The Sparse Overlapping Sets Lasso for Multitask Learning and fMRI analysis

Nikhil Rao

University of Wisconsin - Madison

with
Christopher Cox
Robert Nowak
Tim Rogers

POSTER: SAT46
Can Brains be Crowdsourced?

Q1: What regions of the brain process visual and auditory stimuli?

Q2: Are these regions distributed?

If multiple people participate, can look to leverage data from all subjects to perform better inference.

Typically, data is blurred and mapped to a common brain atlas, losing fine-level information, which may be useful [1]

The Challenge

GOAL:
Identify significant voxels in the brain without trying to match brains, while taking advantage of data from multiple subjects.

Brains of individuals are similar at a coarse level, and different at fine levels.

\[ X_1, X_2, \ldots \] Data
\[ Y_1, Y_2, \ldots \] Responses

minimize \( w_1, w_2, \ldots \) \( \{ \sum_i f(Y_i, X_i, w_i) \} \)

w’s account for the structure in the data.

Can we model the data in a way that allows us to formulate optimization problems that takes into account both the similarities and differences between subjects?
At a glance: Sparse Overlapping Sets Lasso

We formulate the problem as a regularized multitask learning problem that:

1. Accounts for the coarse level similarities
2. Accounts for the fine level dissimilarities
3. Can be solved efficiently
4. Allows us to derive consistency results

Our results specialize to known cases [2,3,4] under certain parameter settings

We obtain results that are comparable to other high dimensional inference procedures

**POSTER: SAT46**

---

