NeuroImage 103 (2014) 130–138

Contents lists available at ScienceDirect

NeuroImage

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

The impact of study design on pattern estimation for single-trial multivariate pattern analysis

Jeanette A. Mumford [a,f,*], Tyler Davis [b], Russell A. Poldrack [c,d,e]

[a] Waisman Laboratory for Brain Imaging and Behavior, WI, USA
[b] Department of Psychology, Texas Tech University, Lubbock, TX, USA
[c] Department of Psychology, University of Texas at Austin, USA
[d] Department of Neuroscience, University of Texas at Austin, USA
[e] Imaging Research Center, University of Texas at Austin, USA
[f] Center for Investigating Healthy Minds at the Waisman Center, University of Wisconsin, Madison, WI, USA

ARTICLE INFO

Article history:
Accepted 11 September 2014
Available online 19 September 2014

Keywords:
fMRI
MVPA
Pattern similarity
Pattern classification
False positive rate

ABSTRACT

A prerequisite for a pattern analysis using functional magnetic resonance imaging (fMRI) data is estimating the patterns from time series data, which then are input into the pattern analysis. Here we focus on how the combination of study design (order and spacing of trials) with pattern estimator impacts the Type I error rate of the subsequent pattern analysis. When Type I errors are inflated, the results are no longer valid, so this work serves as a guide for designing and analyzing MVPA studies with controlled false positive rates. The MVPA strategies examined are pattern classification and similarity, utilizing single trial activation patterns from the same functional run. Primarily focusing on the Least Squares Single and Least Square All pattern estimators, we show that collinearities in the models, along with temporal autocorrelation, can cause false positive correlations between activation pattern estimates that adversely impact the false positive rates of pattern similarity and classification analyses. It may seem intuitive that increasing the interstimulus interval (ISI) would alleviate this issue, but remaining weak correlations between activation patterns persist and have a strong influence in pattern similarity analyses. Pattern similarity analyses using only activation patterns estimated from the same functional run of data are susceptible to inflated false positives unless trials are randomly ordered, with a different randomization for each subject. In other cases, where there is any structure to trial order, valid pattern similarity analysis results can only be obtained if similarity computations are restricted to pairs of activation patterns from independent runs. Likewise, for pattern classification, false positives are minimized when the testing and training sets in cross validation do not contain patterns estimated from the same run.

Introduction

Traditional data analysis approaches in functional magnetic resonance imaging (fMRI) often employ voxel-wise models to identify where in the brain aggregate activation differs between experimental conditions. A more recently developed set of analysis strategies, multivoxel pattern analysis (MVPA), often starts with similar voxel-wise activation estimates, but instead of testing for differences in aggregate activation, focuses on the information contained in the distributed patterns of activation across voxels (Kriegeskorte et al., 2008a; Kriegeskorte, 2011; Haxby et al., 2001; Carlson et al., 2003; Pereira et al., 2009; Norman et al., 2006; Davis and Poldrack, 2013; Haynes and Rees, 2006). Multivariate pattern classification and pattern similarity analyses are two of the most common MVPA strategies. Pattern classifiers test whether an activation pattern can be used to decode the mental state of the subject (Haynes and Rees, 2006; Norman et al., 2006). In pattern similarity analyses, the goal is often not to simply decode mental states, but to examine the geometric relationships between activation patterns for different conditions and stimuli in a task. To this end, pattern similarity analysis involves computing a similarity metric between pairwise activation patterns elicited for different conditions or stimuli, and testing how these pattern similarities relate to psychological states, predictions from cognitive models, or patterns elicited for the same stimuli in non-human primates (Kriegeskorte et al., 2008a; Kriegeskorte, 2011; Kriegeskorte et al., 2008b; Kriegeskorte and Kievit, 2013; Davis and Poldrack, 2013). Despite the increasing popularity of MVPA approaches to fMRI analysis, there have been few systematic studies of how study design impacts results from MVPA.

The present work examines how study design affects estimation of the activation patterns that serve as inputs into MVPA and subsequent Type I error rates. We focus on experimental contexts in which the goal is to accurately estimate separate activation patterns from single

* Corresponding author at: Center for Investigating Healthy Minds at the Waisman Center, University of Wisconsin-Madison, 1500 Highland Ave, Madison, WI 53705.
E-mail address: jeanette.mumford@gmail.com (J.A. Mumford).

http://dx.doi.org/10.1016/j.neuroimage.2014.09.026
1053-8119/

trials within the same run. For example, if 30 exemplars of each of 2 types of stimuli are presented to a subject in a single functional run, the goal is to estimate 60 separate activation patterns that are then input data for a pattern analysis that will attempt to classify or explain the similarity relationships between these exemplars. Single run analyses are common within the pattern similarity framework and have the advantage of saving time and money while collecting data. Although previous work assessed power of single trial parameter estimators within the pattern classification framework in a between-run setting (Turner et al., 2012; Mumford et al., 2012), control of Type I error is more critical since, when not controlled, the resulting statistics are invalid.[1] In the case of pattern similarity, Type I error rates will be quantified for analyses that compare similarity distributions for different pairings of trials from the same run. For pattern classification, Type I error is assessed by testing whether or not the classification accuracy of data generated under the null hypothesis is at chance when the cross validation is performed using trials from the same run.

The primary pattern estimators we examine are the Least Squares All (LSA) and Least Squares Single (LSS) models (Turner et al., 2012; Mumford et al., 2012). Both of these models estimate patterns using a voxelwise general linear model; example design matrices for a single run that presented 5 exemplars each of 2 stimulus types are illustrated in Fig. 1. In the case of LSA, all trials are estimated simultaneously in a single model, using a separate regressor consisting of an impulse (or boxcar) function convolved with a double gamma hemodynamic response function (HRF). This is often referred to as beta-series regression (Rissman et al., 2004) and the parameters, β_1, …, β_{10}, estimate the activation magnitude for each of the 10 trials within a single voxel. These estimates are then aggregated over many voxels to comprise the activation pattern that serves as the input for MVPA. A pitfall of LSA is when trials have a short interstimulus interval (ISI), e.g., less than 3 s between the end of one stimulus and onset of the next stimulus, the regressors become highly correlated, or collinear, which inflates the variance of the resulting parameter estimates. The LSS model reduces this collinearity by using a separate model for each trial, in which the first regressor models the trial of interest and the other two regressors model the remaining trials according to trial type. For example, assuming that the exemplars were images of mammals or reptiles, the first iteration of LSS is modeling the first trial, a mammal, as the first regressor while the other two regressors model the remaining mammals and remaining reptiles, respectively. In this case, only the first parameter estimate is retained in each model and estimates the activation for that individual trial. Previously, LSS has been shown to produce higher classification accuracies than LSA for short ISIs (3–5 s) (Mumford et al., 2012). Although we focus on the LSS and LSA pattern estimators, a third model that simply takes the time point 6 s after stimulus presentation as the pattern estimate (Add6) is also considered.

We will illustrate how temporal autocorrelation and pattern estimation technique, through correlations between regressors, introduce false positive correlations between the activation patterns estimated from the same functional run of BOLD data. These false correlations can then lead to false positive comparisons in pattern similarity distributions or inflated classification accuracies. Surprisingly, even with ISIs as long as 15 s, elevated false positive rates occur in pattern similarity analyses, regardless of which of the three pattern estimators (LSS, LSA, Add6) are used, unless trials are randomly ordered with a unique randomization for each subject. Similar issues arise if the cross validation of a pattern analysis uses only trials from the same functional run. Hence, within-run pattern similarity and classification analyses are not recommended.

After characterizing the pitfalls of single run, or within-run analyses, we examine whether or not between-run analyses offer reasonable solutions. Between-run analyses require multiple functional runs of BOLD data and, in the pattern similarity setting, similarities are only computed between activation patterns that were estimated from independent runs of data. For example, if there are 60 trials per run and two runs, the pattern from the first trial of run 1 would be correlated with all other patterns from run 2, only, in a between-run similarity analysis. For classification, the test and training data sets in the cross validation would comprise patterns from different runs.

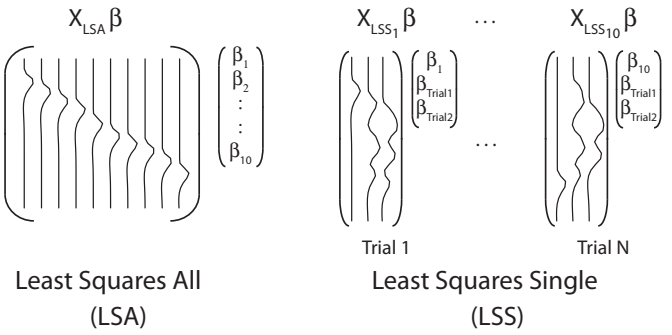

Fig. 1. Model illustration for LSA and LSS. In both cases, trial-specific activations are estimated for each of 10 trials and the model is run in a voxel-wise fashion. The left panel shows Least Squares All (LSA), which estimates all trials simultaneously in a single regression and the estimates β_1, …, β_{10} represent the activation magnitudes for each of the trials. The right panel shows Least Squares Single (LSS) where each trial's activation is estimated in a separate model where the first regressor represents the trial of interest and the two additional regressors model the remaining trials according to trial type. In this case there are 2 trial types. Only the estimates for the first parameter are retained from each model.

The following section contains a theoretical derivation of the LSS- and LSA-based pattern estimates to clearly motivate the impact of study design and temporal autocorrelation on estimated patterns. Both the Methods and Results sections start with pattern similarity analyses and end with classification analyses. Within each analysis setting, within-run approaches are studied first, followed by between-run approaches. This work can be used as a guideline when designing and implementing future MVPA analyses that will yield valid results.

Derivations

The following derivations were used to generate data for the simulation studies and help in understanding the source of invalid statistics for within-run pattern analyses. Before deriving the distributions for the LSS- and LSA-based pattern estimates, we first characterize the distribution of the BOLD time series, Y, within a single voxel. The data follow a multilevel structure where one level (Eq. (2)) describes the trial-specific activations, β, and the other level (Eq. (1)) describes how these trial-specific activations are related to the BOLD time series. Specifically,

$$Y = X_{LSA}\beta + \epsilon_Y, \quad \epsilon_Y \sim N(0, V_Y) \tag{1}$$

$$\beta = \mu + \epsilon_\beta, \quad \epsilon_\beta \sim N\left(0, V_\beta\right). \tag{2}$$

Assume that there are N_{trials} total trials presented, β is a vector of length N_{trials} with a true mean of μ, also a vector of length N_{trials}, and a covariance following the $N_{trials} \times N_{trials}$ matrix, V_β. Since V_β is the true covariance between the trials, it represents the true representational similarity covariance matrix, from which the pattern similarity correlations can be derived. The vector Y is the voxel-wise time series with length N_{tpts} and assumes a mean equal to the product of the trial-specific activation magnitudes, β, and the trial-specific regressors following the LSA design matrix (Fig. 1, left panel). The LSA design matrix consists of a single regressor per trial, where the regressor is constructed by convolving a delta or boxcar function, depending on

[1] In these previous studies, Type I error was assessed and found to be controlled, but this was not reported.