

Discovering signals in fMRI data; a Bayesian nonparametric approach

Project Proposal for STAT308

Ahmed Bou-Rabee, Wanrong Zhu, Zheng Xu, Mo Zhou

February 28, 2017

Introduction

The goal of this project is to formulate and test a method which can be used to adaptively identify clusters of signals in functional magnetic resonance imaging (fMRI) data. Roughly, fMRI measures the change in brain blood flow associated with mental activity [?]. The brain is divided into tiny blocks known as voxels, and the intensity of the blood flow over each voxel is recorded at evenly spaced time intervals while the subject is stimulated. The data is then in the form (voxel, time, intensity of reading). For example, suppose researchers wanted to identify regions of the brain associated with hunger or craving. To aid in this, fMRI readings can be taken while hungry subjects are shown pictures of food.

An advantage of using fMRI is that it's a noninvasive procedure. Due to this, there are many publicly available datasets [?]. However, analyzing fMRI data poses many statistical challenges, one of them being the multiple comparisons problem. Any sampling procedure involves error. Because there are typically hundreds of thousands voxels, it's likely that individual voxels may have high readings, but not be statistically significant. Also, the noise of a reading is typically dependent on location, *i.e.*, it is heteroscedastic. Identifying significant clusters (not just individual voxels) introduces an additional challenge.

Project Goals

For the purposes of our project, we will assume, like [?], that data given to us is of the form (voxel, time, p-value). Given a dataset of this form, we will focus on two questions:

- Can we adaptively identify entire regions of the brain (not just voxels) which are associated with the stimulus while accounting for multiple testing error and heteroscedastic noise?
- There is no response variable in fMRI data, so how can we reliably test our algorithm?

Adaptively identifying regions of interest

Some existing work, e.g., [?], has developed methods which can identify significant voxels and regions in fMRI data. These methods presuppose that regions of interest have already

been identified by biological experts. We would like to focus on identifying these regions adaptively, rather than using predefined ones. One idea we have for doing so involves a non-parametric Bayesian procedure. Specifically, we assume the data is generated according to the following:

1. For each location and time (i, j) , generate a Beta distribution, $\text{Beta}(\alpha_{ij}, \beta_{ij})$, by making a random draw from a prior distribution on Beta distributions, Φ .
2. Generate the p-value at (i, j) by making a random draw from $\text{Beta}(\alpha_{ij}, \beta_{ij})$.

With this assumption, given a set of data, we can form the posterior likelihood and identify what choices of $\text{Beta}(\alpha_{ij}, \beta_{ij})$ maximize the likelihood. (This will be done approximately on a computer using, *e.g.*, MCMC or Gibbs sampling). Then, points with Beta distributions sufficiently close to uniform will be non-significant, while points with right-skewed Beta distributions will likely be significant.

One challenge involves choosing an appropriate Φ . We are supposing that there are clusters in the data, that is, if (i, j) has distribution $\text{Beta}(\alpha_{ij}, \beta_{ij})$ and (w, z) is close to (i, j) , (in some metric, say, normalized Euclidean distance), then it is more likely for (w, z) to also have distribution $\text{Beta}(\alpha_{ij}, \beta_{ij})$. We can encode this information with our choice of Φ .

Testing the algorithm

It is difficult to test the performance of a method on a problem which doesn't have a predefined solution; scientists seek to use fMRI to discover completely new information about the brain. So, we are reduced to heuristics.

In our paper, we will first compare the proposed algorithm against the p-filter, introduced in [?]. The p-filter requires, as input, an fMRI dataset with pre-defined regions of interest. So, one criteria we can test is if our algorithm generates significant clusters which are similar to the significant regions of interest that p-filter produces. In doing so, we must come up with a quantitative way of comparing the outputs of the two algorithms. Once we have this, we will run the comparison procedure on several datasets, including the ones used in the p-filter paper.

We would also like to test the variability of our method. That is, we want our method to output similar significant clusters when run on similar data. To check this, we will use bootstrap. Specifically, given a dataset, we will run our method several times on new datasets which are sampled with replacement from that dataset. Each of these new datasets are statistically similar, so our method should have similar outputs across each run. Here, also, we are faced with the challenge of quantifying what it means for shapeless clusters to be similar.