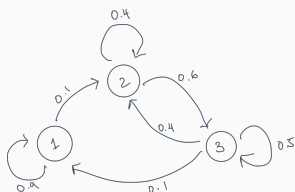


## Markov Chains

CS70 Summer 2016 - Lecture 6B/C

David Dinh  
25 July 2016  
UC Berkeley

### Intuition



A finite Markov chain consists of *states*, *transition probabilities* between states, and an *initial distribution*.

State: where are you now?

Transition probability: From where you are, where do you go next?

Initial distribution: how do you start?

Markov chains are **memoryless** - they don't remember anything other than what state they are.

3

### Agenda

What are Markov Chains? State machine and matrix representations.

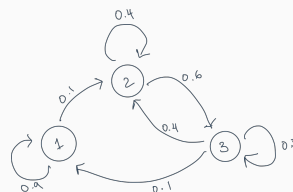
Hitting Time  
Convergence, limiting and stationary distributions



A. A. Markov (1856).

1

### Formally Speaking...



A finite set of states:  $\mathcal{X} = \{1, 2, \dots, K\}$

An initial probability distribution  $\pi_0$  on  $\mathcal{X}$ :  $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$

Transition probabilities:  $P(i, j)$  for  $i, j \in \mathcal{X}$

- $P(i, j) \geq 0, \forall i, j; \sum_j P(i, j) = 1, \forall i$

$\{X_n, n \geq 0\}$  is defined so that:

- $\Pr[X_0 = i] = \pi_0(i), i \in \mathcal{X}$  (initial distribution)
- $\Pr[X_{n+1} = j | X_0, \dots, X_n = i] = P(i, j), i, j \in \mathcal{X}$ .

4

### Motivation

Suppose we flip a coin until we get a three heads in a row. How many coin flips should we expect to do?

Drunkard on an arbitrary graph (remember HW?). When does the drunkard come home?

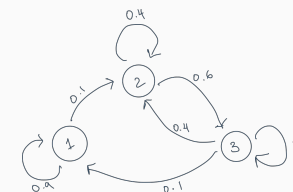
Try solving directly? Problem: conditioning gets really messy.

Need some way to express **state**.

Solution: Markov chains!

2

### One Small (Time)step for a State



At each timestep  $t$  we are in some **state**  $X_t \in \mathcal{X}$ . (random variable.)

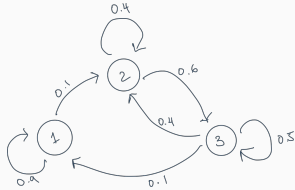
Where do we go next?

$$\Pr[X_{t+1} = j | X_t = i] = P_{i,j}$$

Probability depends on the previous state, but is independent of how it got to the previous state. (It's not independent of states before the previous state - but any dependence is captured in the previous state.)

5

## One Giant Leap with Conditional Probability



At some point we might have a distribution for  $X_t$  - say, it's 1 w.p. 0.2, 2 w.p. 0.3, and 3 w.p. 0.5. Distribution for  $X_{t+1}$ ? Probability that it goes to 1?

$$\begin{aligned}\Pr[X_{t+1} = 1] &= \sum_i \Pr[X_{t+1} = 1 | X_t = i] \Pr[X_t = i] = \sum_i P_{i,1} \Pr[X_t = i] \\ &= 0.9 * 0.2 + 0.1 * 0.3 + 0.1 * 0.5 = 0.23\end{aligned}$$

Rest of distribution for  $X_{t+1}$  can be found similarly.

6

## Linear Algebra Intro

Very quick linear algebra intro:

Matrices: two-dimensional collection of numbers.  $n \times m$  matrix has  $n$  rows,  $m$  columns. Element at  $i$ th row,  $j$ th column denoted  $A_{ij}$ .

$$\begin{bmatrix} 1 & 6 & 7 & 2 \\ 6 & 5 & 6 & 3 \\ 8 & 6 & 2 & 2 \\ 2 & 5 & 3 & 8 \end{bmatrix}$$

Vector: one-dimensional collection of numbers. We deal with **row vectors** -  $n \times 1$  matrices.

$$\begin{bmatrix} 5 & 9 & 3 & 0 \end{bmatrix}$$

7

## Matrix Multiplication

For  $n \times m$  matrix  $A$  and  $m \times p$  matrix  $B$ :

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}$$

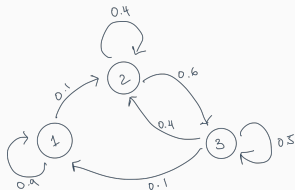
Or for vector  $x$ :

$$(xA)_i = \sum_k x_k A_{ki}$$

$$\begin{bmatrix} 5 & 9 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 6 & 7 & 2 \\ 6 & 5 & 6 & 3 \\ 8 & 6 & 2 & 2 \\ 2 & 5 & 3 & 8 \end{bmatrix} = \begin{bmatrix} 1*5 + 6*9 + 8*3 + 2*0 \\ 6*5 + 5*9 + 6*3 + 5*0 \\ 7*5 + 6*9 + 2*3 + 3*0 \\ 2*5 + 3*9 + 2*3 + 8*0 \end{bmatrix}^T$$

8

## Matrix Markov



Markov chains have a very nice translation to matrices! Transition probabilities form an *transition matrix*  $P$  whose  $i, j$ th entry is  $P_{i,j}$ .

$$P = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0 & 0.4 & 0.6 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$$

Probabilities from a state sum to 1...rows sum to 1... (right) **stochastic matrix**.

9

## Stepping with Multiplication

$$P = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0 & 0.4 & 0.6 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$$

Distributions are vectors. Suppose that  $X_t$  is distributed 1 w.p. 0.2, 2 w.p. 0.3, and 3 w.p. 0.5. Write distribution as vector!

$$\pi_t = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$$

What's the product of  $\pi_t$  and  $P$ ?

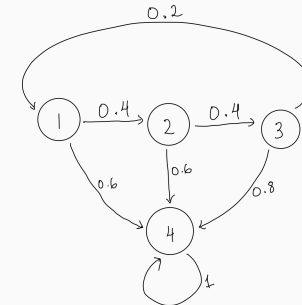
$$\begin{bmatrix} 0.2 * 0.9 + 0.3 * 0 + 0.5 * 0.1 \\ 0.2 * 0.1 + 0.3 * 0.4 + 0.5 * 0.4 \\ 0.2 * 0 + 0.3 * 0.6 + 0.5 * 0.5 \end{bmatrix}^T = \begin{bmatrix} 0.23 & 0.34 & 0.43 \end{bmatrix}$$

This is the distribution of  $X_{t+1}$ .

10

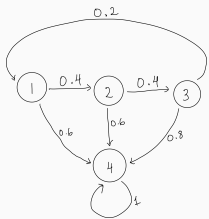
## An Example

California driving test: you get 3 retakes before you have to start the application process all over again. Suppose someone passes a driving test w.p. 0.6, unless it's their final retake, in which case they're more careful and pass w.p. 0.8.



11

## An Example



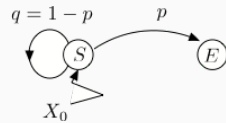
Initial distribution?  $\pi_0 = [1 \ 0 \ 0 \ 0]$   
Transition matrix?

$$T = \begin{bmatrix} 0 & .4 & 0 & .6 \\ 0 & 0 & .4 & .6 \\ .2 & 0 & 0 & .8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

12

## A Simple Example

Let's flip a coin with  $Pr[H] = p$  until we get  $H$ . How many flips, on average?



Let  $\beta(S)$  be the average time until  $E$ , starting from  $S$ . Then,

$$\beta(S) = 1 + q\beta(S) + p \cdot 0.$$

Hence,

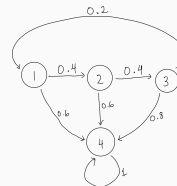
$$p\beta(S) = 1, \text{ so that } \beta(S) = 1/p.$$

Note: Time until  $E$  is  $G(p)$ . We have rediscovered that the mean of  $G(p)$  is  $1/p$ .

14

## Hitting Time

## How Long to Get a Driver's License?



Let  $\beta(S)$  denote expected time to get a driver's license from  $S$ .

$$\beta(1) = 0.6 * 0 + 0.4 * (1 + \beta(2))$$

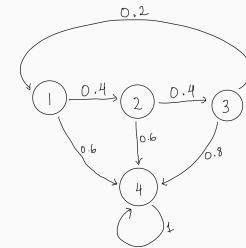
$$\beta(2) = 0.6 * 0 + 0.4 * (1 + \beta(3))$$

$$\beta(3) = 0.8 * 0 + 0.2 * (1 + \beta(1))$$

Solves to  $\beta(1) \approx 0.61$ . Adding the first driving test: 1.61 driving tests.

15

## Motivation



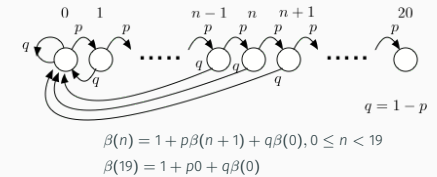
How long does it take to get a driver's license, in expectation?

Generally: given a Markov chain and an initial distribution, how many timesteps do we expect to take before reaching a particular state?

13

## Drunkard's Climb

Our old friend the drunkard to go up a ladder that has 20 rungs. At each time step, he succeeds in going up by one rung with probability  $p = 0.9$ . Otherwise, he falls back to the ground (miraculously unharmed). How many time steps does it take for him to reach the top of the ladder, on average?



$$\beta(n) = 1 + p\beta(n+1) + q\beta(0), 0 \leq n < 19$$

$$\beta(19) = 1 + p \cdot 0 + q\beta(0)$$

$$\Rightarrow \beta(0) = \frac{p^{-20} - 1}{1 - p} \approx 72.$$

See Lecture Note 24 for algebra.

16

Questions?