Block Sparsity and Sampling over a Union of Subspaces

Yonina Eldar and Moshe Mishali

DSP
July 5, 2009
Subspace Sampling

$\theta < 90^\circ$: unique $x$
$\theta \ll 90^\circ$: stable inversion

Oblique projection

$\theta < 90^\circ$
Perfect Reconstruction
Compressed Sensing

- Sample using few measurements

\[ y = A x \]

- Sampling
- Recovery Algorithms
- Guarantees

Unique/Stable mapping?
Compressed Sensing

- Sample using few measurements

\[ y = A x \]

Sampling

\[ y \rightarrow x \]

Unique/Stable mapping?

Recovery Algorithms

- \( \min_{x} ||x||_0 \) s.t. \( y = Ax \)
- \( \min_{x} ||x||_1 \) s.t. \( y = Ax \)
- Orthogonal matching pursuit

Guarantees

- Small RIP constant
- Small coherence
- \( m = O(k(\log(n/k) + 1)) \)
Sparsity Priors

- $x = \begin{array}{c}
\text{Well understood} \\
\text{d columns}
\end{array}$

- $X = \begin{array}{c}
\text{Joint sparsity} \\
\text{Many papers}
\end{array}$

Sparsity = Nonzeros
Sparsity Priors

- Multiband signal:

- Block-sparse signal:

\[ \mathbf{x} = \begin{array}{c}
\hline
\hline
\hline
\hline
\hline
\hline
\end{array} \]

\[ d_1 = 3 \quad d_2 = 4 \quad d_3 = 2 \quad d_4 = 4 \quad d_5 = 2 \]

More general notion of sparsity needed!
Union Sampling

\[ x \in \bigcup_i \mathcal{U}_i \]
where each \( \mathcal{U}_i \) is a subspace

- **Sampling**
- **Recovery Algorithms**
- **Guarantees**

Conditions for Unique/Stable mapping:
- *Lu and Do, 08*
- *Blumensath and Davies, 08*

Goal: Develop stable and efficient recovery algorithms over a union
Outline

- Key observation: Need structured union
- Finite settings: Develop recovery algorithms
  Prove equivalence guarantees
- Infinite union: Intro+Application

Sampling

Conditions for Unique/Stable mapping:
- Lu and Do, 08
- Blumensath and Davies, 08

Recovery Algorithms

- Convex relaxation
- Subspace OMP

Guarantees

- Block RIP
- Block coherence

(Eldar and Mishali) DSP’09 paper
(Eldar and Bolcskei) ICASSP’09 paper
Examples: Unions of Subspaces

\[ x \in \bigcup_{i} \mathcal{U}_i \quad \text{where each } \mathcal{U}_i \text{ is a subspace} \]

- **Sparsity**

  2 - sparse

  \[ \mathcal{U}_1 \ \mathcal{U}_2 \quad \text{...} \quad \mathcal{U}_\ell \quad \ell = \binom{8}{2} \]

- **Multiband signals**

- **Block-Sparsity**

  \[ x = \]

  \[ d_1 = 3 \quad d_2 = 4 \quad d_3 = 2 \quad d_4 = 4 \quad d_5 = 2 \]
Structured Model

\[ \mathcal{U} = \mathcal{A}_1 \oplus \ldots \oplus \mathcal{A}_k \] where \( \mathcal{A}_i \) is selected from a given set \( \{ \mathcal{A}_1, \ldots, \mathcal{A}_m \} \)

Examples:

\[ 0 \cdot \bar{\mathcal{A}} + 0 \cdot \bar{\mathcal{A}} + \alpha \cdot \bar{\mathcal{A}} + 0 \cdot \bar{\mathcal{A}} + \beta \cdot \bar{\mathcal{A}} + 0 \cdot \bar{\mathcal{A}} + \cdots \]

- Standard CS: \( \mathcal{A}_i \) is spanned by \( e_i \). The coefficients are scalars.
Structured Model

\[ \mathcal{U} = A_1 \oplus \ldots \oplus A_k \]  where \( A_i \) is selected from a given set \( \{ A_1, \ldots, A_m \} \)

**Examples:**

\[
\begin{align*}
\begin{array}{c}
0 \\
\end{array}
\end{align*}
\]

\[
\begin{align*}
\begin{array}{c}
\alpha \\
\beta
\end{array}
\end{align*}
\]

\[
\begin{align*}
\begin{array}{c}
\gamma \\
\delta \\
\epsilon
\end{array}
\end{align*}
\]

- **Standard CS**: \( A_i \) is spanned by \( e_i \). The coefficients are **scalars**.

- **Block sparsity**: \( A_i \) is spanned by \( d \) columns of the identity \( I \). The coefficients are **vectors**.

- **Multiband signals**: \( A_i \) is a frequency bin of width \( B \). The coefficients are **sequences**.
Key Result

Any structured union problem can be translated into block sparsity!

- Define a basis for each $\mathcal{A}_i$
- Any $x \in \mathcal{A}_i$ has a representation in terms of a vector $c[i]$ of length $d_i = \dim(\mathcal{A}_i)$
- If $\mathcal{A}_i$ is not in the sum then $c[i] = 0$
- $x$ in the union is represented by:

![Diagram showing block sparsity]

$d_1 = 3$  $d_2 = 4$  $d_3 = 2$  $d_4 = 4$  $d_5 = 2$

- Samples $y_i = \langle a_i, x \rangle$ are equivalent to $y = Dc$

May be continuous  Finite vector
Block Sparsity

\[ y = Dc \quad c - \text{block sparse} \]

- Convex optimization:
  \[ \min \sum_{i=1}^{m} \|c[i]\|_2 \quad \text{s.t.} \quad y = Dc \]
  minimize number of blocks with non-zero energy

- Subspace matching pursuit:
  choose block that best matches the residual

Both recover \( c \) under suitable conditions
Convex Relaxation

- $l_1$ - optimality is based on RIP
- Extend to block-RIP

$$(1 - \delta_k)\|c\|_2^2 \leq \|Dc\|_2^2 \leq (1 + \delta_k)\|c\|_2^2$$

For every block-$k$ sparse $c$ over $\mathcal{I} = \{d_1, \ldots, d_m\}$

Theorem:

If $\delta_{2k} < \sqrt{2} - 1$ then the convex relaxation is exact

(Eldar and Mishali, 08)
**Block-RIP**

- Block RIP constant is typically smaller than standard RIP

- Block RIP condition satisfied with high probability if
  \[ n \approx k \left( \log(m/k) + d \right) \]
  (conventional RIP  \[ n \approx k(d \log(m/k) + d) \])
Example

Block sparsity = 5

Our algorithm improves recovery over standard basis pursuit
Robust Recovery

- Suppose \( c \) is approximately block sparse and measurements are noisy
  \[ y = Dc + z \quad \|z\| \leq \varepsilon \]

- \( \min \sum_{i=1}^{m} \|c[i]\|_2 \quad \text{s.t.} \quad \|y - Dc\| \leq \varepsilon \)

Theorem:

If \( \delta_{2k} < \sqrt{2} - 1 \) then
\[
\|c_0 - c'\|_2 \leq \alpha \|c_0 - c^k\| + \beta \varepsilon
\]

(Eladar and Mishali, 08)

- \( c_0 \) - true vector
- \( c' \) - algorithm output
- \( c^k \) - best block approximation
- \( \alpha, \beta \) - known constants
Block Coherence

- Standard coherence: \( \mu = \max_{i \neq j} \langle d_i, d_j \rangle \)
- Block coherence:
  \[
  \mu_B = \max_{i \neq j} \frac{1}{d} \rho(D^H[i]D[j])
  \]

- \( \rho(A) \) - largest singular value
- \( d \) - block length
- \( D = (D[1] \ D[2]...D[m]) \)
  - \( d \) columns

- Properties:
  - \( 0 \leq \mu_B \leq 1 \)
  - \( \mu_B \leq \mu \) \( \rightarrow \) Improved recovery results
  - Operational meaning: uncertainty relation

(Eladar and Bolcskei, 08)
Recovery Conditions

Theorem:

A block sparse $c$ can be recovered from $y = Dc$ using convex relaxation if $kd < \frac{1}{2}(\mu_B^{-1} + d)$

(Eldar and Bolcskei, 08)

- If block structure is ignored then the condition becomes $kd < \frac{1}{2}(\mu^{-1} + 1)$
  \[ \mu^{-1} \leq \mu_B^{-1} \]
  \[ 1 \leq d \]
- Same conditions ensure recovery with subspace OMP
Sparsity vs. Union Sparsity

**Standard Sparsity**
- $K$ nonzero elements
- Optimization: $l_1$
- Greedy: OMP
- Small RIP
- Small coherence

**Union Sparsity**
- $K$ nonzero blocks/subspaces
- Optimization: mixed $l_2 / l_1$
- Greedy: subspace OMP
- Small block RIP
- Small block coherence
Sparsity vs. Union Sparsity

**Standard Sparsity**
- \( K \) nonzero elements
- Optimization: \( l_1 \)
- Greedy: OMP
- Small RIP
- Small coherence

**Union Sparsity**
- \( K \) nonzero blocks/subspaces
- Optimization: mixed \( l_2 / l_1 \)
- Greedy: subspace OMP
- Small block RIP
- Small block coherence

**Advantage of Union Sparsity:**
- Block coherence and block RIP are smaller than coherence and RIP
  - weaker equivalence conditions
- Empirical performance improvement
Can Treat an Infinite Union!

Application: Multipath Identification

\[ x_T(t) = \sum_{n \in \mathbb{Z}} g(t - nT) \]

Transmitted Signal

Multipath Medium

\[ x_R(t) = \sum_{k=1}^{K} \sum_{n \in \mathbb{Z}} a_k[n] g(t - t_k - nT) \]

Received Signal

Probing rate

\text{Infinite union of infinite dimensional subspaces}

Known pulse shape \( g(t) \) \rightarrow \text{Structure!}
Conclusion

- Efficient recovery for structured union of subspace
- Equivalence and stability using block RIP
- Equivalence using block coherence
- First step: future work should explore other structures

Theory of CS can be extended to subspaces
References


Thank you!