Needs title

Altyn Zhelambayeva[*], Hector G. Flores Rodriguez[†], Hernando Ombao[‡]

Abstract

TODO: Abstract

Keywords: Dimensionality reduction; PCA; Time series

[*]Department of Computer Science, School of Science and Technology, Nazarbayev University, Astana, Kazakhstan; altyn.zhelambayeva@nu.edu.kz

[†]Department of Computer Science, University of California, Irvine, CA 92697, USA; hfloresr@uci.edu

[‡]Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia; hernando.ombao@kaust.edu.sa

1 Introduction

TODO: Intro

2 Methods

Here we can talk about dimensionality reduction for LFP signals and the proposed methods of SPCA and GDPCA. Here we can introduce that we are comparing SPCA vs different lags of GDPCA.

2.1 Performance comparison test for two methods

Since the implementations of two methods differ significantly, there is no exact way of determining which method works better for local field potential signals. The process of SPCA and GDPCA comparison is limited to the identification of the strengths and weaknesses of the two techniques based on certain criterions. We propose several procedures for comparison of their performance when applied to the LFP time series data. First of all, we examine the signal summaries obtained by SPCA and GDPCA for the accuracy of the structural signal reconstruction. This criterion is very important for dimensionality reduction since the factor that structurally resemble the observed LFP signal may allow us to summarize the signals from multiple sources. For this purpose, the factors from two methods were superimposed with observed LFP signals for different number of lags for GDPCA. Secondly, we compare smoothed log periodograms with rectangular smooting window of 11. Periodogram shows the spectral density of the signal and allows the identification the dominant frequencies of a time series (**?**). We also use functional boxplots for finding the quantile boarder curves for log periodograms, comparing median periodograms and inspecting for outlying curves. Then, we compare the two methods based on the original signal reconstruction from principal components. The reconstruction of original signal for SPCA was obtained from linear regression of observed LFP signal against signal summary one, for GDPCA the procedure descibed in (GDPCA paper reference) was performed. This criterion is also very important because it allows us to evaluate the methods based on their ability to reconstruct the observed signal accurately. Finally, we present Mean Squared Errors for the two methods which is the average of the residuals obtained from signal reconstruction from principal components.

2.2 Spectral Principal Component Analysis

High dimensionality of observed LFP data complicates the process of signal interpretation. Many methods were developed to make the exploratory analysis of high dimensional data feasible. One of these methods is Principal Component Analysis (PCA) that was first introduced by Pearson in 1901 and is frequently referred to as conventional Principal Component Analysis. Although the rise and falls in neuronal membrane potentials occur very rapidly, when the sampling rate is high enough to capture the change in membrane potentials (i.e at millisecond scale), there might exist some phase shifts in time between different channels. Conventional PCA ignores the lead-lag structure of the time series by producing the factors which are instantaneous (contemporaneous) linear mixture of the signal. It also does not require much computational resources since only the covariance matrix of zero lag must be decomposed. However, because the temporal dynamics of time series is ignored, the original signal might be completely canceled out. From the time when conventional PCA was first described, many alternative versions that were found to have a better performance for particular types of data were proposed. One of these techniques for dimensionality reduction of time series data is Spectral Principal Component Analysis (SPCA) that was first introduced in (Brillinger). According to (Wang et al.), SPCA considers the lead-lag structure of the time series by using the linear filter of the time series. They also present simulation results where SPCA showed better performance over the conventional PCA when there was a time shift in some electrodes. As discussed in (SPCA paper), spectral principal component $\mathbf{f}(t)$ is a linear convolution of all time series $\mathbf{z}(t) \in \mathbb{R}^n$ and reconstructed time series $\hat{\mathbf{z}}(t)$ is a linear convolution of principal component $\mathbf{f}(t) \in \mathbb{R}^m$, where n and m ($m < n$) are the dimensions of original space of time series and lower dimensional space of principal component, respectively.

2.3 Generalized Dynamic Principal Components for Brain Signals

In comparison to SPCA, we also propose utilizing a generalized approach for obtaining signal summaries. The generalized dynamic principal components (GDPC) procedure, proposed by Peña and Yohai (2016) differs from SPCA as it does not assume any given model and it does not assume a fixed number of factors. In GDPC, there does not need to be a linear or stationary combination of the data, which may better capture nonstationary behavior of our rat brain signals. In this work, the factors are chosen such that the dynamic principal components are optimal in the reconstruction of the original series, where the reconstructed series is optimal with the mean square error (MSE) criterion.

- Description of GDPCA and reference to paper

- Factors of GDPCA with LFP superimposed

3 Results

A couple of sentences to introduce our results

3.1 Log Periodograms

- Compare log periodograms pre

- Compare log periodograms post

3.2 Functional Boxplots

- Introduce functional boxplots with references to papers

- Compare functional boxplots and medians pre

- Compare functional boxplots and medians post

- Compare functional medians pre

3.3 Signal Reconstruction and MSE

- Compare reconstruction and mse pre

- Compare reconstruction and mse post

4 Conclusion

TODO: Conclusion

References

Peña, D. and Yohai, V. J. (2016). Generalized dynamic principal components. *Journal of the American Statistical Association*, 111(515):1121–1131.