Blind Deconvolution and Compressed Sensing

Dominik Stöger

Technische Universität München,
Department of Mathematics,
Applied Numerical Analysis

dominik.stoeger@ma.tum.de

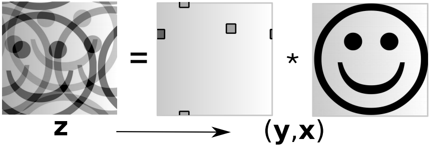
Joint work with Peter Jung (TU Berlin), Felix Krahmer (TU München)

supported by DFG priority program "Compressed Sensing in Information Processing" (CoSIP)

Overview

Overview

Blind Deconvolution



- bilinear inverse problem: $z = B(x, y)$
- ambiguities, constraining x and/or y

Many applications:

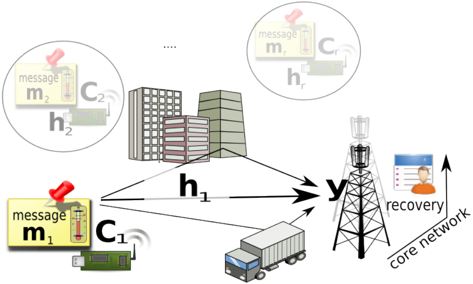
- imaging (blind deblurring)
- radar, e.g., ground penetrating radar (GPR), radar imaging
- speech recognition
- wireless communication

Blind Deconvolution and Demixing

A problem in Wireless Communication:

- r different devices
- device i delivers message m_i
- **Linear encoding:**
  $x_i = C_i m_i$ with $C_i \in \mathbb{R}^{L \times N}$
- **Channel model:**
  $w_i = B_i h_i$, where $B_i \in \mathbb{R}^{L \times K}$
- **Received signal:**

$$y = \sum_{i=1}^{r} w_i * x_i \in \mathbb{R}^L$$



Goal: recover **all** m_i from y

Assumptions on B_i and C_i

$$y = \sum_{i=1}^{r} w_i * x_i = \sum_{i=1}^{r} B_i h_i * C_i m_i$$

- Assume w_i is concentrated on the first few entries, i.e., B_i extends h_i by zeros
- (Our analysis will include more general B_i)
- Choice of C_i is arbitrary $\Rightarrow$ randomize
- Choose C_i to have i.i.d. standard normal entries

Lifting

- There are unique linear maps $\mathcal{A}_i : \mathbb{R}^{K \times N} \to \mathbb{R}^L$ such that for arbitrary h_i and m_i

$$w_i * x_i = B_i h_i * C_i m_i = \mathcal{A}_i \left(h_i m_i^* \right) = \mathcal{A}_i \left(Y_i \right)$$

- $$y = \sum_{i=1}^{r} \mathcal{A}_i \left(h_i m_i^* \right) = \mathcal{A} \left(X_0 \right),$$

  where

$$X_0 = \left(h_1 m_1^*, \cdots, h_r m_r^* \right)$$

- **Low rank matrix recovery problem**

Overview

A convex approach for recovery

- $r = 1$ investigated in [Ahmed, Recht, Romberg, 2012]
- $r \geq 1$ semidefinite program (SDP) [Ling, Strohmer, 2015]

$$\text{minimize} \sum_{i=1}^{r} \| Y_i \|_* \qquad \text{subject to} \sum_{i=1}^{r} \mathcal{A}_i \left(Y_i \right) = y. \qquad \text{(SDP)}$$

- $\| \cdot \|_*$: nuclear norm, i.e., the sum of the singular values
- Recovery is guaranteed with high probability, if

$$L \geq C r^2 \left(K + \mu_h^2 N \right) \log^3 L \log r$$

- coherence parameter: $1 \leq \mu_h = \max_i \frac{\|\hat{h}_i\|_\infty}{\|h_i\|_2} \leq \sqrt{K}$

Linear scaling in r

- Number of degrees of freedom: $r\,(K + N)$
- Recovery guarantee of Ling and Strohmer (up to log-factors)

$$L \geq C r^2 \left(K + \mu_h^2 N\right) \log \cdots$$

- Optimal in K and R, suboptimal in r
- Conjecture by Strohmer: Number of required measurements scales linear in r
- This is supported by numerical experiments

Main result

Theorem (Jung, Krahmer, S., 2016)

Let $\alpha \geq 1$. Assume that

$$L \geq C_\alpha r \left(K \log^2 K + N \mu_h^2 \right) \log^2 L \log \left(\gamma_0 r \right), \quad (1)$$

where

$$\gamma_0 = \sqrt{N \left(\log \left(\frac{NL}{2} \right) \right) + \alpha \log L}$$

and C_α is a universal constant only depending on α. Then with probability $1 - \mathcal{O}\left(L^{-\alpha} \right)$ the recovery program is successful, i.e. there exists X_0 is the unique minimizer of (SDP).

- (Near) optimal dependence on K, N , and r

Overview

Proof overview

Two main steps in the proof:

- Establishing sufficient conditions for recovery
  $\Rightarrow$ approximate dual certificate
- Constructing the dual certificate via Golfing Scheme

Proof overview

Two main steps in the proof:

- Establishing sufficient conditions for recovery
  $\Rightarrow$ approximate dual certificate
- Constructing the dual certificate via Golfing Scheme

- Crucial new ingredient for both steps:
  Restricted isometry property on $2r$-dimensional space

$$T = \Big\{ (u_1 m_1^* + h_1 v_1^*, \cdots, u_r m_r^* + h_r v_r^*) : u_1, \cdots, u_r \in \mathbb{R}^K, v_1, \cdots, v_r \in \mathbb{R}^N \Big\}$$

Intuition for T:
Directions of change when slightly varying the m_i's and h_i's

Restricted isometry property 1

Definition

We say that $\mathcal{A}$ fulfills the restricted isometry property on T for some $\delta > 0$, if for all $X = (X_1, \cdots, X_r) \in T$

$$(1-\delta)\sum_{i=1}^{r}\left\|X_i\right\|_F^2 \le \left\|\sum_{i=1}^{r}\mathcal{A}_i\left(X_i\right)\right\|_{\ell_2}^2 \le (1+\delta)\sum_{i=1}^{r}\left\|X_i\right\|_F^2.$$

[Ling, Strohmer, 2015]: Each operator $\mathcal{A}_i$ acts almost isometrically on

$$T_i = \left\{ u m_i^* + h_i v^* : \ u \in \mathbb{R}^K, \ v \in \mathbb{R}^N \right\}.$$

and $\mathcal{A}_i, \mathcal{A}_j$ are incoherent
$\Rightarrow$ r^2-bottleneck

Restricted isometry property 2

- **Observe:**
  Restricted isometry property for some $\delta > 0$ is equivalent to

$$\delta \geq \sup_{X \in T} \left| \| \sum_{i=1}^{r} \mathcal{A}_i\left(X_i\right) \|_{\ell_2}^2 - \sum_{i=1}^{r} \|X_i\|_F^2 \right|$$
$$= \sup_{X \in T} \left| \| \sum_{i=1}^{r} \mathcal{A}_i\left(X_i\right) \|_{\ell_2}^2 - \mathbb{E}\left[\| \sum_{i=1}^{r} \mathcal{A}_i\left(X_i\right) \|_{\ell_2}^2 \right] \right|.$$

- **Suprema of chaos processes:** The last term can be bounded using results from [Krahmer, Mendelson, Rauhut, 2014].

Overview

Open questions

- Generalization to more general random matrices
- Faster algorithms
- What if only a few number of devices are active? Does one obtain better recovery guarantees?
- Generalization to sparsity assumption on h (instead of a subspace assumption)