Hyper Alignment

Lixiang
May 21, 2018

Outline

- Motivation
- Method
  - Hyper Alignment
  - Canonical Correlation Analysis and others
  - Our method
- Results
- Discussion & future work

Motivation

Data	Stimulus 1	Stimulus 2	Stimulus 3	Stimulus ...	Stimulus m
Sub 1	X(1, 1)	X(1, 2)	X(1, 3)	...	X(1, m)
Sub 2	X(2, 1)	X(2, 2)	X(2, 3)	...	X(2, m)
Sub ...	...	...	...	...	...
Sub n	X(n, 1)	X(n, 2)	X(n, 3)	...	X(n, m)

Expense = size of total stimulus * number of total subjects

Story data-sets: subjects * time points * voxels

Motivation

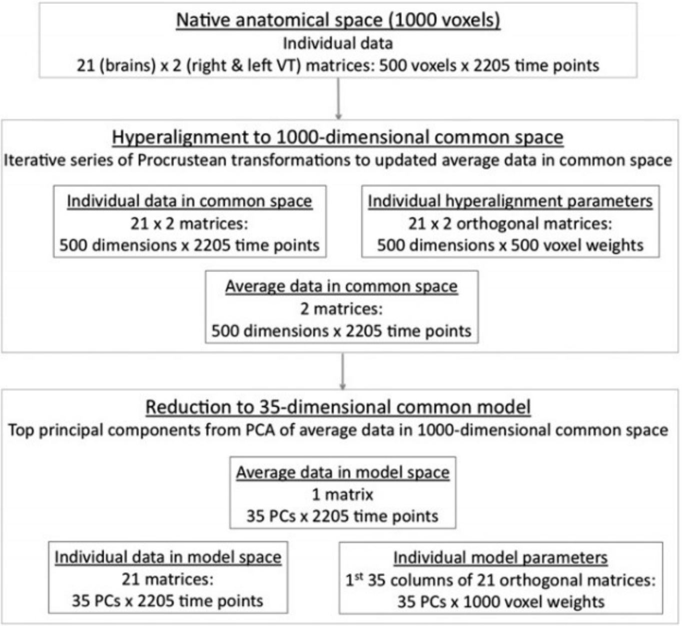
Data	Stimulus 1	Stimulus 2	Stimulus 3	Stimulus ...	Stimulus m
Sub 1	X(1, 1)	X(1, 2)	X(1, 3)		
Sub 2	X(2, 1)		X(2, 3)		X(2, m)
Sub ...	...		...		
Sub n		X(n, 2)		…	X(n, m)

Expense = size of partial stimulus * number of partial subjects

Hypothesis

- Reading a narrative story (or viewing a natural movie) evokes local brain responses that show synchrony across subjects in a large expanse of cortex.
- It should be possible to align patterns of neural response (fMRI) across subjects into a common high-dimensional space.

Hyper Alignment



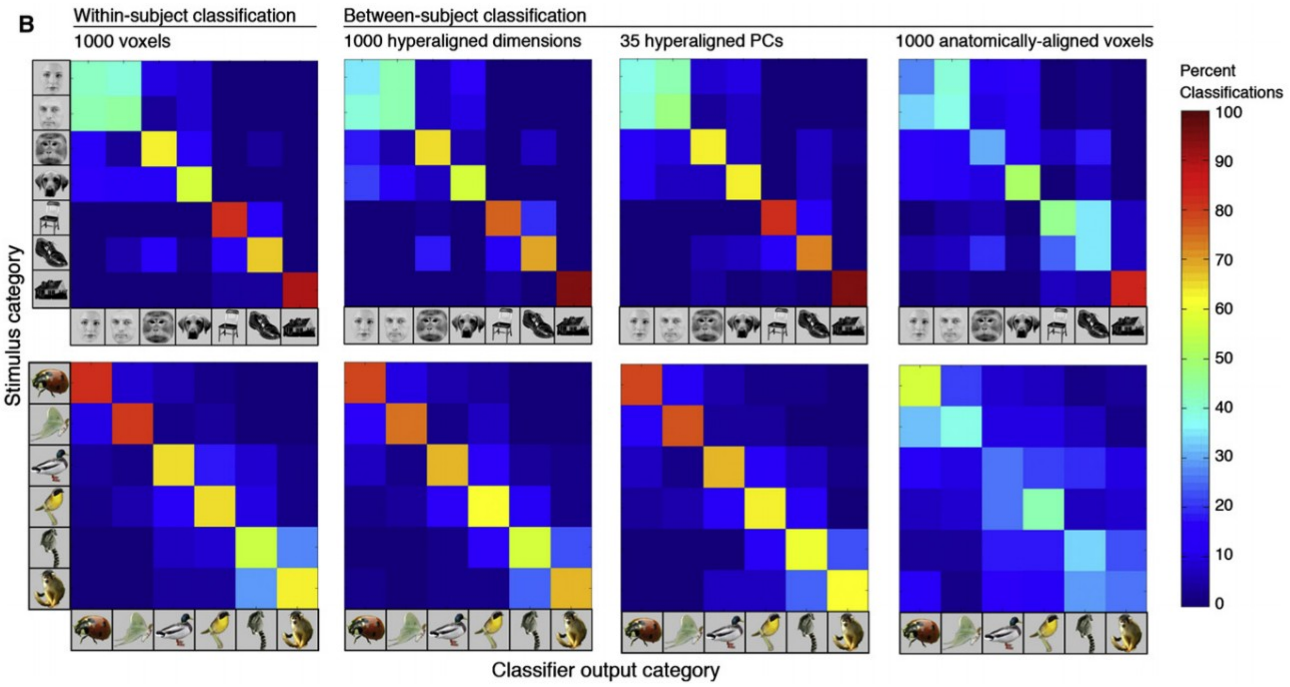
MLE problem:

$$minimize \sum_{i<j} \| X_i R_i - X_j R_j \|_F^2$$

Subject to $R_k^T R_k^T = I$

Matrix R: Procrustean transformations, orthogonal transformation (rotation and reflections) that minimizes the Euclidean distance between two sets of paired vectors.

Hyper Alignment



Classification Accuracy: 60%~70%

Canonical Correlation Analysis

[CCA]

$$\begin{aligned}\text{minimize } & \textstyle\sum_{i<j}\|\mathbf{X}_i\mathbf{R}_i - \mathbf{X}_j\mathbf{R}_j\|_F^2 \\ \text{subject to } & \mathbf{R}_k^T\mathbf{X}_k^T\mathbf{X}_k\mathbf{R}_k = \mathbf{I}\,.\end{aligned}$$

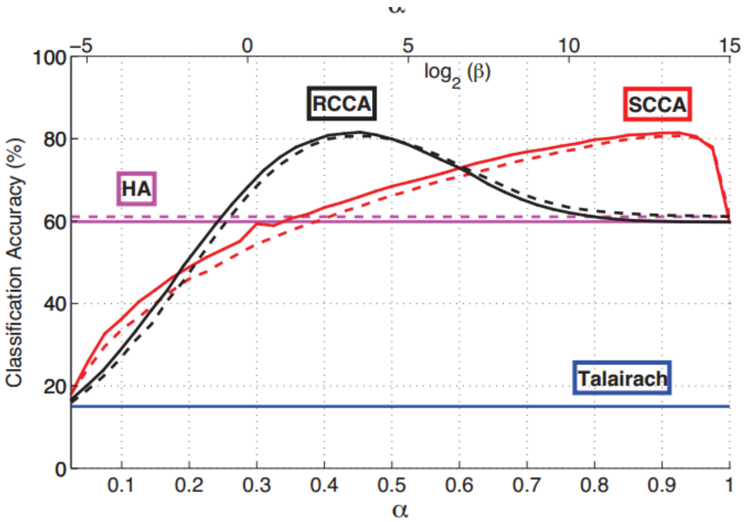
[RCCA]

$$\begin{aligned}\text{minimize } & \textstyle\sum_{i<j}\|\mathbf{X}_i\mathbf{R}_i - \mathbf{X}_j\mathbf{R}_j\|_F^2 \\ \text{subject to } & \mathbf{R}_k^T(\mathbf{X}_k^T\mathbf{X}_k + \beta\mathbf{I})\mathbf{R}_k = \mathbf{I}\end{aligned}$$

[SCCA]

$$\begin{aligned}\text{minimize } & \textstyle\sum_{i<j}\|\mathbf{X}_i\mathbf{R}_i - \mathbf{X}_j\mathbf{R}_j\|_F^2 \\ \text{subject to } & \mathbf{R}_k^T((1-\alpha)\mathbf{X}_k^T\mathbf{X}_k + \alpha\mathbf{I})\mathbf{R}_k = \mathbf{I}\end{aligned}$$

[KCCA] $minimize \sum_{i<j}\|f(X_i)R_i - f(X_j)R_j\|_F^2$



Our method

$$argmin_{R_i,G} \sum_i \|R_i X_i - G\|_F^2$$

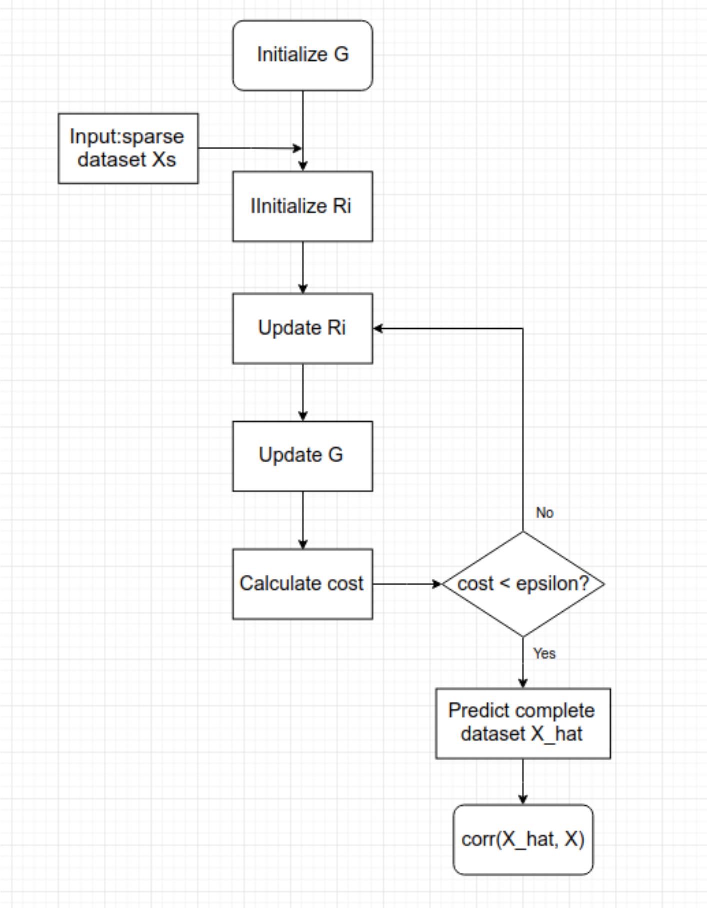
Subject to $R_i R_i^T = I$

$$cost = \sum_i \|R_i X_i - G\|_F^2 + \lambda \sum_i \|R_i R_i^T - T\|_F^2$$

Initialize G: d largest eigenvectors

$$XXT = \sum_i (X_s^T X_s)$$
$$W, V = EVD(XXT)$$
$$G = \sqrt{W[:d]} V[:d]$$

Initialize R: $R_i = G\, Xs_i$



Update Ri: $dR_i = (R_i Xs_i - G) Xs_i^T + 2\lambda (R_i R_i^T - I) R_i$

$$R_{post} = R_{prior} - \beta dR$$

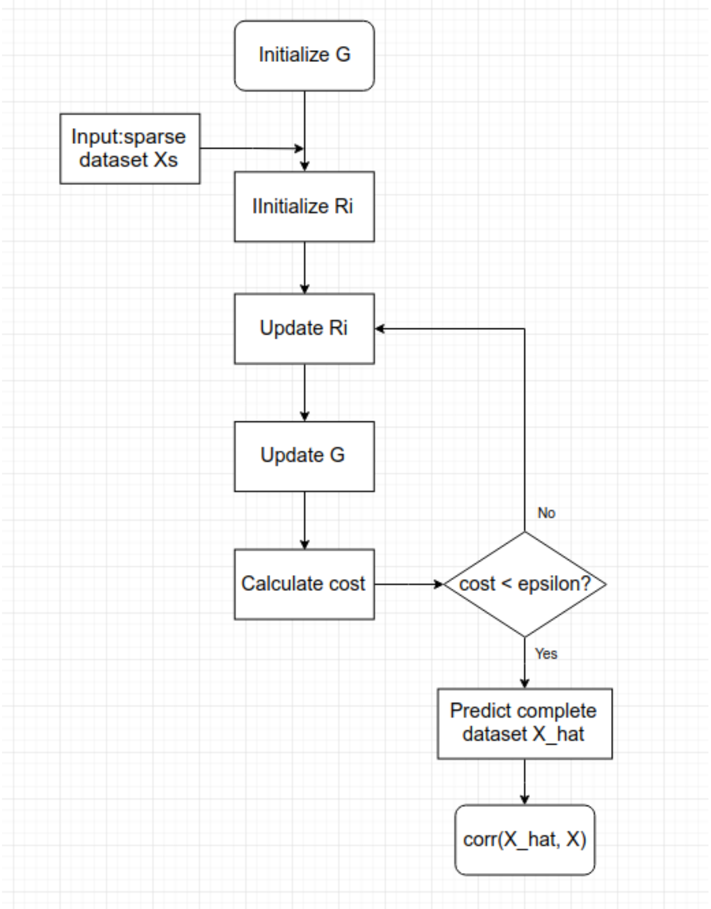
Update G: $G_{tmp} = \sum_i (R_i Xs_i)$

$$G_{post} = \alpha G_{tmp} + (1-\alpha) G_{prior}$$

Predict complete dataset:

$$\hat{X}_i = linear\ regression(R_i X - G)$$

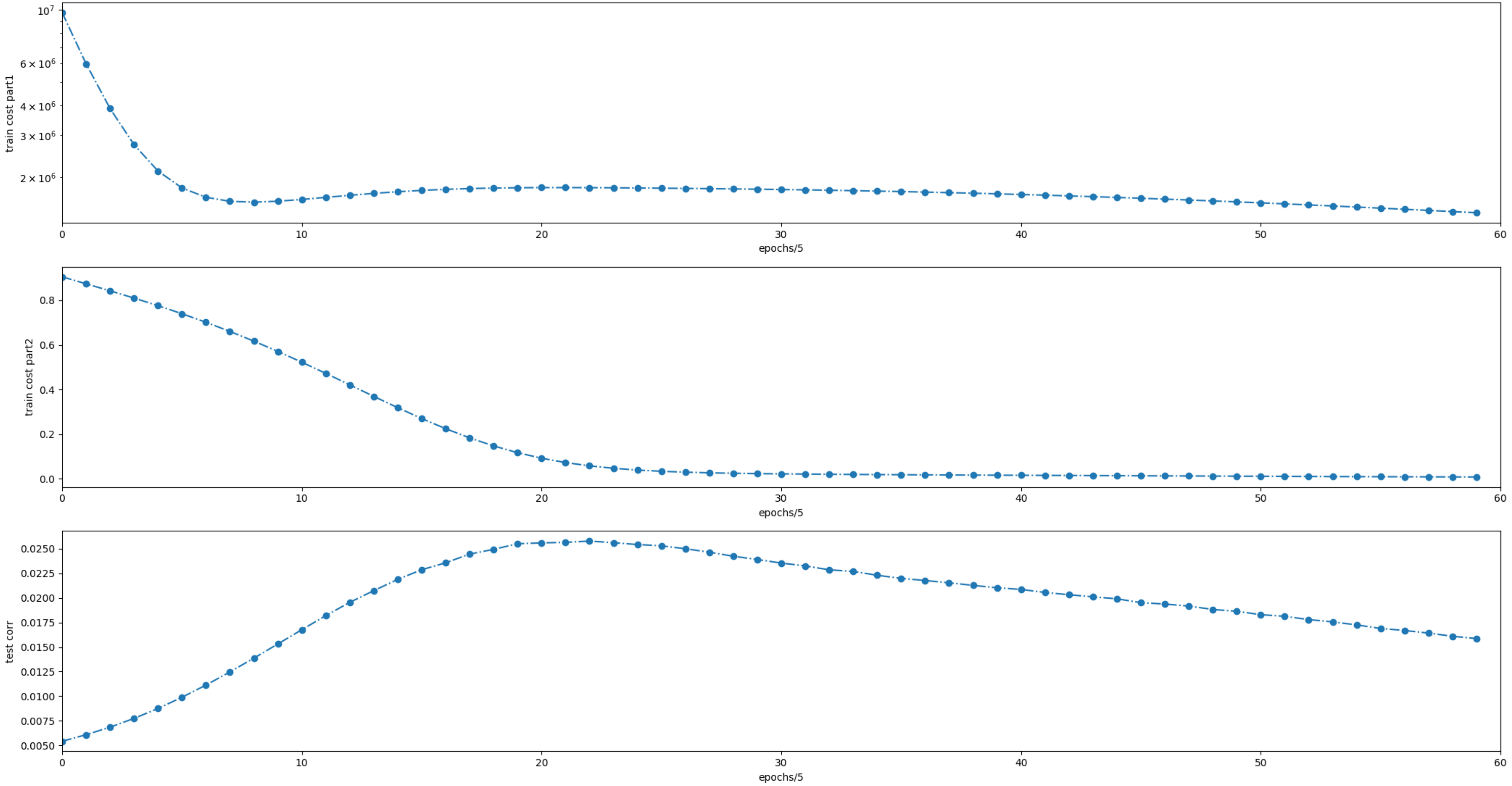
$$corr(X_i, \hat{X}_i)$$



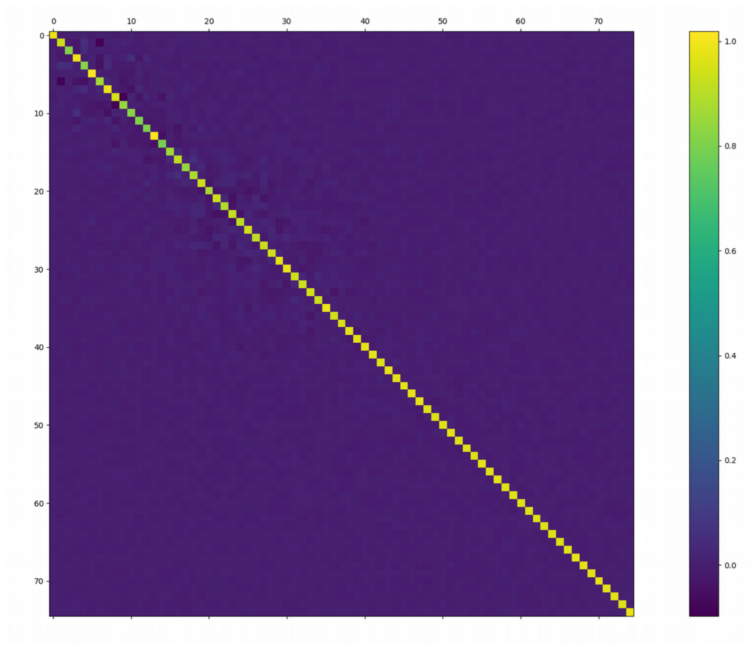
Results: how to sparse full datasets?

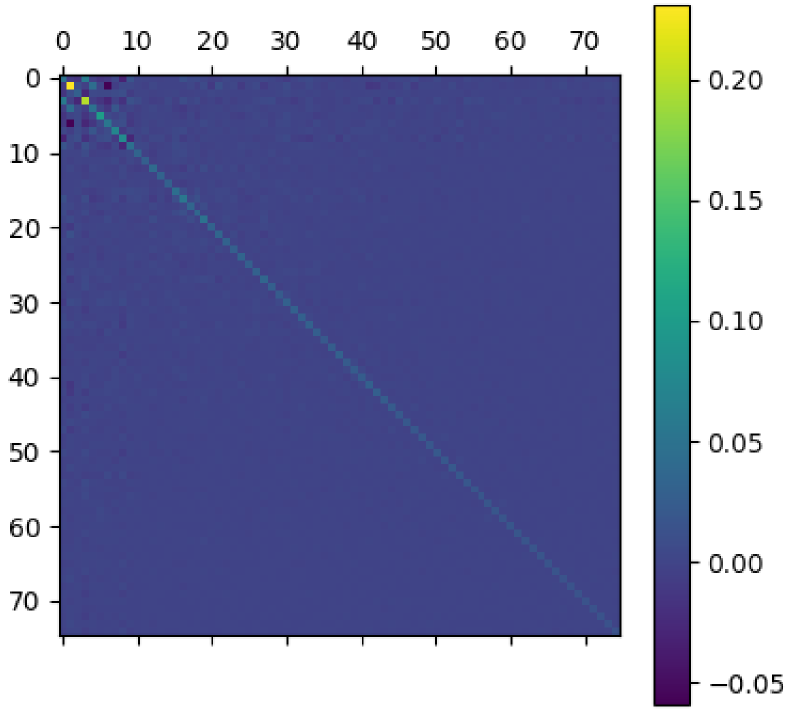
data	Story 1	Story 2	Story 3	Story 4	Story 5	Story 6	Story 7	Story 8	Story 9	Story 10
BG										
JG	Common shared stories									
NNS0										

Results: Learning Curve

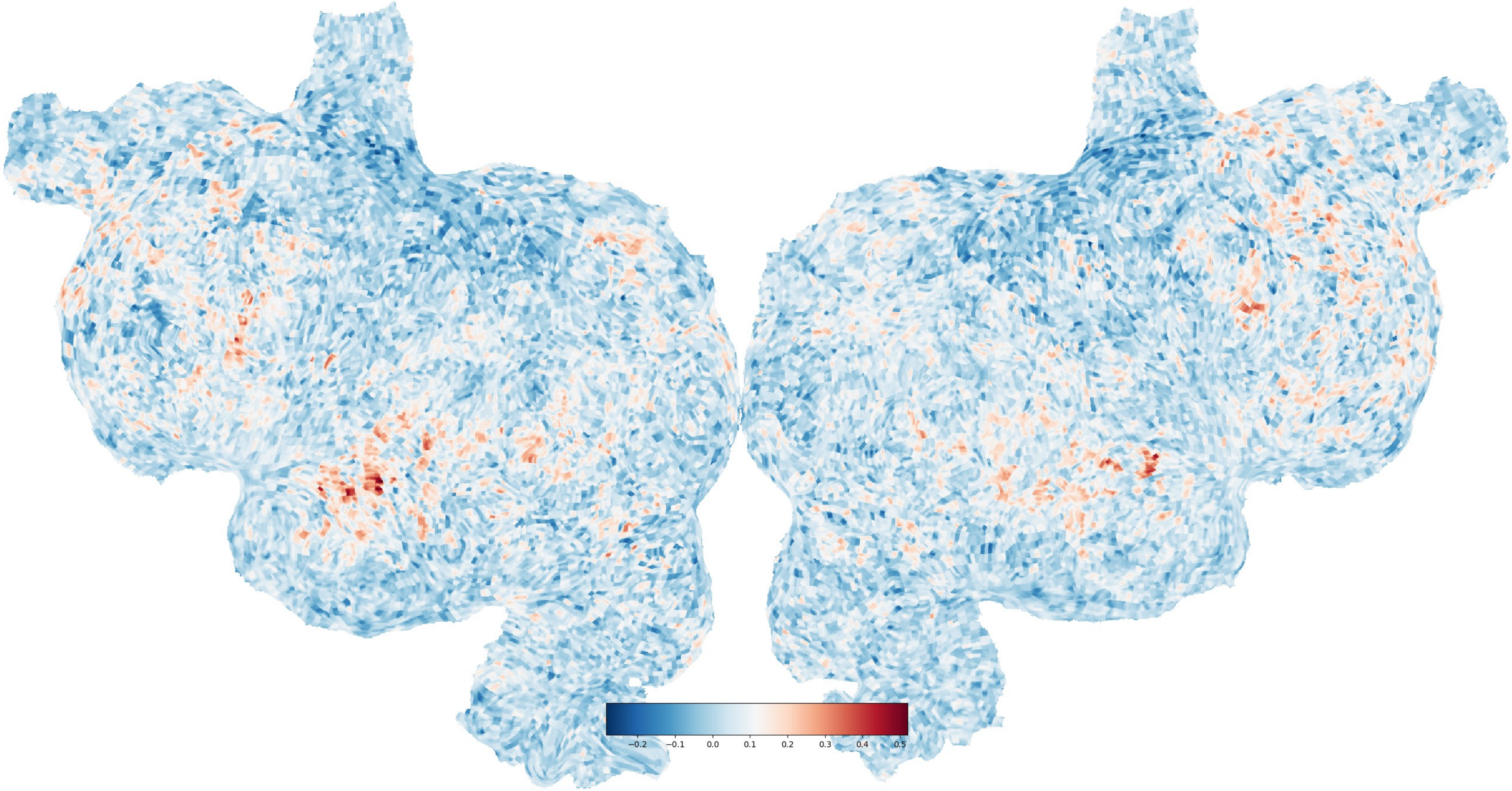


Results: Check Normalized orthogonal R

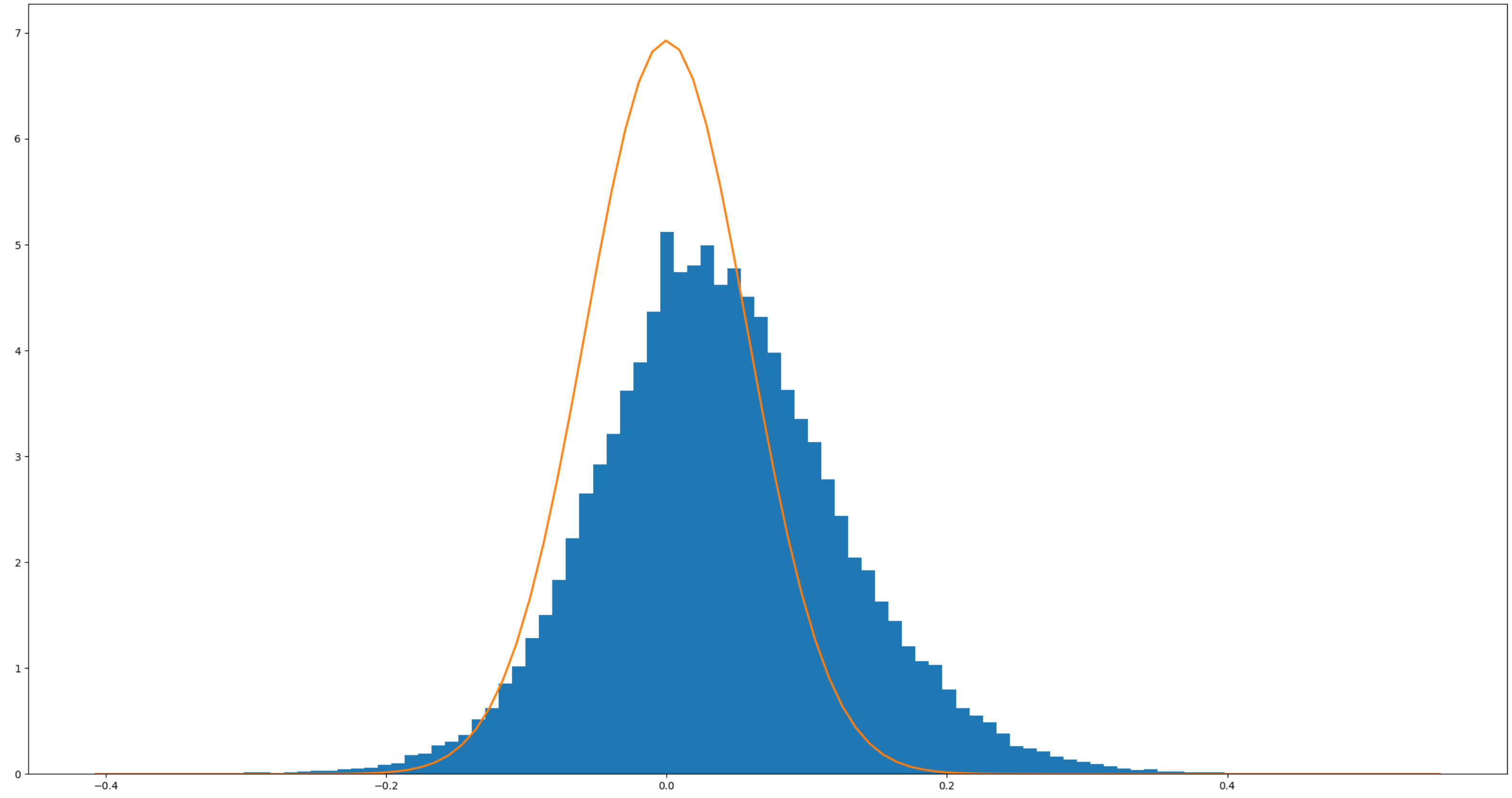




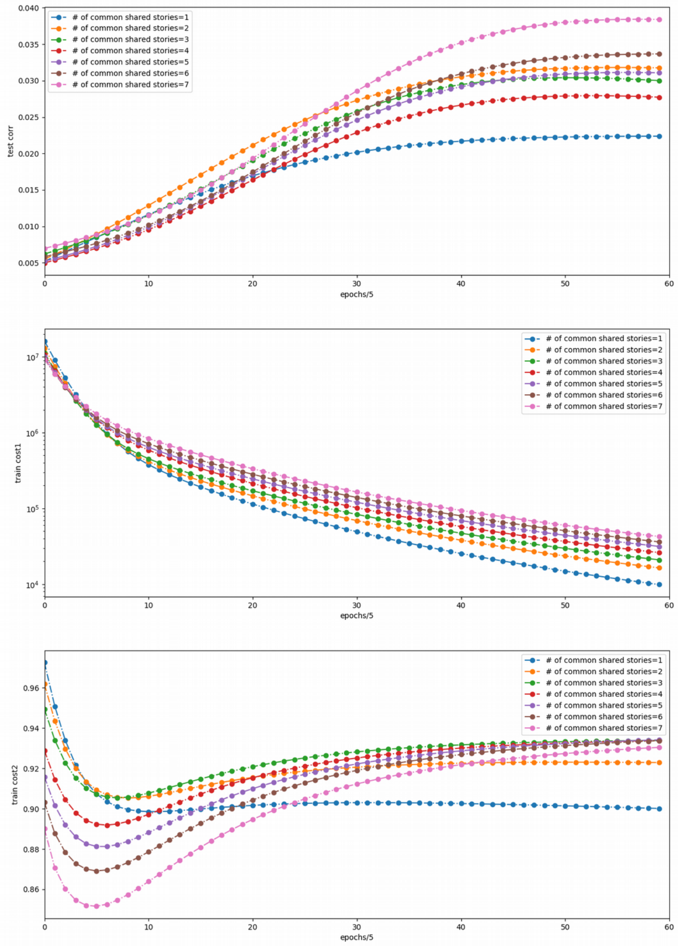
Results: corr plot



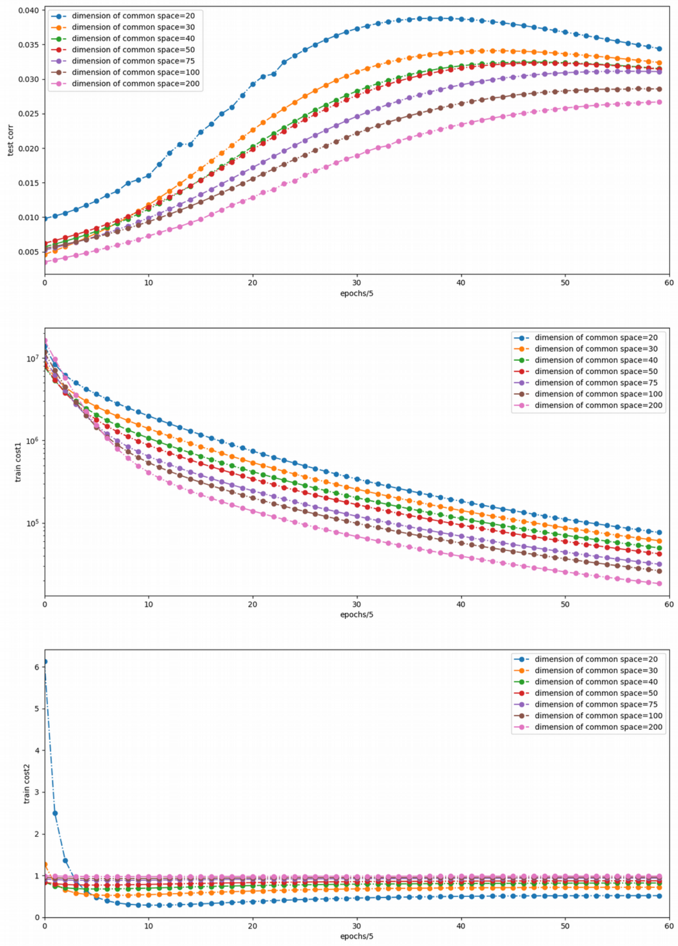
Results: corr hist



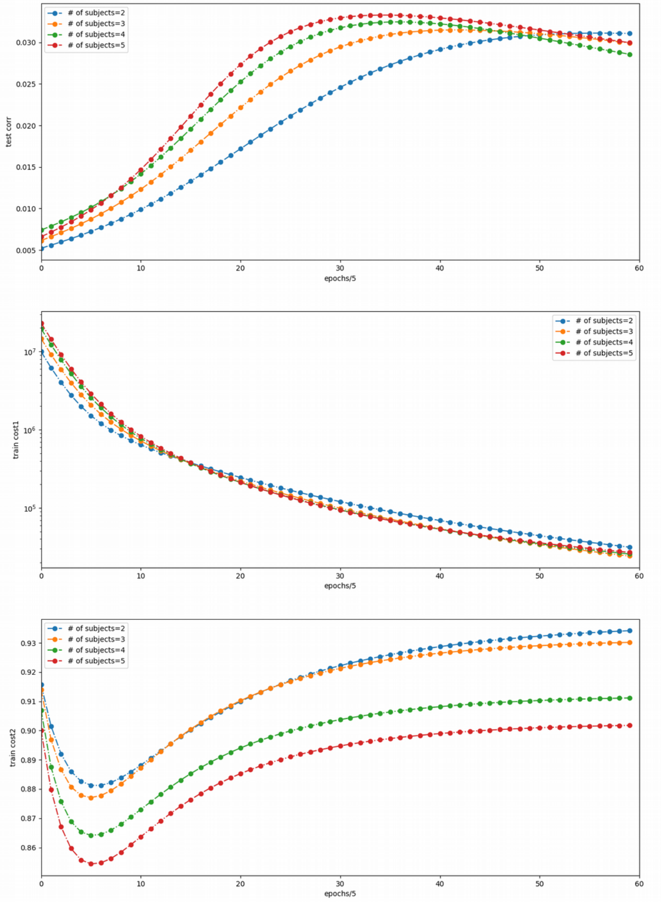
Results: # of common shared stories



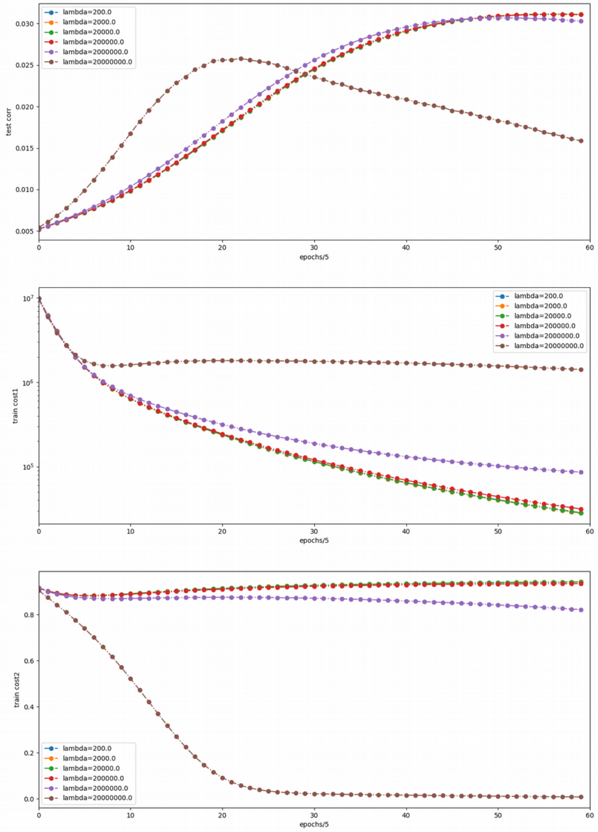
Results: dimension of common subspace

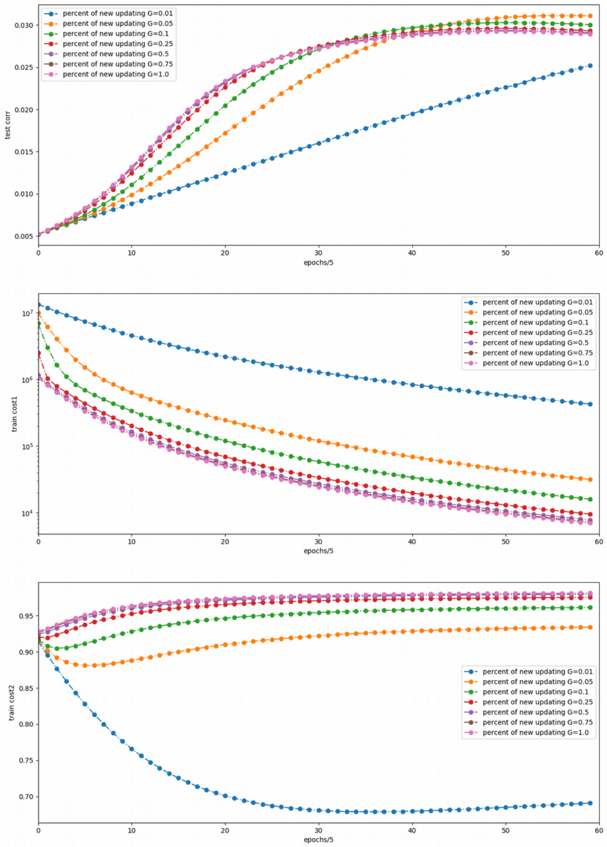


Results: # of subjects



Results: lambdas and alphas





Discussion

- Better way to evaluate our full-completed dataset?
- Matrix G weighted by data quality?
- What’s the more/most efficient way to assign the sparse dataset?

- Ref:
  - A Common, High-Dimensional Model of the Representational Space in Human Ventral Temporal Cortex. James V. Haxby, Peter J. Ramadge
  - Regularized Hyperalignment of Multi-Set fMRI Data. Hao Xu, Peter J. Ramadge
  - Regularized Kernel Canonical Correlation Analysis in Python and Its Applications to Neuroimaging. Natalia Y. Bilenko, Jack L. Gallant