrary the argument holds for any state and the result follows.

Corollary 4.1
Let $(X_0, X_1, ...)$ *be an irreducible and aperiodic Markov chain with state space* $S = \{s_1, ..., s_k\}$ *and transition matrix* P. *Then there extists an* $M < \infty$ *such that* $(P^n)_{i,j} > 0$ *for all* $i, j \in \{1, ..., k\}$ *and all* $n \geq M$.

Interpretation/Consequences:
Give us a property regarding n*-step jump probabilities of aperiodic finite state Markov Chains.*

Idea of proof:
As a corollary of theorem 4.1 it is just a application of that theorem. There is one problem to consider though.
Irreducibility gives us a $m_{i,j} < \infty$ *such that* $(P^{m_{i,j}})_{i,j} > 0$ *for each pair* $s_i, s_j \in S$. *Nothing however states that* $(P^n)_{i,j} > 0$ *for any* $n \neq m_{i,j}$, *and the result we want is a number* $M < \infty$ *such that we have* $(P^n)_{i,j} > 0$ *for all* $n \geq M$. *For this we have to use theorem 4.1 and not just irreducubility, and to do that we need aperiodicity.*

Example 4.2 see page 27 in the course book. Restate this problem as a theorem and prove it :-) It gives us a useful critera for determine if a chain is aperiodic.

Recommended excercise

Problem 4.3 on page 2. About a Markov chain based on chess moves.