Section 8

Sunday, January 13, 2019 12:57 PM

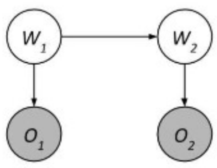
section8

CS188 Fall 2018 Section 8: HMMs + Particle Filtering

1 HMMs

Consider the following Hidden Markov Model.

temporal, unobservable state variables
expect to receive: transition model, sensor model, occasional observation updates



W_1	$P(W_1)$
0	0.3
1	0.7

W_t	W_{t+1}	$P(W_{t+1}\|W_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

W_t	O_t	$P(O_t\|W_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

Suppose that we observe $O_1 = A$ and $O_2 = B$.
Using the forward algorithm, compute the probability distribution $P(W_2|O_1 = A, O_2 = B)$ one step at a time.

1. Compute $P(W_1, O_1 = A)$.

What's the partial JPD of getting both the evidence and each posisble value for state W1?
(Product Rule w/ sensor model)
$P(W_1 = 0, O_1 = A) = P(O_1 = A|W_1 = 0) * P(W_1 = 0) = 0.9*0.3 = 0.27$
$P(W_1 = 1, O_1 = A) = P(O_1 = A|W_1 = 1) * P(W_1 = 1) = 0.5*0.7 = 0.35$

2. Using the previous calculation, compute $P(W_2, O_1 = A)$.

$P(W_2, o_1) = \sum_{x1} P(W2, x1, o1) = \sum_{x1} P(W2, |x1, o1)\, P(x_1, o_1) = \sum_{x1} P(W2, |x1)\, P(x_1, o_1)$
introduce a variable and sum it out, we can refactor into components that we know. Markov property simplifies
$P(W_2 = 0, o_1 = A) = 0.27*0.4 + 0.35*0.8 = 0.388$
$P(W_2 = 1, o_1 = A) = 0.27*0.6+0.35*0.2 = 0.232$

3. Using the previous calculation, compute $P(W_2, O_1 = A, O_2 = B)$.

$P(W_2, o1, o2) = P(o2 \mid W2, o1) * P(W2, o1) = P(o2 \mid W2) * P(W2, o1)$
$P(W_2 = 0, o_1 = A, o_2 = B) = 0.1*0.388 = 0.0388$
$P(W_2 = 1, o_1 = A, o_2 = B) = 0.5*0.232=0.116$

pr
th

just a

4. Finally, compute $P(W_2|O_1 = A, O_2 = B)$.

get actual belief state at W_2, what the FUCK.
$P(W_2 \mid O_1{=}A, O_2{=}B) = P(W_2, O_1{=}A, O_2{=}B) / P(O_1{=}A, O_2{=}B)$

$P(W_2 = 0 \mid O_1{=}A, O_2{=}B) = 0.0388/0.156 = 0.2487$
$P(W_2 = 1 \mid O_1{=}A, O_2{=}B) = 0.116/0.156 = 0.7436$

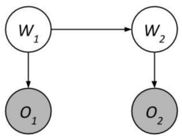
remember, JPDs are the BOMB
$P(O_1 = A, O_2 = B) = \sum_{w_2} P(W2, O1 = A, O2 = B) = 0.0388 + 0.116 = 0.156$

rod rule against o_2, then use the CI rules of sensor model to simplify
e other term we pooped out was the previous belief state JPD

nug of a YUGE jpd

2 Particle Filtering

Let's use Particle Filtering to estimate the distribution of $P(W_2|O_1 = A, O_2 = B)$. Here's the HMM again:



W_1	$P(W_1)$
0	0.3
1	0.7

W_t	W_{t+1}	$P(W_{t+1}\|W_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

W_t	O_t	$P(O_t\|W_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

We start with two particles representing our distribution for W_1.
$P_1 : W_1 = 0$
$P_2 : W_1 = 1$
Use the following random numbers to run particle filtering:

[~~0.22~~, ~~0.05~~, ~~0.33~~, ~~0.20~~, 0.84, 0.54, 0.79, 0.66, 0.14, 0.96]

1. **Observe**: Compute the weight of the two particles after evidence $O_1 = A$.

   do LE weighting based on observed evidence
   $W_1 = P(A \mid 0) = 0.9 \Rightarrow 0.9/1.4 = 0.6429$
   $W_2 = P(A \mid 1) = 0.5 \Rightarrow 0.5/1.4 = 0.3571$

2. **Resample**: Using the random numbers, resample P_1 and P_2 based on the weights.

   For re-sampling, we will re-normalize
   P1: $W_1 = 0$
   P2: $W_1 = 0$

3. **Elapse Time**: Now let's compute the elapse time particle update. Sample P_1 and P_2 from applying the time update.

   P1: $W_2 = 0$
   P2: $W_2 = 0$

4. **Observe**: Compute the weight of the two particles after evidence $O_2 = B$.

   P1: $W_2 = 0$ with weight $P(B|0) = 0.1$
   P2: $W_2 = 0$ with weight $P(B|0) = 0.1$

5. **Resample**: Using the random numbers, resample P_1 and P_2 based on the weights.

   P1: $W_2 = 0$
   P2: $W_2 = 0$

6. What is our estimated distribution for $P(W_2|O_1 = A, O_2 = B)$?

   100% chance of $W_2 = 0$ baby!

   YEEE HAAWWW

3 HMMs (Optional)

Consider a process where there are transitions among a finite set of states $s_1, \cdots, s_k$ over time steps $i = 1, \cdots, N$. Let the random variables $X_1, \cdots, X_N$ represent the state of the system at each time step and be generated as follows:

k possible states we can be in

N timesteps we are interested in

- Sample the initial state s from an initial distribution $P_1(X_1)$, and set $i = 1$
- Repeat the following:
  1. Sample a duration d from a duration distribution P_D over the integers $\{1, \cdots, M\}$, where M is the maximum duration.
  2. Remain in the current state s for the next d time steps, i.e., set

$$x_i = x_{i+1} = \cdots = x_{i+d-1} = s \tag{1}$$

  3. Sample a successor state s' from a transition distribution $P_T(X_t|X_{t-1} = s)$ over the other states $s' \neq s$ (so there are no self transitions)
  4. Assign $i = i + d$ and $s = s'$.

This process continues indefinitely, but we only observe the first N time steps.

(a) Assuming that all three states s_1, s_2, s_3 are different, what is the probability of the sample sequence $s_1, s_1, s_2, s_2, s_2, s_3, s_3$? Write an algebraic expression. Assume $M \geq 3$.

At each time step i we observe a noisy version of the state X_i that we denote Y_i and is produced via a conditional distribution $P_E(Y_i|X_i)$.

(b) Only in this subquestion assume that $N > M$. Let $X_1, \cdots, X_N$ and $Y_1, \cdots, Y_N$ random variables defined as above. What is the maximum index $i \leq N - 1$ so that $X_1 \perp\!\!\!\perp X_N | X_i, X_{i+1}, \cdots, X_{N-1}$ is guaranteed?

(c) Only in this subquestion, assume the max duration $M = 2$, and P_D uniform over $\{1, 2\}$ and each x_i is in an alphabet $\{a, b\}$. For $(X_1, X_2, X_3, X_4, X_5, Y_1, Y_2, Y_3, Y_4, Y_5)$ draw a Bayes Net over these 10 random variables with the property that removing any of the edges would yield a Bayes net inconsistent with the given distribution.

(d) In this part we will explore how to write the described process as an HMM with an extended state space. Write the states $z = (s, t)$ where s is a state of the original system and t represents the time elapsed in that state. For example, the state sequence $s_1, s_1, s_1, s_2, s_3, s_3$ would be represented as $(s_1, 1), (s_1, 2), (s_1, 3), (s_2, 1), (s_3, 1), (s_3, 2)$.

Answer all of the following in terms of the parameters $P_1(X_1), P_D(d), P_T(X_{j+1}|X_j), P_E(Y_i|X_i), k$ (total number of possible states), N and M (max duration).

- What is $P(Z_1)$?

  $P(x_1, t_1) =$

- What is $P(Z_{i+1}|Z_i)$? Hint: You will need to break this into cases where the transition function will behave differently.

  $P(X_{i+1}, t_{i+1} \mid X_i, t_i) =$

- What is $P(Y_i|Z_i)$?

  $P(Y_i \mid X_i, t_i) =$