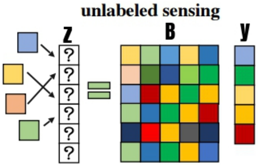
Journal Watch
IEEE Transactions on Signal Processing, March 2018

Akshay Kumar

April 14, 2018

Signal Recovery From Unlabeled Samples

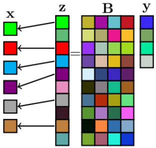


- Overdetermined set of linear equations
- Measurement mismatch via measurement devices
- Elements are not in proper order
- Not all elements of **z** are known
- Few papers in recent times [1] [2] [3]

[1] JU, M Vetterli,"Unlabeled sensing with random linear measurements", IEEE TIT'15
[2] AP, M Wainwright"Linear Regression with Shuffled Data: Statistical and Computational Limits of Permutation Recovery", IEEE TIT'17
[3] D Hu et al., "Linear regression without correspondence", NIPS 2017

Signal Recovery From Unlabeled Samples

- In this paper, authors consider unlabelled, noisy and ORDERED samples



- Problem formulation,

$$\mathbf{x} = \mathbf{S}\mathbf{B}\mathbf{y} + \mathbf{w}$$

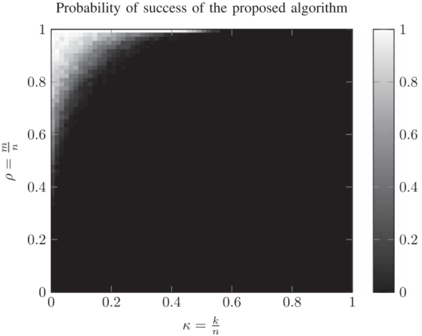
- $\mathbf{B}_{n \times k}$ is a known matrix, $\mathbf{w}$ is noise
- $\mathbf{S}_{m \times n}$ is an unknown matrix with only a single 1 in each row
- Given $\mathbf{B}$ and $\mathbf{x}$ estimate $\mathbf{S}$ and $\mathbf{y}$

Contributions

- Devised a recovery algorithm based on alternating minimization
- For a fixed **S**, the optimal **y** corresponds to least-squares solution of an overdetermined system
- For a fixed **y**, computing the optimal **S** is formulated as Dynamic Programming problem
- They provide theoretical guarantees for stable signal recovery in terms of RIP-type properties of **H**

$$\mathbf{x} = \mathbf{S}\mathbf{B}\mathbf{y} + \mathbf{w} = \underbrace{(\mathbf{y}^T \otimes \mathbf{S})}_{\mathrm{H}} \mathbf{vec}(\mathbf{B}) + \mathbf{w}$$

- $n = 1000$. Averaged over 1000 independent realizations of **S** and **B**



A Compact Formulation for the $\ell_{2,1}$ Mixed-Norm Minimization Problem

Consider the MMV sparse recovery problem,

$$\mathbf{Y}_{m\times L} = \mathbf{A}_{m\times n}\mathbf{X}_{n\times L} + \mathbf{W}_{m\times L}$$

Theorem: The row-sparsity inducing $\ell_{2,1}$ mixed-norm minimization problem

$$\min_{X} \frac{1}{2}\|\mathbf{AX}-\mathbf{Y}\|_F^2 + \lambda\sqrt{L}\|\mathbf{X}\|_{2,1}$$

is equivalent to the convex problem

$$\min_{\mathbf{S}\in \mathrm{D}_+} \mathrm{Tr}((\mathbf{ASA}^{\mathrm{H}} + \boldsymbol{\lambda}\mathbf{I_m})^{-1}\hat{\mathbf{R}}) + \mathrm{Tr}(\mathbf{S}),$$

with $\hat{\mathbf{R}} = \mathbf{YY^H}/\mathbf{L}$ denoting the sample covariance matrix and D_+ describing the set of nonnegative diagonal matrices and,

$$\hat{\mathbf{X}} = \hat{\mathbf{S}}\mathbf{A}^{\mathrm{H}}(\mathbf{A}\hat{\mathbf{S}}\mathbf{A}^{\mathrm{H}} + \boldsymbol{\lambda}\mathbf{I_m})^{-1}\mathbf{Y}.$$

Benefits of SPARse ROW-norm reconstruction (SPARROW)-:

- Low complexity algorithms
- Mixed norm formulation had nL variables, SPARROW has only n variables
- Reduced problem size as SPARROW depends only on $\mathbf{YY^H}(m \times m)$ instead of $\mathbf{Y}(m \times L)$

Few recent papers using the SPARROW to improve complexity,

Low-complexity massive MIMO subspace estimation and tracking from low-dimensional projections
S Haghighatshoar, G Caire - arXiv preprint arXiv:1608.02477, 2016 - arxiv.org

Improved Scaling Law for Activity Detection in Massive MIMO Systems
S Haghighatshoar, P Jung, G Caire - arXiv preprint arXiv:1803.02288, 2018 - arxiv.org

Block-and Rank-Sparse Recovery for Direction Finding in Partly Calibrated Arrays
C Steffens, M Pesavento - IEEE Transactions on Signal …, 2017 - ieeexplore.ieee.org

Joint active device identification and symbol detection using sparse constraints in massive MIMO systems
G Hegde, M Pesavento… - … (EUSIPCO), 2017 25th …, 2017 - ieeexplore.ieee.org

Gridless compressed sensing under shift-invariant sampling
C Steffens, W Suleiman, A Sorg… - Acoustics, Speech and …, 2017 - ieeexplore.ieee.org

Proof:

Core result:

$$\|\mathbf{x}_k\|_2 = \min_{\gamma_k, \mathbf{g}_k} \frac{1}{2}(|\gamma_k|^2 + \|\mathbf{g}_k\|_2^2) \tag{1}$$

$$\text{s.t.} \quad \gamma_k \mathbf{g}_k = \mathbf{x}_k, \tag{2}$$

Any feasible solution must satisfy,

$$\|\mathbf{x}_k\|_2 = \sqrt{|\gamma_k|^2 \|\mathbf{g}_k\|_2^2} \leq \frac{1}{2}(|\gamma_k|^2 + \|\mathbf{g}_k\|_2^2) \tag{3}$$

Equality will hold iff $|\gamma_k| = \|\mathbf{g}_k\|_2$

Hence,

$$\|\mathbf{X}\|_{2,1} = \sum_{k=1}^{K} \|\mathbf{x}_k\|_2 = \min_{\boldsymbol{\Gamma} \in \mathbb{D}, \mathbf{G}} \frac{1}{2}(\|\boldsymbol{\Gamma}\|_F^2 + \|\mathbf{G}\|_F^2) \tag{4}$$

$$\text{s.t.} \quad \mathbf{X} = \boldsymbol{\Gamma}\mathbf{G}, \tag{5}$$

Using above equation,

$$\min_{\boldsymbol{\Gamma}\in\mathbb{D},\mathbf{G}} \frac{1}{2}\|\mathbf{A}\boldsymbol{\Gamma}\mathbf{G}-\mathbf{Y}\|_{\mathsf{F}}^2 + \frac{\lambda\sqrt{L}}{2}(\|\boldsymbol{\Gamma}\|_{\mathsf{F}}^2 + \|\mathbf{G}\|_{\mathsf{F}}^2). \tag{6}$$

For a fixed $\boldsymbol{\Gamma}$, the minimizer of $\hat{\mathbf{G}}$ has a closed form expression

$$\begin{aligned} \hat{\mathbf{G}} &= \left(\boldsymbol{\Gamma}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{A}\boldsymbol{\Gamma} + \lambda\sqrt{L}\mathbf{I}_n\right)^{-1}\boldsymbol{\Gamma}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{Y} \\ &= \boldsymbol{\Gamma}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\left(\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}} + \lambda\sqrt{L}\mathbf{I}_m\right)^{-1}\mathbf{Y}, \end{aligned} \tag{7}$$

Inserting $\hat{\mathbf{G}}$,

$$\min_{\boldsymbol{\Gamma}\in\mathbb{D}} \frac{\sqrt{L}}{2}\Big(\mathrm{Tr}\big((\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}} + \lambda\sqrt{L}\mathbf{I}_m)^{-1}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\big) + \mathrm{Tr}\big(\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathsf{H}}\big)\Big). \tag{8}$$

Upon substituting $\hat{\mathbf{R}} = \mathbf{Y}\mathbf{Y}^{\mathbf{H}}/\mathbf{L}$ and defining nonnegative diagonal matrix $\mathbf{S} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathsf{H}}/\sqrt{\mathbf{L}}$ we get,

$$\min_{\mathbf{S}\in\mathbb{D}_+} \frac{L}{2}\Big(\mathrm{Tr}\big((\mathbf{A}\mathbf{S}\mathbf{A}^{\mathsf{H}} + \lambda\mathbf{I}_m)^{-1}\hat{\mathbf{R}}\big) + \mathrm{Tr}\big(\mathbf{S}\big)\Big). \tag{9}$$

Outlier-Robust Matrix Completion via ℓ_p-Minimization

Goal

- Compute robust solution for noisy (not necessarily Gaussian) Matrix Completion problem

Problem formulation

- $\mathbf{X} = \hat{\mathbf{X}}_{\mathbf{n_1 \times n_2}} + \mathbf{E}$
- Estimate a low-rank matrix, $\hat{\mathbf{X}}$ from few noisy entries $\mathbf{X_\Omega}$

$$\min_{\mathbf{U},\mathbf{V}} ||(\mathbf{UV})_\Omega - \mathbf{X}_\Omega||_p^p \tag{10}$$

where, $\mathbf{U} \in \mathbb{R}^{n_1 \times r}$ and $\mathbf{V} \in \mathbb{R}^{r \times n_2}$

Contributions

- They propose two methods to solve the above problem
- The first uses iterative ℓ_p-regression while second uses ADMM
- Superior to the singular value thresholding, and alternating projection schemes in terms of computational simplicity, statistical accuracy, and outlier-robustness.

Iterative ℓ_p-regression

Use alternating minimization strategy,

$$\mathbf{V^{k+1}} = \arg\min_{\mathbf{U}} ||(\mathbf{U^k V})_\Omega - \mathbf{X}_\Omega||_p^p \tag{11}$$

$$\mathbf{U^{k+1}} = \arg\min_{\mathbf{V}} ||(\mathbf{U V^{k+1}})_\Omega - \mathbf{X}_\Omega||_p^p \tag{12}$$

Focus on one term,

$$\min_{\mathbf{V}} ||(\mathbf{UV})_\Omega - \mathbf{X}_\Omega||_p^p = \min_{\mathbf{V}} \sum_{i,j \in \Omega} |\mathbf{u_i}^T \mathbf{v_j} - \mathbf{X}_{ij}|^p \tag{13}$$

$$= \min_{\mathbf{V}} \sum_{j=1}^{n_2} \sum_{i=1}^{\mathcal{I}_j} |\mathbf{u_i}^T \mathbf{v_j} - \mathbf{X}_{ij}|^p \tag{14}$$

Consider one subproblem,

$$\min_{\mathbf{v_j}} \sum_{i=1}^{\mathcal{I}_j} ||\mathbf{U}_{\mathcal{I}_\mathbf{j}} \mathbf{v_j} - \mathbf{b}_{\mathcal{I}_j}||_p^p \tag{15}$$

Solve the above problem via weighted iterative least-squares

Other interesting papers

- J. Mo, P. Schniter, and R. W. Heath, Jr.,"**Channel Estimation in Broadband Millimeter Wave MIMO Systems With Few-Bit ADCs**"
- R. Zhao, W. B. Haskell, and V. Y. F. Tan, "**Stochastic L-BFGS: Improved Convergence Rates and Practical Acceleration Strategies**"
- I. Bergel and Y. Noam, "**Lower Bound on the Localization Error in Infinite Networks With Random Sensor Locations**"
- F. Van Eeghem, O. Debals, and L. De Lathauwer, "**Tensor Similarity in Two Modes**"
- Y. Gao, H. Vinck, and T. Kaiser, "**Massive MIMO Antenna Selection: Switching Architectures, Capacity Bounds, and Optimal Antenna Selection Algorithms**"
- D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten, "**Symbol-Level Precoding for the Nonlinear Multiuser MISO Downlink Channel**"