Front. Electr. Electron. Eng. China 2011, 6(1): 146–150
DOI 10.1007/s11460-010-0123-x

RESEARCH ARTICLE

Kazushi IKEDA, Daisuke KOMAZAWA, Hiroyuki FUNAYA

Information geometry in neural spike sequences

Abstract An information geometrical method is developed for characterizing or classifying neurons in cortical areas whose spike rates fluctuate in time. The interspike intervals (ISIs) of a spike sequence of a neuron is modeled as a gamma process with a time-variant spike rate, a fixed shape parameter and a fixed absolute refractory period. We formulate the problem of estimating the fixed parameters as semiparametric estimation and apply an information geometrical method to derive the optimal estimators from a statistical viewpoint.

Keywords information geometry, neural spikes, semiparametric estimation, interspike intervals (ISIs)

1 Introduction

The characteristics of neurons in cortical areas have been the subject of recent discussion in the literature. In particular, there has been discussion on the statistical properties of the interspike intervals (ISIs) of spike sequences such as the coefficient of variation, C_V, the skewness coefficient, S_K, the correlation coefficient of consecutive intervals, C_{OR}, and the local variation, L_V [1–5]. In particular, Shinomoto et al. [6] have shown that the local variation with a refractory period, L_{VR}, can almost classify the functions of neurons without any other information.

From the viewpoint of statistical modeling, ISIs can be modeled as a gamma process with a variable rate but a fixed shape parameter [5,7]. Since gamma distributions form a two-dimensional e-flat manifold S from the information-geometrical viewpoint [8–10], the problem of estimating the shape parameter results in a semiparametric estimation [11,12]. These theoretical methods also apply to estimate the refractory period in L_{VR}, which has heuristically been determined to date.

The rest of this paper is organized as follows: Some statistical measures of spike sequences and their properties are introduced in Sect. 2. Sect. 3 formulates the problem of ISIs as a semiparametric estimation and Sect. 4 briefly introduces the information geometry for semiparametric models. The solution of the problem is given in Sect. 5 and Sect. 6 confirms the results by computer simulations. Sect. 7 concludes with some discussion.

2 Statistical measures for ISIs

When a spike sequence is given and its N ISIs are written as $T_1, T_2, \ldots, T_N$, the C_V and S_K measures are defined as the standard deviation of the ISIs divided by the mean of the ISIs and the skewness of the ISI distribution, respectively. That is,

$$C_V = \frac{1}{\bar{T}}\sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\left(T_n - \bar{T}\right)^2}, \tag{1}$$

$$S_K = \frac{\frac{1}{N-1}\sum_{n=1}^{N}\left(T_n - \bar{T}\right)^3}{\left(\frac{1}{N-1}\sum_{n=1}^{N}(T_n - \bar{T})^2\right)^{3/2}}, \tag{2}$$

where

$$\bar{T} = \frac{1}{N}\sum_{n=1}^{N} T_n. \tag{3}$$

The C_V measure expresses the regularity: It takes a low value for a regular spike sequence, one for a sequence of infinite length generated by a fixed Poisson process, and a large value when the process is time-dependent. The S_K measure shows the asymmetry of a sequence: It can be either positive or negative, but it takes two for a sequence of infinite length generated by a stationary Poisson process. However, since they are based on the mean $\bar{T}$ of the ISIs, C_V or S_K will become large when the spike rate is globally modulated, even though the spike

Received July 1, 2010; accepted October 20, 2010

Kazushi IKEDA (✉), Hiroyuki FUNAYA
Nara Institute of Science and Technology, Nara 630-0192, Japan
E-mail: kazushi@is.naist.jp

Daisuke KOMAZAWA
Kyoto University, Kyoto 606-8501, Japan

sequence is locally quasi-regular [1,13]. As a result, they are not suitable for classifying particular neurons such as those found in cortical areas that change their spike rates drastically in a waiting-period task, for example.

To overcome this problem, Shinomoto et al. [5] proposed the L_V measure, defined as

$$L_V = \frac{1}{N-1}\sum_{n=1}^{N-1}\frac{3(T_n - T_{n+1})^2}{(T_n + T_{n+1})^2}, \tag{4}$$

where the factor 3 is taken so that the expectation of L_V becomes 1 when the sequence obeys a stationary Poisson process. Since the L_V measure reflects the stepwise variability of ISIs and does not compare ISIs with different spike rates, L_V can assume a small value, even for a sequence with a time-variant spike rate. They confirmed that C_V undergoes a large change but L_V does not for a sequence generated by a time-dependent Poisson process [13].

Recently, Shinomoto et al. [6] proposed a variant of L_V called L_{VR}, which explicitly includes an absolute refractory period R for each neuron, where the value is common to all neurons. That is,

$$L_{VR} = \frac{1}{N-1}\sum_{n=1}^{N-1}\frac{3(T_n - T_{n+1})^2}{(T_n + T_{n+1} - 2R)^2}, \tag{5}$$

where the refractory period R is determined heuristically from the given data. L_{VR} is shown to classify the functions of neurons in other words, the value of L_{VR}.

3 Statistical model of ISIs

In the literature, ISIs are modeled as a gamma process with a variable rate and a fixed shape parameter. Although the rate can fluctuate continuously, it can be simplified by assuming that the rate parameter is fixed between two spikes. Then, an interspike interval T independently obeys a gamma distribution with a rate $\xi^{(l)}$ and shape parameter κ,

$$q\left(T;\xi^{(l)},\kappa\right) = \frac{\left(\xi^{(l)}\kappa\right)^{\kappa}}{\Gamma(\kappa)}T^{\kappa-1}\exp\left[-\xi^{(l)}\kappa T\right], \tag{6}$$

where l runs from 1 to N.

As for L_{VR}, the absolute refractory period R slightly modifies Eq. (6) to

$$q\left(T;\xi^{(l)},\kappa,R\right) = \frac{\left(\xi^{(l)}\kappa\right)^{\kappa}}{\Gamma(\kappa)}(T-R)^{\kappa-1}\exp\left[-\xi^{(l)}\kappa(T-R)\right]. \tag{7}$$

This is assumed in the existing models that R has a constant value of zero. Therefore, we consider Eq. (7) in the following.

From Eq. (7), the probability distribution of a spike sequence is written as

$$p\left(\{T\};\kappa,R,k(\xi)\right) = \int\prod_{l=1}^{N} q\left(T_l,\xi^{(l)},\kappa,R\right)k\left(\xi^{(l)}\right)\mathrm{d}\xi^{(l)}. \tag{8}$$

This is called a semiparametric model, where two kinds of parameters appear: One is a finite number of parameters of interest, κ and R, and the other is a nuisance parameter, k, which has an infinite degrees of freedom. In other words, estimating the shape parameter κ and the refractory period R is formulated as a semiparametric estimation [11].

4 Information geometry for semiparametric models

We generalize Eq. (8) by assuming that m observations, $\{T^{(l)}\} \equiv \{T_1^{(l)}, T_2^{(l)}, \ldots, T_m^{(l)}\}$, are given for each $\xi^{(l)}$, where $\xi^{(l)}$ is generated from an unknown probability density $k(\xi)$. That is, the distribution of the ISIs in the lth set is described as

$$p\left(\{T^{(l)}\};\xi^{(l)},\kappa,R\right) = \prod_{i=1}^{m} q\left(T_i^{(l)};\xi^{(l)},\kappa,R\right), \tag{9}$$

and that of the sequence is

$$p(\{T\};\kappa,R,k(\xi)) = \int\prod_{l=1}^{N/m} q\left(\{T^{(l)}\},\xi^{(l)},\kappa,R\right)k\left(\xi^{(l)}\right)\mathrm{d}\xi^{(l)}. \tag{10}$$

A semiparametric estimation is known to be solvable using the estimating function method [14,15], where the estimator is the solution of the estimating equation,

$$\sum_{l=1}^{N/m}\sum_{i=1}^{m} f\left(T_i^{(l)},\kappa,R\right) = 0, \tag{11}$$

where $f(T;\kappa,R)$ is an estimating function that satisfies

$$\mathrm{E}_{\kappa,R,k}[f(T;\kappa,R)] = 0, \tag{12}$$

for any κ, R and k. Here $\mathrm{E}_{\kappa,R,k}$ denotes the expectation with respect to the distribution

$$p(T;\kappa,R,k) = \int p(T;\xi,\kappa,R)k(\xi)\mathrm{d}\xi. \tag{13}$$

How can we find good estimating functions? The information geometry [8,9], which sheds light on various fields of information science, gives a basic theory for estimating functions. Amari and Kawanabe [16] show the conditions under which estimating functions exist, how large the set of estimating functions is, and what estimating function is optimal.

We give an intuitive explanation for the theory of estimating functions (see Sect. 3 of Miura et al. [12] and its references for details). Note that in the present paper, the same notation is employed and the superscripts (l) have been omitted for $\xi^{(l)}$ and $T_i^{(l)}$ in the following analysis.

In short, an estimating function extracts the component orthogonal to the nuisance parameters, that includes the elements of the parameters of interest and has no other elements. See Fig. 1, for example, where F_i shows the space in which the information on the parameters of interest is included, F_n shows the space of the nuisance parameters, and F_a shows the orthogonal complement to $F_i \oplus F_n$. Since no assumptions are given for the nuisance parameters, the elements of the estimate in the direction of F_n are of no use. In other words, consideration needs to be made of the space orthogonal to F_n. Hence, we denote by F_e the space where the elements of F_n are removed from F_i. The estimating functions are included in $F_e \oplus F_a$ so that they are not affected by nuisance parameters. Obviously, F_a has no information on the parameters of interest. This means that F_e itself is the optimal space, where the optimal estimating function is included.

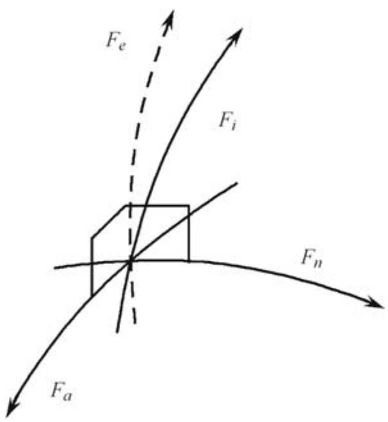

Fig. 1 Illustration of estimating functions

Using the Amari and Kawanabe method, Miura et al. [12] proved that there does not exist any estimating function when $m = 1$ and derived the optimal estimating function when $m > 1$, which leads to the optimal estimator of κ, irrespective of $\xi^{(l)}$. Their estimator has the minimum variance for estimating κ as long as the ISIs obey gamma distributions.

5 Estimating functions for ISIs

Using the theory in the previous section, we derive the optimal estimating function for κ and R in a purely mathematical manner, where the refractory period R is common and constant but unknown, following the analysis in Miura et al. [12].

Substituting Eq. (7), Eq. (9) is written as

$$\prod_{i=1}^{m} q(T_i; \xi, \kappa, R) = \exp\left[\xi \cdot s(\{T\}, \kappa) + r(\{T\}, \kappa, R) - \psi(\kappa, R, \xi)\right], \quad (14)$$

where

$$s(\{T\}, \kappa) = -\kappa \sum_{i=1}^{m} T_i, \quad (15)$$

$$r(\{T\}, \kappa, R) = (\kappa - 1) \sum_{i=1}^{m} \log(T_i - R), \quad (16)$$

$$\psi(\kappa, R, \xi) = -m\kappa \log(\xi\kappa) + m \log \Gamma(\kappa) - m\xi\kappa R. \quad (17)$$

Amari and Kawanabe show that the optimal estimating functions for κ and R are given by

$$u_\kappa^I(\{T\}, \kappa, R) = u_\kappa - \mathrm{E}[u_\kappa | s] \quad (18)$$

$$= \sum_{i=1}^{m} \log(T_i - R) - m\mathrm{E}[\log(T_1 - R)|s], \quad (19)$$

$$u_R^I(\{T\}, \kappa, R) = u_R - \mathrm{E}[u_R | s] \quad (20)$$

$$= (1 - \kappa) \sum_{i=1}^{m} \frac{1}{T_i - R} - m(1 - \kappa)\mathrm{E}\left[\frac{1}{T_1 - R} \Big| s\right], \quad (21)$$

where u_κ^I and u_R^I are defined so that they are orthogonal to any function of s, as seen in Eqs. (18) and (20). The marginal distribution of s and the conditional expectation in the last term of Eq. (19) are given in the Appendix of Miura et al. [12] as being

$$p(s) = \int \delta\left[s + \kappa \sum_{i=1}^{m} T_i\right] \cdot \prod_{i=1}^{m} q(T_i; \xi, \kappa, R) \mathrm{d}T_i k(\xi) \mathrm{d}\xi \quad (22)$$

$$= \int \prod_{i=1}^{m-1} B(i\kappa, \kappa) \left(-\frac{s}{\kappa} - mR\right)^{m\kappa - 1} \frac{(\xi\kappa)^{m\kappa}}{\Gamma(\kappa)^m} \cdot \exp[\xi s + \xi\kappa m R] \frac{k(\xi)}{\kappa} \mathrm{d}\xi, \quad (23)$$

$$\mathrm{E}[\log(T_1 - R)|s] = \int \log(T_1 - R) \delta\left[s + \kappa \sum_{i=1}^{m} T_i\right] \cdot \prod_{i=1}^{m} q(T_i; \xi, \kappa, R) \mathrm{d}T_i k(\xi) \mathrm{d}\xi \frac{1}{p(s)} \quad (24)$$

$$= \log\left[-\frac{s}{\kappa} - mR\right] - \phi(m\kappa) + \phi(\kappa), \quad (25)$$

where δ is the Dirac delta function and $\phi(\kappa)$ is the digamma function defined as

$$\phi(\kappa) = \frac{\Gamma'(\kappa)}{\Gamma(\kappa)}. \tag{26}$$

Similarly, we can derive the conditional expectation in the last term of Eq. (21) as

$$\mathrm{E}\left[\frac{1}{T_1 - R}\middle| s\right] = \int \frac{1}{T_1 - R}\delta\left[s + \kappa\sum_{i=1}^{m} T_i\right] \cdot \prod_{i=1}^{m} q(T_i;\xi,\kappa,R)\mathrm{d}T_i k(\xi)\mathrm{d}\xi \frac{1}{p(s)} \tag{27}$$

$$= \frac{1}{-\frac{s}{\kappa} - mR}\left(1 - \frac{m\kappa - \kappa}{\kappa - 1}\right). \tag{28}$$

In total, the optimal estimating functions are written as

$$u_\kappa^I(\{T\},\kappa,R) = \sum_{i=1}^{m}\log(T_i - R) - m\log\sum_{i=1}^{m}(T_i - R) + m\phi(m\kappa) - m\phi(\kappa), \tag{29}$$

$$u_R^I(\{T\},\kappa,R) = (1-\kappa)\sum_{i=1}^{m}\frac{1}{T_i - R} - \frac{m(1-\kappa)}{\sum_{i=1}^{m}(T_i - R)}\left(1 - \frac{m\kappa - \kappa}{\kappa - 1}\right). \tag{30}$$

Using the estimating functions above, the estimators of κ and R are the solution of the estimating equations,

$$\sum_{l=1}^{N} u_\kappa^I(\{T^{(l)}\},\kappa,R) = 0, \tag{31}$$

$$\sum_{l=1}^{N} u_R^I(\{T^{(l)}\},\kappa,R) = 0. \tag{32}$$

6 Computer simulations

To confirm the validity of the theoretical analysis given above, computer simulations were carried out. In each experiment, κ, R, m and ξ are fixed at 0.5, 0.2, 3 and 1[1], respectively.

Figure 2 shows the estimated κ and R determined by the optimal estimating functions derived above as a function of the number N of observations, where the solid and dashed lines denote the averages of estimated κ and R over 1000 trials and each error bar shows the standard deviation. We can see that the values of κ and R converge to the true values, 0.5 and 0.2, with decreasing deviations, as N increases.

Next, we compared the estimate error of κ of our method to those of the conventional method by Miura et al. with $R = 0.2$ (the true value) and with $R = 0$ (the conventional model). Obviously, the former gives the lower bound of the estimate error in this framework. In Fig. 3 the solid, dashed and dotted lines describe the root-mean-square errors of our method, the lower bound and the conventional model, respectively. We can see that the proposed method has a comparable error to the lower bound and smaller than the conventional model even when examples are few.

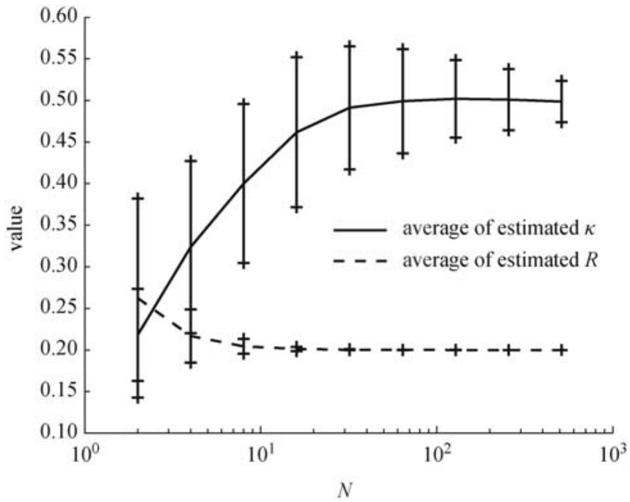

Fig. 2 κ and R of the proposed method versus N

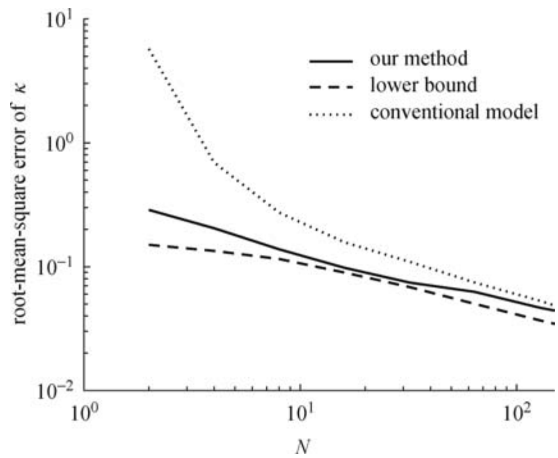

Fig. 3 Comparison of estimated errors of κ

7 Conclusions

In this manuscript, we discuss the statistical models of ISIs. When a spike sequence of a neuron obeys a gamma process with a time-variant spike rate, the information geometry gives the optimal estimating function from the statistical viewpoint. The method is also applicable to the modified model which has an absolute refractory period, and successfully estimates not only the shape parameter but also the refractory period.

1) The estimator does not depend on ξ, and we can assume a constant value.

As shown in Ref. [11], L_V is an approximated estimator of the regularity, and so must be L_{VR}. Since L_{VR} is a strong tool for neuron classification [6], our method will replace L_{VR} in the field in the near future.

Acknowledgements This study was supported in part by Grant-in-Aid for Scientific Research (18300078) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Holt G R, Softky W R, Koch C, Douglas R J. Comparison of discharge variability *in vitro* and *in vivo* in cat visual cortex neurons. Journal of Neurophysiology, 1996, 75(5): 1806–1814
2. Shinomoto S, Sakai Y, Funahashi S. The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. Neural Computation, 1999, 11(4): 935–951
3. Sakai Y, Funahashi S, Shinomoto S. Temporally correlated inputs to leaky integrate-and-fire models can reproduce spiking statistics of cortical neurons. Neural Networks, 1999, 12(7–8): 1181–1190
4. Shinomoto S, Shima K, Tanji J. New classification scheme of cortical sites with the neuronal spiking characteristics. Neural Networks, 2002, 15(10): 1165–1169
5. Shinomoto S, Shima K, Tanji J. Differences in spiking patterns among cortical neurons. Neural Computation, 2003, 15(12): 2823–2842
6. Shinomoto S, Kim H, Shimokawa T, Matsuno N, Funahashi S, Shima K, Fujita I, Tamura H, Doi T, Kawano K, Inaba N, Fukushima K, Kurkin S, Kurata K, Taira M, Tsutsui K, Komatsu H, Ogawa T, Koida K, Tanji J, Toyama K. Relating neuronal firing patterns to functional differentiation of cerebral cortex. PLoS Computational Biology, 2009, 5(7): e1000433
7. Tiesinga P H E, Fellous J M, Sejnowski T J. Attractor reliability reveals deterministic structure in neuronal spike trains. Neural Computation, 2002, 14(7): 1629–1650
8. Amari S. Differential-Geometrical Methods in Statistics. Lecture Notes in Statistics. Berlin: Springer-Verlag, 1985
9. Amari S, Nagaoka H. Methods of Information Geometry. New York: Oxford University Press, 2000
10. Miura K, Shinomoto S, Okada M. Search for optimal measure to discriminate random and regular spike trains. IEICE Technical Report, 2004, NC2004-52
11. Ikeda K. Information geometry of interspike intervals in spiking neurons. Neural Computation, 2005, 17(12): 2719–2735
12. Miura K, Okada M, Amari S. Estimating spiking irregularities under changing environments. Neural Computation, 2006, 18(10): 2359–2386
13. Shinomoto S, Tsubo Y. Modeling spiking behavior of neurons with time-dependent Poisson processes. Physical Review E, 2001, 64(4): 041910
14. Godambe V P. Conditional likelihood and unconditional optimum estimating equations. Biometrika, 1976, 63(2): 277–284
15. Godambe V P. Estimating Functions. Oxford: Oxford University Press, 1991
16. Amari S, Kawanabe M. Information geometry of estimating functions in semiparametric statistical models. Bernoulli, 1997, 3(1): 29–54

Kazushi Ikeda received the B.E., M.E., and Ph.D degrees in mathematical engineering and information physics from the University of Tokyo in 1989, 1991, and 1994, respectively. He was a research associate with the Department of Electrical and Computer Engineering, Kanazawa University from 1994 to 1998. In 1995, he was a research associate of the Chinese University of Hong Kong for three months. From 1998 to 2008, he was with the Graduate School of Informatics, Kyoto University, as an associate professor. Since 2008, he has been a full professor of Nara Institute of Science and Technology. He is the editor-in-chief of Journal of Japanese Neural Network Society, an action editor of Neural Networks, an associate editor of IEEE Transactions on Neural Networks and an associate editor of IEICE Transactions on Information and Systems.

Daisuke Komazawa received the B.E. and M. Informatics degrees from Kyoto University, Kyoto, Japan, in 2007 and 2009, respectively. He is currently working in an industrial company in Japan.

Hiroyuki Funaya received the B.E. and M. Informatics degrees from Kyoto University, Kyoto, Japan, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree from Nara Institute of Science and Technology, Nara, Japan. His research interests include machine learning, statistical inference, and optimization.