Handling correlated and repeated measurements with the smoothed multivariate square-root Lasso

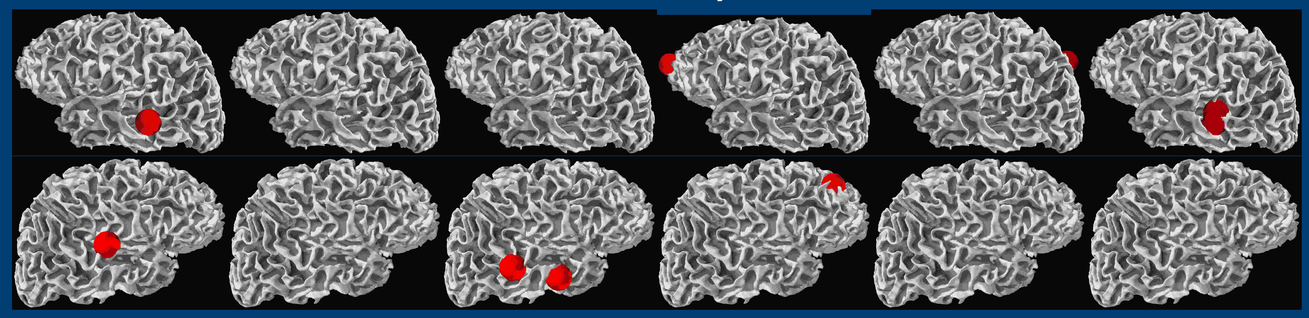
Q. Bertrand[1] M. Massias[1] A. Gramfort[1] J. Salmon[2]

[1]Inria, Univ. Paris Saclay, [2]IMAG, Univ. Montpellier, CNRS

We propose a **convex** concomitant estimator to **jointly** estimate the regression coefficients and the covariance matrix in **high dimensional** linear regression with **correlated Gaussian noise**.

Our estimator **performs better** than existing ones on **synthetic and real data**.

Real data, auditory stimulation



CLaR (ours) SGCL MLER MLE MRCER MTL

arXiv

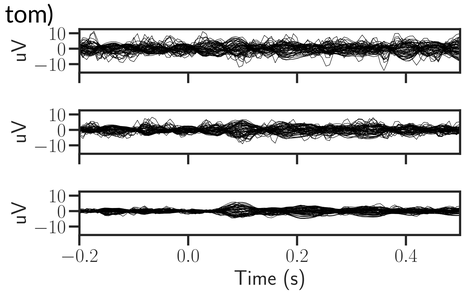
Take a picture to download the full paper

Introduction

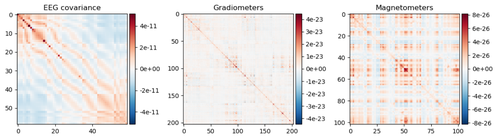
- M/EEG data are **very noisy** (SNR=1)
- it is thus customary to make **several repetitions** of the same experiment
- in order to **average** the signals and increase the signal to noise ratio

Real EEG data

- # of repetitions 5 (top), 10 (middle), 50 (bottom)



- M/EEG data are contaminated with **correlated Gaussian noise**:



Model and notations

Linear Multi-Task setting with correlated Gaussian noise:

- n : # of sensors
- p : # of features
- q : # of tasks/time points
- $X \in \mathbb{R}^{n \times p}$: design matrix
- $B \in \mathbb{R}^{p \times q}$: regression coefficients
- $S \in \mathbb{R}^{n \times n}$: sqrt of the covariance matrix
- $E^{(l)} \in \mathbb{R}^{n \times q}$ noise with i.i.d. normal entries
- $Y^{(l)} \in \mathbb{R}^{n \times q}$: signals

Model: $Y^{(l)} = XB^* + S^* E^{(l)}$, $\forall l \in [r]$

- $\bar{Y} = \frac{1}{r} \sum_l Y^{(l)} \in \mathbb{R}^{n \times n}$ mean of the signals

Our approach: use repetitions

- Concomitant Lasso with Repetitions (CLaR)

$$(\hat{B}, \hat{S}) \in \underset{\substack{B \in \mathbb{R}^{p \times q} \\ S \in \mathbb{S}^n_{++}, S \succeq \underline{\sigma}}}{\arg\min} \frac{\sum_{l=1}^{r} \left\| Y^{(l)} - XB \right\|^2_{S^{-1}}}{2nqr} + \frac{\mathrm{Tr}(S)}{2n} + \lambda \|B\|_{2,1}$$

Prev. : use the mean

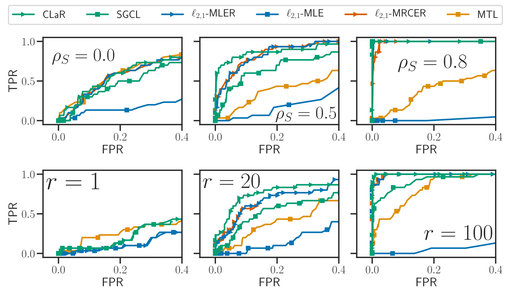
- Multi-Task Lasso (MTL, [1])

$$\hat{B} \in \underset{B \in \mathbb{R}^{p \times q}}{\arg\min} \frac{1}{2nq} \left\| \bar{Y} - XB \right\|^2 + \lambda \|B\|_{2,1}$$

- SGCL ([2])

$$(\hat{B}, \hat{S}) \in \underset{\substack{B \in \mathbb{R}^{p \times q} \\ S \in \mathbb{S}^n_{++}, S \succeq \underline{\sigma}}}{\arg\min} \frac{\left\| \bar{Y} - XB \right\|^2_{S^{-1}}}{2nq} + \frac{\mathrm{Tr}(S)}{2n} + \lambda \|B\|_{2,1}$$

More experiments



Ref. and acknowledgements

Code: `https://github.com/QB3/CLaR`

ERC Starting Grant SLAB ERC-YStG-676943.

*

References

[1] G. Obozinski, B. Taskar, and M. I. Jordan. Joint covariate selection and joint subspace selection for multiple classification problems.

[2] M. Massias, O. Fercoq, A. Gramfort, and J. Salmon. Generalized concomitant multi-task lasso for sparse multimodal regression.