Extending the General Linear Model for analyzing fMRI data to a constrained Multivariate Regression Model

Rajesh R Nandy
Department of Biostatistics, University of North Texas Health Science Center, Fort Worth, USA – Rajesh.Nandy@unthsc.edu

Dietmar Cordes
Cleveland Clinic Lou Ruvo Center for Brain Health, Las Vegas, USA – dietmar.cordes@ryerson.ca

Abstract

Local canonical correlation analysis (CCA) is a multivariate method that simultaneously analyzes the timecourses of a group of neighboring voxels and has been demonstrated to be more sensitive than the conventional univariate general linear model approach. However, unlike the general linear model, an arbitrary linear contrast of the temporal regressors has not been so far incorporated in the CCA formalism. To address the first problem, a multivariate regression model is presented which is a direct extension of univariate general linear model. Mathematically, multivariate regression model (MRM) is equivalent to CCA, but easier to interpret since the framework is similar to general linear model. Arbitrary contrasts can be used in the MRM approach including multivariate contrasts. With multivariate contrasts, it is also possible to test for significance of contrasts on regression coefficients as well as contrasts on voxels. The test for contrasts on voxels is not possible in the univariate framework. Furthermore, a constrained version of MRM is introduced which not only has more sensitivity than univariate general linear model, but also corrects for the potential loss of specificity due to over-fitting in the multivariate model. Quantitative results from simulated and pseudo-real data as well as qualitative results from real data show that constrained MRM can detect activations more accurately for noisy functional MRI (fMRI) data without losing specificity.

Keywords: fMRI data analysis; multivariate regression; constrained canonical correlation analysis; CCA.

1. Introduction

The general linear model is a widely-used functional MRI (fMRI) data analysis method to determine activation in the brain because of its simplicity in both estimation and inference and its greater sensitivity to regional effects than global multivariate analyses [Worsley & Friston 1995]. However, in this classical approach to the analysis of fMRI time-series, each voxel is analyzed as an independent time series. This approach essentially ignores the fact that the signal in one voxel may show statistical dependencies on the signal of a neighboring voxel. Typically, in the univariate analysis, some sort of spatial smoothing (usually Gaussian) is performed to gain sensitivity of the analysis. However, such spatial filtering is uniform over all regions of the brain and a poor smoothing can actually make the analysis less sensitive. Instead, by simultaneously considering the timecourses of a collection of neighboring voxels, complicated spatio-temporal evoked responses can be assessed with a potential increase in sensitivity. It is important to note that the form of the evoked response can differ over a local neighborhood of voxels. This is easily accommodated within a multivariate framework using canonical correlation analysis (CCA), a multivariate extension of ordinary single voxel correlation analysis [Friman et. al. 2001, Nandy & Cordes 2003a]. It has been demonstrated that CCA can provide significant improvement in sensitivity, particularly for fMRI signal with low contrast-to-noise ratio (CNR) [Nandy & Cordes 2003a, Nandy & Cordes 2003b]. CCA can be viewed as an improved form of ordinary spatial filtering where the spatial filter adapts itself depending on the signal from the neighboring voxels.

However, there remain a few issues that have prevented CCA from gaining widespread popularity it deserves. The first issue is the limitation of CCA (as presented so far) to work with specific contrasts of interest. With traditional univariate methods, testing for any contrast of interest is trivial after the univariate analysis with a full set of regressors is performed. There is no need to run separate analyses for different contrasts. Such an approach has so far not been presented within a CCA framework. The second issue, which is equally critical, is spatial specificity. Since the signal from a group of neighboring voxels is simultaneously analyzed, it is tricky to assign any detected activation to any particular voxel. This issue has been at least partially addressed by a constrained CCA approach [Friman et. al. 2003] and an adaptive CCA approach [Nandy & Cordes 2004]. Finally, no formal group analysis method has been presented for CCA and CCA has been primarily restricted to single subject analysis, which is too restrictive.

To address the first two issues related to CCA, we will first present a mathematically equivalent form of CCA, known as multivariate regression model (MRM). This new framework is advantageous for several reasons. It preserves the sensitivity of CCA, but at the same time can also address much better the above mentioned issues with CCA. Since multivariate regression is a direct extension of regression, the results from multivariate regression can be interpreted the same way as in conventional analysis with one added benefit. With this approach, we will not only be able to perform tests with contrasts similar to those in univariate analysis, we will also be able to test multivariate contrasts. In univariate analysis, we can test for significance of contrasts only on the regression coefficients, but not on a set of neighboring voxels. Multivariate contrasts not only allow us to test for significance of contrasts on regression coefficients, but also on a set of neighboring voxels. In this article, we will primarily focus on two aspects. First, we will present appropriate tests that can be used effectively with univariate or multivariate contrasts. This approach, which we refer to as the full model, solves the contrast problem but fails to address the specificity problem.

To address the specificity problem, we will present a constrained version of multivariate regression related to the constrained CCA problem. However the solution to the constrained CCA problem is obtained by imposing the constraint only to the first pair of canonical variates which leads to maximum canonical correlation. This is fine from CCA perspective, since the first pair of canonical variates is indeed the most important and relevant pair of canonical variates. However, the equivalence of CCA and multiple regression depend on all pairs of canonical variates and hence, for the constrained CCA problem, the equivalence of CCA and multiple regression for the full model cannot be directly applied. We will present two different solutions to this specificity problem. The first solution simply uses the first pair of the canonical variates from the solution of the constrained CCA problem and the other pair of canonical variates from the unconstrained CCA solution. The second method involves a reparameterization of the data depending on the contrast of interest on the set of predictors. The advantage of the first method over the second is that just like the unconstrained problem, for the first method, the contrasts can be evaluated as in the unconstrained problem. For the second method, the dataset needs to be reparametrized based on the contrast before the analysis. On the other hand, the second method does not use higher order canonical variates from the unconstrained CCA problem and is more accurate. We will compare the results from both of these methods.

This framework can be extended to the analysis of group fMRI data. However, group analysis using multivariate regression is a topic in its own right and will be presented in a separate article.

2. Theory and methods

Multivariate regression model

The multivariate General Linear Model can be expressed in the following form $\boldsymbol{Y} = \boldsymbol{XB} + \boldsymbol{E}$, where $\boldsymbol{Y}$ $n \times d$ observation matrix, $\boldsymbol{X}$ is $n \times p$ design matrix of rank p, $\boldsymbol{B}$ is $p \times d$ parameter matrix and $\boldsymbol{E}$ is $n \times d$ residual error matrix. E_i, the i-th column of the residual matrix denotes the residual vector for the i-th multivariate observation Y_i, and is assumed to follow multivariate normal distribution with mean 0 and covariance matrix $\boldsymbol{\Sigma}$. E_i and E_j are assumed to be independent for $i \neq j$.

The above representation is analogous to the general linear model in the conventional model for fMRI data analysis in a univariate framework. The only difference in the multivariate framework is that the d columns of the matrix $\boldsymbol{Y}$ denote the timecourses of d neighboring voxels (typically a 3×3 neighborhood in a 2 dimensional slice or a 3×3×3 neighborhood in 3 dimension) instead of the timecourse of a single voxel. Of course for d=1, the model reduces to univariate general linear model. The design matrix $\boldsymbol{X}$ is identical to the design matrix in the univariate model.

Estimation of parameters and hypothesis testing for the full model

The ordinary least squares solution of the multivariate regression model can be expressed as $\widehat{\boldsymbol{B}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$. As mentioned previously, more general forms of contrasts can be tested in the multivariate regression model compared to univariate general linear model. Formally, we can test the following null hypothesis for multivariate contrasts in the most general form:

$\boldsymbol{H_0}: \boldsymbol{C}'\boldsymbol{B}\boldsymbol{A} = \boldsymbol{D}$, where,

$\boldsymbol{C}$ is q×p hypothesis matrix of specified multivariate contrasts
$\boldsymbol{A}$ is d×f transformation matrix of rank f
$\boldsymbol{D}$ is q×f matrix of hypothesized values for the multivariate contrasts.

One can easily see that the matrix $\boldsymbol{C}$, which acts on the rows of the design matrix, is no different from the contrast matrix used in general linear model for fMRI. However, the transformation matrix $\boldsymbol{A}$ is unique to multivariate regression in the sense that it acts on the columns of the design matrix and in univariate general linear model, we do not have multiple columns of contrasts. The transformation matrix $\boldsymbol{A}$ can be interpreted as contrasts on the neighboring voxels which provides us a powerful tool that is simply not available in the univariate framework.

We will now describe the test procedure for the null hypothesis specified above. It can be shown [Timm 1975] that the test can be performed by using functions of eigenvalues of the matrix $\boldsymbol{S_e^*}^{-1}\boldsymbol{S_h}$, where $\boldsymbol{S_h} = (\boldsymbol{C}'\widehat{\boldsymbol{B}}\boldsymbol{A} - \boldsymbol{D})'(\boldsymbol{C}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C})^{-1}(\boldsymbol{C}'\widehat{\boldsymbol{B}}\boldsymbol{A} - \boldsymbol{D})$ is the hypothesis sum of squares and cross-product (SSCP) matrix, $\boldsymbol{S_e} = \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}}\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}}\right)$ is the error SSCP matrix and $\boldsymbol{S_e^*} = \boldsymbol{A}'\boldsymbol{S_e}\boldsymbol{A}$.

The omnibus null hypothesis can then be tested using the Wilks' Lambda statistic $\Lambda(\mathrm{d},\mathrm{q},\mathrm{n}-\mathrm{p}) = \prod_{j=1}^{k}\frac{1}{1+\lambda_j}$, where $k = \min(f,q)$ and λ_j -s are the eigenvalues of $\boldsymbol{S_e^*}^{-1}\boldsymbol{S_h}$.

Equivalence of multivariate regression model and CCA in the full model

So far we have shown that the multivariate regression model is a natural and more general extension of the popular univariate general linear model. However, this generalization also poses a problem. Since we are performing simultaneous regression on a group of voxels, it is not obvious how to assign the location of any detected activation. If the activation is attributed to the group as a whole, we will be losing spatial specificity. In the full model, one would assign the activation to the center voxel, but just as in CCA, it will lead to bleeding of activation which also hurts the spatial specificity. This problem has been addressed in the context of CCA. Since CCA and multivariate regression are mathematically equivalent, we will use the solution for multivariate regression that is equivalent to the solution of the constrained CCA and then apply the contrast, if needed. Due to space limitations, the mathematical derivation of the equivalence is not presented here.

Constrained CCA

As mentioned in the previous section, both CCA and multivariate regression share a spatial specificity problem that may lead to the bleeding artifact. In CCA, the problem can be solved by applying suitable constraints on the canonical variates for the dependent set of variables. For multivariate regression, we only have regression coefficients for the predictors and no such coefficient for the dependent variables. However, due to the equivalence of CCA and multivariate regression, we can first find solution to the constrained CCA problem and then consider the equivalent solution for the multivariate regression problem. Of course, the theoretical distribution of the test statistic for the constrained solution is intractable. Instead, we will use nonparametric methods for hypothesis testing (to be described in the next section). In this section, we will describe the constrained CCA problem and its solution. We will specifically present two natural constraints and examine their relative performances over univariate approaches, but the method can easily be generalized. For the response functions, we use the standard hemodynamic response function (HRF), which is a difference of two Gamma functions. Hence, we impose non-negativity constraint on the coefficients of the response functions. In general, one may or may not impose constraints on the response functions depending on the chosen set of basis functions.

The first constraint is similar in spirit with smoothing. We first impose the restriction that the coefficient for the voxel of interest to have the maximum weight. Second, all voxels that are equidistant from the voxel of interest carry equal weight. Finally, we impose the restriction that the weight of the coefficient for a voxel decreases with increased distance. We refer to this constraint as the symmetric constraint. Although, Gaussian smoothing satisfy all these requirements, what sets CCA apart is the fact that constrained CCA will actually search for the optimal coefficients unlike Gaussian smoothing where coefficients are fixed. To illustrate the constraint, in Figure 1, we have shown a typical 5×5 neighborhood specifying the coefficients for each voxel in the neighborhood. Due to the symmetric nature of the constraint, we refer to this constraint as the symmetric constraint. Mathematically, the constraints on the coefficients can now be expressed as

$$0 \le \alpha_{22} = \alpha_{23} = \alpha_{24} = \alpha_{25} \le \alpha_{14} = \alpha_{15} = \alpha_{16} = \alpha_{17} = \alpha_{18} = \alpha_{19} = \alpha_{20} = \alpha_{21}$$
$$\le \alpha_{10} = \alpha_{11} = \alpha_{12} = \alpha_{13} \le \alpha_{6} = \alpha_{7} = \alpha_{8} = \alpha_{9} \le \alpha_{2} = \alpha_{3} = \alpha_{4} = \alpha_{5} \le \alpha_{1}.$$

α_{23}	α_{17}	α_{11}	α_{16}	α_{22}
α_{18}	α_{7}	α_{3}	α_{6}	α_{15}
α_{12}	α_{4}	α_{1}	α_{2}	α_{10}
α_{19}	α_{8}	α_{5}	α_{9}	α_{14}
α_{24}	α_{20}	α_{13}	α_{21}	α_{25}

Figure 1

The second contrast that we present is for a 3×3 neighborhood and has the property that the center voxel is not only the dominant one, but the coefficient is at least as large as the sum of remaining coefficients. Hence the contrast is referred to as constraint for strong dominance. If the center voxel is labeled as voxel one, mathematically, the constraint can be written as

$$0 \le \alpha_1, \alpha_2, \ldots\ldots\ldots, \alpha_9 \text{ and } \sum_{i=2}^{9} \alpha_i \le \alpha_1 .$$

<u>Distribution of the maximum canonical correlation under constraint</u>

The theoretical distribution of the maximum canonical correlation under constraints is intractable. Hence it is necessary to obtain the distribution empirically using null data. Since resting-state data are not related to the paradigm used in activation data, resting-state data may be considered to be null relative to the paradigm of interest. Hence if the same procedure is implemented on resting state data as in activation data for the same subject, the distribution of the maximum canonical correlation under constraint can be treated as the empirical null distribution. To address the multiple comparison problem, we will be using a semi-parametric method previously developed by us [Nandy & Cordes 2007].

<u>Modified ROC curves</u>

In the results section, we will compare the relative performances of five different analysis methods - single voxel analysis with no smoothing, single voxel analysis with Gaussian smoothing, conventional CCA without correcting for bleeding, constrained CCA with symmetric constraint and constrained CCA with strong dominance constraint. We will use the ROC method, a valuable tool for testing the efficiency of various data analysis methods, which is usually used based on simulated data where the ground truth is known. We have shown in a previous article that the ROC method using simulated data can be modified and applied to real data [Nandy and Cordes, 2003b, Nandy and Cordes, 2004b] yielding so-called “modified ROC curves”.

Imaging

FMRI was performed for 6 normal subjects with IRB approval (according to institutional requirements) in a 3.0T GE HDx MRI scanner equipped with an 8-channel head coil and parallel imaging acquisition using EPI with imaging parameters: ASSET=2, ramp sampling, TR/TE=2sec/30ms, FA= 70deg, FOV=22cmx22cm, thickness/gap=4mm/1mm, 25 slices, resolution 128x128. Two fMRI data sets were obtained for each subject.

The first data set was collected during resting-state where the subject tried to relax and refrain from executing any overt task with eyes closed. The light in the scanner room was shut off in order not to provide any distraction to the subject. Scan duration was 9 min 36sec and included 288 time frames.

The second data set was obtained by performing an event-related motor task involving bilateral finger tapping while the subject was looking at a screen. The motor task lasted for 2 sec and was alternated with a fixation period (serving as a control) of random duration lasting between 2 sec and 10 sec, uniformly distributed. Which task to perform was indicated on the display by the letter 1 for the motor task and 0 for the control task, programmed in EPRIME. Axial slices were collected with 150 time frames giving scan duration of 2.5 minutes.

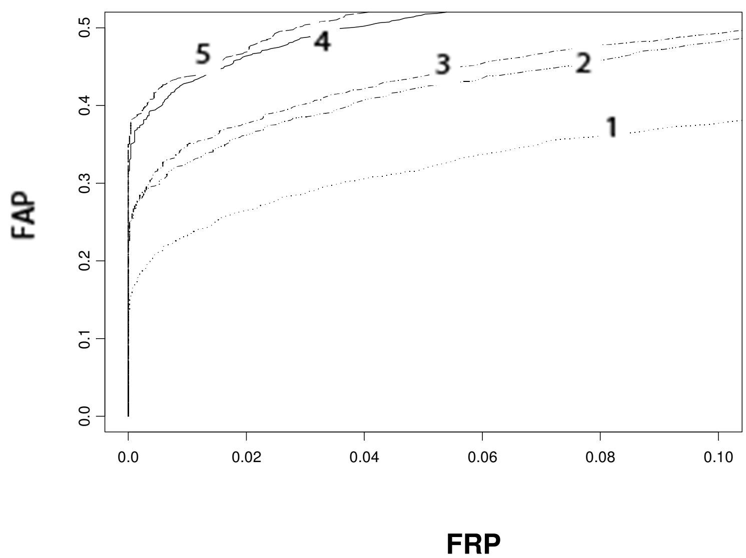

Figure 2. Modified ROC curves for motor data from a single subject

3. Results

We have compared the relative performances of five different analysis methods - single voxel analysis with no smoothing, single voxel analysis with Gaussian smoothing, conventional CCA without correcting for bleeding, constrained CCA with symmetric constraint and constrained CCA with strong dominance constraint using modified ROC method in the previous section for the motor data. In Figure 2, we have presented the modified ROC curves for a single subject. Almost identical results were obtained from the remaining five subjects and hence not presented to avoid repetition. Since in fMRI, high specificity is of utmost significance, we restrict the ROC curves in the region $0 \leq FRP \leq 0.1$. It is evident from the figure that single voxel analysis with no smoothing has the worst performance (curve 1), followed by single voxel analysis with Gaussian smoothing (curve 2), conventional CCA without correcting for bleeding (curve 3), constrained CCA with symmetric constraint (curve 4) and constrained CCA with strong dominance constraint (curve 5). Clearly the constrained CCA methods are the best with the strong dominance constraint slightly edging out the symmetric constraint. The results are similar for the other subjects as well.

4. Discussion

Although we have established a new constrained CCA/multivariate regression method that addresses many of the issues with previously published work on CCA and constrained CCA and also is demonstrably superior in terms of sensitivity to its popular counterparts, the project is far from complete. We will outline some of the key issues which we will address in a series of articles. First, we are developing a hierarchical model on constrained CCA for group analysis. Second, we have restricted ourselves to constraints that can be transformed to a pure non-negativity constraint by a simple linear transformation, for which the exact solution exists. This precludes us from using several constraints, which may be more powerful but may not be transformed to a pure non-negativity constraint by a simple linear transformation. An example of such a relevant constraint is the constraint of weak dominance, where apart from the standard non-negativity constraint on the voxel coefficients, we simply require the coefficient for the center voxel in the neighborhood to be the largest. Preliminary evidence suggests that the constraint of weak dominance is more powerful than the constraint of strong dominance, but finding an efficient algorithm to implement the constraint of weak dominance in practice is a work in progress.

5. Conclusions

To conclude, in this article, we have solved two major issues with regard to the use of CCA in fMRI data analysis. We have used the equivalence of CCA and multivariate regression to make the method applicable to be used with contrasts. Also, we also demonstrated the use of meaningful constraints on coefficients to increase the specificity for a more powerful analysis.

References

Das, S., Sen, P., 1994. Restricted canonical correlations. Lin. Algebra Appl. , 29-47.

Friman, O., Cedefamn, J., Lundburg, P., Borga, M., Knutsson, H., 2001. Detection of neural activity in functional MRI using canonical correlation analysis. Magn. Reson. Med. 45, 323-330.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., 2003. Adaptive analysis of fMRI data. NeuroImage 19, 837-845.

Nandy, R., Cordes, D., 2003a. A Novel Nonparametric Approach to Canonical Correlation Analysis with Applications to Low CNR Functional MRI Data. Magn. Reson. Med. 49, 354-365 .

Nandy, R., Cordes, D., 2003b. Novel ROC-type Method for Testing the Efficiency of Multivariate Statistical Methods in fMRI. Magn. Reson. Med. 49, 1152-1162 .

Nandy, R., Cordes, D., 2004a. Improving the spatial specificity of CCA in fMRI. Magn. Reson. Med. 52, 947-952.

Nandy, R., Cordes, D., 2004b. New Approaches to Receiver Operator Characteristic Methods in Functional Magnetic Resonance Imaging with Real Data using Repeated Trials. Magn. Reson. Med. 52, 1424-1431.

Nandy, R., Cordes, D., 2007. A semi-parametric approach to estimate the family-wise error rate in fMRI using resting-state data. NeuroImage 34, 1562-1576.

Timm, N. H., 1975. Multivariate analysis with applications in education and psychology. Monterey, CA: Brooks/Cole.