"Hide and Seek": Smoothing and the Baum-Welch Algorithm in Hidden Markov Models

André T. Jensen, Asger M. Mikkelsen, Christian Hjort, Thomas Sandø

March 8th 2023

In this paper we review the analysis made in [1]. The concept of Hidden Markov Models will be explained, taking special interest in a method known as smoothing and the Baum-Welch estimating procedure[2]. The algorithms are used in an experiment examining the state of an atom in a optically probed cavity and these results are investigated[1].

1 Introduction

In this paper we review the analysis made in [1]. One type of inverse problem is when a non-observable state X follows a Markov process. The non-observable Markov process is linked to an observable signal s - which can take discrete values - through the signal probabilities $P(s|X=i)$ i.e. the probability of making a certain observation given the state. This is called a Hidden Markov Model (HMM). The parameters of such a model is therefore the number of states, the transition probabilities $P(X_{+1}=j|X=i)$, the prior $P(X=i)$ and the signal probabilities $P(s|X=i)$.

2 Review

In [1] a HMM with three states, corresponding to three atomic states, is proposed to describe single-atom jump dynamics in an optically probed cavity. The observations $\{s_1,...,s_N\}$ are the number of photons hitting a single-photon detector in a time window, and it is proposed that the state of the atom affects this number. A time series of photon counts is shown in fig. 1 where regions of high and low photon counts are easily identified. In regions marked I and II, the low photon count seems to take on slightly different values, which is why a three state model is proposed. [1] uses the following methods to analyze the data.

Smoothing. Given an HMM with known model parameters and a time-series of observations $\{s_1,...,s_N\}$, we would like to determine the posterior probability of being in state i at time t, given all the available observations $P(X_t=i|s_1,...,s_N)$. This analysis is known as smoothing and is used in [1] and derived in [2].
The method gives posterior probability distribution as:

$$P(X_t=i|s_1,...,s_N)=\frac{\alpha_t(i)\beta_t(i)}{\sum_k \alpha_t(k)\beta_t(k)} \tag{1}$$

Please refer to [1] for the equations of $\alpha_t(i)$ and $\beta_t(i)$.

The Baum-Welch algorithm. The process of estimating the a priori unknown transition and signal probabilities is known as the Baum-Welch algorithm and is outlined in [1]. $\gamma_t(i,j)$ is defined as the joint probability of being in state i at time t and being in state j at time $t+1$ given all observations:

$$\gamma_t(i,j)\equiv P\left(X_t=i,X_{t+1}=j \mid s_1,\dots s_N\right) \tag{2}$$

The calculation of $\gamma_t(i,j)$ is described in [1]. This allows us to re-estimate the transition probabilities:

$$P_{est}\left(X_{+1}=j \mid X=i\right)=\frac{\sum_t \gamma_t(i,j)}{\sum_t P\left(X_t=i \mid s_1,\dots s_N\right)}, \tag{3}$$

In a probabilistic way, the denominator represents the total time spent in state i (according to the model) and the numerator represents the number of shifts from state i to state j.
The signal probabilities is re-estimated by:

$$P(s \mid X=i)=\frac{\sum_t P\left(X_t=i \mid s_1,\dots s_N\right)\delta\left(s-s_t\right)}{\sum_t P\left(X_t=i \mid s_1,\dots s_N\right)} \tag{4}$$

In a probabilistic way, the numerator represents the number of times the observation s is made when the model is in state i. With the new transition and signal probabilities (model parameters), one can re-calculate $\alpha_t(i)$ and $\beta_t(i)$ and the process of re-estimating parameters is continued until the parameters converge. The Baum-Welch algorithm is guaranteed to converge to a local ML maximum.

Analysis Using the Baum-Welch and smoothing algorithms $P(X_t = i|s_1, ..., s_N)$ is estimated and shown in fig. 2. It is seen that the model clearly distinguishes between high and low transmission. Also, the models seems to be able to distinguish fairly well between the two states of low transmission, especially in the regions marked I and II in fig. 1. That the model is certain of its state assignment supports the reliability of the model.

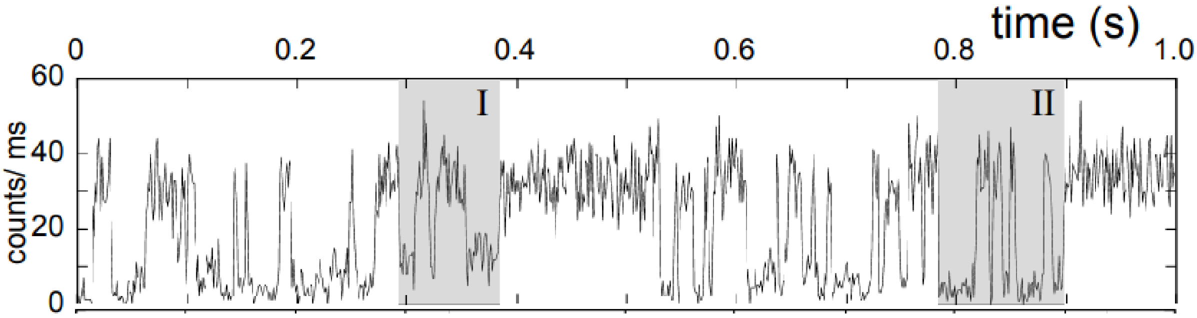

Figure 1: Photon counts binned in 1 ms time intervals. It is proposed that regions of high and low count correspond to different atomic states[1]

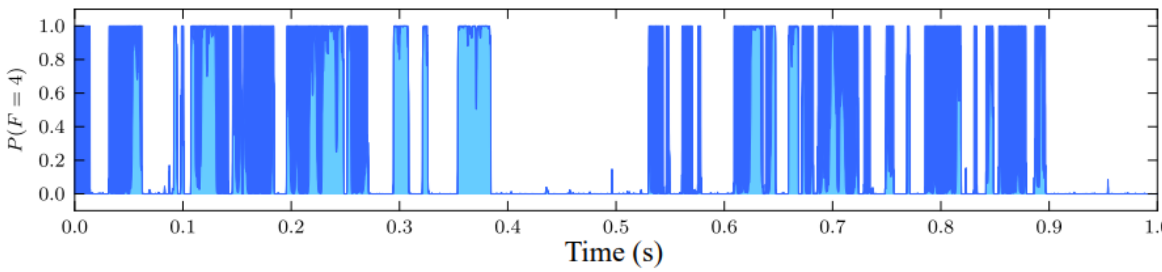

Figure 2: Using smoothing together with the fitting/Baum-Welch algorithm the probabilities for the different atom states are calculated [1]. White, light blue and dark blue corresponds to three different states of the atom. Notice how most of the time only one color has probability close to 1.

3 Conclusion

We have introduced Hidden Markov Models and reviewed its application in [1], where the specific algorithms of Smoothing and Baum-Welch are used. The result of the methods applied to a measurement series was presented.

References

[1] Gammelmark, S., Mølmer, K., Alt, W., Kampschulte, T., and Meschede, D. Hidden markov model of atomic quantum jump dynamics in an optically probed cavity. *Physical Review A 89*, 4 (apr 2014).

[2] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing (Ch.16.3)*. Cambridge University Press, 2007.