[P3-08]

An unbiased estimator of noise correlations under signal drift

Keiji Miura (P)[1,2]

[1] Collaborative Mathematics Research Unit, Graduate School of Information Sciences, Tohoku University
[2] Alliance for Breakthrough between Mathematics and Sciences, PRESTO, JST
E-mail: miura@ecei.tohoku.ac.jp

Abstract— Small correlations in noises (trial-to-trial response variations) can decrease coding capacity by sensory neurons dramatically. Although significant noise correlation has been reported from almost all cortical areas, nonstationarity, such as a drift in signals (mean responses) might engender artificial correlations even if no actual correlation exists. Although attempts to estimate noise correlation under changing environments have been made, they were useful only for specific cases. This paper presents consideration of a bivariate normal distribution for activities of two neurons and advances the proposition of a semiparametric method for estimating its covariance matrix in an unbiased fashion, whatever the time course of the signal is.

Keywords— Noise Correlations, Information Geometry, Population Coding, Multiple Spike Train Data

1 Introduction

Correlations in noise (trial-to-trial variation in response to the same stimulus) can play an important role in information representation in the brain. The degree to which sensory information is represented reliably by neural responses has been characterized by applying an information theoretic approach in a stochastic stimulus–response framework [1].

It has been shown theoretically that correlation in response noises can be a major determinant for coding capacities of sensory information by neurons [1, 2]. Actually, even in a simple homogeneous network with tiny correlation, having more neurons does not help at all [3] (but see also [1, 2]). Therefore, it is extremely important to estimate noise correlation accurately.

Although significant noise correlation has been observed in almost all recorded cortical areas, it has been pointed out that nonstationarity such as a drift in signals (mean responses to a given stimulus) can engender artificial correlation even if no actual correlation exists [4, 5, 6]. Although attempts to estimate noise correlation under changing environments have been made, they were applicable only to specific cases or computationally demanding [4, 5, 6].

For some specific semiparametric statistical models [7], unbiased estimators under arbitrarily changing environments have been obtained in simple, analytically closed forms [8, 9] by using information geometry [10]. As described in this paper, this method is applied to the bivariate normal distribution for activities of two neurons, thereby deriving an optimal estimator of its covariance matrix, which works whatever the signal drift is.

2 Model

I consider a bivariate normal distribution for activities of two neurons:

$$q(X;\xi,\Sigma)=\frac{1}{2\pi|\Sigma|^{\frac{1}{2}}}\exp\{-\frac{1}{2}(\mathbf{X}-\xi)'\Sigma^{-1}(\mathbf{X}-\xi)\}. \quad (1)$$

where $X=\{x,y\}$ and $\xi=\{\xi_1,\xi_2\}$ are vectors. These analyses address the situation in which the covariance matrix Σ is constant whereas the signals ξ can change over time. Especially, when the signals are arbitrarily distributed, but two consecutive signals are the same, the distribution of activities $\{X\}=\{X_1,X_2\}=\{x_1,y_1,x_2,y_2\}$ can be described as a mixed model, as

$$p(\{X\};\Sigma,k)=\int k(\xi)q(X_1;\xi,\Sigma)q(X_2;\xi,\Sigma)d\xi \quad (2)$$

where $k(\xi)$ denotes an unknown distribution of the signals. The only assumption made here is that the consecutive signals have (almost) equal value. That assumption is minimal for successful estimation and realistic as it is satisfied, e.g., when the signal drift is continuous, and preferably, sufficiently slow.

$p(\{X\};\Sigma,k)$ is a semiparametric model [7] because it has both a vector Σ and a function $k(\xi)$ as parameters. It is generally not easy to estimate parameters in semiparametric models because a function space is fundamentally infinite dimensional. Although the maximum likelihood method almost always works optimally for statistical distributions with finite number of parameters, it does not work for semiparametric models because it includes, practically speaking, an infinite number of parameters [9, 11]. The number of parameters $\xi^{(t)}$ (indexed by time t) increases proportionally with the number of observations $\{X^{(t)}\}$. Therefore, the conventional maximum likelihood methods for $\xi^{(t)}$ and Σ yield biased estimates for the semiparametric model, as shown later.

3 Estimators

The goal of these analyses is to estimate the three constant parameters $\Sigma=\{\Sigma_{11},\Sigma_{12}(=\Sigma_{21}),\Sigma_{22}\}$ whatever the signal drift or $k(\xi)$ is. It is obtainable by orthogonalizing statistical parameters [7, 8, 9, 10]. I explicitly derived optimal estimators of the covariance matrix Σ which work whatever $k(\xi)$ is:

$$\hat{\Sigma}_{11}=\frac{1}{N}\sum_{t=1}^{N}\{(x_1^{(t)}-\overline{x^{(t)}})^2+(x_2^{(t)}-\overline{x^{(t)}})^2\} \quad (3)$$

$$\hat{\Sigma}_{22}=\frac{1}{N}\sum_{t=1}^{N}\{(y_1^{(t)}-\overline{y^{(t)}})^2+(y_2^{(t)}-\overline{y^{(t)}})^2\}$$

$$\hat{\Sigma}_{12}=\frac{1}{N}\sum_{t=1}^{N}\{(x_1^{(t)}-\overline{x^{(t)}})(y_1^{(t)}-\overline{y^{(t)}})+(x_2^{(t)}-\overline{x^{(t)}})(y_2^{(t)}-\overline{y^{(t)}})\}$$

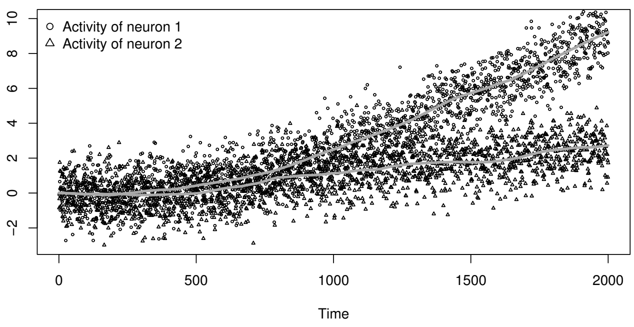

Figure 1: Simulated activities for two different neurons used in the estimation in Fig 2. The thick gray lines denote time-dependent signals ξ generated by the ARIMA(0,2,1) model.

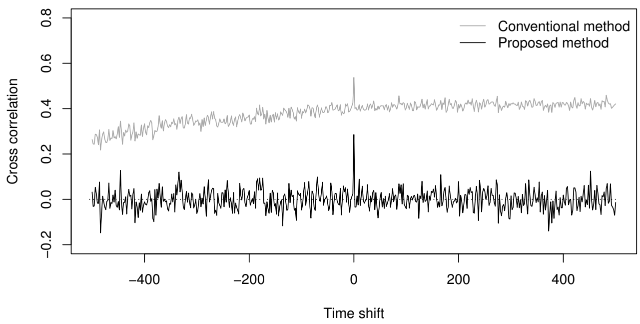

Figure 2: The cross correlation function estimated using the proposed method (black) and the conventional correlation coefficient (gray) for the time shifted data. $\Sigma_{11} = \Sigma_{22} = 1$ and $\Sigma_{12} = 0.3$ were used. A long but single time series was used ($N = 2000$).

Therein, $x_m^{(t)}$ and $y_m^{(t)}$ denote the t-th observation of neural activity. $m(= 1, 2)$ is an index to distinguish two activities for the same $\xi^{(t)}$. The local mean activities were defined by $\overline{x^{(t)}} = (x_1^{(t)} + x_2^{(t)})/2$ and $\overline{y^{(t)}} = (y_1^{(t)} + y_2^{(t)})/2$.

They are unbiased because they are normalized not by dividing by $2(= M)$ but by $1(= M-1)$. Normalization of this type is widely known to guarantee the unbiased nature for Gaussian distributions for fixed signals ξ. Notice that if you estimate $\xi^{(t)}$'s (signals at all time points) and Σ by using the maximum likelihood estimation, the estimators for the covariance matrix Σ are instead divided by $2(= M)$. Consequently, they are biased and always give the half value of those for the proposed method. This result might look surprising, but, the maximum likelihood estimation works only when the number of observation is much larger than the number of parameters [9, 11].

4 Numerical simulation

Here, using numerical simulations, it can be shown that the proposed method works fairly well even if signals ξ drift continuously (violating the assumption that the consecutive two ξ's are exactly the same). In the numerical simulations, first, $\xi_1^{(t)}$ and $\xi_2^{(t)}$ are independently generated by the ARIMA(0,2,1) model whose moving average coefficient is 0.6 [12]. Next, $x^{(t)}$ and $y^{(t)}$ are generated from the normal distribution whose means are the given signals $\{\xi_1^{(t)}, \xi_2^{(t)}\}$ for each time t.

The cross correlation functions for the realization of activities for two neurons in Fig. 1 was computed using the proposed methods and the conventional correlation coefficients, which assumes constant signals ξ. Here, $\rho = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}$ was estimated for the time-shifted data, where y was time-shifted while x was kept.

The proposed methods correctly caused 0 for the time shifted data and $\rho(= 0.3)$ for the simultaneous data as demonstrated by a clear peak (Fig. 2). However, the conventional covariance caused a broad cross correlation function attributable to the temporal correlations in ξ's. Note that broad cross correlation functions have been observed experimentally [6].

Consequently, the proposed method enables estimation of the correlations existing in the simultaneous data independently of the time-dependent signals.

References

[1] L. F. Abbott, P. Dayan: The effect of correlated variability on the accuracy of a population code. Neural Comput., 11 (1999), 91–101.

[2] H. Sompolinsky, H. Yoon, K. Kang, M. Shamir: Population coding in neuronal systems with correlated noise, Phys. Rev. E, 64 (2001), 051904.

[3] E. Zohary, M. N. Shadlen, W. T. Newsome: Correlated neuronal discharge rate and its implications for psychophysical performance, Nature, 370 (1994), 140–35.

[4] A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K. Logothetis, A. S. Tolias: Decorrelated neuronal firing in cortical microcircuits, Science, 327 (2010), 584–7.

[5] A. Renart, J. de la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, K. D. Harris: The asynchronous state in cortical circuits, Science, 327 (2010), 587–90.

[6] W. Bair, E. Zohary, W. T. Newsome: Correlated firing in macaque visual area MT: time scales and relationship to behavior, J Neurosci., 21 (2001), 1676–97.

[7] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, J. A. Wellner: Efficient and adaptive estimation for semiparametric models, Johns-Hopkins University Press, 1993.

[8] S. Amari, M. Kawanabe: Information geometry of estimating functions in semi-parametric statistical models, Bernoulli, 3 (1997), 29–54.

[9] K. Miura, M. Okada, S. Amari: Estimating spiking irregularities under changing environments, Neural Comput., 18 (2006), 2359–86.

[10] S. Amari, H. Nagaoka: Methods of information geometry, American Mathematical Society, 2001.

[11] J. Neyman, E. L. Scott: Consistent estimates based on partially consistent observations, Econometrica, 16 (1948), 1–32.

[12] P. J. Brockwell, R. A. Davis: Introduction to time series and forecasting, Springer-Verlag, 2002.