Identification of Matrix Joint Block Diagonalization

Yunfeng Cai
Congnitive Computing Lab
Baidu Research

April 13 - 15, 2021

- The block diagonal matrix –

$$X = \begin{array}{c} \\ p_1 \\ p_2 \\ \vdots \\ p_\ell \end{array} \begin{array}{c} \begin{array}{cccc} p_1 & p_2 & \cdots & p_\ell \end{array} \\ \begin{bmatrix} X_{11} & 0 & \dots & 0 \\ 0 & X_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & X_{\ell\ell} \end{bmatrix} \end{array}.$$

- The partition $\tau_p = (p_1, p_2, \dots, p_\ell)$ describes the block diagonal structure of X.
- Joint block diagonalization problem (JBDP) – Given a matrix set $\mathcal{C} = \{C_i\}_{i=1}^m$ with $C_i \in \mathbb{R}^{d \times d}$. The JBDP for $\mathcal{C}$ with respect to a partition τ_p is to find a full column rank matrix $A = A(\tau_p) \in \mathbb{R}^{d \times p}$ (τ_p-block diagonalizer) such that all C_i's can be factorized as

$$C_i = A \Sigma_i A^\top = A \,\mathrm{diag}(\Sigma_i^{(11)}, \dots, \Sigma_i^{(\ell\ell)}) A^\top, \quad \forall i,$$

where Σ_i's are all τ_p-block diagonal.
- Blind JBDP (BJBDP) – solve JBDP without knowing τ_p, and $|\,\mathrm{card}(\tau_p)|$ is maximized, i.e., the number of the diagonal blocks is maximized.

- Identification of BJDBP – Let (τ_p, A) be a solution to the BJBDP for $\mathcal{C}$. Let $\widetilde{\mathcal{C}} = \{\widetilde{C}_i\}_{i=1}^m = \{C_i + E_i\}_{i=1}^m$, where $E_i \in \mathbb{R}^{d\times d}$ is a perturbation to C_i for $1 \le i \le m$. Under what conditions, and by what means, we can find a $(\tilde{\tau}_p, \widetilde{A})$ such that

$$\widetilde{C}_i \approx \widetilde{A}\widetilde{\Sigma}_i\widetilde{A}^\top = \widetilde{A}\,\mathrm{diag}(\widetilde{\Sigma}_i^{(11)}, \ldots, \widetilde{\Sigma}_i^{(\ell\ell)})\widetilde{A}^\top, \quad \forall i,$$

  where $\widetilde{\Sigma}_i$'s are all $\tilde{\tau}_p$-block diagonal matrices with $\tilde{\tau}_p \sim \tau_p$, and $\widetilde{A}$ is close to A (up to block permutation and block diagonal scaling).

- Application – independent subspace analysis (ISA), blind source separation, etc.

- A basic model can be stated as

$$\mathbf{x} = A\mathbf{s}.$$

  - $\mathbf{x} \in \mathbb{R}^d$ is the observed mixture;
  - $A \in \mathbb{R}^{d\times p}$ is the unknown mixing matrix and has full column rank;
  - $\mathbf{s} \in \mathbb{R}^p$ is the source signal vector.

- Assumption – Let $\mathbf{s} = \left[\mathbf{s}_1^\top, \ldots, \mathbf{s}_\ell^\top\right]^\top$ with $\mathbf{s}_j \in \mathbb{R}^{p_j}$. Assume that each $\mathbf{s}_j$ has mean 0 and contains no lower-dimensional independent component, all $\mathbf{s}_j$ are independent of each other.

- It holds

$$C_{\mathbf{xx}} = \mathbb{E}(\mathbf{x}\mathbf{x}^\top) = A\mathbb{E}(\mathbf{s}\mathbf{s}^\top)A^\top = AC_{\mathbf{ss}}A^\top,$$

  where $C_{\mathbf{xx}}$, $C_{\mathbf{ss}}$ are the covariance matrices of $\mathbf{x}$ and $\mathbf{s}$, respectively. By assumption, $C_{\mathbf{ss}}$ is τ_p-block diagonal.

- Construct several empirical (time lagged) covariance matrices $\widetilde{C}_i$ of $\mathbf{x}$. The question is that whether we can find $(\tilde{\tau}_p, \widetilde{A})$ by solving the BJBDP for $\{\widetilde{C}_i\}$ such that $\tilde{\tau}_p \sim \tau_p$, and $\widetilde{A}$ is "close" to A? Under what conditions? And how?

- Definition – Given a matrix set $\mathcal{D} = \{D_i\}_{i=1}^m$ with $D_i \in \mathbb{R}^{q\times q}$, define

$$\mathscr{N}(\mathcal{D}) \triangleq \left\{X \in \mathbb{R}^{q\times q} \mid D_i X - X^\top D_i = 0,\ 1 \le i \le m\right\}.$$

- Optimization problem –

$$\begin{aligned} \text{OPT}(\mathcal{D}): \quad & \min_X \operatorname{tr}(X^4), \\ \text{subject to} \quad & X \in \mathscr{N}(\mathcal{D}), \operatorname{tr}(X) = 0, \operatorname{tr}(X^2) = q. \end{aligned}$$

- Theorem – Given a set $\mathcal{D} = \{D_i\}_{i=1}^m$ of q-by-q matrices. Almost surely, it holds

  (I) If $\mathcal{D}$ does not have a nontrivial diagonalizer, then the feasible set of $\text{OPT}(\mathcal{D})$ is empty, i.e., $\text{OPT}(\mathcal{D})$ has no solutions.

  (II) If $\mathcal{D}$ has a nontrivial diagonalizer, then $\text{OPT}(\mathcal{D})$ has a solution X_*. X_* has two distinct real eigenvalues, and the gap between them are no less than two; The eigenvector matrix of X_* is a diagonalizer.

Algorithm and Theoretical Guarantee 5.

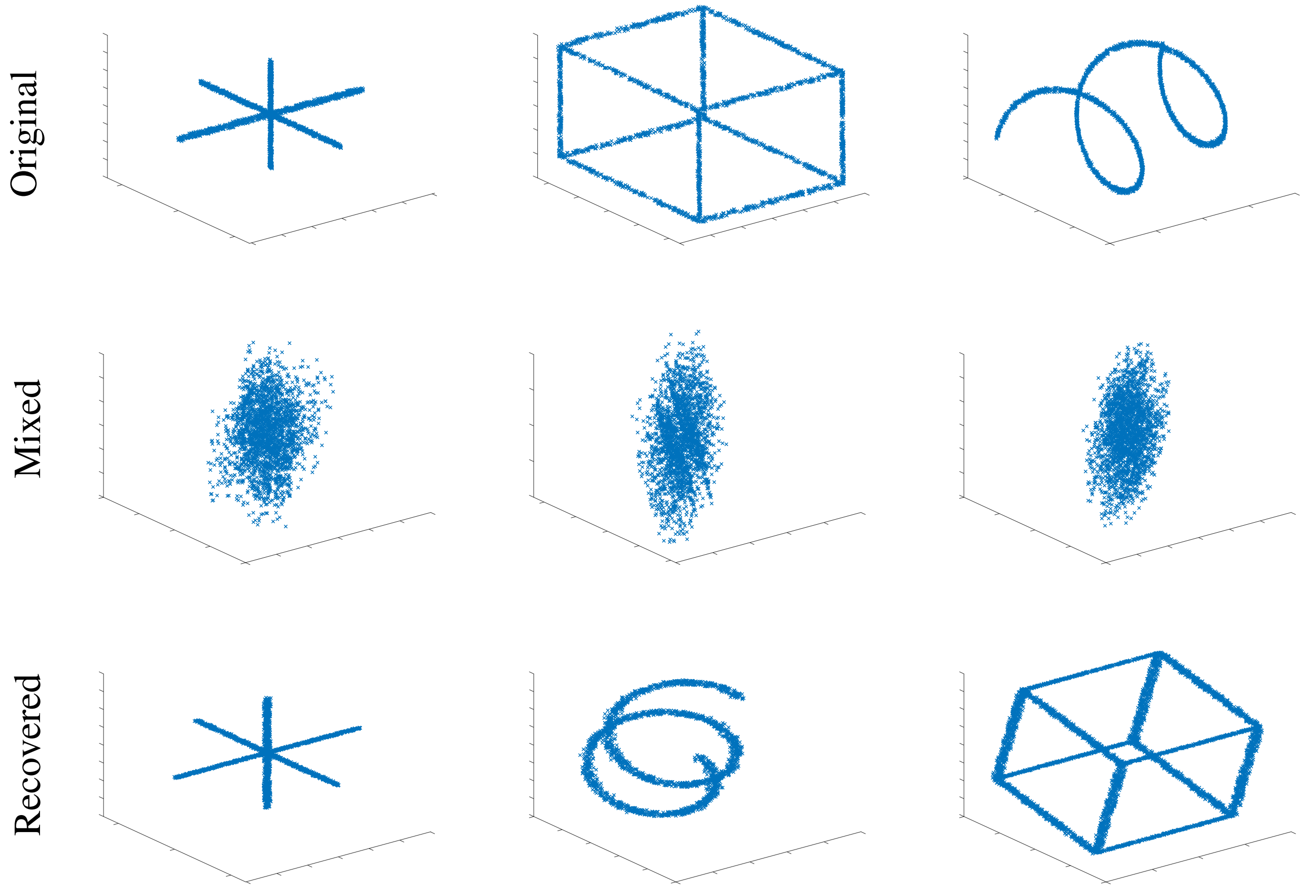
Algorithm 1 Bi-Block Diagonalization (BI-BD)

```
Input: A set $\mathcal{D} = \{D_i\}_{i=1}^m$ of $q$-by-$q$ matrices.
Output: $(\tau_q, Z)$ such that $Z$ is a $\tau_q$-block diagonalizer of $\mathcal{D}$ with $\tau_q = (q_1, q_2)$ or $\tau_q = (q)$.
if feasible set of OPT($\mathcal{D}$) is empty then set $\tau_q = (q)$, $Z = I_q$;
else Solve OPT($\mathcal{D}$), denote the solution by $X_*$;
    Compute $X_* = Y \operatorname{diag}(\Gamma_1, \Gamma_2) Y^{-1}$, where $\Gamma_1 \in \mathbb{R}^{q_1 \times q_1}$, $\Gamma_2 \in \mathbb{R}^{q_2 \times q_2}$, both $\lambda(\Gamma_1)$ and
    $\lambda(\Gamma_2)$ contain only one real number, and the two real numbers are different.
    Set $\tau_q = (q_1, q_2)$, $Z = Y^{-\top}$.
end if
```

Apply Algorithm BI-BD recursively, we can find (τ_p, A), a solution to BJBDP.

Theorem Assume that the BJBDP for $\mathcal{C}$ is uniquely τ_p-block-diagonalizable, and let (τ_p, A) be a solution. Then (τ_p, A) can be identified via our algorithm, almost surely.

Able to handle noisy case as well.



The original source signals, the mixed source signals, and the recovered signals

Thanks for your attention!