Which measure should we use for unsupervised spike train learning?

Antonio R. C. Paiva[1] and Il Park[2]
[1] Scientific Computing and Imaging Institute, University of Utah
[2] Computational NeuroEngineering Laboratory, University of Florida

Abstract

The dynamics of a neural system may not be fully determined by external stimuli, because the neural activity depends on internal states from a wide range of possible causes. For example, bistable dynamics of a single neuron has been observed in vitro via frozen noise injection [1], local field potentials and EEG phase often correlates with response strength, and top-down control such as attention are known to affect responses. In other words, it is nearly impossible to control or observe all the internal variables. Still, we would like to infer these internal states by analyzing the observation variability. Specifically, one can use unsupervised learning techniques such as PCA and clustering [2, 3] to discover the internal states. However, as observed in the kernel methods literature, the usefulness of such methods is tightly associated with the expressiveness of the chosen similarity measure. Hence, we analyze which spike train similarity measure performs best for statistical analysis.

We compare three different measures: the memoryless cross-intensity (mCI) kernel, which is equivalent to a continuous cross-correlation measure, the inner product associated with the b-metric introduced by Houghton [4], and the nonlinear cross-intensity (nCI) kernel [2, 3]. These measures where evaluated with both synthetic spike trains and multi-channel spike trains recorded from an in-vitro neural culture that was electrically stimulated at ten different channels in a random order. These controlled experiments were used so that the accuracy of the result can be evaluated quantitatively. In both datasets, and for both PCA and clustering, mCI performed the worst, and nCI obtained the best results. For example, clustering of the spike train from the in-vitro neural culture recordings into 10 clusters, yielded 60.6% correct separation using mCI and 71.2% using the b-metric' inner product, versus 99.3% using nCI. Our results demonstrate that the nCI measure performs considerably better and is able to nearly perfectly infer the different stimuli from the observed spike trains.

In addition, we analyze why the nCI kernel measure seems to perform so well. We show that the nCI kernel is strictly positive definite and, thus, is a characteristic kernel. This means that two point processes map to different points in the reproducing kernel Hilbert space (RKHS) induced by the kernel. In other words, this kernel ensures that spike trains corresponding to different point processes have distinct means in the RKHS, which allows statistical inference to distinguish their origin. In comparison, the mCI is a product kernel on a linear space and does not have these properties.

References

[1] J.-M. Fellous, P. H. E. Tiesinga, P. J. Thomas, and T. J. Sejnowski Discovering Spike Patterns in Neuronal Responses J. of Neuroscience, 24(12):2989-3001, March 2004.

[2] A. R. C. Paiva, I. Park, and J. C. Principe A reproducing kernel Hilbert space framework for spike train signal processing Neural Comp., 21(2):424449, Feb. 2009.

[3] A. R. C. Paiva, I. Park, and J. C. Principe Inner products for representation and learning in the spike train domain In Karim G. Oweiss, editor, Statistical Signal Processing for Neuroscience, Academic Press, 2010. In-press.

[4] C. Houghton Studying spike trains using a van Rossum metric with a synapse-like filter J. Comp. Neurosci., 26:149155, 2009.