Multivariate analysis of fMRI data by oriented partial least squares

W. S. Rayens[1], A. H. Andersen[2,3]

[1]Statistics, University of Kentucky, Lexington, Kentucky, United States, [2]Magnetic Resonance Imaging and Spectroscopy Center, University of Kentucky, Lexington, Kentucky, United States, [3]Anatomy & Neurobiology, University of Kentucky, Lexington, Kentucky, United States

Abstract

Partial least squares (PLS) has been used in multivariate analysis of fMRI data as a way of incorporating information about the underlying experimental paradigm. In comparison, principal component analysis (PCA) extracts structure merely by summarizing variance and with no assurance that the individual component structures are directly interpretable, nor that they represent salient and useful features. Oriented partial least squares (OrPLS) is a new PLS-like analysis paradigm in which the extracted components can be oriented away from undesirable noise or confounds in the data as well as toward desired targeted structure reflecting the fMRI experiment.

Introduction

Multivariate structure-seeking techniques such as principal component analysis (PCA) are increasingly finding use in the processing of functional imaging data from both PET and MRI experiments [1,2]. These approaches are data-led. Compared to more traditional model-driven analysis techniques, one major advantage of PCA is that it makes no prior assumptions about the partitioning of the data according to some known exogenous experimental design contrast or a predefined set of explanatory basis functions. Principal component analysis seeks to reduce the dimensionality of a multivariate data set by uncovering latent structure to maximally summarize variability. Subsequently, the original spatio-temporal data may be transformed to a (reduced) subspace spanned by a few of the principal eigenvectors of the sample covariance matrix. Due to the inherent orthogonality of the initially computed PC eigenvectors, the interpretation of response component structures that overlap in time may become obscured. Since PCA is not informed by the underlying experimental paradigm, it is quite possible that the response component of primary interest is not the one that summarizes the most variability. In fact, it may be that the desired component is an admixture of several of the raw principal components. Therefore, further rotation through linear transformation of the data can be used to obtain simpler, physiologically more meaningful, interpretations of the components. Toward this end, multivariate analysis by partial least squares (PLS) has been used as an informed, target-directed structure-seeking technique to balance a priori information about stimulus conditions from the underlying experimental paradigm as used in fixed-effect correlation or regression methods and the variance summaries that characterize PCA [3]. The effects of structured noise confounds may be incorporated as constraints in either PCA or PLS, effectively removing a model of the confound (e.g., sinusoids) from the data up front. In PCA, an alternative approach has been to characterize the undesirable confound effects by their correlation matrix. This leads to a generalized eigenvalue problem, where the principal coponents are oriented or steered away from the distribution of confounds, thus maximizing signal to noise. Here we present a new multivariate technique referred to as oriented partial least squares (OrPLS) wherein the structure extracted from the feature space is oriented toward the anticipated response/reference profile implied by the experimental paradigm, while at the same time orienting away from undesirable confounds.

Methods

fMRI was used to map brain activation associated with a bilateral finger tapping task. fMRI data were collected from 24 axial, 3 mm thick slices with an in-plane resolution of 3.438mm × 3.438mm using a gradient echo EPI sequence (TR/TE = 2500/45 ms, FA 90º). 3D EPI volumes were motion corrected using SPM. The experimental paradigm followed a blocked design with alternating periods of finger tapping and rest. A total of 120 image time frames were collected. Initial principal component analysis based on the response in all brain voxels combined with Fourier spectral analysis revealed that most of the variability in the data as a whole was associated with a persistent periodic component centered around 0.13 Hz and unrelated to the task frequency at 0.04 Hz. For the purpose of multivariate analysis and to illustrate the oriented techniques of balancing the influence of signal and noise, we considered a subset of 4002 voxels wherein the response contained both task-related and structured noise confound components. The covariance matrix of the noise confound was estimated from the data matrix $\mathbf{X}$ itself. A subset $\mathbf{X}_1$ containing the measured response in 2754 of the overall 4002 voxels was identified. These voxels were selected on the basis of having small values of the correlation of their response with the box-car shape of the anticipated target profile $\mathbf{y}$ defined by the experimental motor paradigm.

Results

The time-course response profile was estimated using established methods of PCA, oriented PCA, and partial least squares (PLS) as well as with the new method of oriented partial least squares (OrPLS). In all approaches, the estimated response $\mathbf{a}$ was chosen as the first component among a ranked set. The response profiles are shown in Fig. 1, and quantitative global measures of performance are listed in Table 1. As expected, the first PC summarizes the most variability but has low correlation with the box-car target. Oriented PCA yields the highest ratio of signal to noise at the expense of variance summarized. Since PLS incorporates information about the expected box-car response, this technique achieves the highest correlation with the target. A strong resemblance with the box-car reference profile is seen in Fig. 1. Oriented PLS shows consistently good performance on all measures and yields the highest value of the contrast-to-noise ratio for the motor activation paradigm.

Discussion

These results demonstrate the potential of multivariate analysis of fMRI data by oriented partial least squares (OrPLS). This new method combines the target directed aspects of PLS with the abilities of OPCA to steer away from an undesired distribution of noise confounds and achieves superior contrast-to-noise performance. The benefits of OrPLS are likely to be more pronounced in designs with trials of shorter duration where there is greater correlation between physiological confounds and the effects of primary interest.

References

1. Strother, SC, et al., J. Cereb. Blood Flow Metab., 15, 353-360, 1995.
2. Friston, KJ, et al., J. Cereb. Blood Flow Metab., 13, 5-14, 1993.
3. Rayens, WS, et al., Chemometrics, 71, 121-127, 2004.

Table 1: Quantitative performance measures for multivariate analysis methods

	PCA	OPCA	PLS	ORPLS
Variance Summarized	2.95%	2.18%	2.78%	2.61%
Angle with Target	44.53º	59.17º	16.70º	25.44º
Correlation with Target $r = \mathbf{a}^t\mathbf{y}$	0.7128	0.5125	0.9578	0.9030
Data Variance $S_{X,a} = \mathbf{a}^t\mathbf{X}\mathbf{X}^t\mathbf{a}$	4.159	3.066	3.923	3.679
Noise Variance $S_{X1,a} = \mathbf{a}^t\mathbf{X}_1\mathbf{X}_1^t\mathbf{a}$	1.932	1.048	1.825	1.485
Signal-to-Noise Ratio $S_{X,a}/S_{X1,a}$	2.153	2.926	2.150	2.477
Contrast-to-Noise Ratio $r^2 S_{X,a}/S_{X1,a}$	1.094	0.769	1.972	2.020

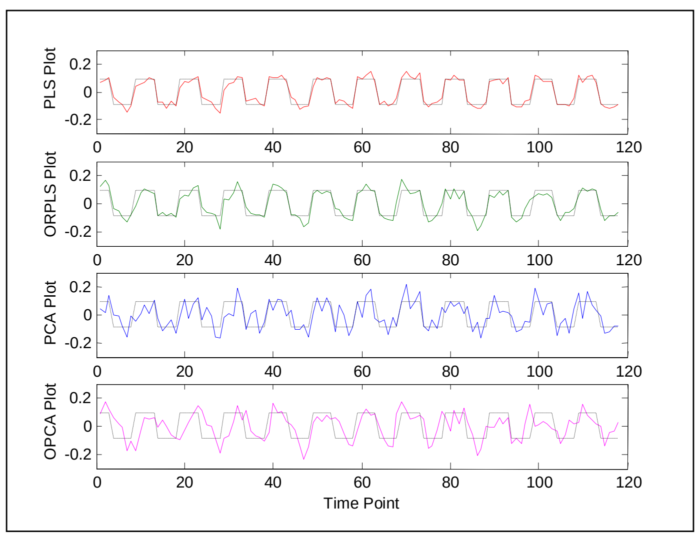

Figure 1: Illustrating the response profiles from multivariate analyses