The title of your project proposal

Jara, Jon jonmigueljara wtsanderson | Shishido, Juan juanshishido | Wu, Paul yuanyiwu | Xu, Wendy berithsu

December 13, 2015

Abstract

Our paper is the *Neural Basis of Loss Aversion in Decision-Making Under Risk* . The experiment attempts to investigate the relationship between neural and behavioral loss aversion. The relationships are observed by neural activity when subjects are exposed to gambles. Our project attempts to run linear models and multi-voxel pattern analysis on the same fMRI data to find areas of the brain sensitive to these gains and loss

1 Introduction

Our research idea is based on the 2007 paper The Neural Basis of Loss Aversion in Decision-Making Under Risk, written by Sabrina M. Tom, Craig R. Fox, Christopher Trepel, and Russell A. Poldrack[2]. Our goal is to reproduce the major part of the paper with possible simplifications and also develop some original analysis on the same data.

A significant portion of the paper is devoted to analyzing neural indicators of loss aversion and correlating that neural loss aversion to behavioral loss aversion. Test subjects are presented with gambles with equal chances of winning and losing and different amounts of money to gain and lose. These gamble offers and the subjects decision of whether or not to accept the gamble are recorded, as well as the subjects neural activities during the tasks. Behavioral risk aversion is measured through modeling the participants decision on the amount of proposed gain and loss of the gambles. On the other hand, neural risk aversion is measured through modeling the participants fMRI data on the amount of proposed gain and loss of the gambles. By analyzing the correlations between the coefficients of these models, the authors find that some parts of the brain are particularly sensitive to loss and that neural activities in the loss sensitive areas are closely related to the extent to which a subject avoids risk in behavioral terms.

We will reproduce this part of the paper using mostly linear regression and logistic models and evaluate our results against the papers results as well as using error analysis. We will also visualize our research results consistently with 2-D plots. Besides, we would like to experiment with Multi-Voxel Pattern Analysis on the fMRI data, which helps detect patterns of brain activity by jointly analyzing multiple voxels at the same time. Our group found that running a conjunction analysis between behavioral and neural coefficients proved to be too difficult so we will be investigating the results separately.

2 Data

2.1 Description

The data set includes 3 sample runs for each subject, with 16 subjects in total. The subjects consist of 9 females and 7 males. For each run, the subject is presented with a series of gambles with different combinations of gains and losses, each combination randomly drawn from a gain/loss matrix of size 16 x 16. For the purpose of analysis, the data are collapsed into a 4 x 4 matrix. The subjects fMRI data generated during the tasks is available for each run.

2.2 Acquiring

A primary goal of ours is to ensure that individuals who wish to replicate or improve upon our work can easily do so. To that end, in terms of data access, we have created a bash script that fetches the data from www.openfmri.org and untars it. It runs only if the data directory doesn't already exist. The script also downloads the checksums and we've written a Python script to check that the hashes match. We will incorporate the Python script in the Makefile for downloading the data.

3 Methods

3.1 Preprocessing

Pre.

3.1.1 Smoothing

After initial plots, we observed that the data was noisy. Smoothing in space reduces the noise by averaging across the independent noise in the voxels, while preserving the signal. Thus, the first step we took in pre-processing the bold image data was to smooth the BOLD data using a Gaussian filter by 2 SDs in all three spatial dimensions. Plots show how smoothing vastly improves the interpetability of the single brain images:

3.1.2 Outliers

The next step was to use RMS of the BOLD fMRI signal to locate and remove outliers to further remove noise from the data. We found that the number of outlying volumes varied greatly from subject to subject and from subject to run with some runs having as many as 100 volumes removed and as little as 0. We plotted RMS values with outliers marked with red markers (similar to what we did in homework 2). Plots between subject 1 run 1 and subject 9 run 3.

3.2 Analysis

3.2.1 Primary

The first task is to reproduce the results of the paper using our own tools. The goal is to find areas of the brain in which neural loss aversion is correlated to behavioral loss aversion. The result would lead to which parts of the brain are responsible for making risky decision. The first step is to fit a logistic regression on the behavioral data using a simple linear regression. We used a combination of pandas and statsmodels packages to fit a logistic regression model on the first subject and all the runs combined. Python modules were used rather than R packages in order to keep the whole project in the same language and for reproducibility. The logistic regression was modeled in this manner:

$$logit(p) = \beta_0 + \beta_1 X_{gain} + \beta_2 X_{loss} + \epsilon$$

With gain and loss values beings the 2 regressors in the regression. Subject 1 had a gain to loss ratio of 2.33, which was inline with gain to loss ratio cited by other sources in the paper (around 2.0). The paper was not specific in how they dealt with multiple runs per subject but we decided that it would be statistically sound to combine the runs for all the patients. In addition, the paper conducted pre-processing steps like smoothing and motion correction that we still need to employ.

The logistic regression provided results inline with the paper (and past "risk-averse") studies and were in the end statistically significant, however, the main goal of neurological study is to observe how these results compare with neural signals and brain activity. This part of the analysis has proven to be less trivial than a simple logistic regression. The challenges stem mostly from working fMRI data in conjunction with the behavioral variables "gain" and "loss". We have tried using simple linear regression at each voxel after convolving the fMRI data, but the results yielded blurry images of brain slices the general locations of brain activity. While these locations give a general idea of which parts of the brain are activated during runs of the experiment, the resulting regressions leaves us with hundreds of beta coefficients that represent activity without specifying whether the activity is due to negative loss response and positive gain response. Further research is needed to perform a similar conjunction analysis that was performed in the paper. If the complexity of separating the different types of signals is beyond the scope of our group, we will move on and use the simple regression model with only one parameter. There is worry that without conjunction analysis we will not see significant results like those found in the paper and it will be hard to compare risk aversions to specific loss/gain values if we can differentiate between the responses. Also, the paper specified specific regions of the brain with large magnitudes of gain and loss response. To simplify our analysis, we could restrict our study to whole brain analysis and see if we find any significant correlation, otherwise we will attempt to segment the same sections as the ones used in the original study.

After modeling the fMRI data and hopefully getting valuable coefficients for loss and gain responses, we will then have to use robust regression to model the values of these neural response coefficients and the behavioral coefficients. Robust regression was used in the paper as a way to decrease the likelihood of any outliers. We might try to simplify this method by using methods learned in class to remove outlier ourselves and then running a simple linear regression. In the end, we believe we can reproduce the results of the paper but we might need to simplify some of the methods, particularly in the process of quantifying neural risk aversion and segmenting risk-averse areas of the brain.

3.2.2 Secondary

As part of our secondary analysis, we are investigating whether it is possible to use Multi-Voxel Pattern Analysis (MVPA) on our data. Norman, Polyn, Detre, and Haxby note that MVPA allows researchers to analyze *vectors* of voxel activity values instead of individual ones[1]. There is greater sentivity with this focus on distributed patterns of activity, making it easier to detect differences between conditions[4]. Typically, linear classifiers, such as Gaussian Naive Bayes or linear support vector machines, are used in MVPA analyses[1]. The features correspond to the brain activity and the label to the cognitive state. MVPA also provides the ability to correlate subject-level classifier estimates across multiple trials[1].

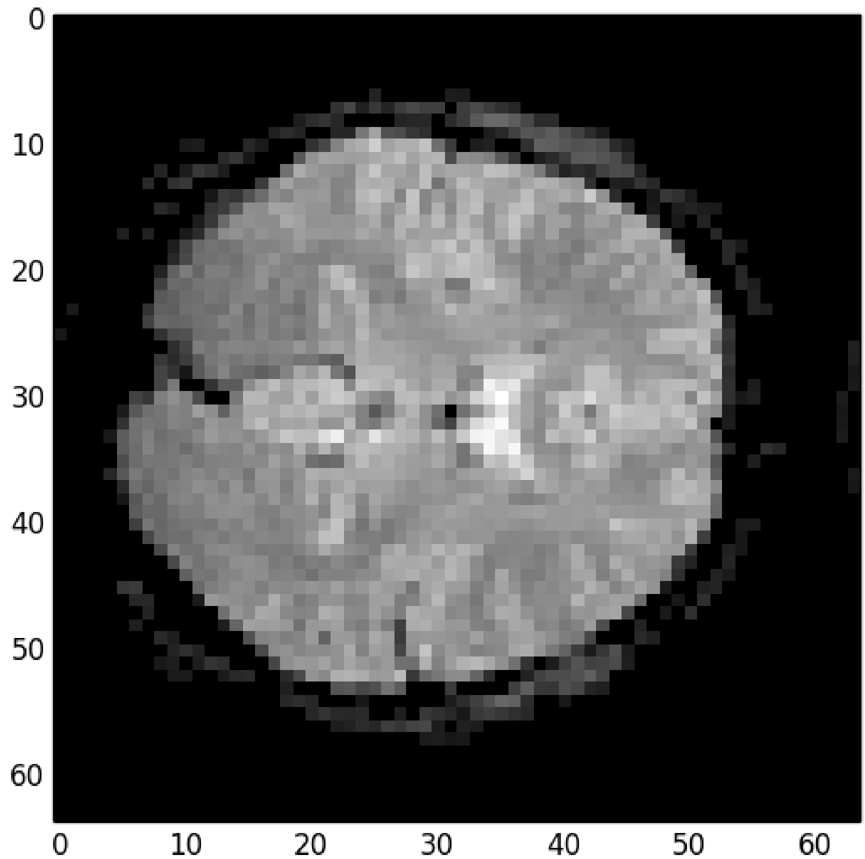
3.3 Tools

We have identified several Python packages we will use for our analysis.

- NumPy
- Nibabel
- Statsmodels
- Scikit-Learn
- Nilearn
- PyMVPA
- Matplotlib
- Seaborn

4 Results

Thus far, we have run some initial analyses on our data. We have performed convolution as well as a linear regression. The following is a plot of a middle slice of the $\hat{\beta}_1$ of a subject's brain.



We also have code for our first pass at the logistic regression using the behavioral data.

5 Discussion

5.1 Challenges

Many of our issues so far have stemmed from the open ended nature of the assignment, resulting in less direction than we are used to. This, combined with the new git workflow, has given us some problems that we are working on overcoming with a more rigid implementation plan that gives everyone a role. We are currently using the workflow as best we can, but will be adding many more Git issues in the coming days in order to better assign work and facilitate multiple pull requests without merge conflicts. We're having good success with Python code when we have a good implementation plan. However, a good deal of our codebase deals with fMRI analysis that we haven't had much experience working with. This has led to difficulty writing effective tests, tests that can inspire full confidence in our implementations. We're working on keeping them simple, using basic modular functions on small datasets to ensure adequate coverage.

5.1.1 Team

We have been working well as a team, and most of our problems are simply efficiency related. Each of us has applicable skills, we're just trying to get them all streamlined. This is coming together through the workflow, mainly in GitHub issues. This allows each of us to work on a different part of the project at the same time. We have had trouble getting everyone together at the same time due to erratic schedules, but have been overcoming this using the workflow.

5.2 Class Concepts

We think that we could have used practice in more advanced fMRI analysis techniques. We've been having some problems with preprocessing, which involves much more detailed work than we thought. We're looking into some extra Python libraries to help sort out some advanced preprocessing techniques, smoothing out confounding variables like respiration and heartbeat. The workflow has been the most helpful focus, but we think we could have used even more practice. Because the workflow facilitates efficient team work, we would have liked to have seen some more group exercises or homework aimed at perfecting the process.

References

[1] K. A. Norman, S. M. Polyn, G. J. Detre, and J. B. Haxby, *Beyond mind-reading: multi-voxel pattern analysis of fmri data*, Trends in Cognitive Sciences, 10 (2006), pp. 424–430.

[2] S. M. Tom, C. R. Fox, C. Trepel, and R. A. Poldrack, *The neural basis of loss aversion in decision-making under risk*, Science, 315 (2007), pp. 515–518.

[3] R. Viviani, G. Grön, and M. Spitzer, *Functional principle component analysis of fmri data*, Human Brain Mapping, 24 (2005), pp. 109–129.

[4] B. Voyager, *Multi-voxel pattern analysis (mvpa)*.

[5] ———, *Spatial smoothing*.