Notes on Tolver:

Necessities of Markov Chains:

1: A space of states: This is just a set of all the states a system can be in.

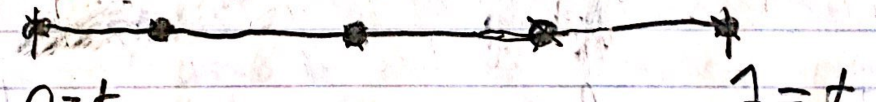
1

2: Probability transition functions: these tell us the probability of a transition from state i to state j. These probabilities are normalized so that, assuming discrete states, there is a 100% chance of a transition occuring; if there are 3 possible states for i to transition to, i, k, and j, then normalization means that $p(i|i) + p(j|i) + p(k|i) = 1$ and so on.

And thats it.

Today we'll be talking about discrete, time-homogenous Markov chains. What this means can be visualized as follows:

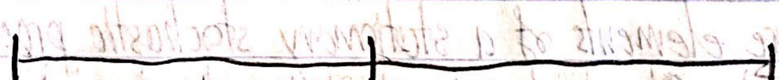
Discrete



Continuous

3

Homogenous

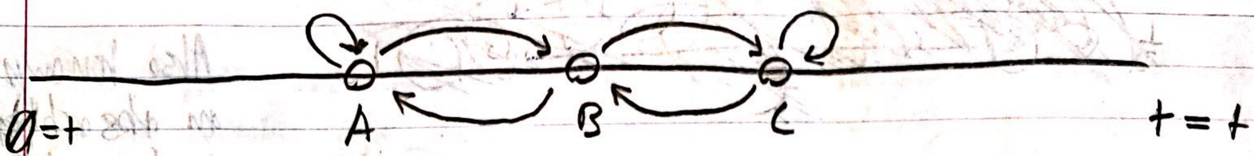


Inhomogenous

$$P(x_n \mid x_{n-1} \ldots x_0) = P(x_n \mid x_{n-1})$$

Something shared by all Markov chains is something called the Markov property, which enforces a kind of memorylessness on the chain where the

2

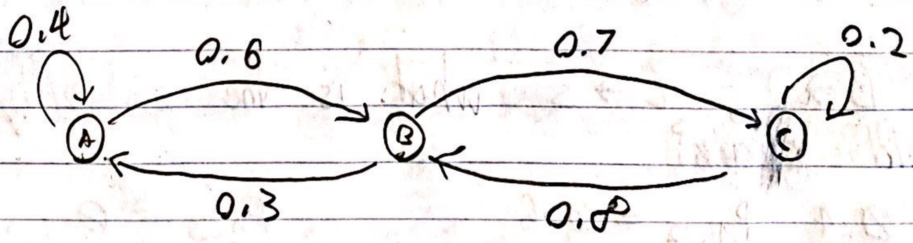


Where B goes is independent of how it got to B, but the behaviour of B is nonetheless different than that of A or C and is in some sense dependent.

4 Before anything else we better review conditional probabilities. The probability of an event B given an event A is written as:

$$P(B \mid A) = \text{Probability that B given A}$$

An example of this could be the above chain:



Probability that state A steps to B $= P(B \mid A) = 0.6$

Probability that state A steps to B and then to A again?

$\hookrightarrow P(B \mid A) \cdot P(A \mid B) = 0.6 \cdot 0.3 = 18/100$

Probability that A goes to B and that C goes to B?

$P(B \mid A) = 0.6$, $P(B \mid C) = 0.8$, $P(A \mid A) P(B \mid C) = 0.32$

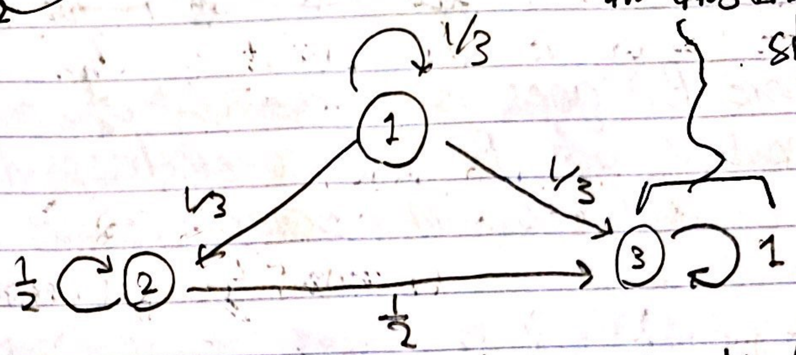
$P(B \mid A) \cdot P(B \mid C) = 0.48$, $P(B \mid A) P(C \mid C) = 0.12$, $P(A \mid A) P(C \mid C) = 0.08$

~~Rules of Combinations and Another Chain~~

~~5: How ... word problem~~

Also known as an absorbing state.

Transition Matrices:



An important part of this directed graph set up is that the diagram information that can be used to assemble transition matrices. The transition matrix of the above chain is given as

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix},$$

Rows sum to one it is a stochastic matrix

$P_{i,j}$ = probability of i jumping to j

If $A \to 1$, $B \to 2$, $C \to 3$ what is the probability matrix of our old chain?

$P_{1,1} = 0.4$ $\quad P_{1,2} = 0.6$ $\quad P_{1,3} = 0$

$P_{2,1} = 0.3$ $\quad P_{2,2} = 0$ $\quad P_{2,3} = 0.7$

$P_{3,1} = 0$ $\quad P_{3,2} = 0.8$ $\quad P_{3,3} = 0.2$

Scanned with CamScanner

$$P = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.3 & 0 & 0.7 \\ 0 & 0.8 & 0.2 \end{bmatrix} \rightarrow$$

What happens if we have an initial probability distribution? If the zeroth step has an initial probability of $\phi(1) = \phi(2) = \frac{1}{2}$, then what is the probability that first step is to 1 followed by 1 to 2 and 2 to 3?

$$P(1|0) \cdot P(2|1) \cdot P(3|2) = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2}$$

If the initial probability distribution of our chain is $\phi(1) = \phi(2) = \phi(3) = \frac{1}{3}$, then what is the likelihood of a $1 \to 2 \to 3$ set of jumps?

$$P(1|0)\,P(2|1)\,P(3|2) = 0.33 \cdot 0.6 \cdot 0.7 = 0.126$$

Theorem 2 in Tolver gives an interesting formula for an N-step transition probability: $\overline{\phi} P^n$ where n is the number of the sight we're intereeted in. This formula gives the likelihoods of the n^{th} site jumping to the others given an set of a certain set of initial distributions. In the given example we have in the text one of the questions is $P(X(2) = i)$ for $i = 1, 2, 3$ the distribution of probabilities is

$$\left[\frac{1}{18}, \ \frac{19}{72}, \ \frac{49}{72} \right]$$

What would $P(X(2) = i)$ be for our chain?

$$\left[\frac{1}{3} \ \ \frac{1}{3} \ \ \frac{1}{3} \right] \begin{bmatrix} \ & \ & \ \\ \ & \ & \ \\ \ & \ & \ \end{bmatrix} = ?$$

Another question we can ask is the likelihood of a certain time being the first time a certain state is visited. The given example is the likelihood of time = 2 being the first time 3 is visited. There are independent ways to reach this from the start in 3 steps and we add each together.

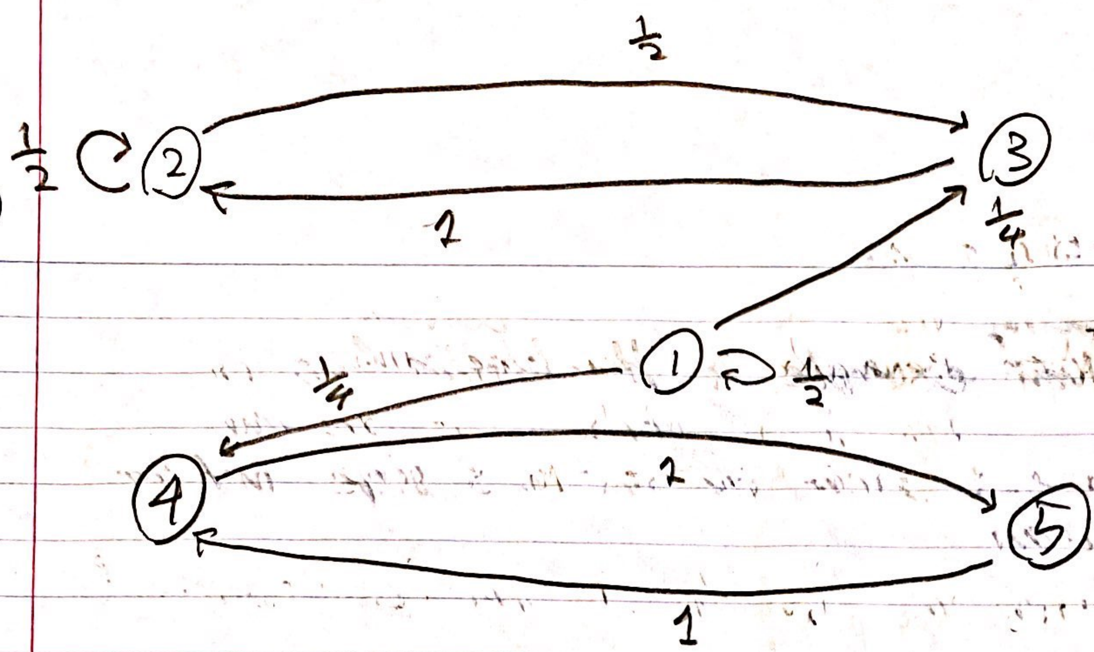
$$P = \phi(1) \cdot P_{1,1} \cdot P_{1,3} + \phi(1) \cdot P_{1,2} \cdot P_{2,3} + \phi(2) \cdot P_{2,2} \cdot P_{2,3} = \frac{19}{72}$$

For us an analogue of this is available if we change the initial probability distribution to $\phi(1) = \phi(2) = \frac{1}{2}$. Now what is $P(T_3 = 2)$?

The anatomy of markov chains:

Can any state access any other?
Are the chains regular?
Communication classes?

Communication classes are a way of cutting up a chain into parts that talk to one another. If we use the convention that a state automatically communicates with itself, then we can partition it into seperate classes. If a state has only one communication class then it is



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

$$\left.\begin{aligned} C_A &= \{2, 3\} \\ C_B &= \{1, 5\} \end{aligned}\right] \sim \text{Closed}$$

$C_C = \{1\} \rightarrow$ not closed since probability of exit is nonzero

A closed communication class is one where once you are in one of those states you just jump from one to another.

$$T_i = \{n > 0 \mid X(n) = i\}$$

Recurrence or hitting time:

Recurrent states are those whose probability of return given sufficiently large steps is equal to 1

$$P(T_i < \infty \mid \phi(0) = i) = 1$$

Transient states are those that may return in a sufficiently long amount of time

$$P(T_i < \infty \mid \phi(0) = i) < 1$$

According to thermodynamics pretty much all microstates are transient states.

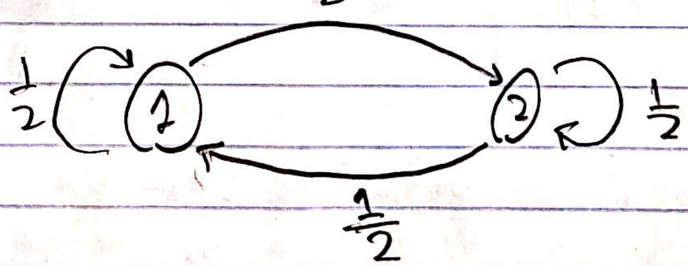
Lets look at our new chain and ask is state 1 transient or recurrent?

$\rightarrow$ Transient: $P(T_1 = \infty \mid \phi(0) = 1) \geq \frac{1}{2} > 0 < 1$

What about 4 and 5? Recurrent: what about 2 and 3?

Look at this chain:

What are 1 and 2?



2 and 3:

2 is Recurrent → why?

3 is a little more difficult. Intuitively, if it starts in 3 it'll jump to 2 and eventually back to 3.

An Introduction to Markov Chains and stationary distributions: We're going to talk about the idea of a stationary matrices and steady-state Markov chains: Lets start with some initial data and a transition matrix P: $\phi = [0.1 \quad 0.9]$, $P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$.

Lets run them together: $\phi P = \phi^1 = [0.62 \quad 0.38]$
Lets do it again: $\phi^1 P = \phi^2 = [0.724 \quad 0.276]$
Lets do it again: $\phi^3 = \phi^2 P = [0.7448 \quad 0.2552]$
and on and on... $\phi^6 = \phi^5 P = [0.749458 \quad 0.2500416]$

And on and on. Notice our state matrix elements seem to be approaching 0.75 and 0.25. What would happen if we run such a matrix?

$$\hat{\phi} = [0.75 \quad 0.25] \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.75 \quad 0.25],$$

Nothing happens, nothing changes and we call this the steady state distribution or the stationary distribu[tion] represented by a stationary matrix.

A markov chain is regular if its trans. matrix is regular. A transition matrix P is regular if some power of P has only positive entries.

Of the following, which are regular matrices?

$A = \begin{bmatrix} 0.2 & 0.8 \\ 0.3 & 0.7 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 0.2 & 0.8 \\ 1 & 0 \end{bmatrix}$

A by inspection $\longrightarrow$ have them investigate the first powers of the others.

Properties of regular Markov chains: for a regular chain's t-matrix P

A) There is a unique stationary matrix S can be found by solving $S \cdot P = S$

Example: $P = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$. Recall $s_1 + s_2 = 1$

$SP = S = [S_1 \quad S_2]$, $S_1 = 0.6s_1 + 0.2S_2 \rightarrow 0.4s_1 = 0.2S_2$

$S_2 = 0.4s_1 + 0.8S_2 \rightarrow$

This is a system of equations and we can solve it for S_1 an[d] S_2. We get $S_2 = 2/3$ and $S_1 = 1/3$

Other 2 properties of regular chains:

B) given an initial-state matrix S_0, this will approach the stationary matrix S.

C) The powers of our transition matrix P, P^n, will approach a matrix $\bar{P}$ if we keep taking higher and higher powers of P. The rows of $\bar{P}$ will be equal to our stationary matrix.

And thats it.

Q is our own Markov chain regular?

At first glance, no.

A related kind of Markov chain is those with absorbing states and non-absorbing states which will in a finite number of states that will jump to the absorbing-stat in a finite number of steps. If this second condition is met, then the chain's transitional matrix will have a limiting matrix and something like a stationary dist.