Latent Factor Methods in Neuroimaging Data

Valerie Bradley, Ana Ignatieva, Yuxi Jiang, James Thornton

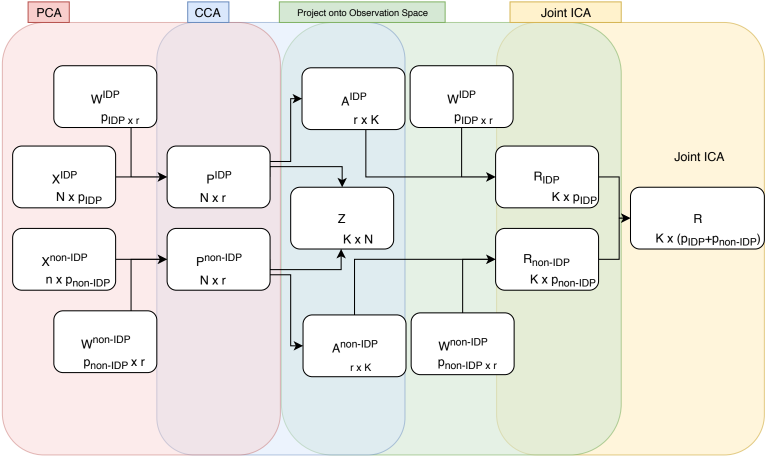
University of Oxford
University of Warwick

16 November 2018

PCA, CCA and ICA

- Observed data $\mathbf{X} \in \mathbb{R}^{n \times d}$
- Linear transformation to underlying latent variable $\mathbf{Y} = \mathbf{XP}$
- $\mathbf{Y} \in \mathbb{R}^{n \times k}$; generally $k << d$
- Choose $\mathbf{P}$ to optimize some objective function $f_{\mathbf{X}}(\cdot)$
  - **PCA**: $f_{\mathbf{X}}(\cdot)$ describes the amount of variance explained by a latent variable
  - **ICA**: $f_{\mathbf{X}}(\cdot)$ describes the independence of latent factors within a data set
  - **CCA**: $f_{\mathbf{X}}(\cdot)$ describes the correlation of latent factors derived from two data sets

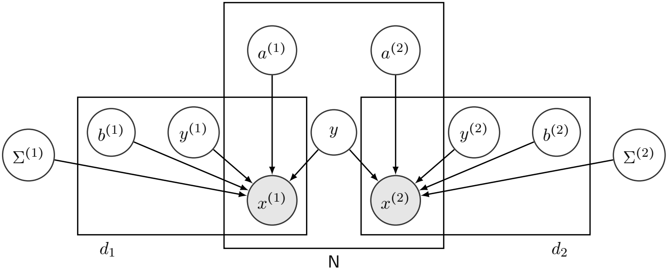
Miller et al. (2016) Review



Bayesian CCA

Generative model:

- $\mathbf{Y} \sim N(0, \mathbf{I})$
- $\mathbf{Y}^{(m)} \sim N(0, \mathbf{I})$
- $\mathbf{X}^{(m)} \sim N(\mathbf{A}^{(m)}\mathbf{Y} + \mathbf{B}^{(m)}\mathbf{Y}^{(m)}, \mathbf{\Sigma}^{(m)})$



Issues

Multiple modalities (views):

$$\Lambda = \begin{bmatrix} A^{(1)} & B^{(1)} & 0 & \cdots & 0 \\ A^{(2)} & 0 & B^{(2)} & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ A^{(m)} & 0 & \cdots & & B^{(m)} \end{bmatrix}$$

Unidentifiable models:

- rotational invariance:

$$\Lambda Q^T Q Y = \Lambda Y, \quad \forall\; Q^T Q = I$$

- scale invariance
- label switching (ignore)

BASS and GFA

BCCA with:

- multiple modalities
- different methods for inducing sparsity via priors
- invariance: non-sparse rotations violate prior structure

GFA:

- *Automatic Relevance Determination (ARD)* prior on Λ
- automatically selects number of LFs
- elements $\rightsquigarrow 0$

BASS:

- hierarchical *Three Parameter Beta* priors on Λ
- global, factor and local shrinkage
- automatically selects number of LFs
- elements $\rightarrow 0$

Data

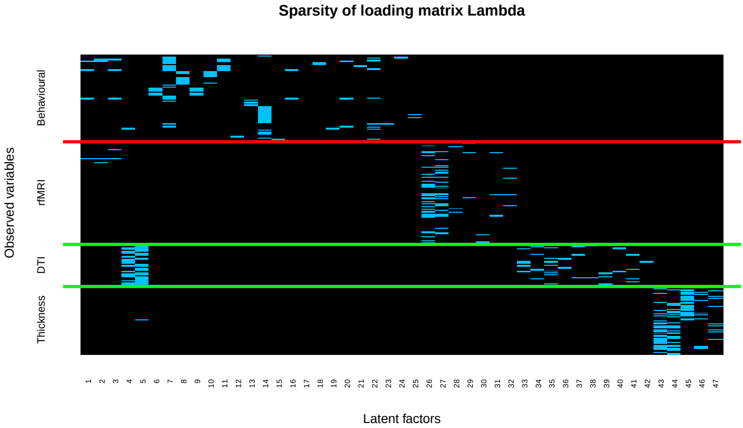
- Brain imaging data:
    - DTI (43 variables)
    - rFMRI (105)
    - Structural thickness (68)
- Behavioural data (90)
- Matched by subject:
    - removed missing data
    - 335 observations

Objective: Use BASS/GFA to look for underlying relationships between different modalities

BASS Results

Imaging + behavioural data:

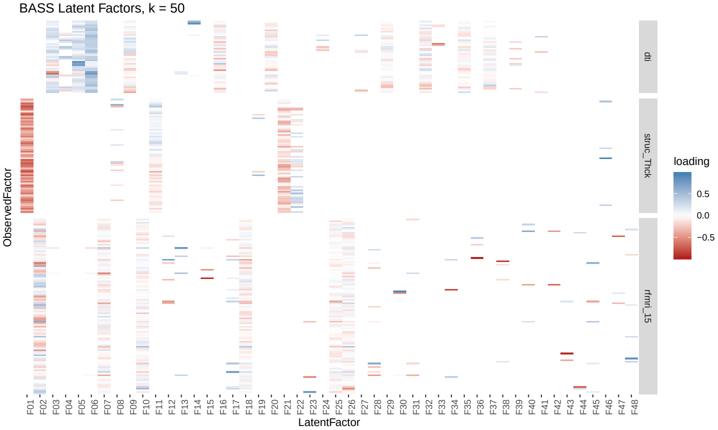
- 2 modalities (imaging + non-imaging)
- initial # latent factors = 50



BASS Results

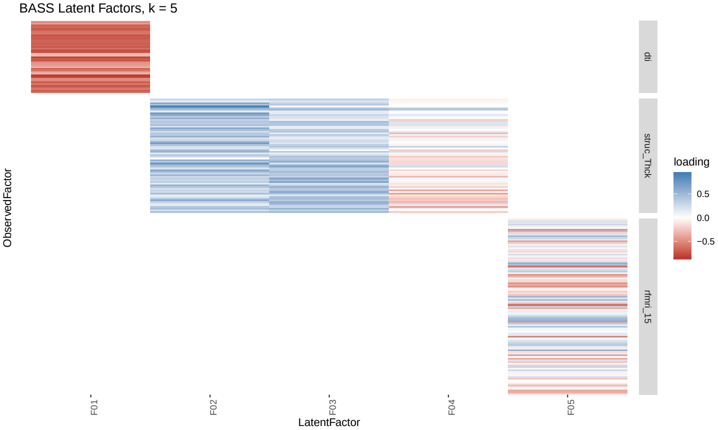
Imaging data only:

- 3 modalities (DTI + rFMRI + structural)
- initial # latent factors = 50



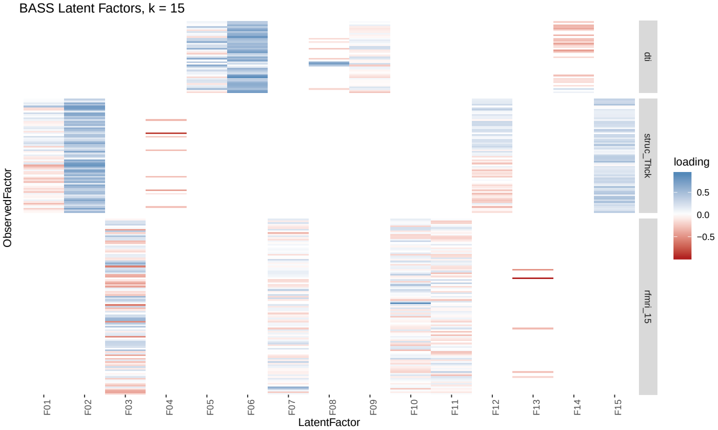
BASS Results - Sparsity

$k = 5$



BASS Results - Sparsity

$k = 15$

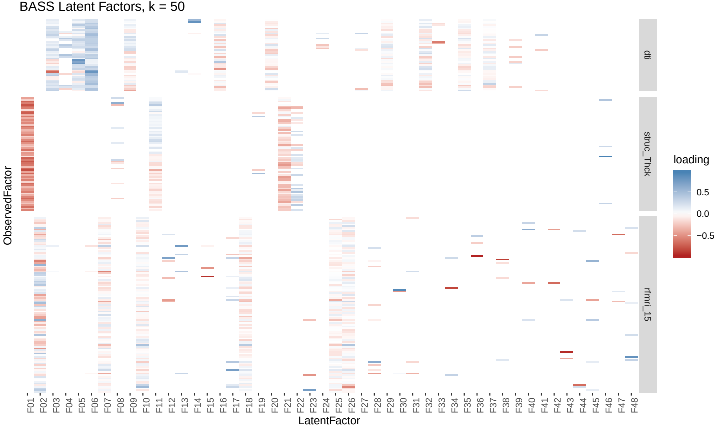


BASS Results - Sparsity

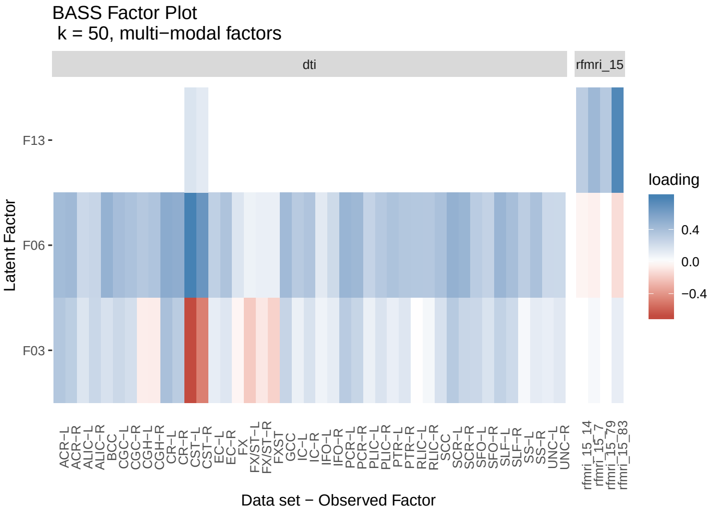
$k = 30$

BASS Results - Sparsity

$k = 50$, only 48 latent factors identified



BASS Results - Multi-modal factors



Cross-Modality Loadings

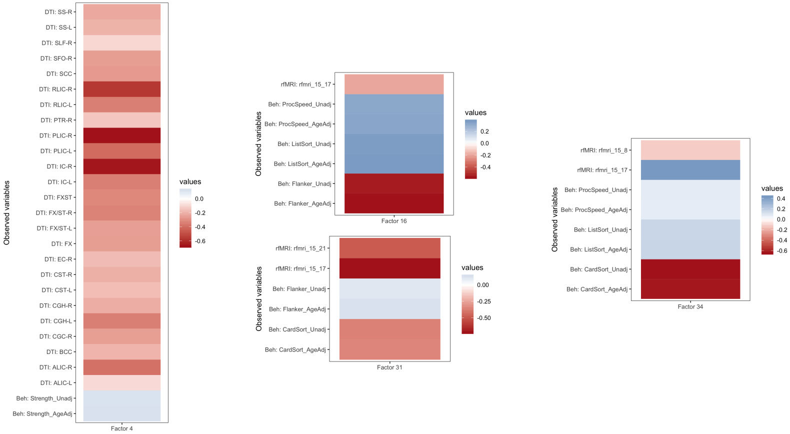
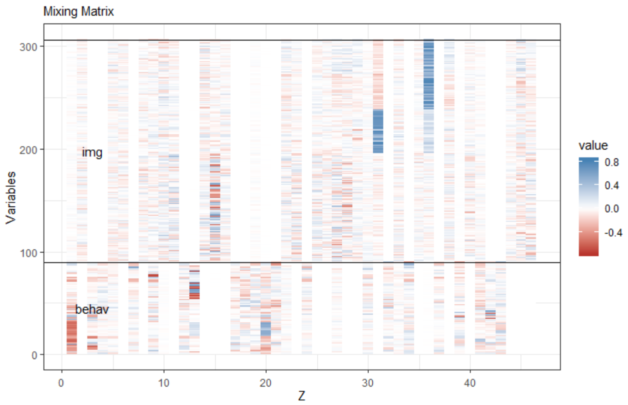


Figure: Heat map showing values of loadings for latent factors 4, 16, 31 and 34 which span both behavioural and imaging modalities

GFA Results

Imaging + behavioural data:

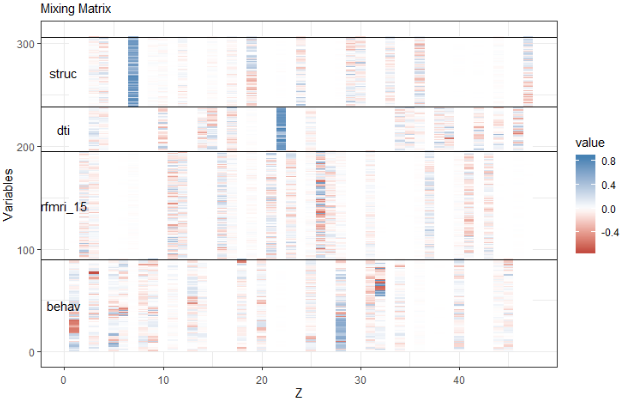
- 2 modalities
- 46 latent factors identified (initial # latent factors $= 50$)
- 9 latent factors active for both modes



GFA Results

Imaging + behavioural data:

- 4 modalities (DTI + rFMRI + structural(thickness) + behavioural)
- 47 latent factors identified (initial # latent factors $= 50$)
- 1 latent factor active for all four modes



Method Comparison

- BASS identified only 3 cross-mode latent factors compared to 9 with GFA (on 2-mode imaging + behavioral applications).
- Similar global sparsity results ($k = 50$ identified similar number of latent factors)
- BASS seems to impose more local and factor sparsity than GFA
- Both were tricky to implement (lack of documentation) and could handle limited amount of data

Conclusion & Extensions

- Able to recover both sparse and dense latent factors, and factors spanning modes
- Both BASS and GFA improve upon PCA-CCA-ICA approach
- Limitations:
    - highly exploratory - hard to interpret and validate
    - implementation and documentation
- Extensions:
    - linear modeling
    - non-linear latent factors

Questions

Thank you! Questions?