Appl. Comput. Harmon. Anal. 36 (2014) 495–507

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Optimized projections for compressed sensing via rank-constrained nearest correlation matrix

Nicolae Cleju

Faculty of Electronics, Telecommunications and Information Technology, "Gheorghe Asachi" Technical University of Iasi, Romania

ARTICLE INFO

Article history:
Received 7 March 2013
Received in revised form 31 July 2013
Accepted 26 August 2013
Available online 4 September 2013
Communicated by Jared Tanner

Keywords:
Acquisition
Compressed sensing
Nearest correlation matrix
Optimization

ABSTRACT

Optimizing the acquisition matrix is useful for compressed sensing of signals that are sparse in overcomplete dictionaries, because the acquisition matrix can be adapted to the particular correlations of the dictionary atoms. In this paper a novel formulation of the optimization problem is proposed, in the form of a rank-constrained nearest correlation matrix problem. Furthermore, improvements for three existing optimization algorithms are introduced, which are shown to be particular instances of the proposed formulation. Simulation results show notable improvements and superior robustness in sparse signal recovery.

1. Introduction

Compressed Sensing (CS) [1] studies the possibility of acquiring a signal x that is a priori known to be sparse in some dictionary D with fewer linear measurements than required by the traditional sampling theorem. In many cases the dictionary D is an orthogonal basis, but we consider here the general case of an overcomplete dictionary.

Consider a signal $x \in \mathbb{R}^n$ that is sparse in some dictionary $D \in \mathbb{R}^{n\times N}$, i.e. x has at least one decomposition γ that has few non-zero coefficients. A number of $m < n$ linear measurements are taken as inner products of x with a set of m projection vectors, arranged as the rows of an acquisition matrix $P \in \mathbb{R}^{m\times n}$

$$y = Px = \underbrace{PD}_{D_e}\gamma. \tag{1}$$

The equation system (1) is undetermined. Under certain conditions on P and D [2], a sufficiently sparse decomposition vector γ is shown to be the unique solution to the optimization problem

$$\hat{\gamma} = \arg\min_{\gamma} \|\gamma\|_{\ell_0} \quad \text{subject to } y = PD\gamma, \tag{2}$$

where $\|\gamma\|_{\ell_0}$ is the number of non-zero elements of the vector γ (the ℓ_0 "norm"). Solving (2) means finding the sparsest decomposition of y in the effective dictionary $D_e := PD$, which is the computational expensive stage of the process, with a large number of algorithms developed for this purpose. After obtaining the approximate decomposition vector $\hat{\gamma}$, the reconstructed signal $\hat{x}$ is obtained as

$$\hat{x} = D\hat{\gamma}. \tag{3}$$

E-mail address: nikcleju@etti.tuiasi.ro.

1063-5203/$ – see front matter
http://dx.doi.org/10.1016/j.acha.2013.08.005

The strict condition $y = PD\gamma$ in (2) is often unrealistic, and therefore a practical version of (2) is

$$\hat{\gamma} = \arg\min_{\gamma} \|\gamma\|_{\ell_0} \quad \text{subject to } \|y - PD\gamma\| \leqslant \epsilon, \tag{4}$$

where ϵ takes into account possible noisy measurements and approximately sparse signals.

Unfortunately, finding the exact solution of the ℓ_0 minimization problem (2) is combinatorial and NP-hard. One of the ways to circumvent this is replacing the ℓ_0 norm with ℓ_1, leading to a tractable convex optimization problem

$$\hat{\gamma} = \arg\min_{\gamma} \|\gamma\|_{\ell_1} \quad \text{subject to } y = PD\gamma, \tag{5}$$

which requires however more strict conditions on P and D to guarantee the uniqueness of the solution. This is known as Basis Pursuit (BP) [3]. The ℓ_1 problem can be converted to a linear program, which is well known in literature and has many efficient solving algorithms available. A second option is to settle with a possibly sub-optimal solution of (2), using a pursuit or thresholding algorithm [4,5] to estimate a solution to (2). In both cases, robustness to noise can be enforced by replacing the strict condition $y = PD\gamma$ with a robust $\|y - PD\gamma\|_2 \leqslant \epsilon$.

The choice of the acquisition matrix P is governed by the principle of incoherence with D: a “good” acquisition matrix has its rows (i.e. the projection vectors) incoherent with the columns of D. Coherence measures the largest correlation between two sets of vectors, and thus incoherence requires a low maximal correlation. Random projections vectors were shown to be a good choice with orthogonal bases [6], since random vectors are incoherent with any fixed basis with high probability. In the overcomplete case, a better acquisition matrix can often be found if one takes into account the correlations between dictionary atoms, since it is not uncommon that dictionaries exhibit significant atom correlation. This is especially true with dictionaries that are learned, i.e. optimized for a particular set of signals. As such, a number of algorithms have been developed for finding optimized projections for signals that are sparse in overcomplete dictionaries [7–9].

This paper proposes modifications for improving three existing algorithms for finding optimized projections. Further, we show that our improvements can be unified in a single formulation based on solving a *rank-constrained nearest correlation matrix* problem [10]. The rest of this paper is organized as follows. In Section 2 we review the main condition for perfect recoverability that underlies most of the considered algorithms. Section 3 presents three state-of-the-art algorithms for finding optimized projections. Improvements for all of them are proposed in Section 4, and we present the proposed unified formulation in Section 5. Simulation results are presented in Section 6. Finally, conclusions are drawn in Section 7.

Throughout this paper we use the following notations. The acquired signal is an n-dimensional vector x, the dictionary is D of size $n \times N, n < N$. A decomposition of x in D is typically denoted as γ, i.e. $x = D\gamma$. The acquisition matrix is P of size $m \times n, m < n$. The product $D_e := PD$ is the effective dictionary, of size $m \times N$. The Gram matrix of D is denoted $G := D^T D$, while the Gram matrix of the effective dictionary D_e is denoted G_e and referred to as *effective Gram matrix*.

2. Acquisition matrices and mutual coherence

A widely used approach to ensure the uniqueness of the solution $\hat{\gamma}$ in (2) or (5) uses the mutual coherence of the effective dictionary $D_e := PD$. The mutual coherence of a dictionary is defined as the maximum absolute value of the inner products of any two of its normalized columns [11]. Thus, the mutual coherence of D_e is the maximum absolute off-diagonal value of the Gram matrix $G_e := D_e^T \cdot D_e$, after normalizing the columns of D_e. The mutual coherence provides a lower bound for the perfect recovery of sparse signals, as shown in Theorem 1 [11–13]:

Theorem 1. *Consider an overcomplete dictionary D with mutual coherence $\mu(D)$ and a signal x such that $x = D\gamma$. If condition (6) is true*:

$$\|\gamma\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(D)}\right), \tag{6}$$

then the following hold:

1. *γ is the sparsest decomposition of x in D, i.e. it is the solution of the optimization problem*

$$\arg\min_{\gamma} \|\gamma\|_0 \quad \textit{subject to } x = D\gamma,$$

2. *γ is recoverable using ℓ_1 minimization [3], i.e. it is also the solution of*

$$\arg\min_{\gamma} \|\gamma\|_1 \quad \textit{subject to } x = D\gamma,$$

3. *γ is recoverable using Orthogonal Matching Pursuit [4].*