Blind Source Separation of Compressively Sensed Signals

Martin Kleinsteuber* and Hao Shen*
Geometric Optimization and Machine Learning Group
Technische Universität München, Munich, Germany
Email: kleinsteuber@tum.de, hao.shen@tum.de

Abstract—We present an approach to simultaneously separate and reconstruct signals from a compressively sensed linear mixture. We assume that the signals have a common sparse representation. The approach combines classical Compressive Sensing (CS) theory with a linear mixing model. Since Blind Source Separation (BSS) from a linear approach combines classical Compressive Sensing (CS) theory with a linear mixing model. Therefore, we assume that the signals have a common sparse representation. The problem of recovering signals from only the mixed observations without knowing the priori information of both the source signals and the mixing process is often referred to as Blind Source Separation (BSS), cf. [1]. Different BSS methods are used in various challenging data analysis applications, such as functional Magnetic Resonance Imaging (fMRI) analysis and microarray analysis. In order to achieve reasonable performance, prominent methods, e.g. Independent Component Analysis (ICA), usually require a large number of observations [2]. Unfortunately, the availability of a large amount of data samples can not be guaranteed in many real applications, due to either cost or time issues.

The theory of compressed sensing (CS), cf. [3] shows that, when a signal is sparse (or compressible) with respect to some basis, only a small number of samples suffice for exact (or approximate) recovery. It is interesting to know that the concept of sparsity has also been used as a separation criterion in the context of BSS [4]. Although a family of efficient algorithms in the probabilistic framework are proposed therein, the scenario with compressively sensed samples has not been studied and thus differs from our approach. In this work, the authors are interested in separating sparse signals which are compressively sampled.

II. PROBLEM DESCRIPTION

For the sake of convenience of presentation, signals are represented as column vectors, instead of the conventional row vectors. The instantaneous linear BSS model is given as follows

\[ Y = SA, \]

where \( S = [s_1, \ldots, s_m] \in \mathbb{R}^{n \times m} \) denotes the data matrix of \( m \) sources with \( n \) samples (\( m \ll n \)), \( A = [a_1, \ldots, a_k] \in \mathbb{R}^{n \times k} \) is the mixing matrix of full rank, and \( Y = [y_1, \ldots, y_k] \in \mathbb{R}^{m \times k} \) represents the \( k \) linear mixtures of \( S \). Here, we consider the scenarios with \( m \geq k \), i.e., the number of observed mixtures is less than or equal to the number of sources. The task of standard BSS is to estimate the sources \( S \), given only the mixtures \( Y \). We refer to [5] for more details.

We assume that all sources \( s_i \in \mathbb{R}^{n} \), for \( i = 1, \ldots, m \), have sparse representations with respect to the same basis, i.e., given \( \Psi \in \mathbb{R}^{n \times n} \) a basis of \( \mathbb{R}^{n} \), referred to as representation basis, each source \( s_i \) is assumed to have a \( \Psi \)-sparse representation with respect to \( \Psi \), denoted by \( x_i \in \mathbb{R}^{n} \), i.e.,

\[ s_i = \Psi x_i, \]  

or more compactly as

\[ S = \Psi X, \]

where \( X = [x_1, \ldots, x_m] \in \mathbb{R}^{n \times m} \).

Now let us take one step further to compressively sample each mixture \( y_i \in \mathbb{R}^{n} \) individually by a sampling basis \( \Phi_i \in \mathbb{R}^{p_i \times n} \) for \( i = 1, \ldots, k \). Then, a compressively sensed observation \( \tilde{y}_i \in \mathbb{R}^{p_i} \) of the \( i \)-th mixture is constructed as

\[ \tilde{y}_i = \Phi_i y_i = \Phi_i \Psi X a_i, \]

We refer to (4) as the compressively sensed BSS (CS-BSS) model.

The task of our work is then formulated as follows: Given the common presentation basis \( \Psi \in \mathbb{R}^{n \times n} \) and the compressively sensed observations \( \tilde{y}_i \in \mathbb{R}^{p_i} \), for \( i = 1, \ldots, k \), together with their corresponding sampling bases \( \Phi_i \in \mathbb{R}^{p_i \times n} \), estimate the mixing matrix \( A \in \mathbb{R}^{n \times k} \) and the sparse representations \( X \in \mathbb{R}^{n \times m} \). Following the well-known argument that the mixing matrix \( A \) is identifiable only up to a column-wise scaling and permutation, without loss of generality, we restrict the mixing matrix \( A \) onto the \( m \times k \) oblique manifold \( OB(m, k) \), which is defined as

\[ OB(m, k) := \left\{ A \in \mathbb{R}^{m \times k} \mid \text{rk}(A) = k, \text{ddiag}(A^T A) = I_k \right\}, \]

where \( I_k \) is the \( k \times k \) identity matrix, and \( \text{ddiag}(Z) \) forms a diagonal matrix, whose diagonal entries are those of \( Z \).

It is unavoidable that, in real applications, observations \( \tilde{y}_i \) are contaminated by noise. In other words, the equalities defined in (4) do not hold in general. In the sense of least squares error, we propose the following cost function

\[ f: OB(m, k) \times \mathbb{R}^{n \times m} \rightarrow \mathbb{R}, \]

\[ f(A, X) := \|X\|_1 + \sum_{i=1}^{k} \lambda_i \| \Phi_i \Psi X a_i - \tilde{y}_i \|^2_2, \]

where the scalars \( \lambda_i \in \mathbb{R}^+ \) weigh the reconstruction error of each mixture individually, and balance these errors against the sparsity term \( \|X\|_1 \). In this work, we provide an analysis of the cost function (6) and propose a geometric conjugate gradient method. The performance of our proposed approach is investigated by numerical experiments.

REFERENCES