Deep Learning for Parameter Recovery from a Neural Mass Model of Perceptual Decision-Making

Emanuele Sicurella (SicurellaE@cardiff.ac.uk)

Jiaxiang Zhang (ZhangJ73@cardiff.ac.uk)
School of Psychology, Cardiff University Brain Research Imaging Centre
Cardiff University, Cardiff, UK

Abstract

In neuroscience, parameter recovery refers to the problem of finding the best parameters of a model for fitting the experimental data. The developing of more biologically plausible computational models of cognition has offered a significant improvement in the predictive power at the cost of a higher complexity posing increasing challenges on parameter recovery. Here, we present a deep learning approach to recover parameters of a two-variables neural mass model simulating evidence accumulation during perceptual decision-making. We show that our algorithm is able to recovery well specific set of parameters but might fail when trying to predict combinations of parameters with a high degree of interaction, *i.e.* parameters that have inherently similar effects on the model's output. Thus, our study suggests that deep learning for parameter recovery should go together with a carefully designed experiment to study the effects of different parameters that are not richly interacting.

Keywords: parameter recovery; deep learning; neural mass model; perceptual decision making; evidence accumulation

Introduction

In cognitive neuroscience, an increasing attention has been put into biologically more realistic models able to capture the cellular and circuit mechanisms instead of relying only on more abstract behavioral models. In general, these models are more complex than behavioral ones meaning that they may have a higher number of parameters and more complex differential equations. Therefore, although they fit the behavioral data very well, it is not straightforward to recover their parameters. Recent works have tackled the problem of parameter recovery from non-differentiable behavioral decision making models like the diffusion model for conflict tasks (White, Servant, & Logan, 2018) or the leaky competing accumulator (Miletić, Turner, Forstmann, & van Maanen, 2017). However, the problem of parameter recovery of biologically realistic models of decision making remains under-represented in literature. In this study, we perform parameter recovery of a neural model using deep learning as it is more computationally efficient and reliable than other global optimization procedures (Stonkute, 2019).

Methods

We simulated the data using the two competing variables neural mass model designed by Wong and Wang (2006) which reproduces the evidence accumulation involved in decision-making during a random dot leftward or rightward motion discrimination task. The two populations interact via mutual inhibitory connections and self excitatory mechanisms. The input is modeled as two different synaptic currents generated by the stimulus and it is stronger at higher dot movement coherence level. The model also considers a common non-selective background input and a noisy current. Here, we used an extension of the model where the common background current is modulated by a factor β to take into account the speed-accuracy trade-off (Standage, Wang, & Blohm, 2014; Heitz & Schall, 2012). The equations describing the output synaptic currents of the two populations l and r are:

$$I_{syn,l} = J_{ll}S_l - J_{lr}S_r + I_l + \beta I_0 + I_{noise,l} \tag{1}$$

$$I_{syn,r} = J_{rr}S_r - J_{rl}S_l + I_r + \beta I_0 + I_{noise,r} \tag{2}$$

where $I_{l,r}$ are the stimulus-generated input currents, $r_{l,r}$ are the firing rates of the neurons of the two populations, $S_{l,r}$ are the gating variables representing the activity of the two populations and $I_{noise,l,r}$ Gaussian noise currents. These variables are described by differential equations which are not reported for simplicity, see (Wong & Wang, 2006) for a full description of the model. For the parameters, $J_{ll} = J_{rr} = 0.2609$nA is the self-excitatory coupling strength, $J_{lr} = J_{rl} = 0.0497$nA the mutual inhibitory coupling strength and β is the modulation of the background current $I_0 = 0.3255$nA.

We focused on recover four parameters: the excitatory and inhibitory coupling strengths J_{ll} and J_{lr}, the background current modulation β and we defined a new parameter τ_{nd} called non-decision time as a simple constant of time added to the reaction time for each decision. Initially, the parameter values have been drawn from the following uniform distributions: $J_{ll} = J_{rr} \in [0.25, 0.30]$nA; $J_{lr} = J_{rl} \in [0, 0.25]$nA; $\beta \in [0.95, 1.05]$; $\tau_{nd} \in [100, 500]$ms.

To determine the intervals we searched for parameter values that gave a biologically plausible outputs. For example, we expected the accuracy to be around 0.5 at 0% coherence and increasing following a sigmoid-like function to saturate at 1 for higher coherence levels (Wong & Wang, 2006; Roitman & Shadlen, 2002).

For each set of parameters, we computed the equations for a total of 10,000 trials. Each trial had a maximum duration of 2.5s for a total of 2,500 iterations of 1ms each. A decision is made if the activity $S_{l,r}$ of one of the two accumulators reached a threshold set at 15Hz. For simplicity, we assumed a positive value for the coherence meaning that the correct direction of

384

the points was always leftward. The output for each trial was then 0 for rightward/wrong decisions, or 1 for leftward/correct decisions. If the trial ended without either accumulators activities reaching the threshold, an output of -1 indicating a non-decision has been generated. If the non-decision rate was higher than 0.2 we discarded the sample. The results were used to compute a set of features describing the performance of the model for each set of parameters. We used the following 13 features: the accuracy calculated as the fraction of correct choices over the total choices made, average reaction times for correct and wrong choices and the 10th, 30th, 50th, 70th and 90th percentiles of the reaction times distributions for wrong and correct choices.

We performed training and hyperparameter optimization on a deep neural network with three fully connected layers using the 13 features as input and the four parameters as outputs (all the Python scripts for data sampling and deep learning training are publicly available at the repository `https://github.com/Emalude/wwparrecdl`.

Results

The model fails to predict J_{ll} values ($R^2 = 0.11$) and performs slightly better for the other parameters with R^2 scores of 0.86, 0.69 and 0.88 for J_{lr}, β and τ_{nd} respectively. We hypothesized that the difficulty of prediction might be due to the concurrency of J_{ll} and β with a similar influence on the outputs of the model. To test this hypothesis, we generated two new batches of samples; in the first, we kept $J_{ll} = 0.2609$ constant. In the second we set $\beta = 1$. We repeated hyperparameter tuning and training separately for the two batches with three parameters as output. The results with the absence of modulation β are shown in Figure 1. The R^2 scores are 0.95, 0.99 and 0.99 for J_{ll}, J_{lr} and τ_{nd} respectively. When J_{ll} is constant, R^2 scores are 0.94, 0.90 and 0.92 for J_{lr}, β and τ_{nd} respectively.

Discussion

Our results suggest that parameters with a similar influence on the output are difficult to recover with high accuracy. Therefore, it is important to carefully design an experiment and combine the parameter selection with previous knowledge about the effects of those parameters and their variability across subjects or conditions. For example, suppose that by analyzing the behavioral outputs of a random dot motion experiment on healthy and pathological subjects, our algorithm recovers a significantly lower self-excitatory synaptic strength in subjects with a certain pathology compared to healthy ones. We could not use this as a final proof of the fact that the pathology lowers J_{ll} because we first need to address the possibility that such difference in the behavioral output is caused by other parameters having a similar effects on the outputs in their range of values (for example β).
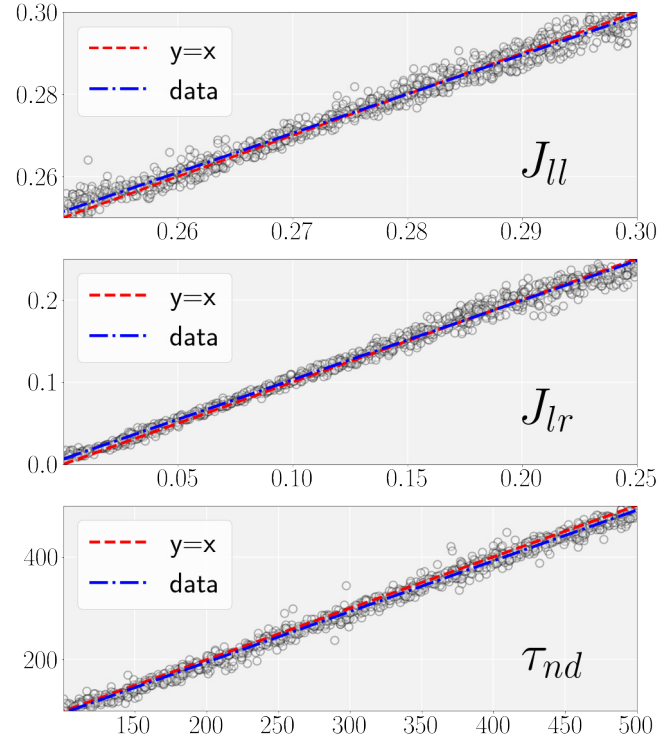
Acknowledgments

Figure 1: Predictions for J_{ll}, J_{lr} and τ_{nd} with other parameters kept constant. Real parameters are on the x-axis and predicted values on the y-axis.

References

Heitz, R. P., & Schall, J. D. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron*, *76*(3), 616–628.

Miletić, S., Turner, B. M., Forstmann, B. U., & van Maanen, L. (2017). Parameter recovery for the leaky competing accumulator model. *Journal of Mathematical Psychology*, *76*, 25–50.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of neuroscience*, *22*(21), 9475–9489.

Standage, D., Wang, D.-H., & Blohm, G. (2014). Neural dynamics implement a flexible decision bound with a fixed firing rate for choice: a model-based hypothesis. *Frontiers in neuroscience*, *8*, 318.

Stonkute, S. (2019). *Deep learning as an alternative to global optimization in diffusion model for conflict tasks*. Doctoral dissertation, School of Psychology, Cardiff University, UK.

White, C. N., Servant, M., & Logan, G. D. (2018). Testing the validity of conflict drift-diffusion models for use in estimating cognitive processes: A parameter-recovery study. *Psychonomic bulletin & review*, *25*(1), 286–301.

Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, *26*(4), 1314–1328.