Neural ensemble states in prefrontal cortex identified using a hidden Markov model with a modified EM algorithm

Gregor Rainer*, Earl K. Miller

Department of Brain and Cognitive Sciences, The Center for Learning and Memory, The RIKEN-MIT Neuroscience Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Accepted 13 January 2000

Abstract

A modification to the standard estimation–maximization algorithm is presented, that allows the identification of neural ensemble states without predetermination of the number of states. Since the actual number of states a neural ensemble goes through in a given time period may vary from trial to trial, this represents a better description than using a predetermined, fixed number of states. The algorithm is used to identify ensemble states of simultaneously recorded neurons in the prefrontal cortex of behaving monkeys while they were freely viewing scenes containing multiple objects. We demonstrate that state transitions are correlated with behavioral and trial events. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Hidden Markov model; HMM; Estimation–maximization algorithm; EM; Electrophysiology

1. Introduction

Traditional methods of analyzing extracellular neural activity are based on averaging. A particular stimulus or behavioral condition is repeated several times, and the response of the neuron is averaged over a given epoch to obtain an estimate of its' activity during that epoch. Responses of single neurons are typically too variable to
allow analysis on a single trial basis. However, simultaneous data from multiple neurons should allow such single trial analysis. Because different neurons generally have different response properties, dynamic ranges and preferred stimuli, it is not advantageous to simply average their activity. Other methods are needed to combine the activity of multiple neurons to obtain a single estimate of ensemble activity. The hidden Markov model (HMM) formalism provides one such method [4,5]. The activity of a set of simultaneously recorded neurons is modeled as a system moving through a predetermined number of stable states. Each state is associated with a set of firing probabilities for each of the neurons. The Baum–Welch algorithm [1,2] (a version of estimation-maximization or EM) can be used to simultaneously estimate maximum likelihood solutions for both the firing probabilities and the probabilities for moving between the states. Besides the neural activity data, the EM algorithm needs to be supplied with initial guesses for the state transition probabilities and the firing probabilities, as well as the number of states. The number of states turns out to be the most important parameter, and the most difficult to determine. The actual number of states in a neural ensemble may vary widely in an epoch of a given length. Determination of the optimum number of states for each trial parametrically is difficult, because more states generally lead to solutions of higher likelihood. This is due to the fact that there are more free parameters and not necessarily because the fit to the empirical data is better. Even though it is possible in principle to determine an optimal state number by penalizing for the number of degrees of freedom, this is computationally intensive and dependent on the choice of penalization. Here, we present a modified version of EM, which does not require the specification of state number prior to execution. The major advantage of this algorithm is that it can capture variations in state number from trial to trial, which is not possible using the standard algorithm. We apply this algorithm to neural ensemble data from the prefrontal cortex of behaving monkeys while they are viewing scenes containing multiple objects. Monkeys were performing a visual short-term memory task that was identical to the task described in [3], except that during the present experiment monkeys were allowed to freely view the sample arrays. In addition, array presentation time was 1000 ms followed by a delay of 1500 ms. For the EM estimation, we used a contiguous 2500 ms time-period, consisting of 500 ms fixation time prior to array onset, 1000 ms of array presentation time and the first 1000 ms of the delay. Importantly, monkeys were not rewarded for gazing at particular objects in the array, they performed eye-movements naturally in the context of the visual memory task. Thus, we consider eye-movements made in the present experiment to be very similar to those made during natural free viewing made outside the laboratory.

2. Results

The EM algorithm and its application to the HMM formalism is fully described in Rabiner [4]. The neural spike trains are first converted to sequences of output symbols. Given a set of \( n \) neurons \( \{1, 2, \ldots, n\} \), the output symbol at each time-bin was chosen to be the number of the neuron which fired an action potential at that
time-bin. In cases, where more than one neuron fired, we arbitrarily chose one of them as the output symbol. While it is possible to assign different output symbols to such collision events and thus capture synchronous activity patterns explicitly, we did not do so because collisions occurred infrequently (usually under 5% of time-bins). Although neural data were collected at 1 ms time resolution, we used 5 ms for the EM estimation. Larger time-bins are advantageous because they save on computer execution time, and thus make the algorithm more efficient. On the other hand, increasing bin size leads to more frequent spike collisions. The bin size was thus chosen to be as large as possible, while keeping the frequency of spike collisions relatively low. Having chosen an alphabet of output symbols, the next and most critical step is to specify the number of states. Before describing the modification we introduced, an example of the standard version of EM (shown in Fig. 1A) is discussed. The top panel shows action potentials from six neurons during the 2500 ms epoch described above. Below are shown eight subsequent iterations of EM using six states. Similar to previous studies, we use only left–right HMMs here, not fully ergodic ones. Left–right HMMs have all state transition probabilities from the \( i \)th to the \( j \)th state \( P_{ij} = 0 \) for \( i < j \). Fully ergodic HMMs do not impose any such restrictions, but for reliable results they typically require more than one input sequence — and thus cannot be easily used for single-trial analyses. As is typical for EM, the algorithm converged in a few iterations. The initial state transition probabilities were chosen to provide broad coverage of the entire time period, so that the final states span the entire time period of 2500 ms. Note that the choice of starting probabilities has a major impact on the final solution. Choosing sojourn probabilities (probabilities for remaining in a given state: \( P_{ii} \)) that are too large results in a general shift of all states to the right of the 2500 ms epoch, and choosing values that are too small causes states to accumulate at the beginning. The initial firing probabilities were set to the same small number for all states and neurons \( p = 10^{-5} \). Results do not depend closely on the exact number chosen, but we did find that choosing the same firing probability for all states is advantageous compared to random initialization, because the choice of state transition probabilities provides enough degrees of freedom and better control over starting conditions. The final output of the algorithm is shown at the bottom, and consists of the previously specified number of states, and the transition times between these states. Note that another output are the firing probabilities for each neuron in each state, which are identical to the mean firing rates in the epochs corresponding to the states (due to the Poisson assumption). The fact that the (local) maximum likelihood solution depends on the starting parameters (in our framework the probabilities \( P_{ij}^{\text{initial}} \)) usually necessitates running the algorithm several times with different (usually random) starting parameters and combining the solutions to obtain an optimal estimate (known as “boosting”). In the simplest case, combination can take the form of choosing the most likely solution. Since no exhaustive search in parameter space is possible, starting from random guesses has obvious drawbacks — it is computationally intensive and there is no guarantee that a globally optimal solution is found. Choosing the number of states can be viewed as an additional parameter that is even more difficult to control for, because (as discussed above) one cannot simply take the most likely solution.
Fig. 1. (A) The standard EM algorithm applied to a single trial for the set of six simultaneously recorded neurons shown in the top panel. Each vertical line represents a single spike (or action potential), each row corresponds to one of the six neurons. Eight iterations of EM are shown using six hidden states, leading to the state transition profile shown in the bottom panel. (B) The modified EM algorithm, which is based on successive 1000 ms segments using only two states each time, applied to the same trial as shown in (A). The output of this algorithm again consists of a state transition profile, shown in the bottom panel.

To circumvent some of these problems, we decided to run the EM algorithm locally on 1000 ms segments of data with just two states. The exact choice of segment length is not critical and guided by prior knowledge what the timescale of interest is — in our case hundreds of milliseconds. This choice is implicitly made also in the standard algorithm by specifying the number of states and duration of the entire segment. The modified algorithm is illustrated in Fig. 1B for the same segment of data described above. We start with a 1000 ms segment at the beginning of the time period, and execute EM on this segment with two states. Note that the soujourn probability for state 1 ($P_{11}$) is chosen such that the initial transition time is near the start of the segment, at around 100 ms. In a few iterations the algorithm finds a maximum likelihood transition time $\tau_1$ at around 740 ms. Once this first transition is found, we select another segment of 1000 msec from the original data, from $[\tau_1, \ldots, \tau_1 + 1000]$
and again run EM on this new segment to find a new transition time $\tau_2$. This procedure is repeated until a transition is found that occurs more than 2500 ms after the beginning of the time period. For EM estimation, we always use 1000 ms segments, and thus the segment may extend beyond the 2500 ms period (e.g. from 2150 to 3150 ms). Once such a transition is found, the algorithm terminates. As with the original algorithm, this leads to a sequence of states, with transition times between them as illustrated at the bottom of Fig. 1B. Note that the number of states is different from the one arbitrarily specified for the previous sequence, and that this number is never explicitly specified in the modified algorithm.

To evaluate the performance of the modified EM algorithm, we examined the relationship of the HMM transitions with behavioral and sensory events on a trial-by-trial basis. Fig. 2 shows the spikes and HMM transitions from Fig. 1, along with the eye movements the monkey made on this particular trial (we show the “eye speed”, which is defined as the magnitude of the difference between successively sampled eye positions). As can be seen, the monkey made two large saccadic eye movements (labelled $s_1$ and $s_2$), and two small microsaccades ($m_1$ and $m_2$) on this particular trial. We note that on this trial, both the onset of the visual response (first HMM transition), and most of the saccades ($s_2$, $m_1$, $m_2$) were accurately identified. Both these correlations were confirmed by running the EM algorithm on all 728 correctly executed trials from this data set: (1) A HMM transition was observed in over 70% of these trials within 200 ms of stimulus onset. (2) HMM transitions occurred within $\pm 100$ ms of saccades more frequently than expected by chance ($P = 1.1 \times 10^{-13}$).
3. Conclusion

We have presented a method that allows the detection of stable states in a set of simultaneously recorded neurons using a hidden Markov model (HMM) without the predetermination of the number of states ahead of time. This allows the model to capture trial-by-trial variations in the experimental data better than previous versions of the algorithm. We show an application to data collected from the prefrontal cortex of behaving monkeys, and demonstrate that HMM transitions computed with our method are correlated with sensory (stimulus presentations) and oculomotor (eye movements) aspects of the task.

References


Gregor Rainer received a M.Sc. degree in semiconductor physics from the University of Vienna, Austria in 1994, and a Ph.D. in systems neuroscience from the Massachusetts Institute of Technology, Cambridge, USA in 1999. He is interested in visual perception, memory and learning, and how different brain regions are involved in these processes. In addition, he works on the application of computational techniques to data collected simultaneously from multiple neurons.

Earl K. Miller is an Associate Professor in the Center for Learning and Memory, Department of Brain and Cognitive Sciences, and RIKEN-MIT Neuroscience Research Center at the Massachusetts Institute of Technology. He received his M.A. and Ph.D. in Psychology and Neuroscience from Princeton University. He is interested in the neural bases of cognitive function.