
ASSIGNMENT-5

HIDDEN MARKOV MODEL

Name: Shivam Yadav

Roll no: 19111054

Branch: Biomedical Engineering

3 NOVEMBER, 2021

1 Introduction

The Hidden Markov Model (HMM) is a statistical Markov model in which the system being represented is considered to be a Markov process with unobservable ("hidden") states X . HMM implies that another process, Y , has behaviour that "depends" on X . The objective is to learn about X by observation of Y . HMM requires that, for each time occurrence n_0 , the conditional probability distribution of Y_{n_0} given the history $X_n = x_n$ less than or equals to n_0 must not depend on $X_n = x_n$ n_0 . Hidden Markov models have been used in thermodynamics, statistical mechanics, physics, chemistry, economics, finance, signal processing, information theory, pattern recognition - including speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges, and bioinformatics. Let X_n and Y_n be discrete-time stochastic processes and n greater than or equals to 1.

2 Examples

2.1 Drawing balls from hidden urns

A hidden Markov process can be seen as a generalisation of the urn issue with replacement in its discrete form. Consider the following scenario: a genie resides in a room that is not visible to the naked eye but contains urns labelled y_1, y_2, y_3 .

2.2 Weather guessing game

The weather, according to Alice, is a discrete Markov chain. "Rainy" and "Sunny" are the two states, but she can't see them immediately. Every day, there's a good possibility Bob will do one of the following activities: "walk, shop, or clean."

3 Structural Architecture

The state space of the hidden variables is discrete in the typical form of HMM, whereas the observations are continuous. Transition probabilities and emission probabilities are the two types of parameters in a hidden Markov model. A categorical or Gaussian distribution can be used to create these parameters. A categorical distribution is used to model the hidden state space, which is considered to consist of one of N potential values. There is a set of emission probabilities governing the distribution of the observed variable at a specific time given the state of the hidden variable for each of the N possible states.

4 Probability of the latent variables

A number of related tasks ask about the probability of one or more of the latent variables, given the model's parameters and a sequence of observations $y(1), \dots, y(t)$.

4.1 Filtering

The aim is to compute the distribution over hidden states of the last latent variable at the end of the series, given the model's parameters and a sequence of observations, i.e. to compute $P(x(t) | y(1), \dots, y(t))$. When the sequence of latent variables is thought of as the underlying states that a process travels through at a series of points in time, with corresponding observations at each point in

time, this job is usually utilised. Then, at the conclusion, it's only natural to inquire about the status of the process.

The forward algorithm can be used to solve this problem quickly.

4.2 Smoothing

This is similar to filtering, except it asks about the distribution of a latent variable in the middle of a sequence, i.e. to compute $P(x(k) \mid y(1), \dots, y(t))P(x(k) \mid y(1), \dots, y(t))P(x(k) \mid y(1), \dots, y(t))P(x(k) \mid y(1), \dots, y(t))$. This can be regarded of as the probability distribution across concealed states for a point in time k in the past, relative to time t , from the perspective described above.

For computing the smoothed values for all hidden state variables, the forward-backward algorithm is a viable choice.

5 Learning

The parameter learning task in HMMs is to find, given an output sequence or a set of such sequences, the best set of state transition and emission probabilities. The task is usually to derive the maximum likelihood estimate of the parameters of the HMM given the set of output sequences. No tractable algorithm is known for solving this problem exactly, but a local maximum likelihood can be derived efficiently using the Baum–Welch algorithm or the Baldi–Chauvin algorithm. The Baum–Welch algorithm is a special case of the expectation-maximization algorithm.

If the HMMs are used for time series prediction, more sophisticated Bayesian inference methods, like Markov chain Monte Carlo (MCMC) sampling are proven to be favorable over finding a single maximum likelihood model both in terms of accuracy and stability.[8] Since MCMC imposes significant computational burden, in cases where computational scalability is also of interest, one may alternatively resort to variational approximations to Bayesian inference, e.g.[9] Indeed, approximate variational inference offers computational efficiency comparable to expectation-maximization, while yielding an accuracy profile only slightly inferior to exact MCMC-type Bayesian inference.

6 Application

HMMs can be applied in many fields where the goal is to recover a data sequence that is not immediately observable (but other data that depend on the sequence are). Applications include:

Computational finance Single-molecule kinetic analysis Cryptanalysis Speech recognition, including Siri Speech synthesis Part-of-speech tagging Document separation in scanning solutions Machine translation Partial discharge Gene prediction Handwriting recognition