PLS (PARTIAL LEAST SQUARES ANALYSIS)

Introduction

Partial Least Squares (PLS) Analysis was first developed in the late 60's by Herman Wold, and works on the assumption that the focus of analysis is on which aspects of the signal in one matrix are related directly to signals in another matrix. It has been developed extensively in chemometrics, and has recently been applied to neuroimaging data. In the application to imaging data, it has been used to identify task-dependent changes in activity, changes in the relations between brain and behaviour, and to examine functional connectivity of one or more brain regions. PLS has similarities to Canonical Correlation in its general configuration, but is much more flexible.

This GUI is the first major release of PLS for neuroimaging. It has been in development for some time, and although this version appears stable, there are always things that can be improved. We are also planning several enhancements, such as univariate testing of means and correlations, advanced plotting routines and higher-order analyses. Please check our website regularly to see if there are updates.

PLS computes a matrix that defines the relation between two (or more) matrices and then analyzes that "cross-block" matrix. In the case of TaskPLS, the covariance between the image dataset and a set of design contrasts can be calculated (an equivalent procedure is described below). The covariance images are analyzed with Singular Value Decomposition (SVD) to identify a new set of covariance images that correspond to the strongest effects in the data. For BehaviourPLS, task-specific correlations of brain activity and behaviour are computed (across subjects), combined into a single matrix, which is then decomposed with SVD. The resultant patterns identify similarities and differences in brain-behaviour relations.

Creating Datamat

Regardless of which type of PLS is to be conducted, the data must be in a form such that all data for all subjects and tasks that are to be analyzed are contained in a single matrix. For image data in general, it is assumed images have been standardized in some manner so that they are all the same shape and size. For PET and MRI data, PLS works best if you use the smallest smoothing filter possible (e.g., no more than twice the voxel size). For ERP data, any filtering, or replacing data in bad channels, etc. must also be done before creating the PLS data matrix.

The 'Session Profile' part of our program loads brain images or ERP waveforms, strings each one out into a vector and then stacks the vectors one on top of the other to make the large data matrix (called **'datamat'** in the PLS programs). Each image (also called subject data file) represents one subject under one condition.

For brain images, the script also eliminates voxels that are zero or non-brain using a threshold, which is specific to each type of image data. Removing the zero and non-brain voxels reduces the size of the datamat considerably and streamlines the computations. Unfortunately, this can restrict the data set if image slices were not prescribed in the same way for all subjects. To make life easy, a "mask" is created based on the voxels that are common for all subjects. After the datamat is reduced, a vector ('**coords**') is generated to remap the reduced datamat into image space again.

PET Images: The code is written assuming you used the SPM99 sterotaxic template with 4x4x4 mm voxels, which creates images having 34 slices each with 40 voxels in the X and 48 voxels in the Y dimensions. For PET scans, the threshold to define brain voxels is 1/4 of the maximum value for a particular subject. The final datamat will have S by C rows and V columns (where S is the number of subjects and C is the number of scans or conditions, V is the number of common brain voxels in the image data set).

ERP: For ERP data, all channels of a subject data file are strung out into a single vector and the vectors are then stacked one on top of the other. The datamat will have S by C rows and E by T columns (where S is the number of subjects, C is the number of conditions, E is the number of channels, and T is the number of time points in the subject data file). If, after creating the datamat, it becomes clear that a particular set of ERP channels is "bad" for most of the subjects, those channels can be eliminated from the analysis using the GUI. As well, single subject data can also be eliminated from the analysis using the GUI.

fMRI: Creating the datamat for fMRI datasets combines the two approaches above. This allows you to run the analysis either on each subject or as a group. Common voxels across subjects and/or runs are identified, then a single row vector is created for each subject, separately for each condition. The datamat will be S by C rows and V by T columns (where S is the number of subjects, C is the number of conditions, V is the number of common voxels, and T is the number of images defined by the user to account for the lag in the hemodynamic response).

TaskPLS: Analysis using Grand Mean Deviation

The TaskPLS is designed to identify whole-brain (scalp) patterns of activity that distinguish tasks. In TaskPLS, a pattern may represent a combination of anticipated effects and some unanticipated ones.

The Grand Mean Deviation analysis is based on representing task means as the deviation around the grand mean computed for each voxel and/or time point. The data are thus averaged within a task, leaving out the within-task variability. (We are exploring the possibility of a "constrained" PLS solution, where a set of apriori contrasts are used to define the solution space).

Next the SVD algorithm is used to get the following three components: brainlv (or salience), singular value (s), and designlv (salience for design). The design scores and brain scores (or scalp scores in ERP) are obtained from the formula below:

design scores LV(n) = designlv
brain scores LV(n) = datamat * brainlv

The saliences for design (designlv) and brain (brainlv) are orthonormal, or standardized. To make comparisons across latent variables easier to visualize, we compute unstandardized saliences. This is accomplished by weighting the saliences by their singular values for the latent variable. All eigenimages and ERP salience plots use the unstandardized saliences.

TaskPLS is run by clicking the 'Run PLS Analysis' button. All results are saved into a specified file. The results will include all data mentioned above, and other useful information.

The TaskPLS results can be displayed by clicking the 'Show PLS Result' button in the main GUI window. For PET and Blocked fMRI, the saliences (eigenimages) and bootstrap ratio images are displayed in a montage that includes all the slices for the LV. For Event-Related fMRI, the results are displayed in a montage as follows: each row represents one lag point, thus the number of rows equals the specified temporal window; each column represents the slices in the image. For ERPs, the LV saliences are displayed as a scalp plot, including only the selected electrodes and epoch. In the results window, you will also find options to display: scatterplots of brain (scalp) scores with design scores, designLV bar plots, and bar plots of the singular values and permutation test results.

Behaviour PLS: Analysis using Behaviour Data

The BehaviourPLS first calculates a correlation vector of behaviour and brain within each task, then stacks these vectors into a single matrix that is decomposed with SVD. Behaviour PLS has the potential to identify commonalities and differences among tasks in brain-behaviour relations.

The behaviour matrix contains one or more behavioural measures that are thought to relate to the measured brain activity. The number of rows in the behaviour matrix and datamat should be the same, with a separate column for each behavioural measure. Since this matrix is created outside of the GUI, it is important that the order of subjects and conditions be identical to the order defined using the GUI to create the datamat.

As for the TaskPLS, the results window initially contains plots of the unstandardized saliences. Within the results window, you can also display scatterplots of brain (scalp) scores with behaviour, bar plots showing the magnitude of the brain-behaviour correlation with confidence intervals, and bar plots of the singular values and permutation

test results. In the brain (scalp) scores by behaviour plots, the linear fit is also plotted to better view the scatter around the correlation.

Tests of Significance

PERMUTATION TEST:

The significance of the latent variable, as a whole, is assessed using permutation tests. We assess the magnitude of the singular values by asking the question: "With any other random set of data, how often is the value for "s" as large as the one obtained originally?" To generate this answer, subjects are randomly reassigned (without replacement) to different conditions, and the PLS is recalculated. Orthogonal procrustes rotation is applied to the resulting BehavLV or DesignLV to correct for reflections and rotations of the resampled data, and the singular values are recalculated based on this rotation (Milan and Whittaker 1995). If the probability of obtaining higher singular values is low, the latent variable is considered to be significant.

For both task and behaviour, 500 permutations are generally sufficient, although probability estimates are typically stable at about 100 permutations.

BOOTSTRAP:

Bootstrap estimation is used to assess the reliability of the brain saliences. In this case, subjects are resampled with replacement. A new datamat, and for BehaviourPLS, a new behaviour matrix are created, and the PLS is recalculated. Thus, unlike for the permutation test, the assignment of subjects to conditions is maintained, but the subjects contributing to task-related effects vary. As for the permutation tests, orthogonal procrustes rotation is applied to the resulting BehavLV or DesignLV to correct for reflections and rotations of the resampled data.

The bootstrap procedure provides an estimate of the standard error for each salience in all latent variables. If the ratio of a salience to its standard error is greater than 2, the salience can be regarded as reliable. (A salience of 2 is roughly equivalent to a z-score of 2 if the distribution is gaussian). The bootstrap estimates serve to assess the contribution of each datapoint to the latent variable structure. The estimates of the standard errors are usually stable after 100 resamplings.

For the BehaviourPLS only, we also use the bootstrap loop to calculate the confidence intervals around the correlation of brain scores (scalp scores) with each behaviour. The brain score-behaviour correlation is calculated for each sample, and the upper and lower limits of the user-specified confidence interval are generated. This distribution is kept as part of the output, so new confidence intervals can be calculated from the command line if needed. Additional behaviour PLS bootstrap output indicates the number of times the bootstrap sample was recalculated because of zero variability in the bootstrap behaviour

matrix (‘countnewboot’) as well as the behaviour/condition combinations that generated the recalculation (‘badbeh’). Occasionally, the resampled behavioural data are skewed, and the resulting confidence intervals do not include the original correlation value . We also provide a very conservative adjustment to the confidence interval calculation that may be informative in those cases (ulcorr_adj; llcorr_adj). However, the correction may not be the most optimal, so use with caution.

PLS is a new method and it can take some time to understand the results. We can be surprised by what it shows us in the data, but can be puzzled as well. We would encourage you to compare your PLS results with other analytic methods available to you. You will find that most of the answers PLS gives you are there in the data, but may not have been obvious on first pass. It does not identify effects that are not there. Finally, since it is a new method, keep in mind that new analytical tools are being developed all the time. We would greatly appreciate hearing from you on what you found with the analysis, what problems you encountered, and any suggestions for modifications to the code or the analytic approach.

Selected References:

PLS for neuroimaging:

Lobaugh, N. J., R. West, et al. (2001). "Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares." Psychophysiology 38(3): 517-30.

McIntosh, A. R., F. L. Bookstein, et al. (1996). "Spatial pattern analysis of functional brain images using Partial Least Squares." Neuroimage 3: 143-157.

McIntosh, A. R., N. J. Lobaugh, et al. (1998). "Convergence of neural systems processing stimulus associations and coordinating motor responses." Cerebral Cortex 8: 648-659.

PLS in other fields

Gargallo, R., C. A. Sotriffer, et al. (1999). "Application of multivariate data analysis methods to comparative molecular field analysis (CoMFA) data: proton affinities and pKa prediction for nucleic acids components." J Comput Aided Mol Des 13(6): 611-23.

Martin, Y. C., C. T. Lin, et al. (1995). "PLS analysis of distance matrices to detect nonlinear relationships between biological potency and molecular properties." J Med Chem 38(16): 3009-15.

Streissguth, A. P., F. L. Bookstein, et al. (1993). The enduring effects of prenatal alcohol exposure on child development: Birth through seven years, a partial least squares solution. Ann Arbor, Michigan, The University of Michigan Press.

Talbot, M. (1997). Partial Least Squares Regression. www.bioss.sari.ac.uk/smart/unix/mplsgxe/slides/intro.htm.

Wold, S., P. Geladi, et al. (1987). "Multi-Way Principal Components and PLS Analysis." Journal of Chemometrics 1: 41-56.

Permutation Tests:

Edgington, E. S. (1980). Randomization tests. New York, Marcel Dekker.

Good, P. (2000). Permutation tests: A practical guide to resampling methods for testing hypotheses. New York, Springer.

Bootstrap:

Efron, B. and R. Tibshirani (1985). The Bootstrap Method for Assessing Statistical Accuracy. Toronto, University of Toronto.

Efron, B. and R. J. Tibshirani (1993). An introduction to the bootstrap. New York, Chapman & Hall.

Milan, L. and J. Whittaker (1995). "Application of the parametric bootstrap to models that incorporation a singular value decomposition." Royal Statistical Society Journal, Series C: Applied Statistics 44(1): 31-49.