Hidden Markov Chains

Dr. Collins Odhiambo

Institute of Mathematical sciences, Strathmore University

codhiambo@strathmore.edu

June 22, 2022

Overview

Hidden Markov Model

- When we can not observe the state themselves but only the result of some probability function(observation) of the states we utilize HMM. HMM is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states.
- Latent or hidden variables in the model are never observed-We may or may not be interested in their values, but their existence is crucial to the model
- Some observations in a particular sample may be missing i.e. Missing information on surveys or medical records or We may need to model how the variables are missing
- A hidden Markov model is a tool for representing probability distributions over sequences of observations. In this model, an observation X_t at time t is produced by a stochastic process, but the state Z_t of this process cannot be directly observed, i.e. it is hidden. This hidden process is assumed to satisfy the Markov property.

Assumptions of HMM

- **Output independence assumption:** Output observation is conditionally independent of all other hidden states and all other observations when given the current hidden state.

$$P(X_t = V_i/Z_t = s_i) = P(X_t = V_i/X_1, X_2, \cdots, X_T, Z_1, Z_2, \cdots, Z_T)$$

$$= B_{ji}$$

- Emission Probability Matrix: Probability of hidden state generating output v_i given that state at the corresponding time was s_j.
- HMM is a statistical model in which the system being modeled are Markov processes with unobserved or hidden states. It is a hidden variable model which can give an observation of another hidden state with the help of the Markov assumption.

Real-world examples of Hidden Markov Models (HMM)

- **Retail scenario:** If you go to the grocery store once per week, it is relatively easy for a computer program to predict exactly when your shopping trip will take more time. The HMM calculates which day of visiting takes longer compared with other days and then uses that information in order to determine why some visits are taking long while others do not seem too problematic for shoppers like yourself.
- **Travel scenario:** By using HMM, airlines can predict how long it will take a person to finish checking out from an airport. This allows them to know when they should start boarding passengers!
- **Medical Scenario:** The HMM are used in various medical applications, where it tries to find out the hidden states of a human body system or organ. For example, cancer detection can be done by analyzing certain sequences and determining how dangerous they might pose for the patient.
- **Marketing scenario:** As marketers utilize a HMM, they can understand at what stage of their marketing funnel users are dropping off and how to improve user conversion rates.

Introduction to Hidden Markov Model

A hidden Markov model is defined by specifying five things:

- Q $=$ the set of states $= \{q_1, q_2, \cdots, q_n\}$
- V $=$ the output alphabet $= \{v_1, v_2, \cdots, v_m\}$
- $\pi(i) =$ probability of being in state q_i at time $t = 0$ (i.e., in initial states)
- A $=$ transition probabilities $= \{a_{ij}\}$, where $a_{ij} =$ Pr[entering state q_j at time $t+1|$ in state q_i at time t]. Note that the probability of going from state i to state j does not depend on the previous states at earlier times; this is the Markov property.
- B $=$ output probabilites $= \{b_j(k)\}$, where $b_j(k) =$ Pr[producing vk at time t — in state q_j at time t]

How an HMM works

- Assume a discrete clock $t = 0, 1, 2, \cdots$
- At each t, the system is in some internal (hidden) state $S_t = s$ and an observation $O_t = o$ is emitted (stochastically) based only on s (Random variables are denoted with capital letters)
- The system transitions (stochastically) to a new state S_{t+1}, according to a probability distribution $P(S_{t+1}|S_t)$, and the process repeats.
- This interaction can be represented as a graphical model (recall that each circle is a random variable, S_t or O_t in this case):
- Markovian assumption

The End