MARKOV CHAINS

GROUP 6

Kwame Nkrumah University of Science and Technology

May 10, 2020

GROUP MEMBERS	INDEX NUMBERS
Achiaa-Frimpong Precious	6745816
Adjei Daniel Owusu	6746716
Agyei Ebenezer	6747916
Akowuah Dennis	6748416
Ankutse Lawrence	6749916
Appiah Doreen	6750516
Obeng Okai Daniel	6759616
Oppong Richard	6760816
Tackie-Otoo Nii Ohene	6763316

Objectives

By the end of this project, it our collective expectation that each and everyone gets to know;

1. Stochastic process
2. Markov chains
3. Characteristics of Markov chains
4. Transition probabilities

Introduction

Modern probability theory studies chance processes for which the knowledge of previous outcomes influences predictions for future experiments. In principle, when we observe a sequence of chance experiments, all of the past outcomes could influence our predictions for the next experiment. For example, this should be the case in predicting a student's grades on a sequence of exams in a course. But to allow this much generality would make it very difficult to prove general results. In 1907, Andrei Andreevich Markov began the study of an important new type of chance process. In this process, the outcome of a given experiment can affect the outcome of the next experiment. This type of process is called a Markov chain.

Main work

Stochastic process

Let τ be a subset of $[0, \infty)$. A family of random variables $\{X_t\}_{t \in \tau}$ is called **stochastic or random process**.
When $\tau = N$, $\{X_t\}_{t \in \tau}$ is said to be a **discrete-time process** and when $\tau = [0, \infty)$, it is called **continuous-time process**. When τ is a singleton say $\tau = \{1\}$, the process $\{X_t\}_{t \in \tau} \equiv X_1$ is really just a single random variable. When τ is finite say $\tau = \{1, 2, ..., n\}$, we have a random vector. Therefore, we can say that a stochastic process is generalization of random vectors. The idea of a stochastic process is very important both in mathematical theory and its applications in *science, engineering, economics etc*. It its used to model a large number of various phenomena where the quantity of interest varies discretely or continuously through time, t in a non-predictable way.

Note, a stochastic process has a Markov property if the conditional probability distribution of future states of the process, that's, conditional on both past and present values depends only on the present state, not on the sequence of events that preceded it.
The values assumed by the random variables $\{X_t\}$ are called states. The set of all possible states forms the state space of the process and this may be discrete or continuous. If the state spece is discrete, the process is referred to as a chain and the states are usually identified with the set of natural numbers $\{0, 1, 2, ...\}$ or a subset of it. An example of a discrete space is the number of customers at a service facility.
An example of a continuous state space is the length of time a customer has been waiting for service.

Main work

Markov chains

A Markov process is a stochastic process whose conditional probability distribution function staisfies the Markov property.
In Markov process, we have;

- Discrete-time Markov chains
- Continuous-time Markov chains

For a discrete-time Markov chains, we observe that its state at a discrete, but infinite, set of times.
Transitions from one state to another can only take place or fail.
The discrete index set τ of the underlying stochastic process by the set of natural numbers $\{0, 1, 2, ...\}$. The successive observations define the random variables $X_0, X_1, X_2, ..., X_n, ...$ at time steps $0, 1, 2, ..., n, ...$ respectively.

So, a discrete-time Markov chain $\{X_n; n = 0, 1, 2, ...\}$ is a staochastic process that satisfies the following relationship called the Markov property;
Prob$\{X_{n+1} = x_{n+1}, X_{n-1} = x_{n-1}, ..., X_0 = x_0\}$
$= Prob\{X_{n+1} = x_{n+1} | X_n = x_n\}$
This means that the probability that the next random variable(X_{n+1}) corresponds to the next state(x_{n+1}) given that the current random variable(X_n) corresponds to the current state(x_n).
Also, we have *transition probability matrix* or *chain matrix*.

$$p(n) = \begin{bmatrix} p_{00}(n) & p_{01}(n) & p_{02}(n) & \dots & p_{0j}(n) & \dots \\ p_{10}(n) & p_{11}(n) & p_{12}(n) & \dots & p_{1j}(n) & \dots \\ p_{20}(n) & p_{21}(n) & p_{22}(n) & \dots & p_{2j}(n) & \dots \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ p_{i0}(n) & p_{i1}(n) & p_{i2}(n) & \dots & p_{ij}(n) & \dots \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{bmatrix}$$

The elements of the matrix p(n) satisfy the following two properties;

$$0 \leq p_{ij}(n) \leq 1$$

for all i,

$$\sum_{\forall j} p_{ij}(n) = 1$$

This means, a matrix that satisfies these properties is called a Markov matrix or stochastic matrix.
A Markov chain is said to time-homogenous if for all states i and j we have,
$Prob\{X_{n+1} = j|X_n = i\} = Prob\{X_{n+m+1} = j|X_{n+m} = i\}$
for n=0,1,2,.. and $m \geq 0$

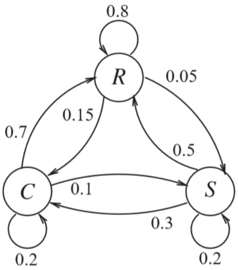
A good mental image to have about continuous-time Markov chain is simply a discrete time Markov chain in which transitions can happen at any time.

A homogenous Markov chain,
$p_{ij} = Prob\{X_1 = j|X_0 = i\} = Prob\{X_2 = j|X_1 = i\} = Prob\{X_3 = j|X_2 = i\} = ...$
In this case, $p_{ij}(n)$ is replaced with p_{ij}, since transition no longer depends on n, Where n is an infinite set of time step.
On the other hand, a Markov chain is said to be time-nonhomogenous if
$p_{ij}(0) = Prob\{X_1 = j|X_0 = i\} \neq Prob\{X_2 = j|X_1 = i\} = p_{ij}(1)$
Considering an example of a homogenous discrete-time Markov chain on a weather model that describes the daily weather pattern in Belfast, Northern Ireland.
Three weather patterns are being considered; rainy, cloudy, and sunny.
These three weather conditions describe the three states of Markov chain;
state 1(R) represents rainy day;
state 2(C) represents cloudy day;
state 3(S) represents sunny day.

Main work

Markov chains

The weather is observed daily.
On any given rainy day, the probability it will rain the next day is estimated at 0.8*; the probability that the next day will be cloudy is* 0.15*, while the probability that tomorrow will be sunny is only* 0.05*.*
Similarly, the figure below shows the probabilities assigned when a particular day is cloudy or sunny.



$$p = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$$

The elements in P represent conditional probabilities.
For example, element P_{32} tells us that the probability that tomorrow is cloudy, given that today is sunny, is 0.3.

Characteristics of Markov chain

- No matter how the process arrived at its present state, the possible future states are fixed.
  Which means, the probability of transitioning to any particular state is dependent solely on the current state and time elapsed.
- The *state space*, or set of all possible states, can be anything:numbers, weather conditions, letters etc.

Main work

Transitions

Transitions are the changes of the state of the system.
The probabilities associated with various state changes are called *transition probabilities*
A *transition matrix* or *transition probability matrix* describes the probabilities of a particular transitions and an initial state across the state space.

Example

Consider a Markov chain with three possible states **1**,**2** and **3** and the following transition probabilities

$$p = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

a. Find $P(X_4 = 3|X_3 = 2)$
b. Find $P(X_3 = 1|X_2 = 1)$
c. Find $P(X_0 = 1, X_1 = 2)$, if $P(X_0 = 1) = \frac{1}{3}$

Solution

(a)
By definition,
$P(X_4 = 3|X_3 = 2) = p_{23} = \frac{2}{3}$
(b)
By definition,
$P(X_3 = 1|X_2 = 1) = p_{11} = \frac{1}{4}$
(c)
$P(X_0 = 1, X_2 = 1) = P(X_0 = 1)P(X_1 = 2|X_0 = 1)$
$= \frac{1}{3} . p_{12}$
$= \frac{1}{3} . \frac{1}{2} = \frac{1}{6}$

Conclusion

Markov chains are an important mathematical tool in stochastic processes. In orther words, They are extremely applicable in many industrial, biological and many other fields.
Markov chains also help in minimizing time spent on analyzing data and thereby maximize work output.

References

- Kemeny, John G.; Hazleton Mirkil; J. Laurie Snell; Gerald L. Thompson (1959). *Finite Mathematical Structures* (1st ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc. Library of Congress Card Catalog Number 59-12841. Classical text. cf Chapter 6 *Finite Markov Chains* pp. 384ff.
- Seneta, E. *Non-negative matrices and Markov chains*. 2nd rev. ed., 1981, XVI, 288 p., Softcover Springer Series in Statistics. (Originally published by Allen & Unwin Ltd., London, 1973)