Demixing Sparse Signals via Convex Optimization

Yi Zhou*, Yingbin Liang*
*Department of EECS, Syracuse University

Signal Mixture Model

Signal $z_0 \in \mathbb{R}^n$ is a mixture of $x_0, y_0 \in \mathbb{R}^n$:

$$z_0 := x_0 + y_0$$

1. Image feature decomposition, image denoising, signal separation;
2. How to formulate a proper demixing model?
3. What is the theoretical performance guarantee of the model?

Model Formulation

Sparse signal components: orthonormal basis $\Psi, \Phi \in \mathbb{R}^{d \times n}$

- $\theta_x := \Psi x_0$, $\theta_y := \Phi y_0$ (card($\theta_x$), card($\theta_y$)) \leq n.

Convex Optimization Model: for some $\lambda > 0$

$$\min_{x, y} \left\{ \|y\|_1 : \theta_x := \Psi x, \theta_y := \Phi y \right\} \text{ s.t. } x + y = z_0$$

(P)

1. Seek for feasible decomposition with minimum $\ell^1$ norm (convex);
2. When is (P) exact, i.e., $(x_0, y_0)$ be its unique solution pair?

Review of Existing Work

Separability relies on the coherence between $\Psi, \Phi$.

Mutual Coherence [1]:

$$\mu(\Psi, \Phi) := \max \langle \phi, \psi \rangle.$$

Let $\text{supp}(\theta_x)$ be fixed and $\text{supp}(\theta_y)$ be uniformly random. Then (P) is exact w.h.p provided that

$$\text{card}(\theta_x) + \text{card}(\theta_y) \leq O \left( \left\lceil \log n \right\rceil \right).$$

- In general, $\mu \in \left[ \frac{1}{\sqrt{m}}, 1 \right]$.
- Mutual coherence barrier: $\mu(\Psi, \Phi) = 1$.

Cluster Coherence [2]:

$$\mu(\Psi, \Phi) := \max \sum_{i=1}^k \left| \langle \phi_i, \psi \rangle \right|.$$ (P) is exact in the asymptotic regime (i.e., for all $j \to n$) nearly perfectly provided that the corresponding cluster coherence vanishes.

- Characterize asymptotic exactness;
- We want a quantifiable local exactness condition.

Local Subspace Coherence

Assumption 1.

The signal components $x_0, y_0$ satisfy:
1. $\text{supp}(\theta_x)$ is fixed, while $\text{supp}(\theta_y)$ satisfies $P(j \in \text{supp}(\theta_y)) \sim \text{Bernoulli}(\mu)$;
2. $\text{sign}(\theta_x), \text{sign}(\theta_y)$ take values from $\{-1, 1\}$ with equal probability.

Signal Subspace:

$$X := \text{span}\{\phi_j, j \in \text{supp}(\theta_x)\}, \quad Y := \text{span}\{\phi_j, j \in \text{supp}(\theta_y)\}.$$ (2)

- $x_0 \in X$, and $y_0 \in Y$; $P_X, P_Y$ projections.

Definition 1. Local Subspace Coherence

The local subspace coherence between basic vectors $\{\phi_j\}$ and the subspace $X$ is

$$\mu(X, \phi_j) := \|P_X \phi_j\|_2, \quad j \in [n].$$

- Measures how aligned is $\phi_j$ with subspace $X$;
- For $X := \bigoplus_{i=1}^d X_i$, it holds that $\mu^2(X, \phi_i) = \sum_{i=1}^d \mu^2(X_i, \phi_i)$.

Main Result

Theorem 1. Exactness of (P)

Suppose Assumption 1 hold and set $\lambda = \frac{1}{\sqrt{\log n}}$ in (P). Then $(x_0, y_0)$ is the unique minimizer of (P) with probability at least $1 - n^{-\sqrt{\log n}}$, provided that for all $j \in [n]$

$$1 - p_j \geq C \log(X, \phi_j) \log^2 n,$$ (4)

where $C$ is a universal positive constant.

- essentially, $P(\phi_j \notin Y_j) = \mu(X, \phi_j)$;
- makes $X, Y$ be incoherent and hence distinguishable.

Illustrative Examples

Example 1: $\Psi = f$, $\Phi = I$.
- Coherence pattern: $|\langle \phi, \psi \rangle| = \frac{1}{\sqrt{m}}$.
- local subspace coherence $\mu(X, \phi_j) = \sqrt{\frac{\text{card}(\theta_x)}{\text{card}(\theta_y)}} - \sqrt{\frac{\text{card}(\theta_y)}{\text{card}(\theta_x)}} \leq O \left( \left\lceil \log \frac{n}{\text{card}(\theta_x)} \right\rceil \right)$ (5)

Example 2: $\Psi = \Psi_f$, $\Phi = f$.
- Coherence pattern: $|\langle \phi, \psi \rangle| \leq \frac{1}{\sqrt{m}} \log \frac{n}{\text{card}(\theta_x)}$;
- local subspace coherence $\sum_{j=1}^n \mu(X, \phi_j) \leq \frac{\text{card}(\theta_x)}{\log \frac{n}{\text{card}(\theta_x)}}$.

$$\sum_{j=1}^n \mu(X, \phi_j) \leq O \left( \frac{\sqrt{\log n}}{\log \frac{n}{\text{card}(\theta_x)}} \right).$$ (6)

Numerical Experiment

Fig. 2: Comparison of success region between (a) $\theta_x$ be uniformly at random and (b) $\theta_x$ be adapted to local subspace coherence.

References