Reasoning Under Uncertainty Over Time

Alice Gao

Lecture 16

Readings: R & N 15.1 to 15.3

Based on work by K. Leyton-Brown, K. Larson, and P. van Beek

Outline

Learning Goals

By the end of the lecture, you should be able to

- Construct a hidden Markov model given a real-world scenario.
- Perform filtering, prediction, smoothing and derive the most likely explanation given a hidden Markov model.

Inference in a Changing World

So far, we can reason probabilistically in a static world.
However, the world evolves over time.

Applications:

- weather predictions
- stock market predictions
- patient monitoring
- robot localization
- speech and handwriting recognition

The Umbrella World

You are the security guard stationed at a secret underground installation. You want to know whether it's raining today, but your only access to the outside world occurs each morning when you see the director coming in with, or without, an umbrella.

States and Observations

- The world contains a series of time slices.
- Each time slice contains a set of random variables, some observable, some not.
  $\mathbf{X}_t$ the un-observable variables at time t
  $\mathbf{E}_t$ the observable variables at time t

What are the observable and unobservable random variables in the umbrella world?

The transition model

How does the current state depend on the previous states?

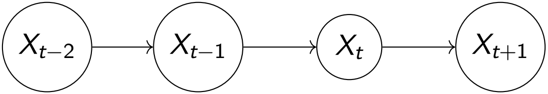
$$P(\mathbf{X}_t | \mathbf{X}_{t-1} \wedge \mathbf{X}_{t-2} \wedge \mathbf{X}_{t-3} \wedge \cdots \wedge \mathbf{X}_1)$$

Problem: As t increases, the number of previous states is unbounded. The conditional probability distribution can be unboundedly large.

Solutions: Make the Markov assumption — the current state depends on only a finite fixed number of previous states.

K-order Markov processes

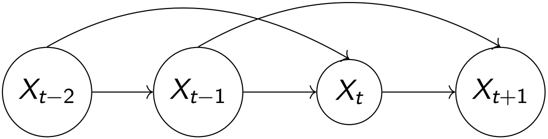
First-order Markov process:



The transition model:

$$P(\mathbf{X}_t | \mathbf{X}_{t-1} \wedge \mathbf{X}_{t-2} \wedge \mathbf{X}_{t-3} \wedge \cdots \wedge \mathbf{X}_1) = P(\mathbf{X}_t | \mathbf{X}_{t-1})$$

Second-order Markov process:

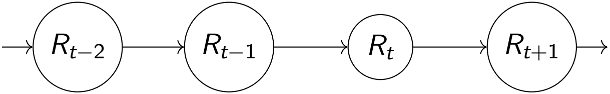


The transition model:

$$P(\mathbf{X}_t | \mathbf{X}_{t-1} \wedge \mathbf{X}_{t-2} \wedge \mathbf{X}_{t-3} \wedge \cdots \wedge \mathbf{X}_1) = P(\mathbf{X}_t | \mathbf{X}_{t-1} \wedge \mathbf{X}_{t-2})$$

Transition model for the umbrella world

The future is independent of the past given the present.



The transition model:

$$P(R_t | R_{t-1} \land R_{t-2} \land R_{t-3} \land \cdots \land R_1) = P(R_t | R_{t-1})$$

Stationary Process

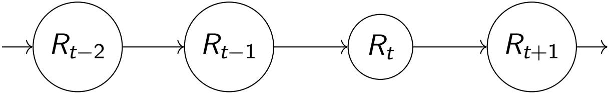
Is there a different conditional probability distribution for each time step?

Stationary process:

- The dynamics does not change over time.
- The conditional probability distribution for each time step remains the same.

Transition model for the umbrella world

$$P(R_t|R_{t-1}) = 0.7$$
$$P(R_t|\neg R_{t-1}) = 0.3$$



Sensor model

How does the evidence variable $\mathbf{E}_t$ for each time step t depend on the previous and current state variables?

Sensor Markov assumption: Any state is sufficient to generate the current sensor values.

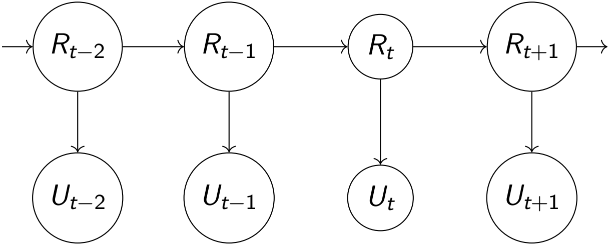
$$P(\mathbf{E}_t|\mathbf{X}_t \land \mathbf{X}_{t-1} \land \cdots \land \mathbf{X}_1 \land \mathbf{E}_{t-1} \land \mathbf{E}_{t-2} \land \cdots \land \mathbf{E}_1)$$
$$= P(\mathbf{E}_t|\mathbf{X}_t)$$

Complete model for the umbrella world

$$P(R_1) = 0.5$$

$$P(R_t | R_{t-1}) = 0.7$$
$$P(R_t | \neg R_{t-1}) = 0.3$$

$$P(U_t | R_t) = 0.9$$
$$P(U_t | \neg R_t) = 0.2$$



Hidden Markov Model

- A Markov process
- The state variables are unobservable.
- The evidence variables, which depend on the states, are observable.

Common Inference Tasks

- Filtering: the posterior distribution over the most recent state given all evidence to date.
- Prediction: the posterior distribution over the future state given all evidence to date.
- Smoothing: the posterior distribution over a past state, given all evidence to date.
- Most likely explanation: find the sequence of states that is most likely to have generated all the evidence to date.

Filtering

Given $x_{t-1} = P(R_{t-1}|U_1 \wedge \cdots \wedge U_{t-1})$,
how do we compute $x_t = P(R_t|U_1 \wedge \cdots \wedge U_t)$?

Examples: $P(R_1 = r_1|U_1)$ and $P(R_2 = r_2|U_1 \wedge U_2)$

CQ: Filtered Estimate

CQ: What is $P(R_1 = t | U_1 = t)$?

(A) 0.518

(B) 0.618

(C) 0.718

(D) 0.818

(E) 0.918

Filtered Estimate $P(R_2 = r_2 | U_1 = u_1 \wedge U_2 = u_2)$

$$
\begin{aligned}
&P(R_2 = r_2 | U_1 = u_1 \wedge U_2 = u_2) \\
=&\alpha P(U_2 = u_2 | R_2 = r_2) \sum_{r_1} P(R_2 = r_2 | R_1 = r_1) P(R_1 = r_1 | U_1 = u_1)
\end{aligned}
$$

Forward recursion:

- From $P(R_1 = r_1 | U_1 = u_1)$ to $P(R_2 = r_2 | U_1 = u_1 \wedge U_2 = u_2)$
- From $P(R_2 = r_2 | U_1 = u_1 \wedge U_2 = u_2)$ to $P(R_3 = r_3 | U_1 = u_1 \wedge U_2 = u_2 \wedge U_3 = u_3)$
- ...

CQ: Filtering

CQ: Consider $P(U_2|R_2 \land U_1)$.
Which one of the following simplifications is valid?

(A) $P(U_2|R_2 \land U_1) = P(U_2|R_2)$

(B) $P(U_2|R_2 \land U_1) = P(U_2|U_1)$

(C) $P(U_2|R_2 \land U_1) = P(U_2)$

(D) None of (A), (B), and (C) is a valid simplification.

CQ: Filtering

CQ: Consider $P(R_2|R_1 \land U_1)$.
Which one of the following simplifications is valid?

(A) $P(R_2|R_1 \land U_1) = P(R_2|R_1)$

(B) $P(R_2|R_1 \land U_1) = P(R_2|U_1)$

(C) $P(R_2|R_1 \land U_1) = P(R_2)$

(D) None of (A), (B), and (C) is a valid simplification.

Forward Recursion for Filtering

$$
\begin{aligned}
&P(R_t | U_1 \wedge \cdots \wedge U_t) \\
=&\alpha P(U_t | R_t) \sum_{R_{t-1}} P(R_t | R_{t-1}) P(R_{t-1} | U_1 \wedge \cdots \wedge U_{t-1})
\end{aligned}
$$

Prediction

Given $P(R_{t+k}|U_1 \wedge \cdots \wedge U_{t-1})$,
how do we compute $P(R_{t+k+1}|U_1 \wedge \cdots \wedge U_{t-1})$?

Forward Recursion

$$P(R_{t+k+1}|U_1 \wedge \cdots \wedge U_{t-1})$$
$$= \sum_{R_{t+k}} P(R_{t+k+1}|R_{t+k}) P(R_{t+k}|U_1 \wedge \cdots \wedge U_{t-1})$$

Smoothing

For $1 \leq k < t$,

$$
\begin{aligned}
&P(R_k|U_1 \wedge \ldots U_t) \\
=&\alpha\, P(R_k|U_1 \wedge \cdots \wedge U_k)\, P(U_{k+1} \wedge \cdots \wedge U_t|R_k)
\end{aligned}
$$

Forward Recursion

$$
\begin{aligned}
&P(R_t|U_1 \wedge \cdots \wedge U_t) \\
=&\alpha P(U_t|R_t) \sum_{R_{t-1}} P(R_t|R_{t-1}) P(R_{t-1}|U_1 \wedge \cdots \wedge U_{t-1})
\end{aligned}
$$

Backward Recursion

$$
\begin{aligned}
&P(U_{k+1} \wedge \cdots \wedge U_t|R_k) \\
=&\sum_{R_{k+1}} P(U_{k+1}|R_{k+1})\, P(U_{k+2} \wedge \cdots \wedge U_t|R_{k+1})\, P(R_{k+1}|R_k)
\end{aligned}
$$

Most likely explanation

Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Construct a hidden Markov model given a real-world scenario.
- Perform filtering, prediction, smoothing and derive the most likely explanation given a hidden Markov model.