CS 486 — Lecture 14: Hidden Markov Models

1 Umbrella Model

- Suppose you are in some underground facility and you want to know if it is raining, but you can only tell if it is raining when someone comes in with or without an umbrella.
- Note that we assume this is noisy data — maybe they come in with an umbrella but it isn't raining, or VV.
- We could model this world with a series of time slices; each time slice contains a set RVs, some observable and some not observable.
- X_t represents an un-observable variable; in this case it is representing whether it rains ($R_1, R_2, \dots$).
- E_t represents an observable variable; in this case it is $U_1, U_2, \dots$ representing whether the director is carrying an umbrella.
- In general, every state may depend on all previous states:

$$P(X_t | X_{t-1} \wedge X_{t-2} \wedge \cdots \wedge X_1)$$

- A problem with this is that as t increases, the conditional probability distribution gets too large.
- Our solution is to make the current state depend on a fixed number of states.

1.1 Stationary Model

- A first-order Markov process is defined such that each state depends on the previous the one previous state.
- So $P(X_t | dots) = P(X_t | X_{t-1})$.
- A second-order Markov process allows each state to depend on the two previous states.
- So now, $P(X_t | \dots) = P(X_t | X_{t-1} \wedge X_{t-2})$.
- The Markov assumption states that the future is independent of the past given the present.
- The stationary process model assumes two things:
  1. The dynamics does not change over time.
  2. The conditional probability distribution for each time step remains the same.
- By modelling as a stationary process, we simplify the world. Some advantages of doing this is that it's simple and sometimes, natural choice (usually dynamics don't change IRL).
- If the dynamics do change, often it means there is another variable that once incorporated, will no longer change the dynamics!
- Only need a finite number of parameters for an infinite model.

1.2 Sensor Model

- How does the evidence variable, E_t, at time t, depend on the previous and current states?
- The sensor Markov assumption: each state is sufficient to generate its observations.
- That is:

$$P(E_t | X_t \wedge \cdots \wedge X_1 \wedge E_{t-1} \wedge \cdots \wedge E_1) = P(E_t | X_t)$$

2 Inference Tasks

- Common inference tasks:
  - Filtering — which state am I in right now?
  - Prediction — which state will I be in tomorrow?
  - Smoothing — which state was I in yesterday?
  - Most likely explanation — which sequence of states is most likely to have generated the observations?
- Note that a HMM is a Bayesian network, so variable elimination algorithm can be used!

2.1 Filtering

- Given the observations up to today, which state am I in today?
- For example, let's say we observed:
  - Day 1: $P(R_1|u_1)$
  - Day 2: $P(R_2|u_{1:2})$
  - $\ldots$
  - Day n: $P(R_2|u_{1:n})$

  where $u_{1:n} = u_1 \wedge u_2 \wedge \cdots \wedge u_n$.