SOS LASSO: A NEW METHOD FOR FINDING DISTRIBUTED REPRESENTATIONS IN FMRI DATA

CHRISTOPHER COX, NIKHIL RAO, ROBERT NOWAK, TIMOTHY ROGERS   UNIVERSITY OF WISCONSIN, MADISON

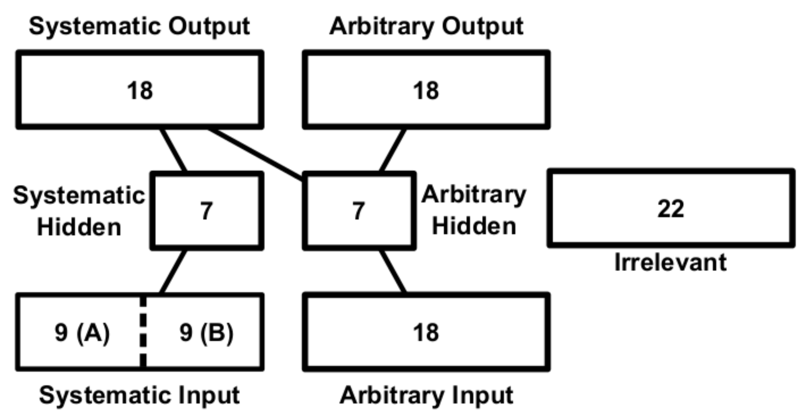
Introduction

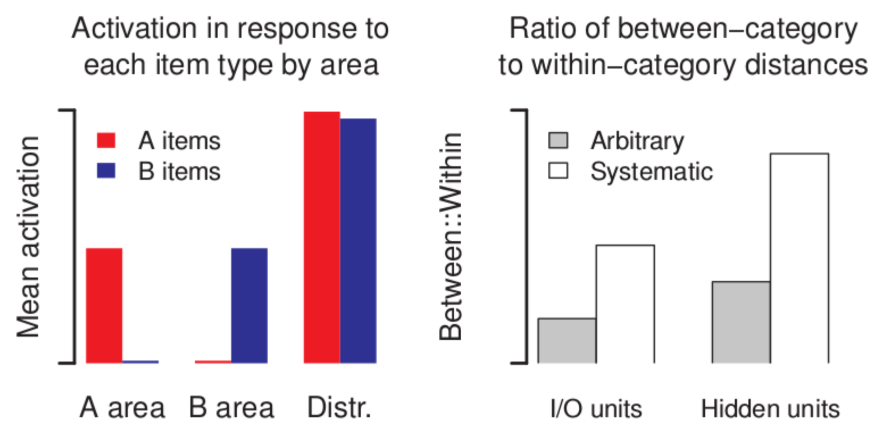
- **PDP models** have motivated many influential hypotheses.
- **Distributed representations** are at the heart, responsible the learning, responding, generalizing, and predicting that the models are capable of.
- However, there is limited neural evidence for distributed representations, e.g. from fMRI.
- There may be a disconnect between methods of analysis and representational assumptions of PDP that **systematically overlook** distributed patterns in the brain.
- **SOS LASSO** is an optimization technique that is sensitive to distributed representations, *similarly located* in *samples of subjects*.

Data Generation & Method

Data were generated by training an auto-encoder neural network.

- Two areas specified to be A and B selective.
- One area placed between systematic input and output units.
- Categorization is possible based on either region, but the information is represented very differently.

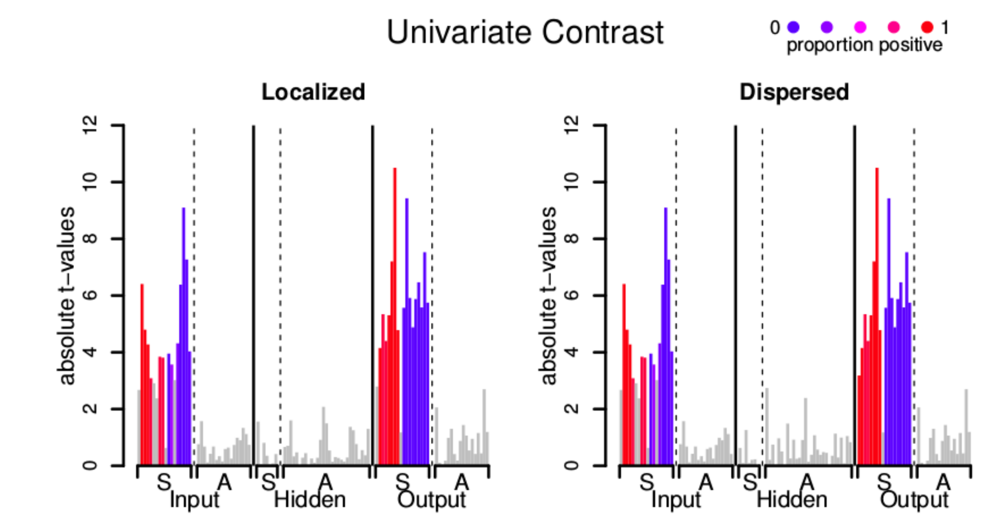




Methods vary in their representational assumptions, and will label different patterns of data as "important". We explore this by applying a range of methods to these data, to obviate the consequences of these assumptions.

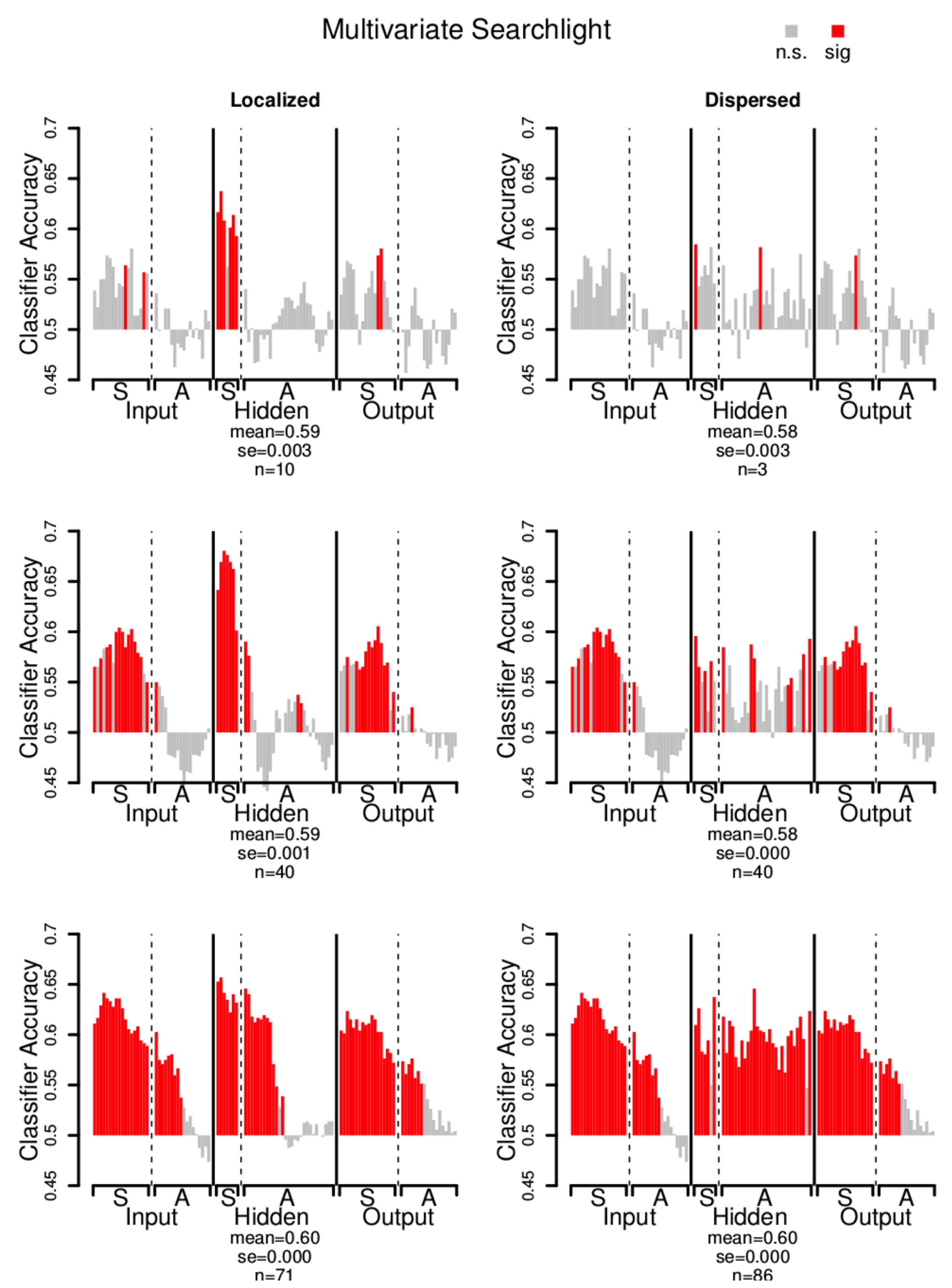
UNIVARIATE

Seeks consistent, localized *activation*.



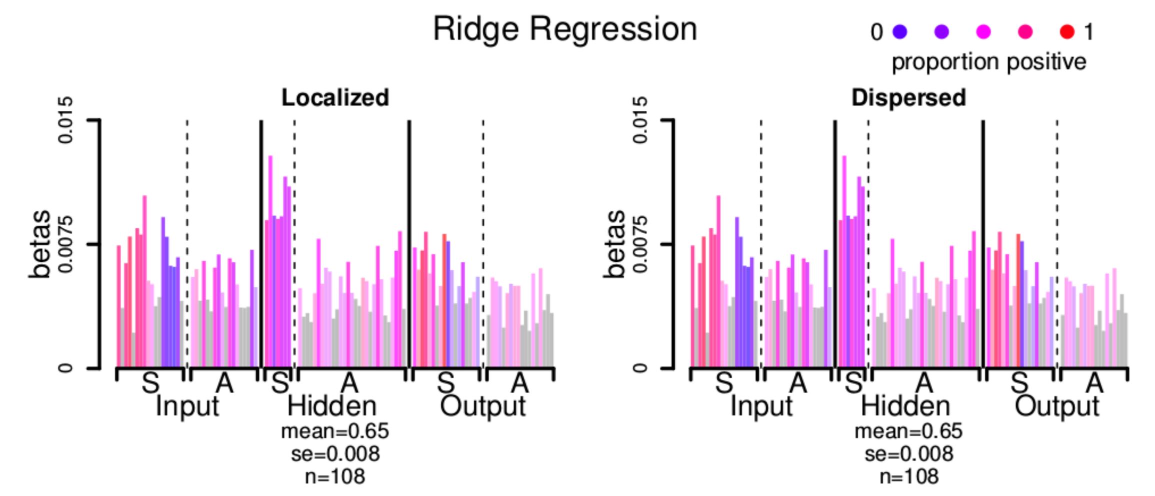
SEARCHLIGHT

Seeks consistent, localized *information*.



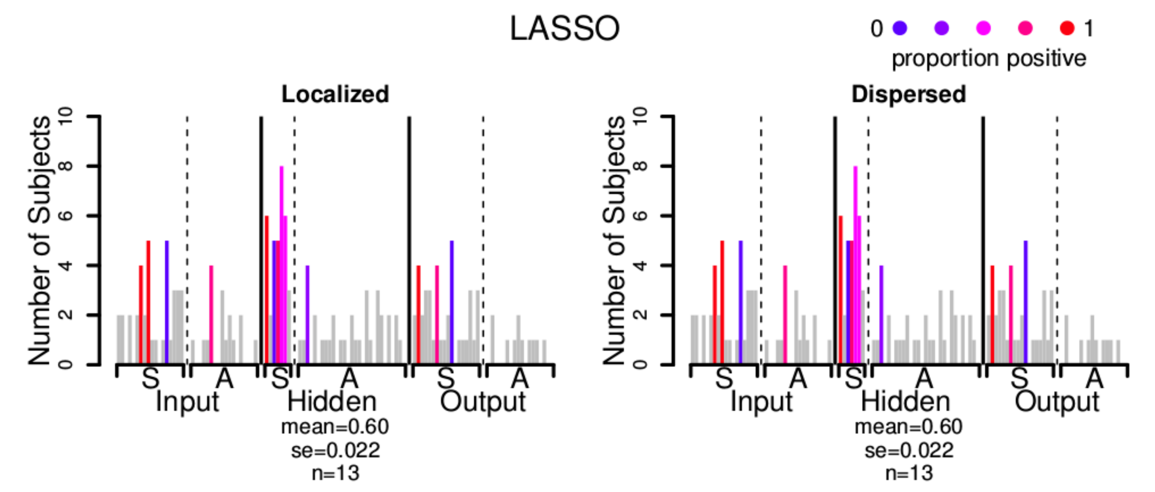
RIDGE REGRESSION

$h(\beta) = \sum \beta_j^2$ Seeks best solution using *all* voxels.



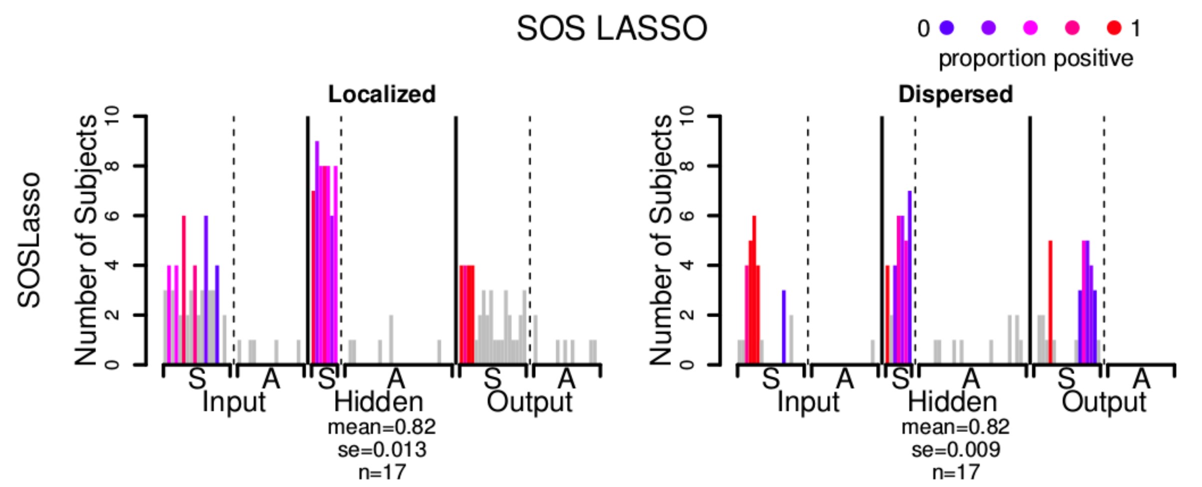
LASSO

$h(\beta) = \sum |\beta_j|$ Seeks best solution using *fewest* voxels.



SOS LASSO

Seeks best solution using fewest voxels and fewest groups, in individuals *and* in the sample.

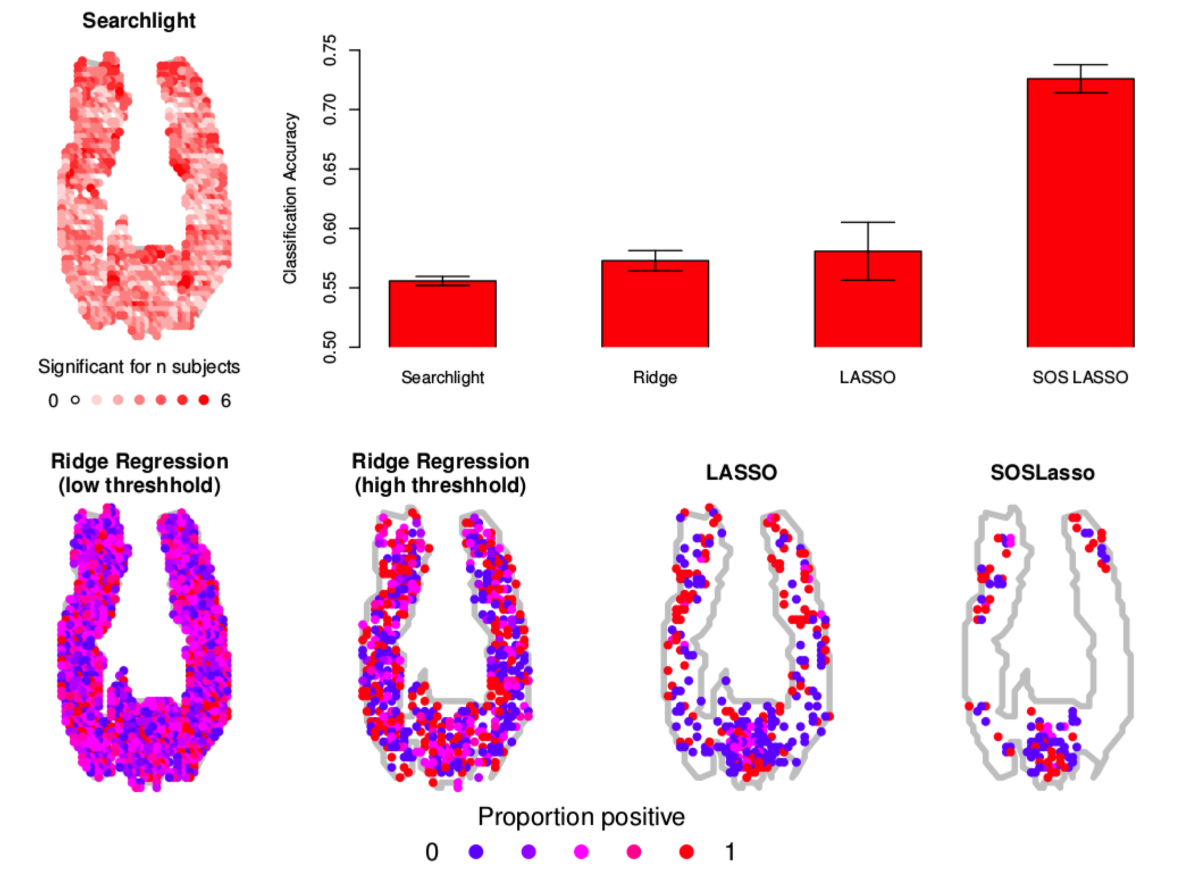


INTRODUCING REGULARIZED REGRESSION

$$\arg\min_{\beta} \sum_{i=1}^{n} \log(1 + e^{-\bar{y}_i X_i \beta}) + \lambda h(\beta)$$

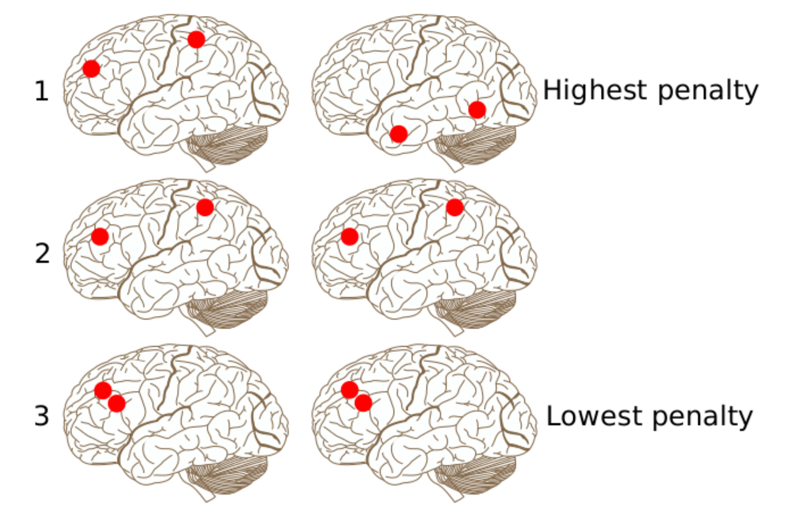
Regularized (logistic) regression involves adding a penalty that is a function of the solution β. The nature of this penalty encourages different β to be found. The severity of this penalty is scaled by λ; ordinary regression when $\lambda = 0$.

fMRI APPLICATION



SOS LASSO PENALTY

$$h(\beta) = \sum \left(\alpha \sqrt{\sum w_G^2} + (1 - \alpha) \sum |w_G| \right)$$



SUMMARY (ASSUMPTIONS)

Assumption	U	SL	L	R	SOS
Local representations within an individual	✓	✓			✓†
Local representations between individuals	✓	✓			✓†
Consistent representation between individuals	✓				
Independence	✓				
Representation is sparse			✓		✓
Representation is redundant	✓			✓	

† SOS LASSO assumes "locality" in a very different way than the univariate and searchlight methods. *Similarity* is defined *a priori*.