Neural Networks 102 (2018) 67–77

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Multineuron spike train analysis with R-convolution linear combination kernel

Taro Tezuka [*]

Center for Artificial Intelligence Research, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Japan
Faculty of Library, Information, and Media Science, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Japan

ARTICLE INFO

Article history:
Received 19 March 2017
Received in revised form 14 November 2017
Accepted 20 February 2018
Available online 6 March 2018

Keywords:
Spike trains
Neural coding
Kernel methods
Gaussian process regression

ABSTRACT

A spike train kernel provides an effective way of decoding information represented by a spike train. Some spike train kernels have been extended to multineuron spike trains, which are simultaneously recorded spike trains obtained from multiple neurons. However, most of these multineuron extensions were carried out in a kernel-specific manner. In this paper, a general framework is proposed for extending any single-neuron spike train kernel to multineuron spike trains, based on the R-convolution kernel. Special subclasses of the proposed R-convolution linear combination kernel are explored. These subclasses have a smaller number of parameters and make optimization tractable when the size of data is limited. The proposed kernel was evaluated using Gaussian process regression for multineuron spike trains recorded from an animal brain. It was compared with the sum kernel and the population Spikernel, which are existing ways of decoding multineuron spike trains using kernels. The results showed that the proposed approach performs better than these kernels and also other commonly used neural decoding methods.

- A novel way of extending any single spike train kernel to multineuron spike trains is proposed.
- Subclasses of a kernel having fewer parameters are introduced, making optimization more feasible.
- Gaussian process regression with the kernel outperformed other existing decoding methods.

1. Introduction

A spike train is a sequence of time points at which a neuron generates an action potential. It is often regarded as the fundamental element of information processing in the brain (Grün & Rotter, 2010). Not only is it essential for understanding biological neural networks, but it is also essential for constructing artificial neural networks since it is the form of information representation used by the brain and likely to have some advantages. Many artificial neural networks use spike trains to represent information, for example, in pulsed neural networks.

To decode information content from biological spike trains, various spike train metrics (distances) have been proposed (Houghton & Victor, 2011; Naud, Gerhard, Mensi, & Gerstner, 2011; van Rossum, 2001; Victor & Purpura, 1996). These metrics have also been extended to multineuron spike trains (that is, a simultaneously recorded set of spike trains) obtained from multiple neurons (Aronov, 2003; Houghton & Kreuz, 2013; Kreuz, Chicharro, Houghton, Andrzejak, & Mormann, 2012). Multineuron spike trains are also called "parallel spike trains" (Grün & Rotter, 2010).

Although metrics are most commonly used in spike train analysis, there is yet another approach to defining a similarity measure between spike trains, namely, to define a symmetric positive-definite (SPD) kernel. Its merit is that kernel methods such as Gaussian process regression can be applied (Rasmussen & Williams, 2006). Kernel methods are popular in machine learning since they can extend multivariate analysis on vector spaces for arbitrary data structures. Although the analysis can be conducted on spike trains by transforming them into finite dimensional vectors through binning, the process usually results in information loss. For this reason, many spike train kernels have been proposed (Eichhorn, Tolias, Zien, Kuss, Rasmussen, Weston, et al., 2004; Fisher & Banerjee, 2010; Paiva, Park, & Príncipe, 2009; Park, Seth, Rao, & Príncipe, 2012; Shpigelman, Singer, Paz, & Vaadia, 2003).

While spike train metrics are often used both for single-neuron and multineuron spike trains, spike train kernels are seldom extended to multineuron spike trains. Among these few examples, Shpigelman et al. used the Spikernel and kernel autoregressive-moving average (kernel-ARMA) to analyze multineuron spike trains (Shpigelman, Lalazar, & Vaadia, 2008; Shpigelman, Singer, Paz, & Vaadia, 2005). Park et al. proposed the product kernel and the direct sum kernel as means of extending spike train kernels to the multineuron case (Park, Seth, Paiva, Li, & Príncipe, 2013). Li et al. applied the sum kernel to real data (Li, Brockmeier, Choi,

---

[*] Correspondence to: Faculty of Library, Information, and Media Science, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Japan.
E-mail address: tezuka@slis.tsukuba.ac.jp.

https://doi.org/10.1016/j.neunet.2018.02.013
0893-6080/© 2018 Elsevier Ltd. All rights reserved.

Table 1
Notation.

Single-neuron spike train kernel	
i: sample index	N: number of samples in dataset
$s^{(i)}$: ith single-neuron spike train	t_ι: time that ι-th spike was generated
τ: time	η: filtering function
σ: parameter for Laplace filtering	$\lambda_{\eta,s}$: function obtained by convolving s with η
k: kernel	S: set of all possible single-neuron spike trains
Multineuron spike train kernel	
$\mathbf{x}^{(i)}$: ith multineuron spike train	$\mathbf{x}_m^{(i)}$: spike train from mth neuron in $\mathbf{x}^{(i)}$
M: number of neurons	k_e: elementary kernel
C: coefficient matrix	P: SPD coefficient matrix
A: rectangular matrix	Δ: diagonal matrix with non-neg. components
$W_{A,\Delta}$: factor analysis matrix	γ: off-diagonal component of P
k_C: linear combination kernel	k_P: RCLC kernel
k_γ: OPRC kernel	ψ_e: parameters of elementary kernel k_e
Evaluation	
$y^{(i)}$: dependent variable	h: index for test sample
Y: set of all conditions	Y_0: set of conditions in training data
Ω: set of all samples	$\mathfrak{S}_O$: data used for hyperparameter optimization
$\mathfrak{S}_R$: data used for regression	$\mathfrak{r}$: ratio for hyperparameter optimization
θ: direction in degrees	
Decoding methods	
ϕ_m: firing rate of neuron m	υ_m: preferred direction of neuron m
ν_m: number of spikes for neuron m	ω: coefficients for spline
ϱ: knot location for spline	λ_T: time-varying rate
ρ: regularization coefficient	

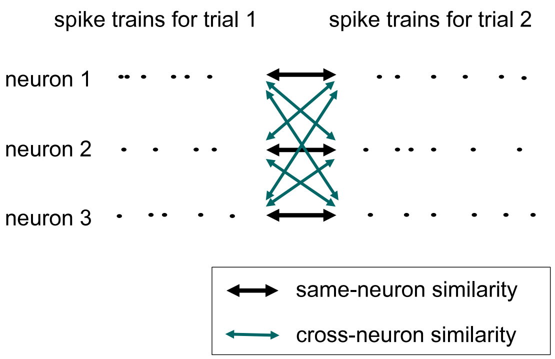

Fig. 1. Dependence in spike generation among different neurons in multineuron spike trains.

Francis, Sanchez, & Príncipe, 2014; Li, Park, Brockmeier, Chen, Seth, Francis, Sanchez, & Príncipe, 2013).

A more general approach is taken in this paper. Namely, a special case of Haussler's R-convolution kernel (Haussler, 1999) is applied to multineuron spike trains in a way that it naturally extends a kernel defined on single-neuron spike trains (Tezuka, 2014). Also, some special subclasses of the R-convolution spike train kernel are explored. Kernels in these subclasses have fewer parameters and are therefore more suited for optimization when the size of observed data is limited.

One of such parameters relates to the amount of collaborative coding by neurons. If the information is coded independently among neurons, using similarity between spike trains generated by different neurons, as illustrated in Fig. 1, will not add much to the decoding performance. It has been pointed out, however, that relative spike timings between different neurons are meaningful. In that case, cross-neuron similarity will help decode information. Many existing multineuron spike train metrics have parameters that express the amount in which neurons are interchangeable (Houghton & Kreuz, 2013; Houghton & Victor, 2011; Kreuz et al., 2012). One of the subclasses of the R-convolution spike train kernels proposed in this paper also has one parameter that represents the amount of similarity in spike trains between two different neurons.

In essence, a multineuron spike train is a set of labeled time points. In other words, it is a sample obtained from a marked point process. The R-convolution spike train kernel is, therefore, applicable to other data with the same structure, for example, failures in a production line, customers visiting an e-commerce site, sporadic events in global finance, and outbreaks of natural disasters. The R-convolution linear combination (RCLC) kernel proposed in this paper is applicable in these domains as well. Variables used in the paper are summarized in Table 1.

2. Related work

Numerous distances (metrics) have been proposed for single-neuron spike trains because they can be applied to various machine learning tasks including classification and clustering (Houghton & Kreuz, 2013; Houghton & Victor, 2011; Kreuz et al., 2012; Naud et al., 2011; van Rossum, 2001; Victor & Purpura, 1996). Several of these distances have been extended to the multineuron case (Aronov, 2003; Aronov, Reich, Mechler, & Victor, 2003; Houghton & Sen, 2008). Another similarity measure between spike trains that is not a distance is the symmetric positive-definite (SPD) kernel, which is an abstraction of the inner product. One advantage of using an SPD kernel is that it provides access to the full armory of kernel methods including SVM and Gaussian process regression (Schölkopf & Smola, 2001). Some of the existing similarity measures for spike trains are summarized in Table 2. They can be classified into two approaches. One is based on *embedding*, where spike trains are embedded into a high-dimensional vector space with a metric. Another method is *edit distance*, which measures the cost of editing one of two spike trains to match the other one (Houghton & Victor, 2011). The van Rossum distance and the Victor–Purpura distance are notable examples of these two approaches, respectively.

2.1. Spikernel

Shpigelman et al. proposed the Spikernel, where spike trains are converted into a vector of binned spike rates and the similarity is