



CS 68: BIOINFORMATICS

Prof. Sara Mathieson
Swarthmore College
Spring 2018



Outline: Mar 30

- Markov chains
- Hidden Markov Models (HMMs)

Notes:

- CR/NC/W deadline today
- I am around this afternoon
- Faculty candidate talk today right after class

Markov Chains

Conditional probability

- Idea of conditional probability

$$P(A, B) = P(A)P(B|A)$$

Conditional probability

- Idea of conditional probability

$$P(A, B) = P(A)P(B|A)$$

- Bayes' Theorem

$$P(A)P(B|A) = P(B)P(A|B)$$

conditional prob example

u = bring umbrella
 r = rain
 s = sun

only 2 options.

$W \in \{r, s\}$
 \uparrow
random variable

values

given

$$P(u|r) = 0.9$$

$$P(u|s) = 0.2$$

$$\left. \begin{array}{l} P(r) = 0.4 \\ P(s) = 0.6 \end{array} \right\} \text{sum to } 1$$

Q: what is the prob of umbrella?

$$P(u) = \sum_{w \in \{r, s\}} P(u, w) = P(u, r) + P(u, s) \quad \text{cond. prob.}$$
$$= P(r)P(u|r) + P(s)P(u|s)$$

marginalizing over weather

$$= (0.4)(0.9) + (0.6)(0.2)$$

$$= 0.36 + 0.12 = \boxed{0.48}$$

Markov Chains

- Markov assumption: current state only depends on the previous state

$$P(z_i | z_0, z_1, \dots, z_{i-1}) = P(z_i | z_{i-1})$$

Markov Chains

- Markov assumption: current state only depends on the previous state

$$P(z_i | z_0, z_1, \dots, z_{i-1}) = P(z_i | z_{i-1})$$

- This allows us to simplify the probability of observing a Markov chain:

$$P(z_0, z_1, \dots, z_L) = P(z_0) \prod_{i=1}^L P(z_i | z_{i-1})$$

Markov Chains

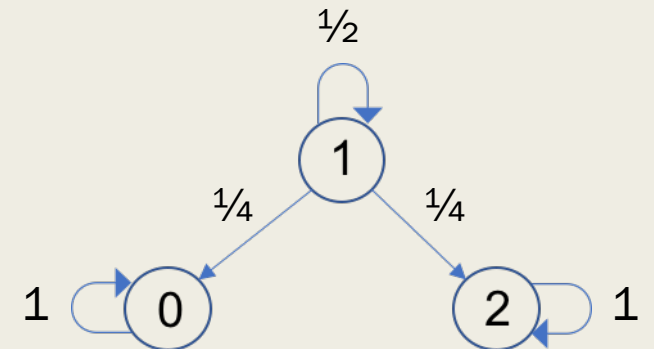
- Markov assumption: current state only depends on the previous state

$$P(z_i | z_0, z_1, \dots, z_{i-1}) = P(z_i | z_{i-1})$$

- This allows us to simplify the probability of observing a Markov chain:

$$P(z_0, z_1, \dots, z_L) = P(z_0) \prod_{i=1}^L P(z_i | z_{i-1})$$

- Note the difference between the state diagram (right) and an observed state sequence



Markov Chains

- Markov assumption: current state only depends on the previous state

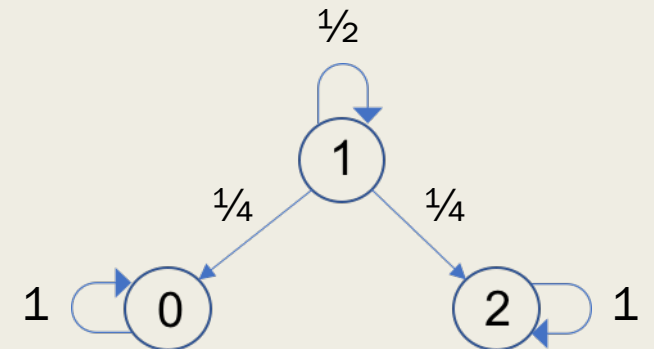
$$P(z_i | z_0, z_1, \dots, z_{i-1}) = P(z_i | z_{i-1})$$

- This allows us to simplify the probability of observing a Markov chain:

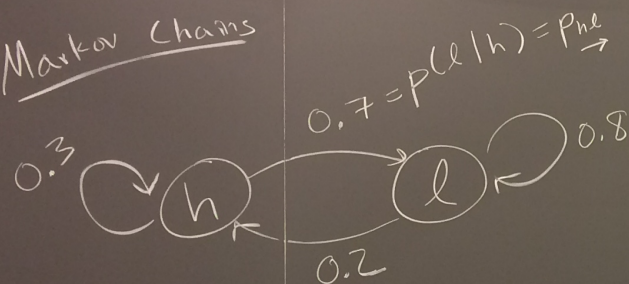
$$P(z_0, z_1, \dots, z_L) = P(z_0) \prod_{i=1}^L P(z_i | z_{i-1})$$

- Note the difference between the state diagram (right) and an observed state sequence

Note that the sum of
outgoing probabilities
should be 1



Markov chains



at gene

Q: what % of nights on avg are spent in low sleep? (π_l)

$$[\pi_h = 1 - \pi_l]$$

stationary distribution

- how much time do we spend in each state on average?

$$\pi_j = \sum_{i \in X} \pi_i P_{ij}$$

↑
state space

Solve for π_l

$$\pi_l = \pi_l P_{ll} + \pi_h P_{hl}$$

$$\pi_l = \pi_l 0.8 + (1 - \pi_l) 0.7$$

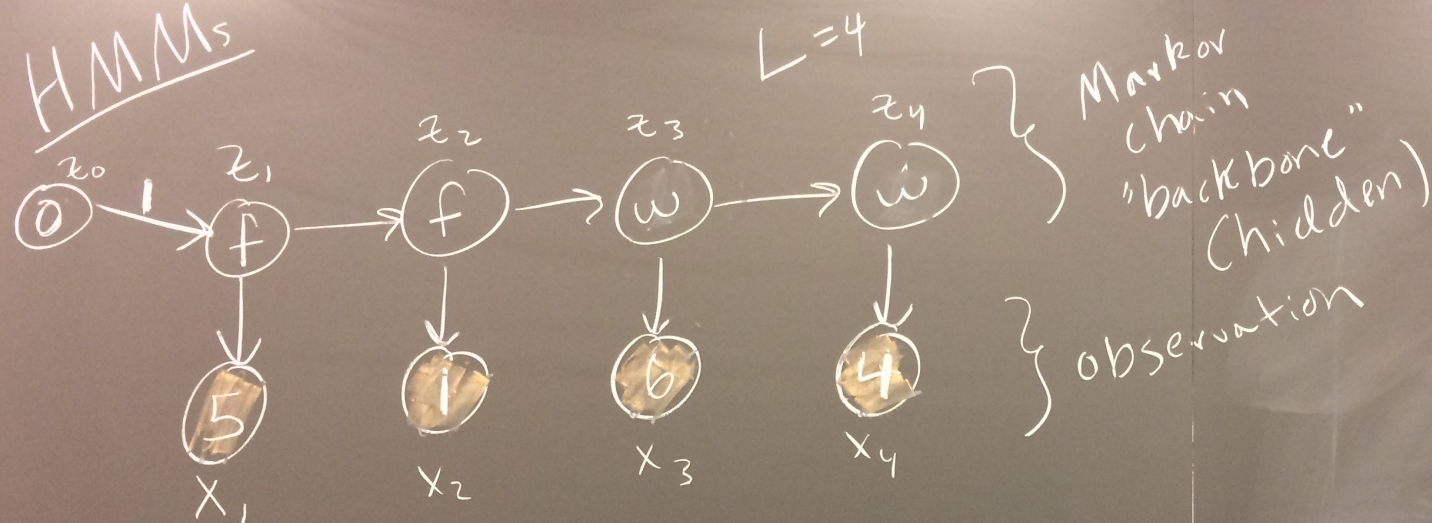
$$-0.7 = (-1 + 0.8 - 0.7)\pi_l$$

$$\pi_l = \frac{0.7}{0.9} = \boxed{\frac{7}{9}}$$

$$\boxed{\pi_h = \frac{2}{9}}$$

Hidden Markov Models

HMMs



Joint Prob of \vec{x}

$$P(\vec{x}, \vec{z})$$

transition probabilities

$$a_{kl} = P(z_i = l | z_{i-1} = k), \quad a_{fw} = 0.05$$

emission probabilities

$$e_k(b) = P(\overset{\text{obs}}{x_i = b} | \overset{\text{hidden}}{z_i = k})$$

Joint
prob of \vec{x} & \vec{z}

$$P(\vec{x}, \vec{z}) = \prod_{i=1}^L a_{z_{i-1} z_i} \cdot e_{z_i}(x_i)$$

product

example

$$a_{fw} = 0.05$$

$$P(\vec{x}, \vec{z}) = \underbrace{(a_{of} \cdot e_f(5))}_{i=1} \underbrace{(a_{ff} \cdot e_f(1))}_{i=2} \underbrace{(a_{fw} \cdot e_w(6))}_{i=3} \underbrace{(a_{ww} \cdot e_w(4))}_{i=4} = (1 \cdot \frac{1}{6}) (0.95 \cdot \frac{1}{6}) (0.05)(\frac{1}{2})(0.9 \cdot \frac{1}{10})$$

Occasionally
Dishonest
Casino

