Journal Watch
IEEE Transactions on Signal Processing (April 01, 2018)
IEEE Transactions on Information Theory (May, 2018)

Lekshmi Ramesh

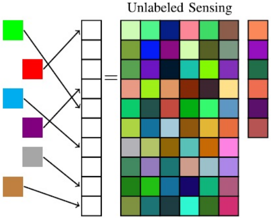
Indian Institute of Science
Bangalore

June 2, 2018

- Setup: System of linear equations

$$\mathbf{y} = A\mathbf{x}$$

where $A \in \mathbb{R}^{m \times n}$, when observations are unlabeled
We have access to all entries of $\mathbf{y}$, but not their labels



- Main result: When A has iid entries drawn from a continuous distribution, then $\mathbf{x}$ can be recovered exactly w.p. 1, without knowledge of labels of $\mathbf{y}$, if $m \geq 2n$

- An example.
  - Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then $\mathbf{y} = \begin{bmatrix} x_2 \\ x_1 \\ x_1 \end{bmatrix}$.
  - In this case, $\mathbf{x}$ can be recovered exactly, by counting the number of times an entry appears in the unordered observations.
- However, it can be shown that there exist $A \in \mathbb{R}^{3\times 2}$ for which different $\mathbf{x}$ give rise to the same unordered observation.
- Unique recovery of $\mathbf{x}$ is guaranteed if A is a $2n \times n$ matrix with iid entries from a continuous probability distribution.
  Proof constructive, uses a combinatorial algorithm

- Setup: Noisy linear model with unknown permutation

  $$\mathbf{y} = \Pi^* A \mathbf{x}^* + \mathbf{w}$$

  where $A \in \mathbb{R}^{n \times d}$, Π^* is an unknown permutation, $\mathbf{x}^*$ is the unknown signal and $\mathbf{w}$ is additive Gaussian noise

- Random design setting is considered: A has iid $\mathcal{N}(0, 1)$ entries

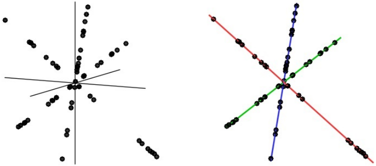
- Contribution: Conditions on n, d and SNR under which Π^* is exactly/approximately recoverable

- Upper bound obtained by analyzing the ML decoder

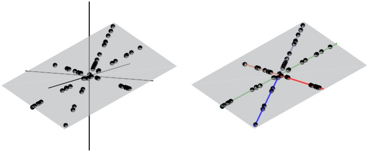
$$(\hat{\Pi}_{ML}, \hat{\mathbf{x}}_{ML}) = \underset{\Pi \in \mathcal{P}, \mathbf{x} \in \mathbb{R}^d}{\arg\min} \|\mathbf{y} - \Pi A \mathbf{x}\|_2^2$$

- No polynomial-time algorithm known for permutation recovery
- Lower bound obtained by connecting to decoding over Gaussian channel

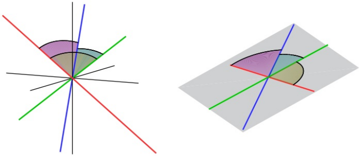
- Goal: Study restricted isometry property of Gaussian random matrices for the compression of two subspaces
- Motivation: Compressive Subspace clustering
  - Label data points that are assumed to be drawn from a union of subspaces model
  - Only a compressed version of the data points is available
- Main result
  The *distance* between two subspaces remains almost unchanged whp after random projection if the ambient dimension after projection is sufficiently large.



(a) Data points from three clusters



(b) Susbspace clustering after compression



(c) Principal angles decrease after compression

- Main result

Consider a set of L subspaces $\mathcal{X}_1, \ldots, \mathcal{X}_L \in \mathbb{R}^N$, each of dimension no more than d. If these subspaces are projected onto $\mathbb{R}^n$ by a Gaussian random matrix $\Phi \in \mathbb{R}^{n \times N}$,

$$\mathcal{X}_i \xrightarrow{\Phi} \mathcal{Y}_i := \{\mathbf{y} : \mathbf{y} = \Phi\mathbf{x},\ \mathbf{x} \in \mathcal{X}_i\}, \quad i \in [L],$$

and $d << n < N$, then we have

$$(1-\epsilon)D^2(\mathcal{X}_i, \mathcal{X}_j) \le D^2(\mathcal{Y}_i, \mathcal{Y}_j) \le (1+\epsilon)D^2(\mathcal{X}_i, \mathcal{X}_j), \quad \forall i, j$$

with probability at least

$$1 - \frac{2dL(L-1)}{(\epsilon - \frac{d}{n})^2 n},$$

where D denotes the projection F-norm distance between subspaces.

- Subspace clustering: Label non-linearly separable data generated from a union of subspaces model, when no ground truth is available
  - Current methods are computationally complex
  - This paper deals with performing clustering when a sketched version of the data is given
- Model

$$\mathbf{x}_i = C^j \mathbf{y}^j + \boldsymbol{\mu}^j + \mathbf{v}_i,\ \forall\ \mathbf{x}_i \in S^j,$$

where
$\{\mathbf{x}_i\}_{i=1}^N$ are D-dimensional data vectors
$S_1, \ldots, S_k$ are the k subspaces of dimension $d_1, \ldots, d_k$
$C^j \in \mathbb{R}^{D \times d_j}$ is a basis for S^j

- Goal is to find the subspace assignment vector π_i for each $\mathbf{x}_i$ under the constraints $\pi_{ij} \geq 0$ and $\sum_j \pi_{ij} = 1$
    - Hard clustering: $\pi_i \in \{0, 1\}^k$, membership to a single subspace
    - Soft clustering: $\pi_i \in [0, 1]^k$, an entry of π_i can be thought of as probability of belonging to a subspace
- Sketch-SC algorithm
    - A modification of the standard sparse-SC algorithm

$$\begin{aligned} \underset{Z}{\text{minimize}}\, & \|Z\|_1 + \frac{\lambda}{2}\|X - XZ\|_2^2 \\ \text{s.t.}\ \ & Z^\top \mathbf{1} = \mathbf{1},\ \operatorname{diag}(Z) = \mathbf{0} \end{aligned}$$

Final labels obtained by post processing Z

- Sketch-SC
  - Generate JLT mtarices R_1 and R_2
  - Form sketched data matrix $\tilde{X} = R_1 X$
  - Form dictionary $\tilde{B} = \tilde{X} R_2$
  - Solve

$$\underset{A}{\text{minimize}}\, \|A\| + \frac{\lambda}{2} \|\tilde{X} - \tilde{B} A\|_2^2$$

Other interesting papers

- Sparse Activity Detection for Massive Connectivity. *Z. Chen, F. Sohrabi, and W. Yu*
- Tradeoffs Between Convergence Speed and Reconstruction Accuracy in Inverse Problems. *R. Giryes, Y. C. Eldar, A.M. Bronstein, and G. Sapiro*
- High-Dimensional MVDR Beamforming: Optimized Solutions Based on Spiked Random Matrix Models. *L. Yang,M. R.McKay, and R. Couillet*
- Phase Retrieval With Random Gaussian Sensing Vectors by Alternating Projections. *I. Waldspurger*
- Minimax Lower Bounds for Noisy Matrix Completion Under Sparse Factor Models. *A. V. Sambasivan and J. D. Haupt*