MVPA BY CROSS-VALIDATED MANOVA

Carsten Allefeld

v2, 2015-1-12

This text documents the Matlab implementation for SPM of the method introduced by Carsten Allefeld and John-Dylan Haynes, "Searchlight-based multi-voxel pattern analysis of fMRI by cross-validated MANOVA", NeuroImage, 89:345–357, 2014.

PREREQUISITES

The cross-validated MANOVA is based on a ("first-level") multivariate General Linear Model. This model has to be specified and estimated in SPM before using these functions.

Estimation of the model is necessary in order to access SPM's estimates of various fMRI data properties, especially the temporal correlation of the residuals. The functions use the generated `SPM.mat` file and the data files referenced therein, as well as the mask image (`mask.hdr` and `mask.img`). The beta images generated during estimation can be deleted.

The functions may work with other versions, but have been specifically developed and tested with SPM8 and Matlab 7.11–8.1 (R2010b–R2013a).

INTERFACE

The main interface is given by the function

```
cvManovaSearchlight(dirName, slRadius, Cs, permute)
```

which computes the cross-validated MANOVA on a searchlight. `dirName` is the name of the directory where the `SPM.mat` file is located, `slRadius` is the radius of the searchlight in voxels, and `Cs` is a cell array whose elements are contrast matrices. `permute` specifies

whether permutation values should be computed and defaults to `false`.

Simple ("t-like") contrasts are specified as a *column vector*, complex ("F-like") contrasts as a matrix of several columns. Please note that this is the transpose of the format used in the SPM user interface.

The rows of a contrast matrix correspond to the model regressors for each session *separately*, i.e. other than in SPM the contrast should not be explicitly replicated for several sessions. Instead, the program performs the replication internally, assuming that (at least the leading) regressors for each session model the same effects. If there are fewer rows in a contrast matrix than there are regressors for a session, the matrix is zero-padded.

The searchlight radius r is interpreted such that every voxel is included for which the distance from the center voxel is *smaller than or equal* to the radius. This means that $r = 0$ leads to a searchlight size of 1 voxel, $r = 1$ to 7 voxels, $r = 2$ to 33 voxels, and so on. This definition may differ from the one used in other implementations of MVPA algorithms and in publications. Note that it is possible to use fractional values for r.

The result of the analysis are estimates of a multivariate measure of effect size, the pattern discriminability D, which is intended as a drop-in replacement for the conventional measure of classification accuracy. Statistical parametric maps of D are written to images with filenames of the form

```
spmD_C####_P####.nii
```

enumerating all contrasts and permutations, in the same directory as the `SPM.mat` file. Additionally, an image of the numbers of voxels contained in each searchlight is written to `VPSL.nii`.

To ease the specification of contrasts, the utility function `contrasts` can be used to generate contrast matrices for all main effects and interactions of a factorial design, in a form suitable for use with `cvManovaSearchlight`.

For example, `Cs = contrasts([2 3])` results in

```
Cs = { [ 1  1  1 -1 -1 -1]'
       [ 1 -1 -0  1 -1 -0 ; -0  1 -1 -0  1 -1]'
       [ 1 -1 -0 -1  1  0 ; -0  1 -1  0 -1  1]' };
```

For further documentation, please refer to the help texts in the `m`-files.

REMARKS

— The functions are optimized for the computation of several contrasts (and permutations) in one run. One call of `cvManovaSearchlight` with several contrasts will take substantially less time than several calls for each contrast separately.

— The function reads the complete data set into memory. The analysis should therefore be run on a computer with a sufficient amount of main memory, and using other memory-intensive programs at the same time should be avoided.

— The estimation of D is based on the GLM residuals and therefore depends on a properly specified model. That means that all effects that are known to systematically occur should be included in the model. Because sub-effects can be selected through the mechanism of constrasts, it is neither necessary nor advisable to use different GLMs as the basis of different MVPA analyses.

– The fMRI model specification should include the modeling of temporal autocorrelation in order to correctly estimate the pattern distinctness. For this, the option "serial correlations" has to be set to the (default) value `AR(1)`.

— Depending on the data set, it may be possible to perform the analysis for searchlight radii of up to 5 or 6. However, a large radius leads to very long computation times as well as decreased numerical precision. The recommended radius is 3, resulting in a searchlight of 123 voxels.

ROI ANALYSIS

The publication Allefeld and Haynes (2014) describes cross-validated MANOVA only for searchlight analyses. This package contains additional *experimental* code for region–of–interest analyses. The function

```
[D, p] = cvManovaRegion(dirName, region, Cs, lambda, permute)
```

performs ROI-based analysis on a set of voxels, specified by the parameter `region` in the form of a logical 3d-volume.

Other than in the case of searchlight analysis, an ROI may contain so many voxels that adequate estimation of the error covariance matrix is no longer possible. This problem is here solved by shrinkage regularization towards the covariance matrix diagonal. If the regularization parameter (shrinkage strength) `lambda` is not specified, a near-optimal value is estimated from the data using the method of Schäfer and Strimmer (2005).

Since the result of ROI analysis is not an image but a scalar, it is directly returned by the function.

REVISIONS

– v2, 2015-1-12: experimental code for ROI analysis

– v1, 2013-12-20: initial release