A CLUSTERING BASED FRAMEWORK FOR DICTIONARY BLOCK STRUCTURE IDENTIFICATION

Ender M. Ekşioğlu

Electronics and Communications Engineering Department, Istanbul Technical University, Istanbul, Turkey

The Question

If we are given a data set which is block-sparse over a known dictionary, how can we learn the underlying block structure?

Introduction

Utilizing an overcomplete dictionary to sparsely represent signals leads to efficient representations. The nonzero coefficients might occur in blocks, in which case the signals will be referred to as block-sparse signals. Attempts at block-sparse representation algorithms include the greedy Block Orthogonal Matching Pursuit (BOMP) [1], combined ℓ_2/ℓ_1 optimization and extensions of iterative sparse coding algorithms to model-based sparse representation. There have been some recent attempts at dictionary learning for block-sparse signals [2].

A common theme underlying both block-sparse signal representation and block-sparse dictionary design is the need for the dictionary block structure which defines how the atoms of the dictionary are grouped together into blocks. Hence, one important step in block-sparse signal processing is the recovery of the underlying block structure for a given block-sparsifying dictionary and data set pair.

The Solution

We propose a method for capturing the block structure of the atoms included in a block-sparsifying dictionary from the block-sparse data. The proposed method utilizes clustering at a certain stage, and we also develop a hierarchical agglomerative clustering algorithm suitable for use at this clustering stage. The clustering framework developed in this paper allows the use of different proximity measures between blocks, and it allows the use of standard clustering algorithms from literature.

Block-Sparsity Intro

The distribution of dictionary atoms to blocks can be given by an assignment vector $\boldsymbol{\Gamma} \in \mathbb{R}^K$. The set of the indices of atoms included in the j block of $\boldsymbol{\Gamma}$ will be denoted by $\boldsymbol{\Omega}_j^{\boldsymbol{\Gamma}}$, where the size of block j will be denoted as the cardinality $|\boldsymbol{\Omega}_j^{\boldsymbol{\Gamma}}|$. A vector $\boldsymbol{w}$ is said to be block k-sparse over $\boldsymbol{\Gamma}$, if its non-zero components occur in only k of the total B blocks, hence if $\|\boldsymbol{w}\|_{\boldsymbol{\Gamma}} = \|\boldsymbol{w}\|_{2,0}^{\boldsymbol{\Gamma}} = k$. A noise-free formulation for the block-sparse signal representation problem of signal vector $\boldsymbol{x}$ over a dictionary $\mathbf{D}$ and block structure $\boldsymbol{\Gamma}$ can be given as follows.

$$\hat{\boldsymbol{w}} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \|\boldsymbol{w}\|_{\boldsymbol{\Gamma}} \text{ s.t. } \boldsymbol{x} = \mathbf{D}\boldsymbol{w} \qquad (1)$$

Block-Sparse Signal Representation with Block Structure Identification

We define the block-sparse representation with block structure identification problem as follows.

$$\{\hat{\boldsymbol{\Gamma}}, \hat{\mathbf{W}}\} = \underset{\boldsymbol{\Gamma}, \mathbf{W}}{\operatorname{argmin}} \sum_{n=1}^{N} \|\boldsymbol{w}_n\|_{\boldsymbol{\Gamma}} \quad \text{s.t. } \mathbf{X} = \mathbf{D}\mathbf{W} \text{ and } |\boldsymbol{\Omega}_j^{\boldsymbol{\Gamma}}| \le s, j \in \boldsymbol{\Gamma} \qquad (2)$$

Main Algorithm

Input: $\mathbf{D}$, $\mathbf{X} = \{\boldsymbol{x}_n\}_{n=1}^N$

Goal: $\{\hat{\boldsymbol{\Gamma}}, \hat{\mathbf{W}}\} = \underset{\boldsymbol{\Gamma}, \mathbf{W}}{\operatorname{argmin}} \sum_{n=1}^{N} \|\boldsymbol{w}_n\|_{\boldsymbol{\Gamma}}$ s.t. $\mathbf{X} = \mathbf{D}\mathbf{W}$ and $|\boldsymbol{\Omega}_j^{\boldsymbol{\Gamma}}| \le s, j \in \boldsymbol{\Gamma}$.

1: Initialize $\mathbf{W}$ as the solution of the regular sparse representation problem: $\mathbf{W}^{(0)} = \underset{\mathbf{W}}{\operatorname{argmin}} \sum_{n=1}^{N} \|\boldsymbol{w}_n\|_0$ s.t. $\mathbf{X} = \mathbf{D}\mathbf{W}$

2: Find optimally block-sparsifying $\boldsymbol{\Gamma}$ for constant $\mathbf{W}^{(0)}$:

$$\hat{\boldsymbol{\Gamma}} = \underset{\boldsymbol{\Gamma}}{\operatorname{argmin}} \sum_{n=1}^{N} \|\boldsymbol{w}_n^{(0)}\|_{\boldsymbol{\Gamma}} \text{ s.t. } |\boldsymbol{\Omega}_j^{\boldsymbol{\Gamma}}| \le s, j \in \boldsymbol{\Gamma} \qquad (3)$$

3: Find optimally block-sparse $\mathbf{W}$ for constant $\hat{\boldsymbol{\Gamma}}$:

$$\hat{\mathbf{W}} = \underset{\mathbf{W}}{\operatorname{argmin}} \sum_{n=1}^{N} \|\boldsymbol{w}_n\|_{\hat{\boldsymbol{\Gamma}}} \text{ s.t. } \mathbf{X} = \mathbf{D}\mathbf{W} \qquad (4)$$

Framework for Block Structure Identification

The second step (3) of Alg.1 is where the optimally block-sparsifying block structure for a given representation matrix $\mathbf{W}$ should be found. Let us define what we call as the sparse representation indicator matrix, $\mathbf{I}_W = \mathcal{I}\{\mathbf{W}\}$. $\mathcal{I}\{\cdot\}$ is an indicator function which acts elementwise on the argument matrix. The identification of the block structure for the atoms can be simplified to clustering the rows of $\mathbf{I}_W$ into groups.

Block Structure Identification Algorithm

Input: $\mathbf{W} = \{\boldsymbol{w}_n\}_{n=1}^N$, and block size s.

Goal: Find optimal $\hat{\boldsymbol{\Gamma}} = \underset{\boldsymbol{\Gamma}}{\operatorname{argmin}} \sum_{n=1}^{N} \|\boldsymbol{w}_n\|_{\boldsymbol{\Gamma}}$.

1: Form the sparse representation indicator matrix $\mathbf{I}_W = \mathcal{I}\{\mathbf{W}\}$.

2: Apply clustering algorithm on the rows of $\mathbf{I}_W$, $\{\boldsymbol{i}^j\}_{j=1}^K$.

3: Form the block structure corresponding to the clustering of the rows of $\mathbf{I}_W$.

Clustering

Block structure identification algorithm as described by Alg.2 necessitates clustering of the K rows of the sparse representation indicator matrix $\mathbf{I}_W$ into groups. The steps for a hierarchical agglomerative algorithm which outputs clusters with at most s elements are given in Alg.3.

Examples for similarity distance metrics and linkage schemes are described below.

Hamming similarity:

$$\mathrm{d}(\boldsymbol{i}^{j_1}, \boldsymbol{i}^{j_2}) = \sum_{n=1}^{N} v_n \text{ where } v_n = \begin{cases} 1, & \boldsymbol{i}^{j_1}[n] = \boldsymbol{i}^{j_2}[n] \\ 0, & \boldsymbol{i}^{j_1}[n] \ne \boldsymbol{i}^{j_2}[n] \end{cases}.$$

Group average (GA) linkage:

$$\mathrm{sim}\{\boldsymbol{\Omega}_{m_1}, \boldsymbol{\Omega}_{m_2}\} = \frac{1}{|\boldsymbol{\Omega}_{m_1}||\boldsymbol{\Omega}_{m_2}|} \sum_{j_1 \in \boldsymbol{\Omega}_{m_1}, j_2 \in \boldsymbol{\Omega}_{m_2}} \mathrm{d}(\boldsymbol{i}^{j_1}, \boldsymbol{i}^{j_2}).$$

Clustering Algorithm

Hierarchical Agglomerative Clustering with maximal cluster size s (HAC-s)

Input: K elements to cluster $\{\boldsymbol{i}^j\}_{j=1}^K$ and maximal cluster size s.

Goal: Group the K elements into clusters $\boldsymbol{\Omega}_m$ such that there are at the most s elements in each cluster.

1: Initialize $\boldsymbol{\Omega}_m \leftarrow \{m\}, m = 1, 2, \ldots, K$.

$$\mathbf{A}(m_1, m_2) = \begin{cases} \mathrm{sim}\{\boldsymbol{\Omega}_{m_1}, \boldsymbol{\Omega}_{m_2}\} & m_1 \ne m_2 \\ 0 & m_1 = m_2 \end{cases}$$

2: **while** $\mathbf{A}$ is not all zeroes matrix **do**

3: Find the two clusters with similarity $\max(\mathbf{A})$ and join them into a single cluster.

4: Update the similarity matrix as

$$\mathbf{A}(m_1, m_2) = \begin{cases} \mathrm{sim}\{\boldsymbol{\Omega}_{m_1}, \boldsymbol{\Omega}_{m_2}\} & |\boldsymbol{\Omega}_{m_1}| + |\boldsymbol{\Omega}_{m_2}| \le s \\ 0 & |\boldsymbol{\Omega}_{m_1}| + |\boldsymbol{\Omega}_{m_2}| > s \end{cases}$$

5: **end while**

Simulations

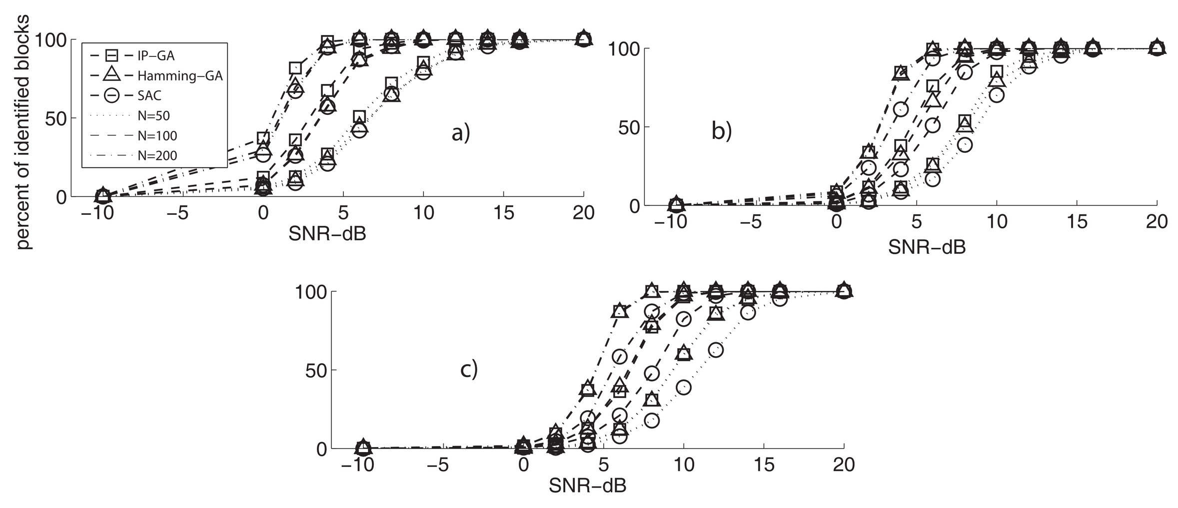

Figure 1: Percent of correctly identified blocks. The graphs present the block structure identification performance for HAC-s with IP-GA and Hamming-GA similarity measures and the SAC algorithm. a) $k = 4$ b) $k = 6$ c) $k = 8$.

References

[1] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, June 2010.

[2] K. Rosenblum, L. Zelnik-Manor, and Y. C. Eldar, "Dictionary optimization for block-sparse representations," *ArXiv e-prints*, http://arxiv.org/abs/1005.0202.