Encoding Neurons' Communication: A Statistical Approach to Analyze Interactive Neural Spike Trains

Ruiwen Zhang Feng Liu

Abstract

Neuroscience experiments and neural spike train data have special features that present novel and exciting challenges for statistical researches. Several standard statistical procedures, widely used in other fields of science have found their way into mainstream application in neuroscience data analysis. Given the firing times of an ensemble of neurons, an integrate several inputs and fire model is introduced based on the conditional intensity function approach. This is different from the existing methods where the intensity function is approximated by discretization with the sampling intervals chosen arbitrary. In this paper, we model the log conditional intensity function directly by employing a polynomial spline function for the target or response spike train and a tensor product of splines for the peer or predictor spike trains. The parameters are defined by those used in constructing the polynomial splines, and they will be estimated by the maximum likelihood method. The statistical properties of this procedure will be evaluated using a simulated experiment. Our model captures the underlying spontaneous firing of the target as well as the stimulus inputs from its peers, and both in continuous time.

Key Words: Point process, Log conditional intensity function, Maximum likelihood, Spline regression, Interacting neurons.

1. Introduction

The conditional intensity function is a history-dependent generalization of the rate function for a neural spike train. The parametric estimation of the conditional intensity function is a good start, however the strong assumption of the underlying probability model turns to be inappropriate in most cases. The field of nonparametric estimation has been advanced with the use of computing technology, we focus on the numerical approximation using spline functions.

In this paper, the discretization and the parametric component will be relaxed by modeling the log conditional intensity function directly using polynomial spline functions. For the target neuron, a cubic spline is used to model its effect in the spirit of HEFT (Kooperberg, Stone and Truong, 1995); while linear splines and their tensor products are used to model the associated neurons or peers' effects. This method is similar to the HARE methodology for hazard regression (Kooperberg, Stone and Truong, 1995) with the exception that is a time-dependent covariate approach. The advantage here is that the flexibility of the model will allow us to examine directly the impacts of other neurons within the neural network, either stimulating or inhibiting. This model specification will be described in detail in Section 2.

2. Conditional Log-intensity Function

For one particular neuron, it has the spontaneous firing rate depending on its own natural characteristics. Also, the firings of other neurons (also called peers) within the neural network have impacts on the target neuron. Suppose there are M such peers and their firing times before time t are given by $\mathbf{x}(t) = (x_1(t), .., x_M(t))$. Let T be a positive random variable whose distribution may depend on the peers $\mathbf{x}(t)$.

Let $\lambda(\cdot|\mathbf{x}(s), s > 0)$ denote the conditional intensity function of T given $\mathbf{x}(s)$ so that

$$\lambda(t|\mathbf{x}(t))\, dt = Prob(T \in (t, t + dt)|\mathbf{x}(t)).$$

Let $\alpha(\cdot|\mathbf{x}(t))$ denote the log conditional intensity function. To simplify the discussion, we assume the conditional intensity function at time t depends only on the value of the covariates up to that time; that is, we assume that $\lambda(t|\mathbf{x}(s), s > 0) = \lambda(t|\mathbf{x}(s), 0 \leq s \leq t)$ and hence that $\alpha(t|\mathbf{x}(s), s > 0) = \alpha(t|\mathbf{x}(s), 0 \leq s \leq t)$.

In this paper, we model the log conditional intensity function via a spline function $\mu(t) = \sum_{i=1}^{p} \theta_i B_i(t)$ and a vector of M predetermined functions, $\boldsymbol{\beta} \cdot \mathbf{h}(\mathbf{x}(t))$, corresponding to the covariates for the affects of the interacting neurons.

3. Maximum Likelihood Estimation

Given the interspike interval (ISI) data with the vector of covariates $\mathbf{x}(t)$ and a set $B_1, ..., B_p$ of basis functions, we will estimate the model coefficients by maximum likelihood.

The partial log-likelihood can be written as

$$\begin{aligned}\phi(t|\mathbf{x}(t)) &= \log \lambda(t|\mathbf{x}(t)) - \int_0^t \lambda(s|\mathbf{x}(s))ds \\ &= \alpha(t|\mathbf{x}(t)) - \Lambda(t|\mathbf{x}(t)).\end{aligned} \tag{1}$$

We take the partial derivatives of $\phi(\cdot)$ to examine its concavity and maximize the log likelihood. It is easily derived to see that the second derivative is always definite nonpositive.

The log-likelihood function is a sum for all N consecutive ISIs or called waiting time, denoted as $\{u_k\}_{k=1}^{N}$, then the log-likelihood function is

$$l(\boldsymbol{\theta}, \beta_1) = \sum_{k=1}^{N} \alpha(u_k|\mathbf{x}(u_k)) - \sum_{k=1}^{N} \int_0^{u_k} \lambda(s|\mathbf{x}(s))ds \tag{2}$$

The maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$ and $\widehat{\beta_1}$ is given by maximizing the log-likelihood, $l(\boldsymbol{\theta}, \beta_1)$.

4. Simulation Study

We follow the principle ideas in "integrate-and-fire" model by Lapicque (Abbott, 1999). Stimulus from peers in the form of current inputs are accumulative until the very next firing of target cell, after that triggering, the stimulus before would have no impact on any coming events, like resetting the clock and counting from zero. In the meantime, voltage of membrane naturally decays along the time, referred to as "leaky integrate-and-fire", so that the impact of peer firing diminish gradually. Based on those neurobiological supports, we proposed a model of interspike interval (ISI) data which follows Section 2 while the function of covariates is decreasing over time.

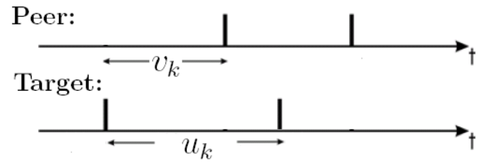

Figure 1: Target and peer spike trains are recorded simultaneously. Between two consecutive spikes of the target, the distance is our observation and u_k is for the interval ending at the k^{th} spike. The peer may fire within the interval, and then we record the distance from the beginning of u_k to the peer firing time as our covariate v_k.

For two interacting neural spike trains, the target spike train has N spikes or N ISIs, $\{u_k\}_{k=1}^{N}$, then each spike of the peer belongs to a specific ISI of the target spike train and only have its affect within that interval. So, for any u_k, we denote the distance form the beginning of the interval to a particular spike of peer within the interval as v_k which is then the covariate for the k^{th} observation (See Figure 1). For the case that the peer has more than one time firing occurred within the interspike interval (ISI) of the target, we only consider the last one as the covariate.

We simulated 100 spike trains and each of which has 650 spike times given a deterministic peer spike train. In Figure 2 and Figure 3, the estimators of θ_1 to θ_4 corresponding to knot locations $\{0.2, 0.4, 0.6, 0.8\}$ and the estimators of β_1 and β_2 are shown respectively after 100 simulations.

REFERENCES

Abbott, L. F. (1999). Lapique's introduction of the integrate-and-fire model neuron. *Brain Research Bulletin* **50**(5/6), 303-304.

Brillinger, D. R. (1988). Some statistical methods for random process data from seismology and neurophysiology. *The Annals of Statistics* **16**(1), 1-54.

Brillinger, D. R. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cybernetics* **59**, 189-200.

Brillinger, D. R. and Segundo, J. P. (1979). Empirical examination of the: threshold model of neuron firing. *Biological Cybernetics* **35**, 213-220.

Brown, E. N., Barbieri, R., Ventura, V. R. E., and Frank, L. M. (2001). The time-rescaling theorom and its application to neural spike train data analysis. *Neural Computation* **14**, 325-346.

Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C. and Wilson, M. A. (1998). A statistical paradigm for neuralspike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *Journal of Neuroscience*. **18**, 7411C7425.

Kooperberg, C., Stone, C. J., and Truong, Y. K. (1995a). Hazard regression. *Journal of the American Statistical Association* **90**, 78-94.

Kooperberg, C., Stone, C. J., and Truong, Y. K. (1995b). The L_2 rate of convergence for hazard regression. *Scandinavian Journal of Statistics* **22**, 143-157.

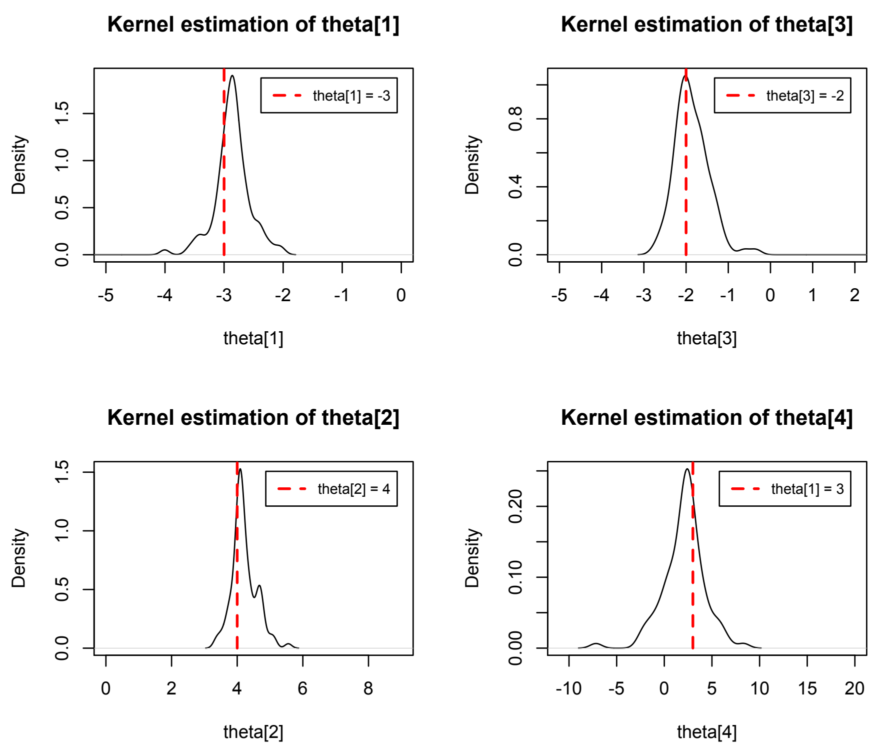

Figure 2: Kernel density estimations of θ's from 100 simulations. In the setup, the spline knot locations are fixed as $\{0.2, 0.4, 0.6, 0.8\}$, and two covariates are considered in the model. The estimates have unbiased means and less variations.

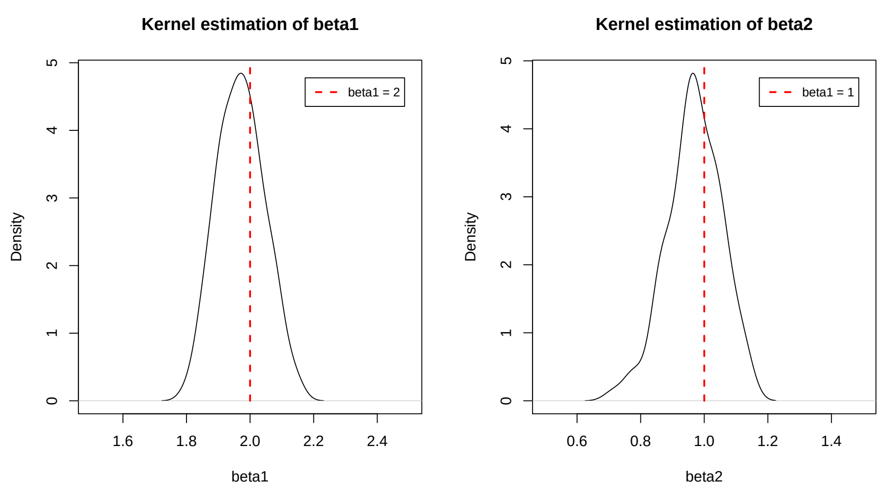

Figure 3: Kernel density estimations of β_1 and β_2 from 100 simulations. In the setup, the spline knot locations are fixed as $\{0.2, 0.4, 0.6, 0.8\}$, and two covariates are considered in the model. The estimates have unbiased means and less variations.