HMMs

Dreycey Albin

November 2019

The key need for Hidden Markov Models (HMMs) was a for creating probabilistic models that can make inferences into sequence data. The most widely used example is for speech processing, but this technique can be applied to many systems. HMMs are perfect for modeling inherently noisy data, especially in the case of time varying or sequential data. The primary idea is that there is a state connected to each observable state, and there is a most probable (or likely) sequence of hidden states for a given sequence of observable, and vice-versa [2].

They key object or solution for HMMs is the idea of hidden states and observable states, and the idea of transition frequencies. These together allow for the stochastic model to be built. In addition, three techniques allowed for inferences to be made using HMMs, and the paper describes each of these: (1) How well does a model fit the observable sequence?; (2) Given an observable sequence, what is the most likely state sequence?; (3) How can the parameters of the model be learned? [2].

One limitation is how quick the probabilities tend towards zero in the forward and backward computations. This algorithm allows the maximum a posteriori to be found, but these probabilities approach zero geometrically. Of note, it is assumed that using the logs of the probabilities could be an answer to this problem. Another problem is if there is finite training data. This will cause certain probabilities to go to zero, just because they're not found in the training data [2].

An interesting idea from the paper is the idea of continuous vectors for the observable sequence rather than the discreet set of observables usually used. These are really interesting because there may be multiple features for each observable. This is interesting, because it is harder for me to conceptualize, and because it looks to outperform using the discreet model in the case of isolated word recognition [2].

The paper by Sean Eddy describes the use of Hidden Markov Model based profiles for proteins. These are essentially HMMs that can be used for multiple alignment of many different proteins. This out performs the current profiles being used at the time because a multiple sequence alignment of protein is not required (as is usually used/required for this problem), the model can be trained off of the sequences alone [1].

References

[1] Sean R Eddy. Hidden markov models. *Current opinion in structural biology*, 6(3):361–365, 1996.

[2] Lawrence R Rabiner and Biing-Hwang Juang. An introduction to hidden markov models. *ieee assp magazine*, 3(1):4–16, 1986.