Optimization Algorithms for Sparse Representations and Applications

Pando Georgiev[†], Fabian Theis[†‡] and Andrzej Cichocki[†]
[†] Brain Science Institute, RIKEN
Lab. for Advanced Brain Signal Processing
2-1, Hirosawa, Wako-shi, Saitama, 351-0198, Japan
[‡] Institute of Biophysics, University of Regensburg D-93040 Regensburg, Germany
georgiev,cia@bsp.brain.riken.go.jp, fabian@theis.name

Abstract

We consider the following *sparse representation problem*, which is called Sparse Component Analysis: identify the matrices $\mathbf{S} \in \mathbb{R}^{n \times N}$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$ $(m \leq n < N)$ uniquely (up to permutation of scaling), knowing only their multiplication $\mathbf{X} = \mathbf{AS}$, under some conditions, expressed either in terms of $\mathbf{A}$ and sparsity of $\mathbf{S}$ (*identifiability* conditions), or in terms of $\mathbf{X}$ (*Sparse Component Analysis* conditions). A crucial assumption (sparsity condition) is that $\mathbf{S}$ is *sparse of level* k in sense that each column of $\mathbf{S}$ has at most $m - k$ nonzero elements $(k = 1, 2, ..., m - 1)$.

We present two type of optimization problems for such identification. The first one is used for identifying the mixing matrix $\mathbf{A}$: this is a typical clustering type problem aimed to finding hyperplanes in $\mathbb{R}^m$ which contain the columns of $\mathbf{X}$. We present several optimization algorithms for this clustering problem together with modifications for large scale problems.

The second type of problem is those of identifying the source matrix $\mathbf{S}$. This corresponds to finding a sparse solution of a linear system (a difficult combinatorial problem), for which we propose a new algorithm.

Applications include Blind Signal Separation of under-determined linear mixtures of signals in which the sparsity is either given a priori, or obtained with some preprocessing techniques as wavelets, filtering, etc. A typical example is a mixture of images, for which the Haar wavelets give very good sparsification. Very suitable for our method are the EEG data sets. In such a way we can recover some brain signal features. Another applications include separation of special types of nonlinear mixtures, convolutive mixtures, nonlinear clustering problems, etc., after reduction of the problem to a linear one in a high dimensional space.