

UiO : **Department of Mathematics**
University of Oslo

Lecture 1. Introduction to Markov chains

STK2130 – Modelling av stokastiske
prosesser



Anton Yurchenko-Tytarenko



NIELS HENRIK ABEL
1802 – 1829

MATEMATIKER, BERØMT FOR
BANEBRYTENDE ARBEIDER INNEN
LIGNINGSTEORI, UENDELIGE REKKER
OG ELLIPTISKE FUNKSJONER



Discrete time stochastic processes

Definition

A discrete time stochastic process is a sequence $\{X_0, X_1, X_2, \dots\}$ of random variables with values in some set $\mathcal{S}$.

- The process is sometimes denoted by $\{X_n : n \geq 0\}$ or by $\{X_n\}$.
- The index n is interpreted as time, and is a non-negative integer.
- If $\mathcal{S}$ is finite or countable (e.g. $\{0, 1\}$ or $\mathbb{N}$), the process has a discrete (state) space.
- The elements of $\mathcal{S}$ are referred to as states. Thus, X_n is the *state* of the process at time n , $n = 0, 1, 2, \dots$

Remark

In simple cases $X_0, X_1, \dots$ are independent. However, in this course we will focus on cases where $X_0, X_1, \dots$ are *dependent*.

Markov chains

Definition

A discrete time, discrete space stochastic process $\{X_0, X_1, X_2, \dots\}$ is called a *time homogeneous Markov chain* if its *transition probabilities* satisfy the following:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

for all states $i_0, i_1, \dots, i_{n-1}, i, j \in S$ and all $n \geq 0$.

- The relation between conditional probabilities above means that “*the future*” value X_{n+1} of the process depends on the “*history*” of the process $X_0, \dots, X_n$ *only through the value in the previous time point* X_n .

Notation for transition probabilities and one-step transition probability matrix

Notation:

- $P_{i,j} := \mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0)$.
- $P_{i,j}$ can be interpreted as probability of getting to state j given that the previous value of the process is i .
- One-step transition probability matrix:

$$\mathbf{P} = [P_{i,j}]_{i,j \in \mathcal{S}},$$

where $P_{i,j} \geq 0$ for all $i, j \in \mathcal{S}$, and $\sum_{j \in \mathcal{S}} P_{i,j} = 1$ for all $i \in \mathcal{S}$.

Example 1. Forecasting the Weather

Consider a Markov chain with state space $\mathcal{S} = \{0, 1\}$, where 0 represents rain and 1 represents no rain. By X_n we denote a random variable which represents presence/absence of the rain on day n .

We assume that:

- If it rains today, then it will rain tomorrow with probability α ,
- If it does not rain today, then it will rain tomorrow with probability β .

Mathematically, this can be expressed as follows:

$$P_{0,0} = \mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = \alpha, \quad n = 0, 1, 2, \dots$$

$$P_{1,0} = \mathbb{P}(X_{n+1} = 0 \mid X_n = 1) = \beta, \quad n = 0, 1, 2, \dots$$

This implies that we also have:

$$P_{0,1} = \mathbb{P}(X_{n+1} = 1 \mid X_n = 0) = 1 - \alpha, \quad n = 0, 1, 2, \dots$$

$$P_{1,1} = \mathbb{P}(X_{n+1} = 1 \mid X_n = 1) = 1 - \beta, \quad n = 0, 1, 2, \dots$$

Example 1. Forecasting the Weather

Remark

- 1 In this example, we assume that the transition probabilities *do not change over time*, i.e., that the process is *time-homogeneous*.
- 2 Since we have assumed that $\{X_n\}$ is a Markov chain, this implies that given the weather today, the weather tomorrow is independent of the weather in past days.

The one-step transition probability matrix for $\{X_n\}$ is given by:

$$\mathbf{P} = \begin{pmatrix} P_{0,0} & P_{0,1} \\ P_{1,0} & P_{1,1} \end{pmatrix} = \begin{pmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

Example 2. A Communications System

A communications system transmits the digits 0 and 1. Each digit transmitted must pass through several stages, at each of which there is a probability p that the digit entered will be unchanged when it leaves.

We let X_n denote the digit entering the n -th stage.

Then $\{X_n : n \geq 0\}$ is a Markov chain with state space $\mathcal{S} = \{0, 1\}$, and with a transition probability matrix:

$$\mathbf{P} = \begin{pmatrix} P_{0,0} & P_{0,1} \\ P_{1,0} & P_{1,1} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}$$

Example 3. Mood swings

A person (Gary) has three mood states: cheerful, so-so and glum. The “mood” state space is denoted by $\mathcal{S} = \{0, 1, 2\}$, where 0 represents cheerful, 1 represents so-so and 2 represents glum.

We let X_n denote Gary’s mood state at day n , $n = 0, 1, 2, \dots$, and assume that $\{X_n\}$ is a Markov chain. Thus, Gary’s mood state tomorrow depends only on his mood state today, not on previous days.

The one-step transition probability matrix for $\{X_n\}$ is assumed to be:

$$\mathbf{P} = \begin{pmatrix} P_{0,0} & P_{0,1} & P_{0,2} \\ P_{1,0} & P_{1,1} & P_{1,2} \\ P_{2,0} & P_{2,1} & P_{2,2} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}$$

Example 3. Mood swings

$$\mathbf{P} = \begin{pmatrix} P_{0,0} & P_{0,1} & P_{0,2} \\ P_{1,0} & P_{1,1} & P_{1,2} \\ P_{2,0} & P_{2,1} & P_{2,2} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}$$

Remark

- 1 The highest numbers in this transition probability matrix are $P_{0,0}$ and $P_{2,2}$. Thus, if Gary is cheerful or glum one day, the most likely mood state the next day will be the same mood state.
- 2 The lowest numbers in this transition probability matrix are $P_{2,0}$ and $P_{0,2}$. Thus, a change in mood state directly from cheerful to glum or the other way around are the least likely changes to occur.

Example 4. Transforming a Process into a Markov Chain

Suppose that whether or not it rains today depends on previous weather conditions *through the last two days*:

- If it rained today and yesterday, then the probability of rain tomorrow is 0.7.
- If it rained today but not yesterday, then the probability of rain tomorrow is 0.5.
- If it did not rain today but it rained yesterday, then the probability of rain tomorrow is 0.4.
- If it did not rain today and yesterday, then the probability of rain tomorrow is 0.2.

Warning!

If we denote X_n the weather on day n , $\{X_n\}$ is *NOT* a Markov chain! Why?

Example 4. Transforming a Process into a Markov Chain

In order to represent the weather in our example as a Markov chain, we need to perform a trick that is sometimes called *Markovian lift*. The idea is to consider the weather on *two consecutive days* together. Namely,

- State space $S = \{RR, \bar{R}R, R\bar{R}, \bar{R}\bar{R}\}$ where:
 - RR = Rain yesterday, Rain today,
 - $\bar{R}R$ = No rain yesterday, Rain today,
 - $R\bar{R}$ = Rain yesterday, No rain today,
 - $\bar{R}\bar{R}$ = No rain yesterday, No rain today.
- By X_n denote the weather on days n and $n - 1$.

Assume that $X_n = RR$, i.e. it rained on days $n - 1$ and n . Then, for X_{n+1} only two states are possible: RR (if it rains on day $n + 1$) or $R\bar{R}$ (if it does not rain on day $n + 1$). By the above assumptions, we get:

$$\mathbb{P}(X_{n+1} = RR \mid X_n = RR) = 0.7, \quad \mathbb{P}(X_{n+1} = R\bar{R} \mid X_n = RR) = 1 - 0.7 = 0.3.$$

Example 4. Transforming a Process into a Markov Chain

Now it is easy to see that the process X_n representing the weather on two consecutive days becomes Markovian! We can calculate its transition probabilities:

$$\mathbb{P}(X_{n+1} = RR \mid X_n = RR) = 0.7, \quad \mathbb{P}(X_{n+1} = R\bar{R} \mid X_n = RR) = 1 - 0.7 = 0.3,$$

$$\mathbb{P}(X_{n+1} = RR \mid X_n = \bar{R}R) = 0.5, \quad \mathbb{P}(X_{n+1} = R\bar{R} \mid X_n = \bar{R}R) = 1 - 0.5 = 0.5,$$

$$\mathbb{P}(X_{n+1} = \bar{R}R \mid X_n = R\bar{R}) = 0.4, \quad \mathbb{P}(X_{n+1} = \bar{R}\bar{R} \mid X_n = R\bar{R}) = 1 - 0.4 = 0.6,$$

$$\mathbb{P}(X_{n+1} = \bar{R}R \mid X_n = \bar{R}\bar{R}) = 0.2, \quad \mathbb{P}(X_{n+1} = \bar{R}\bar{R} \mid X_n = \bar{R}\bar{R}) = 1 - 0.2 = 0.8.$$

Thus, we have the following transition probability matrix:

$$\mathbf{P} = \begin{pmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{pmatrix}$$

where the row and column order is: $RR, \bar{R}R, R\bar{R}, \bar{R}\bar{R}$.

Example 5. A Random Walk

Let ξ_n , $n \geq 1$, be independent identically distributed random variables such that

$$\xi_n = \begin{cases} 1, & \text{with probability } p, \\ -1, & \text{with probability } 1 - p. \end{cases}$$

Consider $X_0 = 0$ and $X_n = \sum_{k=1}^n \xi_k$. Then $\{X_n\}$ is a Markov process called (*Bernoulli*) *random walk*.

- State space: $\mathcal{S} = \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
- For any $i \in \mathcal{S}$: $P_{i,i+1} = p$, $P_{i,i-1} = 1 - p$.
- If $X_0 = 0$, then X_n is odd if n is odd and even if n is even.

A random walk is a stochastic model for the position of a particle “walking” on a straight line: at each point of time it either takes one step to the right with probability p or one step to the left with probability $1 - p$.

Example 6. A Gambling Model

Consider a gambler who initially has a fortune of M NOK. At each play of the game, either wins 1 NOK with probability p or loses 1 NOK with probability $1 - p$. If we suppose that our gambler quits playing either when he goes broke or he attains a fortune of N NOK, then the gambler's fortune is a Markov chain having transition probabilities

$$P_{i,i+1} = p, \quad P_{i,i-1} = 1 - p, \quad i = 1, 2, \dots, N - 1, \\ P_{0,0} = P_{N,N} = 1.$$

This Markov chain is quite similar to the random walk which starts at M with one important exception: if the process reaches either state 0 or state N , it “gets stuck” there. Such states are called *absorbing*.

Example 7. Automobile Insurance

The annual automobile insurance premium depends on the claim history of the client. Each policyholder is given a positive integer-valued state which represents the “reliability” of a client; the annual insurance premium is a function of this state (*i.e. the more accidents the client gets into, the more he has to pay as a premium*). A policyholder's state changes from year to year in response to the number of claims made by that policyholder in the previous year (see table below).

State	Annual Premium	Next state if			
		0 claims	1 claim	2 claims	3 or more claims
1	200	1	2	3	4
2	250	1	3	4	4
3	400	2	4	4	4
4	600	3	4	4	4

Example 7. Automobile Insurance

The client state X_n during year n can be modelled via a Markov chain (as the state next year depends only on the state this year). According to the table on the previous slide, the transition from state to state depends only on the number of claims.

Denote Y_n the random variable that represents number of claims on a year prior to n (so that X_n is determined by Y_n). We assume that $Y_1, Y_2, \dots$ are independent identically distributed random variables such that

$$\mathbb{P}(Y_n = k) = a_k, \quad k = 0, 1, 2, \dots$$

Now we can calculate:

- If $X_n = 1$, then
 - $X_{n+1} = 1$ only if $Y_n = 0$, i.e. $\mathbb{P}(X_{n+1} = 1 \mid X_n = 1) = a_0$,
 - $X_{n+1} = 2$ only if $Y_n = 1$, i.e. $\mathbb{P}(X_{n+1} = 2 \mid X_n = 1) = a_1$,
 - $X_{n+1} = 3$ only if $Y_n = 2$, i.e. $\mathbb{P}(X_{n+1} = 3 \mid X_n = 1) = a_2$
 - $X_{n+1} = 2$ only if $Y_n \geq 3$, i.e. $\mathbb{P}(X_{n+1} = 4 \mid X_n = 1) = 1 - a_0 - a_1 - a_2$.

Example 7. Automobile Insurance

By similar arguments we find that the transition probability matrix of the process is:

$$\mathbf{P} = \begin{pmatrix} a_0 & a_1 & a_2 & 1 - a_0 - a_1 - a_2 \\ a_0 & 0 & a_1 & 1 - a_0 - a_1 \\ 0 & a_0 & 0 & 1 - a_0 \\ 0 & 0 & a_0 & 1 - a_0 \end{pmatrix}$$

UiO : Department of Mathematics

University of Oslo



Anton Yurchenko-Tytarenko



Lecture 1. Introduction to Markov chains

STK2130 – Modelling av stokastiske prosesser

