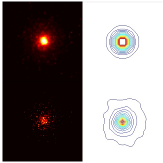
Recovering Simultaneously Structured Data via Non-Convex Iteratively Reweighted Least Squares

Christian Kümmerle
University of North Carolina at Charlotte

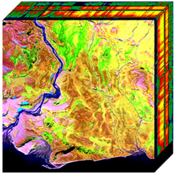
Johannes Maly
LMU Munich
Munich Center for Machine Learning

NeurIPS 2023

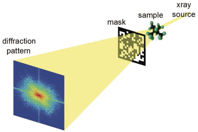
Motivation: Recovery of Simultaneously Structured Data



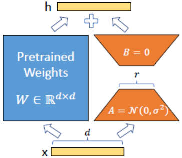
Sparse Blind Deconvolution



Hyperspectral Imaging



Phase Retrieval



Parameter-Efficient Machine Learning

Problem: Given (linear) observation map $\mathcal{A}$, observations $\mathbf{y}$, reconstruct $\mathbf{X}_* \in S_1 \cap S_2$ from

$$\mathbf{y} = \mathcal{A}(\mathbf{X}_*) + \eta \in \mathbb{R}^m,$$

where S_1 and S_2 are two **heterogenous subsets of parsimonious structure**.

Challenges of Recovery of Simultaneously Structured Data

Problem: Given (linear) observation map $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^m$, observations $\mathbf{y} \in \mathbb{R}^m$, reconstruct $(n_1 \times n_2)$ matrix $\mathbf{X}_* \in S_1 \cap S_2$ from

$$\mathbf{y} = \mathcal{A}(\mathbf{X}_*) + \eta \in \mathbb{R}^m,$$

where S_1 and S_2 are two **heterogenous subsets of parsimonious structure**.

Fundamental Challenges:

- What is the minimal sample complexity that makes the problem reliably solvable (dependent on the complexities of S_1 and S_2)?
- What are computationally efficient algorithms that achieve that?

Focus of this work:

- Simultaneously low-rank and row-sparse matrices $\mathbf{X}_* \in \mathbb{R}^{n_1 \times n_2}$, e.g.,
  - $S_1 = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2} : \operatorname{rank}(\mathbf{X}) \leq r\}$ and
  - $S_2 = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2} : \|\mathbf{X}\|_{2,0} \leq s_1\}$.
- Propose a **computationally efficient algorithm** that achieves **state-of-the-art data efficiency**
- Establish **local convergence analysis** that applies for minimal sample complexity $m = \Omega(r(s + n_2) \log(en_1/s))$.

Our Contributions

- Combine non-convex, continuous surrogate objective with smoothing strategy to formulate a tailored iteratively reweighted least squares (IRLS) algorithm.
- Overcomes negative results by **(Oymak et al. 2015)** for related **convex** surrogate modeling.
- Prove locally quadratic convergence rate of proposed IRLS under restricted isometry property on $S_1 \cap S_2$.
- Introduce ``self-balancing'' of smoothed surrogate objective for multiple parsimonious structures.

Our Contributions

- Combine non-convex, continuous surrogate objective with smoothing strategy to formulate a tailored iteratively reweighted least squares (IRLS) algorithm.
- Overcomes negative results by **(Oymak et al. 2015)** for related **convex** surrogate modeling.
- Prove locally quadratic convergence rate of proposed IRLS under restricted isometry property on $S_1 \cap S_2$.
- Introduce ``self-balancing'' of smoothed surrogate objective for multiple parsimonious structures.

References:

(1) Christian Kümmerle, Johannes Maly

Recovering Simultaneously Structured Data via Non-Convex Iteratively Reweighted Least Squares,

NeurIPS 2023, `https://openreview.net/pdf?id=50hs53Zb3w`.

(2) Samet Oymak, Amin Jalali, Maryam Fazel, Yonina Eldar, Babak Hassibi

Simultaneously structured models with application to sparse and low-rank matrices

IEEE Transactions on Information Theory 61.5 (2015): 2886-2908

(3) Christian Kümmerle, Claudio Mayrink Verdun

A Scalable Second Order Method for Ill-Conditioned Matrix Completion from Few Samples

ICML 2021.