Hidden Markov Models

CISC 5800
Professor Daniel Leeds

Representing sequence data

- Spoken language
- DNA sequences
- Daily stock values

Example: spoken language

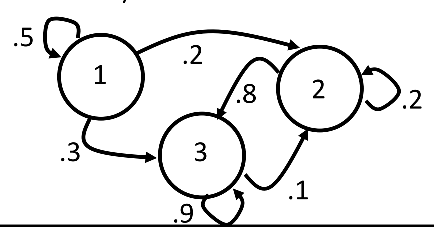
F?r plu? fi?e is nine

- Between F and r expect a vowel: “aw”, “ee”, “ah”; NOT “oh”, “uh”
- At end of “plu” expect consonant: “g”, “m”, “s”; NOT “d”, “p”

Markov Models

Start with:

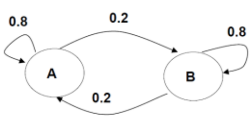
- *n* states: $s_1, \dots, s_n$
- Probability of initial start states: $\Pi_1, \dots, \Pi_n$
- Probability of transition between states: $A_{i,j} = P(q_t=s_i \mid q_{t-1}=s_j)$



A dice-y example

$$\Pi_A = 0.3, \Pi_B = 0.7$$

- Two colored die



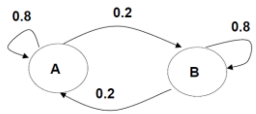
- What is the probability we start at s_A?
- What is the probability we have the sequence of die choices:

  s_A, s_A?

- What is the probability we have the sequence of die choices:

  s_B, s_A, s_B, s_A?

A dice-y example



- What is the probability we have the die choices s_B at time t=5

$$\Pi_A = 0.3, \Pi_B = 0.7$$

- Dynamic programming: find answer for q_t , then compute q_{t+1}

State\Time	t_1	t_2	t_3
s_A	0.3		
s_B	0.7		

$$p_t(i) = \sum_j p(q_t = s_i | q_{t-1} = s_j) p_{t-1}(j)$$

$p_t(i) = P(q_t{=}s_i)$ -- Probability state i at time t

6

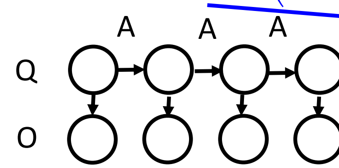
Hidden Markov Models

- Actual state q “hidden”
- State produces visible data o: $\phi_{i,j} = P(o_t = x_i | q_t = s_j)$

Probability observe value x_i when state is s_j

- Compute

$$P(\boldsymbol{O}, \boldsymbol{Q} | \boldsymbol{\theta}) = p(q_1|\pi) \left(\prod_{t=2}^{T} p(q_t | q_{t-1}, \boldsymbol{A}) \right) \left(\prod_{t=1}^{T} p(o_t | q_t, \boldsymbol{\phi}) \right)$$



ϕ

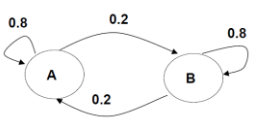
Probability of state sequence

Probability of observation sequence, given states

8

Deducing die based on observed “emissions”

Each color is biased



o	$P(o \mid s_A)$	$P(o \mid s_B)$
1	.3	.1
2	.2	.1
3	.2	.1
4	.1	.2
5	.1	.2
6	.1	.3

Intuition – balance transition and emission probabilities

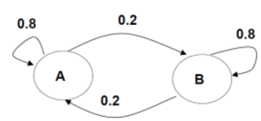
Observed numbers: 554565**2**54556 – the 2 is probably from s_B

Observed numbers: 554565**2**13321 – the 2 is probably from s_A

9

Deducing die based on observed “emissions”

Each color is biased



o	$P(o \mid s_R)$	$P(o \mid s_B)$
1	.3	.1
2	.2	.1
3	.2	.1
4	.1	.2
5	.1	.2
6	.1	.3

- We see: 5 What is probability of o=5 , q=B (blue)
- We see: 5, 3 What is probability of **o**=5,3 | **q**=B, B?

10

Goal: calculate most likely states given observable data

$$\boxed{\arg\max_Q P(Q \mid O)} = \arg\max_Q \frac{P(O \mid Q)P(Q)}{P(O)}$$

$$= \boxed{\arg\max_Q P(O \mid Q)P(Q)}$$

Define and use $\delta_t(i)$

$$\delta_t(i) = \max_{q_1 \dots q_{t-1}} p(q_1 \dots q_{t-1} \wedge q_t = s_i \wedge O_1 \dots O_t)$$

$\delta_t(i)$: max possible value of $P(q_1,..,q_t,o_1,..,o_t)$ given we insist $q_t=s_i$

Find the most likely path from q_1 to q_t that

- $q_t=s_i$
- Outputs are $o_1, \dots, o_t$

Viterbi algorithm: $\delta_t(i)$

$\delta_1(i) = \Pi_i P(o_1 | q_1 = s_i) = \Pi_i \phi_{o_1,i}$

$\delta_t(i) = P(o_t | q_t = s_i) \max_j \delta_{t-1}(j) P(q_t = s_i | q_{t-1} = s_j) =$
$\boldsymbol{\phi_{o_t,i} \max_j \delta_{t-1}(j) A_{i,j}}$

$P(Q^* | O) = \text{argmax}_Q P(Q|O) = \text{argmax}_i \delta_t(i)$

Viterbi algorithm: bigger picture

Compute all $\delta_t(i)$'s

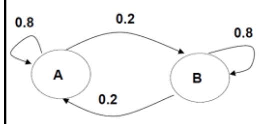
- At time t=1 compute $\delta_1(i)$ for every state i
- At time t=2 compute $\delta_2(i)$ for every state i (based on $\delta_1(i)$ values)
- …
- At time t=T compute $\delta_T(i)$ for every state i (based on $\delta_{T-1}(i)$ values)

Find states going from t=T back to t=1 to lead to max $\delta_T(i)$

- Now find state j that gives maximum value for $\delta_T(j)$
- Find state k at time T-1 used to maximize $\delta_T(j)$
- …
- Find state z at time 1 used to maximize $\delta_2(y)$

Viterbi in action: observe "5, 1"

$\Pi_A = 0.3, \Pi_B = 0.7$



o	$P(o\|s_A)$	$P(o\|s_B)$
1	.3	.1
2	.2	.1
3	.2	.1
4	.1	.2
5	.1	.2
6	.1	.3

$\delta_2(A)$:

$\delta_2(B)$:

	t=1 (o_1=5)	t=2 (o_2=1)
$q_t=s_A$	.3x.1 = .03	
$q_t=s_B$	.7x.2 = .14	

Parameters in HMM

Initial probabilities: π_i

Transition probabilities $A_{i,j}$

Emission probabilities $\phi_{i,j}$

How do we learn these values?

20