A Group Comparison in FMRI Data Using a Semiparametric Model Under Shape Invariance

Arunava Samaddar[1], **Cheolwoo Park**[1], **Nicole A. Lazar**[1], **Christopher J. Helms**[1], **Jennifer E. McDowell**[2] **and Brooke Jackson**[2]

Department of Statistics, University of Georgia[1]; Department of Psychology, University of Georgia[2]

Abstract

The objective of this work is to compare differences of groups of subjects in brain activation changes across two sessions associated with practice-related cognitive control in multiple regions of interest of the brain. In functional magnetic resonance imaging (fMRI) data analysis, it is often challenging to directly compare brain signals from different groups of subjects due to low signal-to-noise ratio in fMRI data. Using the property that brain signals in regions of interest may contain a similar pattern across subjects in a task-related experiment, we develop a semiparametric approach under shape invariance to quantify and test the differences in sessions and practice groups. We estimate the commonly shared function using local polynomial regression, and estimate shape invariant model parameters using evolutionary optimization methods. We apply hypothesis testing procedures on the scale parameter to determine whether the practice effect is present for different practice groups and whether a difference exists among groups before and after the practice in multiple regions of interest. It is shown that the brain signal shows attenuation at the post-practice session for the task-related practice group only and also shows the difference among the practice groups in some of the regions tested.

Introduction

Two main objectives in this work are to statistically test i) which practice groups and ROIs show attenuation at the post-practice session and ii) which ROIs show a difference among the three practice groups. We propose a semiparametric approach under shape invariance and accomplish both objectives in a unified framework. The shape invariance model assumes that individual regression curves contain a certain pattern and can be obtained from a common shape function by linear transformations of the axes (Kneip and Engel, 1995). We extend the work of Kneip and Engel (1995) in a single group setting to the comparison of multiple sessions and groups through the parameters in the model.

Data Introduction

The data set we explore contains BOLD activation signals from 64 subjects (31 Healthy and 33 with Schizophrenia).The tasks were presented in 5 blocks of antisaccade trials. After the initial run, the subjects were split into two practice groups, antisaccade (15 Healthy and 16 with Schizophrenia) and prosaccade (16 Healthy and 17 with Schizophrenia). Each practice group member practiced his/her assigned task once a day, for four days, excluding weekends. Once the four practice days were completed, all subjects repeated the antisaccade run in the scanner.

Methodology

Shape Invariance Model

Suppose that there are L groups and n time points. Let S_l represent the number of subjects for the l^{th} group ($L = 1, \cdots, L$). We assume that the bivariate data (t_{ijk}, Y_{ijk}) satisfy the following regression model within each ROI individually :

$$Y_{ijkl} = f_{ikl}\left(t_{ijkl}\right) + \epsilon_{ijkl}, \quad i = 1, \cdots, S_l; j = 1; \cdots, n, k = 1, 2$$

In our example, $k = 1$ denotes the pre-practice session and $k = 2$ the post-practice session, and Y_{ijkl} is the average value over the voxels in a given ROI for subject i, time point j, session k and group l. The ϵ_{ijkl}s are the error term with mean zero, and the design points can be simplified as $t_{ijkl} = t_j = j$ in this example.

Finally, the f_{ikl}s denote unknown smooth regression functions, and we assume that the functions contain a pattern shared by all subjects but their amplitudes might be different for different sessions and groups. We attempt to quantify this difference and test its statistical significance.

Pairwise Attenuation Test

The average signal in most of the ROIs for all groups at both sessions has roughly the same box-car pattern, following the stimulus presentation trail. In order to take advantage of this common feature, we propose to postulate a semiparametric shape invariance model (Kneip and Engel, 1995): for each group l,

$$f_{ikl}\left(\theta_{ik,3,l}t + \theta_{ik,2,l}\right) = \theta_{ik,1,l}\phi(t) + \theta_{ik,4,l}$$

where $\theta_{ik,1,l}, \theta_{ik,3,l} \in \mathbb{R}^2_+$ and $\theta_{ik,2,l}, \theta_{ik,4,l} \in \mathbb{R}^2$ are unknown true parameters for subject i and session k, and phi_l is an unknown common regression function shared across subjects and sessions for the lth group. In this work, we set $\theta_{ik,3,l} = 1$ and $\theta_{ik,4,l} = 0$.

The vertical scale parameter $\theta_{ik,1,l}$ is a key component of our statistical inference procedure for detecting attenuation and groupwise differences. Hence, we focus on the vertical scale parameter $\theta_{ik,1,l}$ and horizontal shift parameter $\theta_{ik,2,l}$.

In order to test the attenuation at post-practice session for the lth group, the following hypotheses are considered: $H_0 : \theta_{1,1,l} = \theta_{2,1,l} vs H_a : \theta 1, 1, l > \theta 2, 1, l$.

We use the pairwise t, sign, and Wilcoxon signed-rank tests, which cover both parametric and nonparametric tests, in our simulation and real data analysis.

Estimating Parameters

We estimate the horizontal shifting and vertical scale parameters.

1. Compute an initial estimate of $\hat{\phi}_l^{(0)}$of ϕ by

$$\hat{\phi}_l^0 = \frac{1}{2S_l}\sum_{k=1}^{2}\sum_{i=1}^{S_l}\hat{f}_{ik,l}\left(t + \hat{\theta}_{ik,2,l}^{(0)}\right)$$

2. Obtain $\hat{\theta}_{ik,1,l}^{(0)}$ and $\hat{\theta}_{ik,2,l}^{(0)}$ by solving

$$\min_{\theta_{ik,1,l},\theta_{ik,2,l}} \int \left[\hat{f}_{ik,l}(t + \theta_{ik,2,l}) - (\theta_{ik,1,l}\phi_{k,l}^{m-1}(t))\right]^2 dt$$

3. Update

$$\hat{\phi}_l^{(m)} = \frac{1}{S_l}\sum_{i=1}^{S_l}\hat{f}_{ik,l}(t + \hat{\phi}_{ik,2,l}^{(m)})$$

and iterate $m = 1, \cdots, M$ until the minimum is achieved in step 2. We have used GenSA and PSO to generate $\hat{\theta}_3^{;k}$ based on minimization of ψ and selected a final $\hat{\theta}_3^{;k}$ for sufficiently small change in ψ from one iteration to the next.

Simulation

We simulate a BOLD signal $x(t)$ by modeling the box-car pattern of on and off stimuli. For each subject, an amplitude a_r is used to control sigma amplitude, and time t is shifted by adding $d_{j,r}\, U(-2, 2)$:

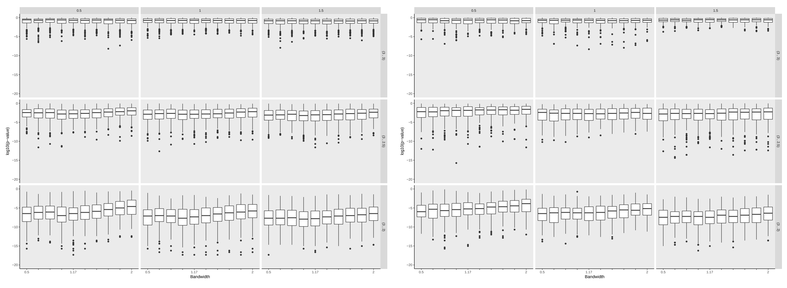
$$x_r(t_j) = a_r\frac{4}{\pi}\sum_{k=1}^{10^3}\frac{sin(2\pi(2k-1)0.05(t_j + d_{j,r}))}{2k-1}.$$

We add noise from $N(0, \sigma^2_{noisei,r})$. We generate the data as described above with twelve subjects at pre- and post-practice sessions using the five sets of amplitudes. We simulate data for pre and post practice sessions using amplitudes:

Table 1: Simulation parameters

pre	3	3	3
post	3	2.5	2

Table 2: Pairwise wilcoxon test results (GenSA & PSO)



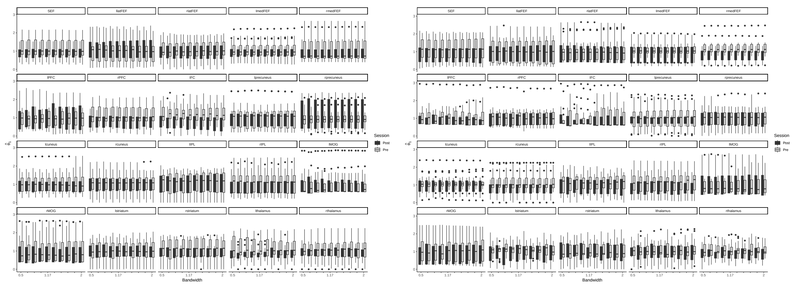
(a) GenSA (b) PSO

Each row represents a different amplitude set and each column represents a different SNR. Looking from left to right, we observe a separation with the increase of SNR. The choice of bandwidth does not appear to affect the results much. This result confirms that the statistical significance in attenuation becomes stronger as the difference in the amplitudes between pre- and post-sessions increases. We also observe a smaller p-value as the SNR increases as expected.

Read Data Analysis

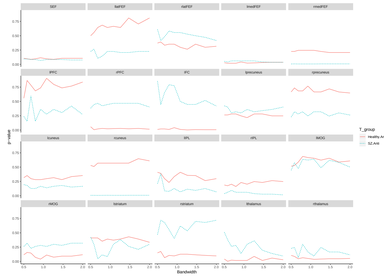
In this subsection, we analyze the data described earlier using the proposed estimation and test procedures. In what follows we show boxplots of estimates θ_1 values for each ROI and the corresponding p-values.

Table 3: Boxplots of estimated vertical scale parameters for each ROI (GenSA)



(a) Healthy (b) SZ

Table 4: P-values of paired t-tests for each ROI (GenSA and PSO)



P-values for pairwise attenuation tests

It can be seen that the healthy antisaccade group tends to yield lower p-values (stronger evidence of attenuation) than the other group particularly in the IFC, PFC-r, thalamus and medFEF regions under the GenSA estimation for a large range of bandwidths.

Conclusion

We propose the semiparametric approach under shape invariance model to test attenuation between two sessions in fMRI data. We demonstrate through simulated and real examples that the proposed approach detects attenuation between two sessions and groupwise differences. The pairwise attenuation tests show that there is convincing evidence of significant attenuation at the post-practice for the antisaccade practice group in the SEF, PFC-R, FC-l and thalamus regions.

This work was supported by National Science Foundation under Grant No. NSF IIS-1607919