Angela Andreella, and Livio Finos

angela.andreella@unive.it

Procrustes analysis for high dimensional

Introduction

The Procrustes problem is aimed at matching matrices using similarity transformations by minimizing their Frobenius distance. It allows comparison of matrices with dimensions defined in an arbitrary coordinate system.

The extension from the spatial context to a more general and high-dimensional one is a theoretical challenge that needs adequate attention:

- **Ill-posed problem** $\rightarrow$ any rotation completely changes the interpretation of the results;
- **Polynomial time complexity** with respect the number of variables to be aligned.

Perturbation model

Let $\{\boldsymbol{X}_i \in \mathbb{R}^{n\times m}\}_{i=1,\dots,N}$ be a set of matrices to be aligned:

$$\boldsymbol{X}_i = \alpha_i(\boldsymbol{M}+\boldsymbol{E}_i)\boldsymbol{R}_i^\top + \boldsymbol{1}_n^\top \boldsymbol{t}_i \quad \text{subject to} \quad \boldsymbol{R}_i \in \mathcal{O}(m),$$

where $\boldsymbol{E}_i \sim \mathcal{MN}_{n,m}(0, \boldsymbol{\Sigma}_n, \boldsymbol{\Sigma}_m)$.
Consider the SVD of $\boldsymbol{X}_i^\top \boldsymbol{\Sigma}_n^{-1} \boldsymbol{M} \boldsymbol{\Sigma}_m^{-1} = \boldsymbol{U}_i^\top \boldsymbol{D}_i \boldsymbol{V}_i$:

- $\hat{\boldsymbol{R}}_i = \boldsymbol{U}_i^\top \boldsymbol{V}_i$;
- $\hat{\alpha}_{i\boldsymbol{R}_i} = ||\boldsymbol{\Sigma_m}^{1/2} \hat{\boldsymbol{R}}_i^\top \boldsymbol{X}_i \boldsymbol{\Sigma}_n^{-1/2}||_F^2 / tr(\boldsymbol{D}_i)$.

Problem:

- If $\boldsymbol{Z} \in \mathcal{O}(m)$, then $\{\hat{\boldsymbol{R}}_i \boldsymbol{Z}\}_{i=1,\dots,N}$ still valid maximum likelihood solutions.
- Consider $\boldsymbol{X}_i \in \mathbb{R}^{n\times m}$, if $n < m$, then the ML estimate for $\boldsymbol{R}_i$ is not unique.

ProMises model

Consider the perturbation model, we impose to $\boldsymbol{R}_i$ the von Mises-Fisher distribution:

$$f(\boldsymbol{R}_i) = C(\boldsymbol{F}, k)\exp(tr(k\boldsymbol{F}^\top \boldsymbol{R}_i))$$

where $C(\boldsymbol{F}, k)$ is a normalizing constant, $\boldsymbol{F} \in \mathbb{R}^{m\times m}$ is the location matrix parameter, and $k \in \mathbb{R}^+$ is the concetration parameter.

- the posterior distribution $f(\boldsymbol{R}_i | k, \boldsymbol{F}, \boldsymbol{X}_i)$ is a conjugate distribution to the von Mises–Fisher prior distribution with location posterior parameter $F^\star = \boldsymbol{F}^\star = \boldsymbol{X_i}^\top \boldsymbol{\Sigma}_n^{-1} \boldsymbol{M} \boldsymbol{\Sigma}_m^{-1} + k\boldsymbol{F}$;
- Consider the SVD of $\boldsymbol{X}_i^\top \boldsymbol{\Sigma}_n^{-1} \boldsymbol{M} \boldsymbol{\Sigma}_m^{-1} + k\boldsymbol{F} = \boldsymbol{U}_i^\top \boldsymbol{D}_i \boldsymbol{V}_i$:
  - $\hat{\boldsymbol{R}}_i = \boldsymbol{U}_i^\top \boldsymbol{V}_i$;
  - $\hat{\alpha}_{i\boldsymbol{R}_i} = ||\boldsymbol{\Sigma_m}^{1/2} \hat{\boldsymbol{R}}_i^\top \boldsymbol{X}_i \boldsymbol{\Sigma}_n^{-1/2}||_F^2 / tr(\boldsymbol{D}_i)$.
- If $\boldsymbol{F}$ has full rank, the MAP estimates for $\boldsymbol{R}_i$ are unique.

Efficient ProMises Model

The Efficient ProMises approach projects the matrices $\boldsymbol{X}_i$ into an n-lower-dimensional space using the semi-orthogonal transformation from the thin SVD of $\boldsymbol{X}_i$, i.e., $\boldsymbol{Q}_i \in \mathbb{R}^{m\times n}$.

$$\max_{\boldsymbol{R}_i \in \mathcal{O}(m)} tr(\boldsymbol{R}_i^\top \boldsymbol{X}_i^\top \boldsymbol{\Sigma_n}^{-1} \boldsymbol{X}_j \boldsymbol{\Sigma}_m^{-1} + k\boldsymbol{F}) =$$

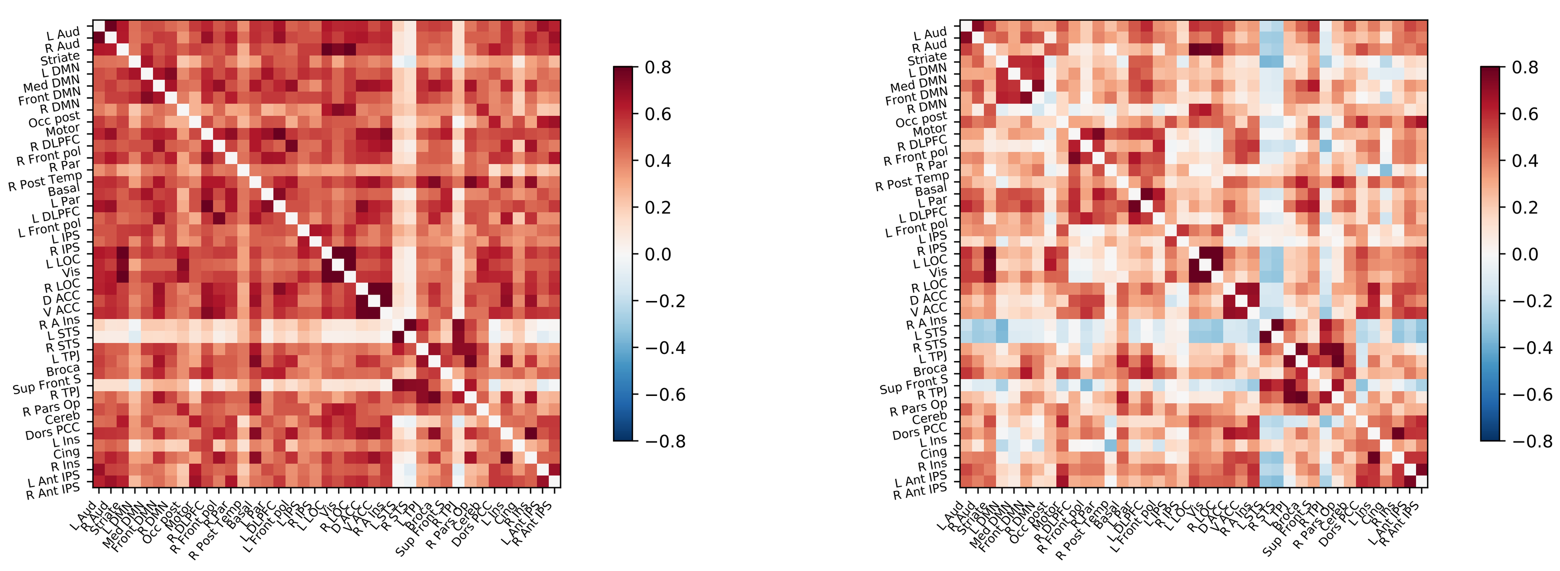
$$\max_{\boldsymbol{R}_i^\star \in \mathcal{O}(n)} tr\{\boldsymbol{R}_i^{\star\top} (\boldsymbol{Q}_i^\top \boldsymbol{X}_i^\top \boldsymbol{\Sigma}_n^{-1} \boldsymbol{X}_j \boldsymbol{\Sigma}_m^{-1} \boldsymbol{Q}_j^\top + k\boldsymbol{F}^\star)\},$$

We align the $\{\boldsymbol{X}_i \boldsymbol{Q}_i\}_{i=1,\dots,N}$ matrices instead of $\{\boldsymbol{X}_i\}_{i=1,\dots,N}$, and then we project back the aligned matrices by $\boldsymbol{Q}_i^\top$.

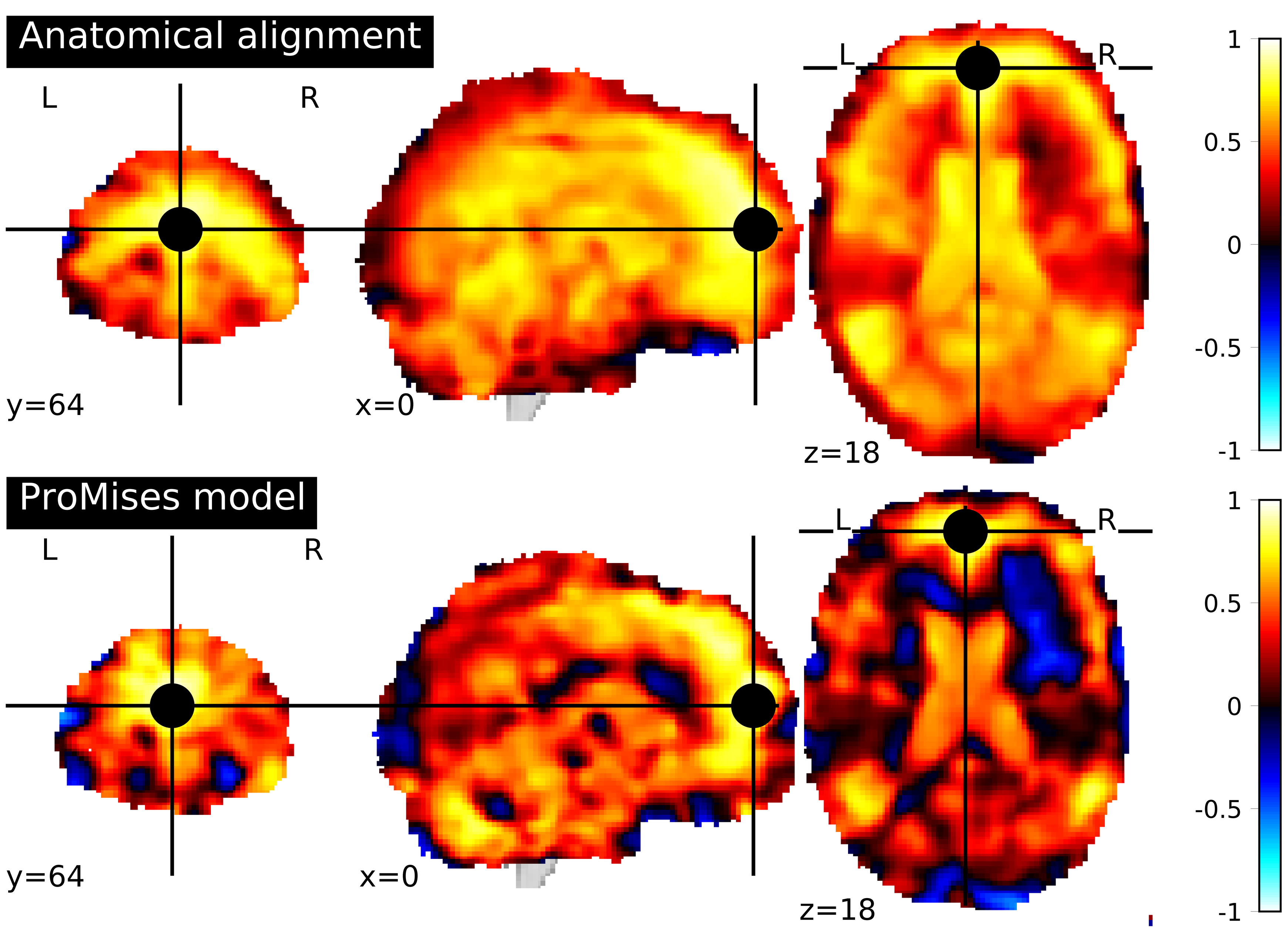
fMRI Data Application

The study consists of neural activations of 18 subjects passively listening to vocal (i.e., speech) and nonvocal sounds.
We performed region of interest and seed-based correlation analysis. The seed-based correlation map shows the level of functional connectivity between a seed and every voxel in the brain, whereas the region of interest analysis expresses the functional correlation between predefined regions of interest coming from a standard atlas.

fMRI Data Application



fMRI Data Application



Take home messages

The issues of the perturbation model —non-uniqueness, critical interpretation, and inapplicability when $n << m$— are completely surpassed by our Bayesian extension:

- Unique and interpretable orthogonal transformations;
- Its efficient approach extends the applicability to high-dimensional data
- the matrix $\boldsymbol{F}$ addresses the estimate of the optimal rotations, which is favorable in the analysis of fMRI data because, in this context, the variables have a spatial anatomical location.

References