ETC3430: Financial mathematics under uncertainty

Dan Zhu

Monash Business School
Monash University, Australia

Semester 1, 2021

March 2, 2022

Outline

When my first child started in daycare, I started to register the outcome of a stochastic variable with two possible outcomes

- ill
- ok

Consecutive recordings of the health state of a child made every morning is an excellent example of a sample of a discrete-time stochastic process. The sampling regime is discrete because I do not register the health state continuously at any time point but only once a day. The process is stochastic (in contrast to deterministic) because I never know with certainty whether the child will be ill or healthy on the following morning.

A sample path

The sample of the health state on the first 17 days (also referred to as the sample path) is given below

ok, ok ,ok ,ok ,ok ,ill ,ill ,ill ,ill ,ill ,ok ,ok ,ok ,ok ,ok ,ok ,ill ,...

A stochastic process is a mathematical model for a sequence of random variables. The model should allow us to compute the probability of various events associated with random phenomena.

Questions to Answer

- What is the average time between two periods of illness?
- How many days should I expect to be off work to take care of a sick child within the next year?

One of the aims of this course is to learn how to build mathematical models for random events that allows us to compute the answer to the type of questions listed above.

Markov Chains in a Netshell

Within the class of stochastic processes, one could say that Markov chains are characterised by the dynamical property that they never look back. The way a Markov chain continues tomorrow is affected by where it is today but independent of where it was yesterday or the day before yesterday. As long as we know the present value of a Markov chain, our prediction about the future behaviour of the the process does not change if we get additional information about past recordings of the process.

Another example: Gambling

Suppose that you start with \$10, and you wager \$1 on an unending, fair, coin toss indefinitely, or until you lose all of your money. Let X_t represents the number of dollars you have after n tosses; then the sequence is a Markov process. If I know that you have \$12 now, you will either have \$11 or \$13 after the next toss with even odds. One can't improve the guess the added knowledge that you started with \$10, then went up to \$11, down to \$10, up to \$11, and then to \$12. The fact that the knowledge of earlier tosses does not improve the guess showcases the Markov property, the memoryless property of a stochastic process.

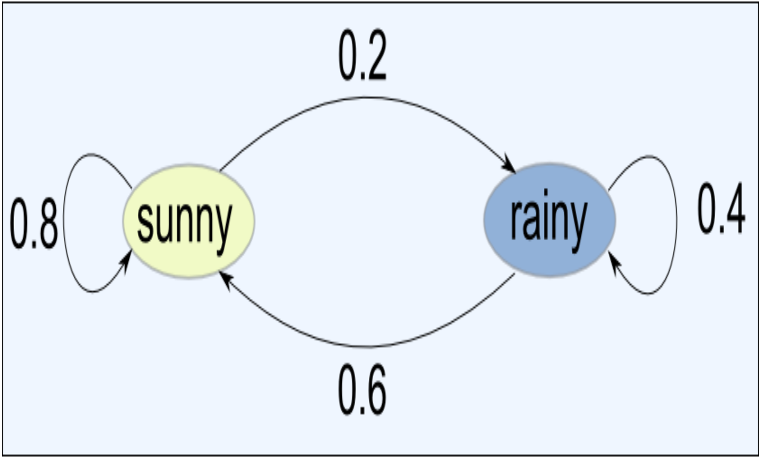
In class Exercise

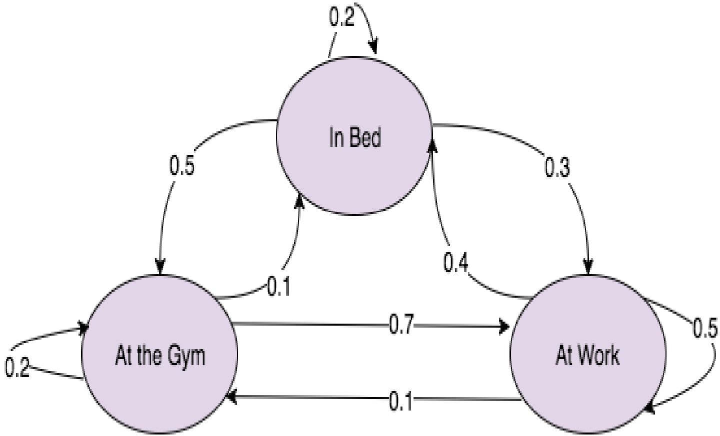
Come up with your own example where the Markov chain is appropriate.

Transition Diagrams

We advocate for visualising the dynamics of a Markov chain whenever possible. The transition graphs with nodes (or vertices) represent the Markov chains' states, and edges represent transitions.

Weather




0.2
In Bed
0.5
0.3
0.1
0.4
0.7
0.5
0.2
At the Gym
At Work
0.1

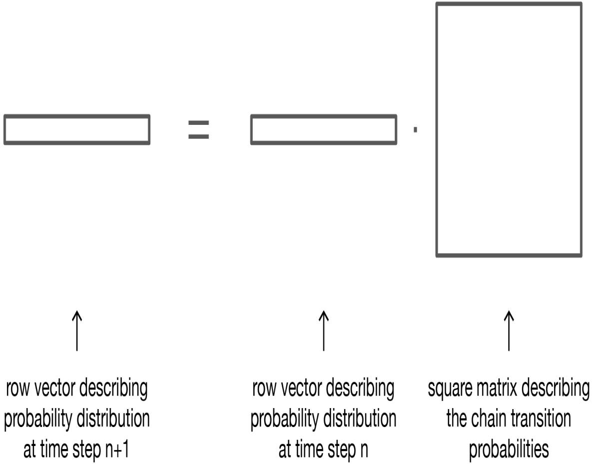
Rat in the open maze

Example (Rat in the open maze)

Consider a rat in a maze with four cells, indexed $1-4$*, and the outside (freedom), indexed by 0 (that can only be reached via cell 4). The rat starts initially in a given cell and then takes a move to another cell, continuing to do so until finally reaching freedom. At each move (transition), the rat, independent of the past, is equally likely to choose from among the neighbouring cells,i.e., the rat does not learn from past mistakes.*

Five minute exercise: Draw the transition diagram for this example!

Transition


row vector describing probability distribution at time step n+1
row vector describing probability distribution at time step n
square matrix describing the chain transition probabilities

Markov Property Recall

Consider a discrete time stochastic process $Z_n \in \mathbb{S}$, it is said to be *Markov* if for all $n > 0$, the distribution of Z_{n+1} is determined by the state of Z_n, and does not depends on the past values Z_j for $0 \le j < n$, mathematically

$$\mathbb{P}(Z_{n+1} = j | Z_n = i_n, Z_{n-1} = i_{n-1}, ...Z_0 = i_0) = \mathbb{P}(Z_{n+1} = j | Z_n = i_n).$$

Note: Independent increment is a sufficient condition for Markov property, but not all markov chains have independent increments!

Rat in the open maze

Let X_n denotes the cell visited right after the nth move

- discrete state space $\mathbb{S} = 0, 1, 2, 3, 4$
- a discrete time $\mathbb{J} = 0, 1, 2, 3....$
- Markov.

The Joint and Marginal Distributions

Due to Markov Property, it is easy to show that

$$\begin{aligned}&\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, ...Z_0 = i_0)\\=&\mathbb{P}(Z_0 = i_0)\prod_{j=1}^{n}\mathbb{P}(Z_j = i_j | Z_{j-1} = i_{j-1})\end{aligned}$$

and

$$\mathbb{P}(Z_1 = i) = \sum_{j\in\mathbb{S}}\mathbb{P}(Z_1 = i | Z_0 = j)\mathbb{P}(Z_0 = j).$$

Transition Matrix

Definition

A discrete time Markov chain is said to be time homogeneous if the transition probability

$$\mathbb{P}(Z_n = j | Z_{n-1} = i), \quad i, j \in \mathbb{S}$$

is independent of $n \in \mathbb{N}$.

Transition Matrix

Definition

Let $d = \dim(\mathbb{S})$*, the probability of transition for a discrete time time homogeneous Markov chain with discrete state space* $\mathbb{S}$ *is a transition matrix,* $\mathbb{P} \in [0,1]^{d \times d}$ *such that*

$$\mathbb{P}_{i,j} = \mathbb{P}(Z_n = j | Z_{n-1} = i) \;\; i, j \in \mathbb{S}.$$

- Since $\sum_{j \in \mathbb{S}} \mathbb{P}(Z_n = j | Z_{n-1} = i) = 1$ for all i, we have $\sum_{j \in \mathbb{S}} \mathbb{P}_{i,j} = 1$, i.e. the sum each row of the transition matrix is one.
- Absorbing state: $\mathbb{P}_{k,k} = 1$.

Rate in the maze

The transition matrix of the rat in the maze is

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 \\ 1/3 & 0 & 1/3 & 1/3 & 0 \end{bmatrix}$$

The absorbing state is the first one.

Five minute exercise, write down the transition matrix for the other two examples.

Transition Matrix in General

In general, a discrete time Markov chain can be described the transition matrix $\mathbb{P}^{m,n}$ such that

$$\mathbb{P}^{m,n}_{i,j} = \mathbb{P}(X_n = j | X_m = i)$$

The one step case is

$$\mathbb{P}^{m,m+1}_{i,j} = \mathbb{P}(X_{m+1} = j | X_m = i).$$

When the one step transition is independent of m, the time index, we have the homogeneous case.

What is the distribution of X in the future?

Consider the process at time n, X_n, the probability of $X_n = j$ is

$$\mathbb{P}(X_n = j) = \sum_{i \in \mathbb{S}} \mathbb{P}(X_0 = i)\mathbb{P}(X_n = j | X_0 = i)$$

Let π_0 denote the initial distribution of X_0 , i.e. a vector that each element $\pi_{0,j} = \mathbb{P}(X_0 = j)$, the above expression becomes

$$\mathbb{P}(X_n = j) = \left(\pi_0 \mathbb{P}^{0,n}\right)_j$$

the jth element of the $1 \times \dim(S)$ vector $\pi_0 \mathbb{P}^{0,n}$.

The Probability distribution: $X_n \sim \pi_n$

Hence, we can be obtained via

$$\pi_n = \pi_0 \mathbb{P}^{0,n}.$$

In general, if we know the distribution at time m, the markov property implies that

$$\pi_n = \pi_m \mathbb{P}^{m,n} \text{ for } n > m.$$

This is to say that given the distribution of the process at time m and the transitional probabilities, we know the distribution of the markov process in any time $m > n$.

The Chapman-Kolmogorov equation is an identity relating the joint probability distributions of different sets of coordinates on a stochastic process. It facilitates the analysis of long term behaviour of Markov chains.

Theorem

The transition probabilities of a discrete time Markov chain obey the CK equations

$$\mathbb{P}_{i,j}^{m,n} = \sum_{k \in \mathbb{S}} \mathbb{P}_{i,k}^{m,l} \mathbb{P}_{k,j}^{l,n}$$

for all $i, j \in \mathbb{S}$ and all integer times $m < l < n$.

This is to say

$$\mathbb{P}^{m,n} = \mathbb{P}^{m,l} \mathbb{P}^{l,n}.$$

Quick Proof

The logic behind this is rather simple, that

$$
\begin{aligned}
&\mathbb{P}(X_n = j|X_m = i) \\
=&\sum_{k\in\mathbb{S}} \mathbb{P}(X_n = j, X_l = k|X_m = i) \text{ Disjoint Events} \\
=&\sum_{k\in\mathbb{S}} \mathbb{P}(X_n = j|X_l = k, X_m = i)\mathbb{P}(X_l = k|X_m = i) \text{ Conditional Prob} \\
=&\sum_{k\in\mathbb{S}} \mathbb{P}(X_n = j|X_l = k)\mathbb{P}(X_l = k|X_m = i) \text{ Markov}
\end{aligned}
$$

Rate in the Maze

For the rat in the maze chain,

$$\mathbb{P}^{(2)}_{1,1} = \mathbb{P}(X_2 = 1 | X_0 = 1)$$

denote the probability that the rat, starting initially in cell 1, is back in cell 1 two steps later. Clearly, this can happen only if the rat goes to cell 2 then back to cell 1, or goes to cell 3 then back to cell 1, yielding

$$\begin{aligned} &\mathbb{P}(X_1 = 2, X_2 = 1 | X_0 = 1) + \mathbb{P}(X_1 = 3, X_2 = 1 | X_0 = 1) \\ =&0.25 + 0.25 = 0.5. \end{aligned}$$

Time-Homogeneous Markov chain $\mathbb{P}^{(n)}$

Given a time-Homogeneous Markov chain $\{X_n\}$ with transition matrix $\mathbb{P}$, it is of interest to consider the analogous n-step transition matrix, $\mathbb{P}^{(n)}$

$$\mathbb{P}^{(n)}_{i,j} = \mathbb{P}(X_{n+m} = j | X_m = i) \text{ for all } m$$

a n-step transition probability, denotes the probability that n time units later the chain will be in state j given it is now (at time m) in state i.

Time-Homogeneous Markov chain $\mathbb{P}^{(n)}$

Lemma

The n-step transition matrix

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

That is $\mathbb{P}^{(n)}$ is equal to $\mathbb{P}$ multiplied by itself n times.

The proof of above lemma is by induction using CK equations.

Rat in the maze

$$\mathbb{P}^{(2)}_{1,1} = 0.5$$

$$\mathbb{P}^{(2)} = \mathbb{P} \times \mathbb{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ 1/6 & 0 & 5/12 & 5/12 & 0 \\ 1/6 & 0 & 5/12 & 5/12 & 0 \\ 1/3 & 1/3 & 0 & 0 & 1/3 \end{bmatrix}$$

The Probability distribution: $X_n \sim \pi_n$

The CK equation allows us to break down long term transition into step-wise transitions

$$\pi_n = \pi_0 \mathbb{P}^{0,1} \mathbb{P}^{1,2}, ..., \mathbb{P}^{n-1,n}$$

Time-homogeneity further simplifies this

$$\pi_n = \pi_0 \mathbb{P}^n.$$

Example: No Claim Discount Policy

Example

A Motor insurance company grants its customers with no discount, 25% discount, and 60% discount. The transition probability is given by

$$\mathbb{P} = \begin{bmatrix} 0.25 & 0.75 & 0 \\ 0.25 & 0 & 0.75 \\ 0 & 0.25 & 0.75 \end{bmatrix}$$

Given in 2015, the distribution of customers in the three states are 0.8, 0.1and 0.1 respectively, the 2019 distribution is

$$(0.8, 0.1, 0.1)\mathbb{P}^4 = (0.1023, 0.2262, 0.6715)$$

Time Inhomogeneous Markov Chain

The CK equation implies that

$$\mathbb{P}^{m,n} = \mathbb{P}^{m,m+1}\mathbb{P}^{m+1,m+2}, \dots \mathbb{P}^{n-1,n}.$$

However, a $\mathbb{P}^{k,k+1}$'s are **not** the same, the above expression can not be further simplified, hence time inhomogeneous Markov Chain are harder to analyze.

Reversionary annuity

Example

While a husband is alive he pays into a scheme which will make regular payments to his wife after he is dead. The four states:

$$\{H\&W \quad alive, H \quad alive, W \quad alive, neither \quad alive\},$$

and transitions are clear, but again probabilities are age-dependent.

NCD

Example

We can also construct a time inhomogeneous version of NCD to reflect the changes in traffic condition and customer cohorts, hence the yearly transition matrix

$$\mathbb{P}^{t,t+1} = \begin{bmatrix} p_{0,0}(t) & p_{0,1}(t) & 0 \\ p_{1,0}(t) & 0 & p_{1,2}(t) \\ 0 & p_{2,1}(t) & p_{2,2}(t) \end{bmatrix}$$

that is the state transition becomes a matrix function of time.

The Stationary Distribution

Definition

A probability distribution on $\mathbb{S}$, i.e. a family π such that $(\pi_i)_{i\in\mathbb{S}}$ in $[0,1]$ such that

$$\sum_{i\in\mathbb{S}} \pi_i = 1$$

is said to be stationary if, starting at $X_0 \sim \pi$, it turns out $X_1 \sim \pi$ at time , i.e.

$$\pi = \pi\mathbb{P}.$$

A stationary distribution of a Markov chain is a probability distribution that remains unchanged in the Markov chain as time progresses, i.e. once it is reach, it will not change subsequently.

Solve for the stationary distribution

In general, we need to solve for

$$\pi = \pi \mathbb{P}$$

n unknown with n equations. In addition, the sum of π equals 1.

Example (NCD)

$$\mathbb{P} = \begin{bmatrix} 0.25 & 0.75 & 0 \\ 0.25 & 0 & 0.75 \\ 0 & 0.25 & 0.75 \end{bmatrix}$$

We can construct three equations with three unknowns

$$0.25\pi_0 + 0.25\pi_1 = \pi_0$$

$$0.75\pi_0 + 0.25\pi_2 = \pi_1$$

$$0.75\pi_1 + 0.75\pi_2 = \pi_2$$

$$\pi_0 + \pi_1 + \pi_2 = 1.$$

Solving the system give $\pi = (0.0769, 0.2308, 0.6923)$.