Modelling subject variability in large scale fMRI data

Christoffer Hjort[1], *Alessandro Montemurro*[1], *Edoardo Nemni*[1]

[1]*DTU Compute, Technical University of Denmark*

DTU Compute
Department of Applied Mathematics and Computer Science

Introduction

The main objective of this study is relate functional connectomes to behavioral traits using the subject specific scores obtained from a variant of the constrained PARAFAC2 model. The data used are taken from the **Human Connectome Project** (HCP), which is a project to construct a map of the complete structural and functional neural connections in vivo within and across individuals. The goal after fitting the variant of the PARAFAC2 model is studying if the extracted matrices can predict behavioral traits from the HCP in order to define how brain activity relates to behavioral data. In particular whether any specific patterns of brain connectivity are associated with personality defined using the **Five Factor Model (NEO-FFI)** by Costa & McCrae. The data analyzed are obtained using indipendent component analysis (ICA) parcellation of the resting-state fMRI data from HCP.

Methods

Consider a neuro-imaging matrix $\mathbf{X}^{(s)} \in \mathbb{R}^{V \times T}$, where V is the number of brain regions for subject s and T is the number of time samples.

The connectivity between voxels can me modelled with
(1) Baseline: $V \times V$ **covariance** matrix $\mathbf{X}^{(s)}\mathbf{X}^{(s)\top}$ (or with the *correlogram*, if the means of the voxels are subtracted);
(2) **low-rank decomposition models**.
For the second approach, **PARAFAC2** model can be used. In its general formulation, PARAFAC2 with K components has the form

$$\mathbf{X}^{(s)} = \mathbf{A}\mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)} + \mathbf{E}^{(s)}, \qquad s = 1, \ldots, S$$

For our purpose, we must constrain the decomposition:

- $\mathbf{P}^{(s)}$ is orthogonal;
- $\mathbf{A}$ is fixed to be the identity, such that we have one factor per brain region;
- $\mathbf{C}^{(s)}$ is diagonal and all the diagonal elements are positive.

$\mathbf{C}^{(s)}$, $\mathbf{R}$, $\mathbf{P}^{(s)}$, $\mathbf{Q}^{(s)}$ are estimated following an **Alternating Least Squares** (ALS) approach: alternatively, a certain loss function is minimized w.r.t one of the matrices, keeping the other fixed.

Making different assumptions on the noise, different cost functions and updates for the matrices are obtained:

- **No noise estimation:** The loss function is

$$\mathcal{L} = \sum_{s=1}^{S} ||\mathbf{X}^{(s)} - \mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)}||_F^2 \quad = \quad \sum_{s=1}^{S} \mathrm{tr}\left[(\mathbf{X}^{(s)} - \mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)})(\mathbf{X}^{(s)} - \mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)})^\top\right]$$

and the corresponding updating rules are

i) $\mathbf{P}^{(s)^{new}} = \mathbf{V}^{(s)}\mathbf{U}^{(s)\top} = (\mathbf{U}^{(s)}\mathbf{V}^{(s)\top})^\top, \quad \text{for } \mathbf{U}^{(s)}\mathbf{\Sigma}^{(s)}\mathbf{V}^{(s)\top} = \text{s.v.d}\left(\mathbf{X}^{(s)\top}\mathbf{C}^{(s)}\mathbf{R}\right)$

ii) $[\mathbf{C}^{(s)}]_{ii} = \dfrac{[\mathbf{R}\mathbf{P}^{(s)}\mathbf{X}^{(s)\top}]_{ii}}{[\mathbf{R}\mathbf{R}^\top]_{ii}}$ *iii)* $\mathbf{R}^{new} = \dfrac{\sum_s \mathbf{C}^{(s)\top}\mathbf{X}^{(s)}\mathbf{P}^{(s)\top}}{\sum_s \mathbf{C}^{(s)\top}\mathbf{C}^{(s)}}$

- **Noise estimation:** Define $\mathbf{Q}^{(s)} = \mathrm{diag}\left(\frac{1}{\sigma_{(v,t)}}\right)$, $v = 1 \ldots V$. The i^{th} diagonal element $\sigma_{i,t}$ represents the standard deviation of the error for region i at time t. Suppose the noise is zero-mean Gaussian with covariance matrix $\mathbf{Q}^{(s)}\mathbf{Q}^{(s)\top}$. Matrices have to be estimated again with ALS, following a maximum likelihood approach.
The log-likelihood function is

$$\log\mathcal{L} = \sum_{s=1}^{S}\sum_{v=1}^{V}\sum_{t=1}^{T}\left[-\frac{1}{2}\log(2\pi\sigma_{v,s}^2) - \frac{1}{2\sigma_{v,s}^2}\left(\mathbf{X}^{(s)}(v,t) - (\mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)})(v,t)\right)^2\right]$$

The updates are

i) $\mathbf{Q}^{(s)^{new}} = \left[\frac{1}{T}\sum_{t=0}^{T}\left(\mathbf{X}^{(s)}(v) - (\mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)})(v)\right)^2\right]^{-\frac{1}{2}}$

ii) $\mathbf{P}^{(s)^{new}} = \mathbf{V}^{(s)}\mathbf{U}^{(s)\top} = \left(\mathbf{U}^{(s)}\mathbf{V}^{(s)\top}\right)^\top$
where $\mathbf{U}^{(s)}\mathbf{\Sigma}^{(s)}\mathbf{V}^{(s)\top} = \text{s.v.d}\left(\mathbf{X}^{(s)\top}\mathbf{Q}^{(s)}\mathbf{Q}^{(s)\top}\mathbf{C}^{(s)}\mathbf{R}\right)$

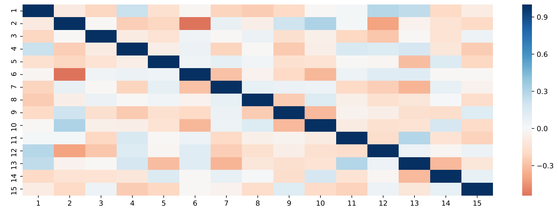
iii) $\mathbf{C}^{(s)^{new}}_{ii} = \dfrac{[\mathbf{R}\mathbf{P}^{(s)}\mathbf{X}^{(s)\top}\mathbf{Q}^{(s)}\mathbf{Q}^{(s)\top}]_{ii}}{[\mathbf{Q}^{(s)}\mathbf{Q}^{(s)\top}\mathbf{R}\mathbf{R}^\top]_{ii}}$ *iv)* $\mathbf{R}^{new} = \dfrac{\sum_s \mathbf{C}^{(s)\top}\mathbf{Q}^{(s)}\mathbf{Q}^{(s)\top}\mathbf{X}^{(s)}\mathbf{P}^{(s)\top}}{\sum_s \mathbf{C}^{(s)\top}\mathbf{Q}^{(s)}\mathbf{Q}^{(s)\top}\mathbf{C}^{(s)}}$

The correlogram $\mathbf{X}^{(s)}\mathbf{X}^{(s)\top}$ and the matrix $\mathbf{C}^{(s)}$ are fed into a regularized linear regression algorithm for **prediction** and into a regularized logistic regression for **classification**.

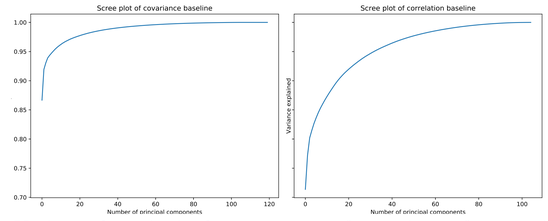
Experiments

For the experiments $\mathbf{X}^{(s)} \in \mathbb{R}^{15\times4800}$. That is, 15 brain regions each sampled with 4800 time points for any subject, and the data is comprised of 1003 subjects. This high dimensional data is reduced by modelling it with the models specified in the Methods section.

- **Covariance baseline:** Covariance of the 15 brain regions for each subject, i.e. the upper triangle and diagonal of the covariance matrix. This gives 120 features for each subject.
- **Correlation baseline:** Correlation of the 15 brain regions for each subject, i.e. the upper triangle of the correlation matrix. This gives 105 features for each subject.



Correlogram of the 15 brain regions for a random subject.



Variance by principal components of covariance features (left), and variance by principal components of correlation features (right).

- **PARAFAC2 no noise estimation:** Decomposes each $\mathbf{X}^{(s)}$ into the matrices $\mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)}$. The diagonal elements of $\mathbf{C}^{(s)}$ contains scores for each brain region of subject s, and these scores are used as features. These extracted features $\mathbf{C}^{(s)}$ gives 15 features for each subject.
- **PARAFAC2 noise estimation:** Decomposes each $\mathbf{X}^{(s)}$ into the matrices $\mathbf{C}^{(s)}\mathbf{R}\mathbf{P}^{(s)} + \mathbf{Q}^{(s)}$. The diagonal elements of $\mathbf{C}^{(s)}$ contains scores for each brain region of subject s, while $\mathbf{Q}^{(s)}$ contains region and subject specific noise. Both $\mathbf{C}^{(s)}$ and $\mathbf{Q}^{(s)}$ are used as features, and each gives 15 features for each subject.

Using Ridge Regression, the extracted features from each model are used to classify or predict behavioural traits about the subjects.

Results

	Covariance		Correlation		PARAFAC2	
	No PCA	PCA	No PCA	PCA	No noise	Noise
Gender (Classification)	82%	82%	82%	82%	70%	70%
Age (Classification)	44%	48%	46%	45%	45%	44%
Personality Extroversion (Regression)	1.06	1.05	1.05	1.04	1.02	1.02

Conclusion

- The scores obtained by the variant of the PARAFAC2 model can be used to predict behavioral traits from the HCP data.
- The classification based on the PARAFAC2 scores performs better than a random prediction, but worse than the baseline.
- The classification using PARAFAC2 scores performs better than the majority class classification. Regression using the PARAFAC2 scores is not better than predicting using the mean.
- Using the variant of the PARAFAC2 model cannot guarantee a better result compared to an easy approach as the correlogram.

References

[1] Madsen, K. H., Churchill, N. W., & Mrup, M. (2016). Quantifying functional connectivity in multi-subject fMRI data using component models.

[2] Kiers, H. A. L., Ten Berge, J. M. F., & Bro, R. (1999). PARAFAC2-Part I. A direct fitting algorithm for the PARAFAC2 model.

[3] Nielsen, S. F. V., Mrup, M. (2018). Modeling Subject Variability in Large Scale Functional Magnetic Resonance Imaging Data