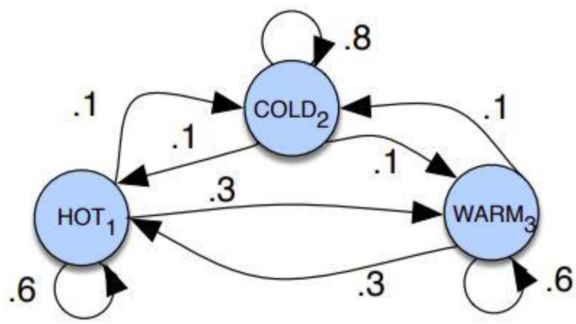
Hidden Markov Models

Pushkin Rathore 18111043

15th September, 2021

The HMM is based on augmenting the Markov chain. A Markov chain is a model that tells us something about the probabilities of sequences of random variables, states, each of which can take on values from some set. These sets can be words, or tags, or symbols representing anything, like the weather. A Markov chain makes a very strong assumption that if we want to predict the future in the sequence, all that matters is the current state. The states before the current state have no impact on the future except via the current state. It's as if to predict tomorrow's weather you could examine today's weather but you weren't allowed to look at yesterday's weather.



(a)

A Markov chain for weather (a), showing states and transitions. A start distribution π is required; setting π = [0.1, 0.7, 0.2] for (a) would mean a probability 0.7 of starting in state 2 (cold), probability 0.1 of starting in state 1 (hot), etc.

More formally, consider a sequence of state variables q1, q2, …... qi. A Markov model embodies the Markov assumption on the probabilities of this sequence: that Markov assumption when predicting the future, the past doesn't matter, only the present. Figure (a) shows a Markov chain for assigning a probability to a sequence of weather events, for which the vocabulary consists of HOT, COLD, and WARM. The states are represented as nodes in the graph, and the transitions, with their probabilities, as edges. The transitions are probabilities: the values of arcs leaving a given state must sum to 1.

The reason it is called a Hidden Markov Model is because we are constructing an inference model based on the assumptions of a Markov process. The Markov process assumption is simply that the "future is independent of the past given the present". In other words, assuming we know our present state, we do not need any other historical information to predict the future state. To make this point clear, let us consider the scenario below where the weather, the hidden variable, can be hot, mild or cold and the observed variables are the type of clothing worn. The arrows represent transitions from a hidden state to another hidden state or from a hidden state to an observed variable.

Notice that, true to the Markov assumption, each state only depends on the previous state and not on any other prior states.

Markov and Hidden Markov models are engineered to handle data which can be represented as 'sequence' of observations over time. Hidden Markov models are probabilistic frameworks where the observed data are modelled as a series of outputs generated by one of several (hidden) internal states. Markov models are developed based on mainly two assumptions.

Limited Horizon assumption: Probability of being in a state at a time t depend only on the state at the time (t-1).

$$P\ (z_t, z_{t-1}, z_{t-2}, \dots . z_1) = P(z_t \,/\, z_{t-1})$$

That means state at time t represents enough summary of the past reasonably to predict the future. This assumption is an Order-1 Markov process. An order-k Markov process assumes conditional independence of state z_t from the states that are k + 1-time steps before it.

Stationary Process Assumption: Conditional (probability) distribution over the next state, given the current state, doesn't change over time.

$$P\ (z_t \,/\, z_{t-1}) = P(z_2 \,/\, z_1) \quad \text{---------->}\ t \in 2 \dots \mathrm{T}$$

That means states keep on changing over time but the underlying process is stationary. the Markov property specifies that the probability of a state depends only on the probability of the previous state, but we can build more "memory" into our states by using a higher order Markov model. In an nth order Markov model:

$$P(x_i \mid x_{i-1}, x_{i-2}, \dots, x_1) = P(x_i \mid x_{i-1}, \dots, x_{i-n})$$