Project Alpha Progress Report

Kent Chen, Rachel Lee, Ben LeRoy, Jane Liang, Hiro Udagawa

November 12, 2015

Background

The Paper

- from OpenFMRI.org (ds009)
- "The Generality of Self-Control" (Jessica Cohen, Russell Poldrack)

The Data

- BART study with event-related neurological stimulus and 24 subjects.
- 24 subjects, 3 conditions per subject

Data fetching and preprocessing

- Set up a make file to download and decompress data
- Wrote a loop to get hashes of all files in all subdirectories of data belonging to one group member, saved the dictionary of hashes to a JSON file
- “make validate”

Initial analysis

Convolution

- Worked with problems with event-related stimulus model

Smoothing

- Convolution with a Gaussian filter (scipy module)

Linear regression

- Multiple and single regression with stimulus (all conditions and seperate)

Initial analysis

Hypothesis testing

- General t-tests on β values
- Across suject analysis

Time series

- ARIMA(1,1,1) model

PCA

- Modeling against
- SVD

Our plan

Initial

- Analysis to perform: multiple subjects, time series, PCA, multiple testing
- Using only BART study for feasibility

Goal

- Trying to reproduce methods, but it won't all be the same

Simplification steps

- They used a lot of packaged software, we don't have those. Also not familiar with some of their methods.

Our plan

Issues we have encountered/discussed

- Convolution/time intervals
- Multiple comparisons

Method of validating models

- t-tests, RSS, permutations

Our process

Most difficult aspect of project?

- working with fMRI data, moreso than Git workflow

Ill-defined assignment?

- having the freedom to make decisions on what direction to take

Success in overcoming these obstacles?

- (work-in-progress)

Our process (cont'd)

Issues with working as a team?

- 5 people means it's hard to find time to meet in person

Most useful parts of class?

- Git workflow

Least helpful?

- fMRI.

Our process (cont'd)

What do we need to successfully complete the project?

- try our best to reproduce as much as possible
- if time allows, explore new approaches

Difficulty of making work reproducible?

- making sure that stuff works for both Python 2 and 3. Travis is a pain, but testing is important.

Potential topics to cover in class in the future

- Overview of brain / neuroanatomy?
- More linear regression (ANOVA)? PCA? The mathematics or the implementation?
- Machine learning (classification, prediction, cross-validation)?
- Permutation tests (and maybe bootstrap)?
- Software tools (Git, Make, Python, statmodels, etc.)
- Technical writing and scientific visualization?
- Advanced topics (regularized regression, selective inference)