

1. FINITE MARKOV CHAINS

1.1. **Concept and examples.** On the first day I explained the concept behind finite Markov chains, gave the definition and two examples. But first I explained how you convert a higher order difference equation into a first order matrix equation. When we randomize this process we get a finite Markov chain.

1.1.1. *reduction to first order.* I used the Fibonacci sequence as an example to illustrate how higher order equations can be reduced to first order equations in more variables. The Fibonacci sequence is the sequence

$$1, 1, 2, 3, 5, 8, 13, \dots$$

given by the second order difference equation

$$f(n) = f(n-1) + f(n-2).$$

To convert this to first order you let

$$g(n) := f(n-1).$$

Then $f(n-2) = g(n-1)$ and the original equation becomes:

$$f(n) = f(n-1) + g(n-1).$$

Thus $(f(n), g(n))$ depends only on $(f(n-1), g(n-1))$ and the relation is given by the matrix equation:

$$(f(n), g(n)) = (f(n-1), g(n-1)) \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

I explained it like this: You have to make your decision about what to do tomorrow based on the information you have today. You can only use the information that you had yesterday if you *recorded it*. Thus, every day, you need to record the important information, either on paper or in your computer, otherwise it is lost and won't be available tomorrow.

The Fibonacci sequence, in this first order form, looks like this:

n	0	1	2	3	4	5
$f(n)$	1	1	2	3	5	8
$g(n)$	0	1	1	2	3	5

So, on Day 4, the information you have is today's number 5 and the record you kept of yesterday's number 3. You add these to get tomorrow's number 8 and you record the number 5 so that you have still have it tomorrow. Each day you look at the information you get that day and the information that was recorded from the past. So, this process is realistic and makes sense.

1.1.2. *concept.* “A stochastic process is a random process which evolves with time.” This definition is too broad for a careful, complete mathematical analysis, especially at the beginning.

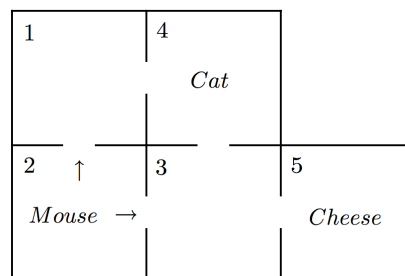
We want to start with simple models that we can analyze and understand completely. Then we will go to more and more general models adding complexities one step at a time.

A Markov chain is a stochastic process which has four simplifying assumptions:

- (1) There are only finitely many states. For example, in the Kermack-McKendrick model, there were only 3 states: S, I, R . I also used the example of the Brandeis campus. If we made the movement of people on campus into a Markov process then the set of states would be the buildings (plus one for the outside). Your exact location, for example which room you were in, is disregarded.
- (2) Time is discrete. Time is a nonnegative integer (starting at $t = 0$). For example, for movement of people on campus, people are only allowed to move from building to building on the hour. Or, we only record or notice which building people are in at 1pm, 2pm, 3pm, etc.
- (3) You forget the past. What happens at time $n + 1$ depends only on the situation at time n . Which building you are in at 2pm depends only on which building you were in at 1pm. If you add more states (more variables), you can keep track of information from the past and still satisfy the “forget the past” rule.
- (4) Rules of movement do not change with time. If, at 2pm, everyone in building 2 move to building 5 then the same thing will happen at 3pm, 4pm, etc. The Fibonacci sequence or any first order recurrence has this property.

I used two examples to illustrate these principles.

1.1.3. *mouse example.*¹



¹from “Markov Chains ... ” by Pierre Brémaud

In this example, a mouse is randomly moving from room to room. The cat and cheese do not move. But, if the mouse goes into the cat's room, he never comes out. If he reaches the cheese he also does not come out.

This will be a Markov chain with the following details.

- (1) There are 5 states (the five rooms). I numbered them: 1,2,3,4,5.
- (2) The mouse moves in integer time, say every minute.
- (3) The mouse does not remember which room he was in before. Every minute he picks an adjacent room at random, possibly going back to the room he was just in.
- (4) The probabilities do not change with time. For example, whenever the mouse is in room 3 he will go next to room 2,4 or 5 with equal probability.

The mouse moves according to the *transition probabilities*

$$p(i, j) = \mathbb{P}(\text{the mouse goes to room } j \text{ when he is in room } i).$$

These probabilities form a matrix called the *transition matrix*:

$$P = (p(i, j)) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

I pointed out the two important properties of this matrix:

- (1) Every row adds up to 1. This is because the mouse has to go somewhere or stay where he is. When all of the possibilities are listed and they are mutually exclusive, the probabilities must add up to 1.
- (2) The entries are nonnegative and at most 1 (because they are probabilities).

1.1.4. *students example.* ²

Each year, the students at a certain college either flunk out, repeat the year or go on to the next year with the following probabilities:

$$\begin{aligned} p &= \mathbb{P}(\text{flunking out of school}) \\ q &= \mathbb{P}(\text{repeating a year}) \\ r &= \mathbb{P}(\text{passing to the next year}) \end{aligned}$$

The first step is to determine what are the states. Then find the transition matrix. Later we can answer other questions, such as: What is

²from "Finite Markov Chains" by Kemeny and Snell

the probability that a Freshman will eventually graduate? and how long will it take?

There are 6 states: the student is either

- (1) Freshman
- (2) Sophomore
- (3) Junior
- (4) Senior
- (5) graduated
- (6) flunked out

The transition matrix is

$$P = (p(i, j)) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} q & r & 0 & 0 & 0 & p \\ 0 & q & r & 0 & 0 & p \\ 0 & 0 & q & r & 0 & p \\ 0 & 0 & 0 & q & r & p \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

It is important to notice that the rows add up to 1:

$$p + q + r = 1.$$

This means there are no other possibilities except for the three that were listed.

1.1.5. *definition.* Here is the precise mathematical definition.

Definition 1.1. A *finite Markov chain* is a sequence of random variables X_0, X_1, \dots which take values in a finite set S called the *state space* so that, for all $n \geq 0$ and all values of x_0, x_1, \dots, x_n , we have:

$$\mathbb{P}(X_{n+1} = x \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x \mid X_0 = x_n)$$

The $S \times S$ matrix P with entries

$$p(x, y) := \mathbb{P}(X_1 = y \mid X_0 = x)$$

is called the *transition matrix*.

The probability equation can be broken up into two steps:

$$\mathbb{P}(X_{n+1} = x \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x \mid X_n = x_n)$$

$$\mathbb{P}(X_{n+1} = x \mid X_n = x_n) = \mathbb{P}(X_1 = x \mid X_0 = x_n)$$

The first equation says that what happens at time $n + 1$ depends only on the state at time n and not on the state at previous times. The second equation says that the transition probabilities are the same at time n as they were at time 0.