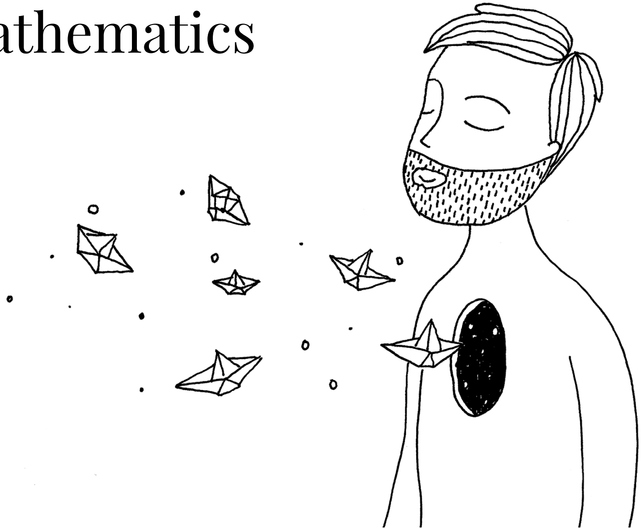
4509 - Bridging Mathematics

Markov Chains



PAULO FAGANDINI

NOVA SCHOOL OF BUSINESS & ECONOMICS

Discrete-time Markov chains

1. Time is indexed by an integer variable, say n.
2. At period n, the **state** of the chain is denoted by X_n..
3. $\mathcal{S}$ is a finite set of possible states, then $X_n \in \mathcal{S}$.
4. We will allow for m different states, then $\mathcal{S} = \{1, 2, \dots, m\}$, for $m \in \mathbb{N}$.

Discrete-time Markov chains

Definition

Markov Chain The Markov chain is described in terms of its **transition probabilities** p_{ij}: whenever the state happens to be i, there is probability p_{ij} that the next state is equal to j:

$$p_{ij} = P(X_{n+1} = j | X_n = i), \quad i, j \in \mathcal{S}$$

with $p_{ij} \geq 0$ and $\sum_{j=1}^{m} p_{ij} = 1 \ \forall i$.

Note: the probability does not depend on time, nor anything else than the present state.

How to specify then a Markov Model?

- Identify:
    1. $\mathcal{S}$ the set of states.
    2. the set of possible transitions, (i, j) where $pij > 0$
    3. the values for those p_{ij}
- The Markov chain specified by this model is a sequence of r.v.s $X_0, X_1, X_2, \ldots$, that can take values in $\mathcal{S}$, and which satisfy:

$$P(X_{n+1} = j | X_n = i, \{X_\nu = i_\nu\}_{\nu=0}^{n-1}) = p_{ij}$$

for any n, and any $i, j \in \mathcal{S}$, and all possible sequences $i_0, \ldots, i_{n-1}$ of earlier states.

It is convenient to sort all these probabilities in a two-dimensional array like this:

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

This is called the **Transition Probability Matrix**. This matrix is defined as having in each row i and column j the probability of transitioning from state i to state j.

Example, Bertsekas and Tsitsiklis (2008)

Alice is taking a probability class and in each week, she can be either up-to-date or she may have fallen behind. If she is up-to-date in a given week, the probability that she will be up-to-date (or behind) in the next week is 0.8 (or 0.2, respectively) . If she is behind in the given week, the probability that she will be up-to-date (or behind) in the next week is 0.6 (or 0.4, respectively) . We assume that these probabilities do not depend on whether she was up-to-date or behind in previous weeks, so the problem has the typical Markov chain character (the future depends on the past only through the present)

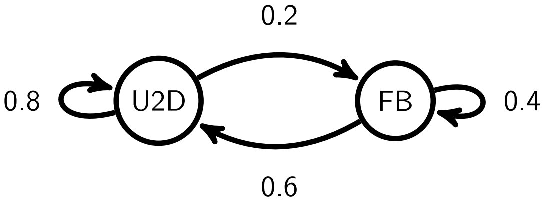
Let 1 be the state of being up-to-date and 2 that she fell behind.

The transition probabilities: $p_{11} = 0.8 \quad p_{12} = 0.2 \quad p_{21} = 0.6 \quad p_{22} = 0.4$

The transition probability matrix:

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$

The transition probability graph:



We have said that the state today depends only on the state in the previous period. This is true, however, we can get around this constraint.

Consider the following example:

1. A working machine can be working the next day with probability p, and be broken with probability $1 - p$.
2. A broken machine can be working the next day with probability q, and remain broken with probability $1 - q$.

However, what happens if the machine cannot be fixed for, say, 4 straight days? Maybe we need to buy a new one. To model this we can introduce new states to our system.

1. A working machine can be working the next day with probability p, and be 1-day broken with probability $1 - p$, and zero for n-days broken for $n > 1$.
2. A 1-day broken machine can be working the next day with probability q, and become 2-day broken with probability $1 - q$, and zero for n-days broken for $n \neq 2$.
3. A 2-days broken machine can be working the next day with probability r, and become broken for 3 days with probability $1 - r$, and zero for n-days broken for $n \neq 3$.
4. A 3-days broken machine can be working the next day with probability s, and become broken for 4 days with probability $1 - s$, and zero for n-days broken for $n \neq 4$.
5. A 4-days broken machine can be working with probability 1, and zero for all the other broken states.

Definition (n-Step Transition Probabilities)

Let $r_{ij}(n)$ represent the probability that the state after n time periods will be j, given that the current state is i.

$$r_{ij}(n) = P(X_n = j | X_0 = i)$$

Proposition (Chapman-Kolmogorov)

The n-step transition probabilities can be generated by the recursive formula:

$$r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)p_{kj}, \quad \textit{for } n > 1, \ \textit{and all } i, j$$

starting with

$$r_{ij}(1) = p_{ij}$$

Note that this is an element of the following matrix:

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}^n$$

This “realization” allow us to be able to ask and answer some interesting questions:

- What can we say about limits? What happens as $n \to \infty$?
- The dependence of the state at n over the initial state becomes smaller as n increases.
- What can we say qualitatively about the behavior of this markov chain?

Consider the transition matrix for the example we just saw:

$$A = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \quad A^2 = \begin{bmatrix} 0.7600 & 0.2400 \\ 0.7200 & 0.2800 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0.7520 & 0.2480 \\ 0.7440 & 0.2560 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 0.7504 & 0.2496 \\ 0.7488 & 0.2512 \end{bmatrix} \quad A^5 = \begin{bmatrix} 0.7501 & 0.2499 \\ 0.7498 & 0.2502 \end{bmatrix}$$

$$A^6 = \begin{bmatrix} 0.7500 & 0.2500 \\ 0.7500 & 0.2500 \end{bmatrix} \quad A^7 = \begin{bmatrix} 0.7500 & 0.2500 \\ 0.7500 & 0.2500 \end{bmatrix}$$

Note how as $n \to \infty$ $r_{ij}(n)$ goes to a limit that does not depend on the initial state.

Definition (Accessible state)

A state j is accessible from a state i if $\exists n \in \mathbb{N}$ such that the n-step transition probability $r_{ij}(n)$ is positive, i.e., if there is positive probability of reaching j, starting from i, after some number of periods.

Definition (Recurrent state)

Let $A(i)$ be the set of states that are accessible from i. We say that i is **recurrent** if $\forall j \in A(i) \Rightarrow i \in A(j)$.

Definition (Transient state)

A state is called **transient** if it is not recurrent.

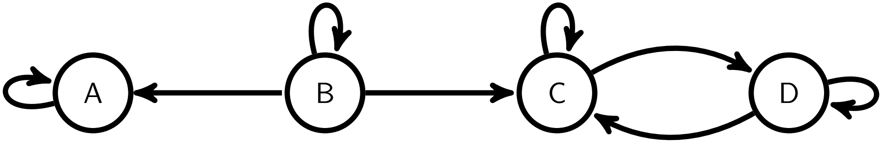
Corollary

A recurrent state will be visited an infinity amount of times.

Corollary

A transient state will be visited a finite amount of times.

Which of the following nodes are transient and which are recurrent?



Recurrent | Transient | Recurrent | Recurrent

Definition (Recurrent class)

If i is a recurrent state, the set of sattes $A(i)$ that are accessible from i form a **recurrent class** (or simply a class), meaning that states in $A(i)$ are all accessible from each other, and no state outside $A(i)$ is accessible from them.

Steady state behavior

When we talk about steady state in Markov Chains, it is not the "state" that is steady, but the probabilities of arriving to a certain state, remember the example we had before?

$$\pi_j = P(X_n = j), \quad \text{when } n \text{ is large.}$$

Theorem (Steady-State Convergence Theorem)

Consider a Markov chain with a single recurrent class, which is periodic. Then, the states j are associated with steady-state probabilities π_j that have the following properties:

1. *For each j, we have*
$$\lim_{n\to\infty} r_{ij}(n) = \pi_j, \quad \forall i$$
2. *The π_j are the unique solution to the system of equations below:*
$$\pi_j = \sum_{k=1}^{m} \pi_k p_{kj}, \quad j = 1, \ldots, m,$$
$$1 = \sum_{k=1}^{m} \pi_k$$
3. *We have*
$$\pi_j = 0 \quad \text{, for all transient states } j$$
$$\pi_j > 0 \quad \text{, for all recurrent states } j$$

Note that the steady-state probabilities add up to 1...
Therefore these form a probability distribution on the state space, this is called the **stationary distribution** of the chain.

Definition (Balance Equations)

The equations

$$\pi_j = \sum_{k=1}^{m} \pi_k p_{kj}, \quad j = 1, \ldots, m,$$

are called **balance equations**, and they are a direct consequence of the first part of the Steady-State Convergence Theorem, and the Chapman-Kolmogorov equation.

Definition (Normalization Equation)

The equation

$$sum_{k=1}^{m} \pi_k = 1$$

is known as the **normalization equation.**

Example

Consider our original example: $p_{11} = 0.8 \quad p_{12} = 0.2 \quad p_{21} = 0.6 \quad p_{22} = 0.4$
Clearly, on the limit $r_{ij} \to \pi_j$ if this converges, then the balance equations say:

$$\pi_1 = \pi_1 p_{11} + \pi_2 p_{21} \quad \pi_2 = \pi_1 p_{12} + \pi_2 p_{22}$$

Which, replacing, become:

$$\pi_1 = 0.8\pi_1 + 0.6\pi_2 \quad \pi_2 = 0.2\pi_1 + 0.4\pi_2$$

Solving, we obtain $\pi_1 = 3\pi_2$ in both equations, which together with the normalization equation $\pi_1 + \pi_2 = 1$ lead us to:

$$\pi_1 = 0.75$$
$$\pi_2 = 0.25$$