

Introduction of the matlab toolbox: COMMON

Gan Xinjun, Xu Bing and Feng Jianfeng
Fudan University, CCSB.

1 Introduction

This package is a matlab toolbox for identifying common links between two datasets especially for fMRI data, with the help of estimate for the 2-dimensional false discovery rate, which is defined by

$$\text{fdr2d}(z_1, z_2) = \pi_0 \frac{f_0(z_1, z_2)}{f(z_1, z_2)} \quad (1)$$

where f_0 is the null distribution and f is observed distribution with mixture form, see [2], Z_1 and Z_2 are the statistics for Dataset 1 and Dataset 2. In this package, we use two-sample t statistics.

To single out the "common" links (significant in both two datasets), the following procedure is employed,

1. With the help of 2-D fdr approach, taking q (default 0.1) as threshold, find out the significant subset denoted by \mathcal{A} .
2. Implementing 1-D method (fdr1d) on the two datasets respectively with a threshold α (default 0.5), obtain the significant subset \mathcal{B} from the first dataset and \mathcal{C} from the second dataset.
3. Identify the common links $\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}$.

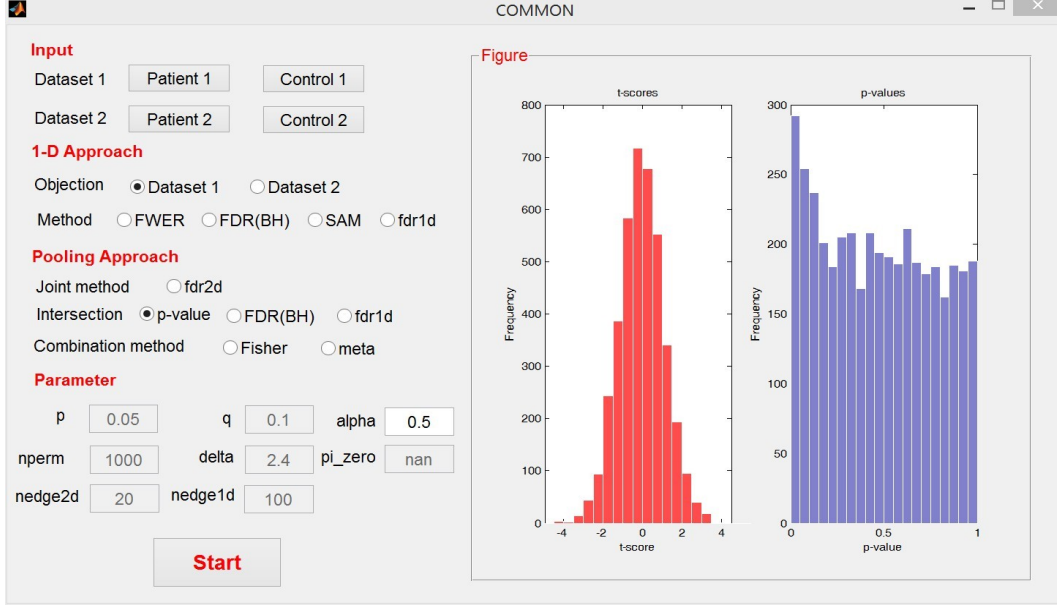
2 The main interface

In matlab command window, typing

```
1 COMMON
```

will call the GUI.

The main interface is as follows, which consists of four parts: Input, 1-D approach, Pooling approach and the figure.



Input. As shown in the figure, you can import two datasets. Each dataset consists of two groups: Patient and Control. The format of fMIR data for each group is 3-D array in matlab with the following form

$$\text{Time} \times 90 \times \text{sample} \quad (2)$$

The length of the second dimension of the data is 90, corresponding the size of region in the brain.

1-D approach. This is for the analysis of 1-D case using some popular multi-comparison correction methods. First, in Objection step the user must specify the dataset and then the histogram of the t-scores and p-values will be shown instantly in the figure part. Second, in Method step, the user should choose an approach, FWER (Bonferroni bound), FDR (Benjamini-Hochberg), SAM (significance analysis of microarray) and fdr1d. Third, in Parameter step, set the parameters for different methods.

Pooling approach. This is the crucial step for identifying 'common' links. In this part, 2-D approach are distinguished into 3 categories: Joint method (fdr2d), Intersection and Combination method. The Joint method corresponds the procedure mentioned before. For more details, see [3].

1. Joint method (fdr2d). Use the joint density of the Datasets to estimate the local false discovery rate of 2-D case. The contours of 2-D fdr will be shown and the common links between the Datasets will be singled out using the procedure introduced before.

2. Intersection. These methods use the intersection of 1-D approach as output.
3. Combination method. These methods first perform a combination of the statistics of the Datasets, like Fisher's method or meta analysis, see [3], and then employ the `fdr1d` approach for the combination.

All the results will be saved as `'result.mat'` automatically.

References

- [1] Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.
- [2] Efron, B., 2010. *Large-scale inference: empirical Bayes methods for estimation, testing, and prediction*. Cambridge University Press.
- [3] Gan XJ, Xu B, Ji XX, Lu WL Nichols T, Feng JF. A Statistical Approach for Trans-attribution Model. Submitted.
- [4] Ploner, A., Calza, S., Gusnanto, A., Pawitan, Y., 2006. Multidimensional local false discovery rate for microarray studies. *Bioinformatics* 22, 556-565.
- [5] Tusher, V.G., Tibshirani, R., Chu, G., 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences* 98, 5116-5121.
- [6] Yudi Pawitan and Alexander Ploner (). *OCplus: Operating characteristics plus sample size and local fdr for microarray experiments*. R package version 1.32.0.