Signal Processing 144 (2018) 341–351

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Nonconvex penalties with analytical solutions for one-bit compressive sensing

Xiaolin Huang[a], Ming Yan[b,*]

[a] Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, and MOE Key Laboratory of System Control and Information Processing, Shanghai 200240, PR China

[b] Department of Computational Mathematics, Science and Engineering and Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 27 May 2017
Revised 19 August 2017
Accepted 21 October 2017
Available online 23 October 2017

Keywords:
One-bit compressed sensing
Non-convex penalty
Analytical solutions

ABSTRACT

One-bit measurements widely exist in the real world and can be used to recover sparse signals. This task is known as one-bit compressive sensing (1bit-CS). In this paper, we propose novel algorithms based on both convex and non-convex sparsity-inducing penalties for robust 1bit-CS. We consider the dual problem, which has only one variable and provides a sufficient condition to verify whether a solution is globally optimal or not. For positive homogeneous penalties, a globally optimal solution can be obtained in two steps: a proximal operator and a normalization step. For other penalties, we solve the dual problem, and it needs to evaluate the proximal operators for many times. Then we provide fast algorithms for finding analytical solutions for three penalties: minimax concave penalty (MCP), ℓ_0 norm, and sorted ℓ_1 penalty. Specifically, our algorithm is more than 200 times faster than the existing algorithm for MCP. Its efficiency is comparable to the algorithm for the ℓ_1 penalty in time, while its performance is much better than ℓ_1. Among these penalties, sorted ℓ_1 is most robust to noise in different settings.

Published by Elsevier B.V.

1. Introduction

Analog-to-digital converting (ADC) is a necessary process in digital processing, and the choice of the bit-depth is an important issue. The extreme case is to use one-bit measurements, which enjoy many advantages, e.g., they can be implemented by one low power comparator running at a high rate. Mathematically, one-bit compressive sensing (1bit-CS) is to recover a K-sparse vector $\mathbf{x} \in \mathbb{R}^n$ ($\|\mathbf{x}\|_0 \leq K$) from m one-bit quantized measurements

$$y_i = \operatorname{sgn}(\mathbf{u}_i^\top \mathbf{x} + \varepsilon_i), \tag{1}$$

where $\mathbf{u}_i \in \mathbb{R}^n$ is the ith sensing vector, ε_i is the noise in the measurement, and the function sgn returns 1 for a positive number and -1 otherwise. The sensing system and measurements are represented by $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m]$ and $\mathbf{y} = [y_1, y_2, \ldots, y_m]^\top$, respectively. Due to the low power and high sampling rate, one-bit measurements have been applied in the estimation of frequency, phase, and direction of arrival (DOA) [1–3]. For example, in the DOA estimation, a radar with one-bit measurements has a higher scan speed than others. One-bit measurements are also attractive in distributed networks [4,5], where the use of one-bit measurements largely reduces the communication load.

If the underlying signal is sparse, then sparsity pursuit techniques can help signal recovery, which is similar to the regular compressive sensing. Therefore, since its proposal by Boufounos and Baraniuk [6], 1bit-CS has attracted much attention in both the signal processing society [7–10] and the machine learning society [11–14]. Because the one-bit information has no capability to specify the magnitude of the original signal, we assume $\|\mathbf{x}\|_2 = 1$ without loss of generality (there is also some work on norm estimation, see, e.g., [15]), and 1bit-CS can be explained as finding the sparest vector on the unit sphere that coincides with the observed signs, i.e,

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & \|\mathbf{x}\|_0, \\ \text{subject to} \quad & y_i = \operatorname{sgn}(\mathbf{u}_i^\top \mathbf{x}), \quad \forall i = 1, 2, \ldots, m, \\ & \|\mathbf{x}\|_2 = 1. \end{aligned} \tag{2}$$

This is an NP-hard problem, and several algorithms are developed to approximately solve it or its variants [6,7,16,17]. The constraint in (2) does not tolerate noise or sign flips, and it may exclude the real signal from the feasible set. Additionally, the feasible set may be empty, and there is no solution for (2). One way to deal with noise and sign flips is to replace the constraint $y_i = \operatorname{sgn}(\mathbf{u}_i^\top \mathbf{x})$ by a loss function. For example, the one-sided ℓ_1 loss and the one-sided

* Corresponding author.
E-mail addresses: xiaolinhuang@sjtu.edu.cn (X. Huang), yanm@math.msu.edu (M. Yan).

https://doi.org/10.1016/j.sigpro.2017.10.023
0165-1684/Published by Elsevier B.V.

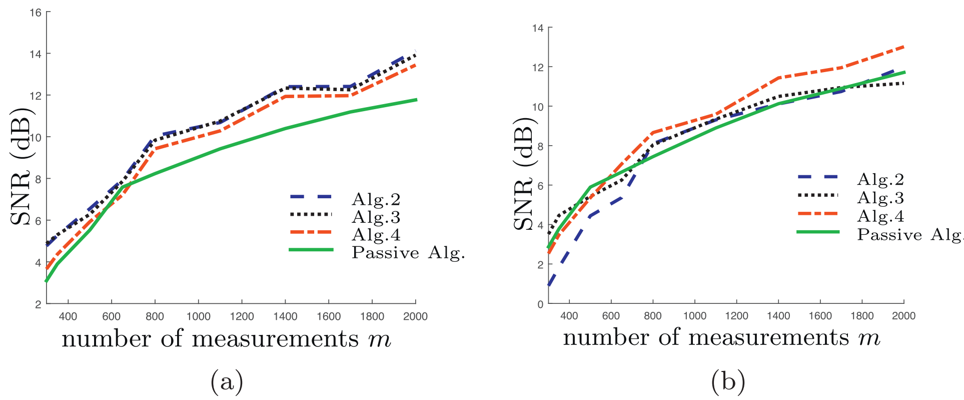

Fig. 1. Recovery performance for different numbers of measurements: MCP minimization (blue dashed line), ℓ_0 minimization (black dotted line), sorted ℓ_1 penalty (red dash-dotted line), and ℓ_1 minimization (green solid line). In this experiment $n = 1000, K = 15, s_n = 10$, and sign flip ratio is 10%. (a) using the ideal parameters; (b) using parameters selected by 10-fold cross-validation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ℓ_2 loss are considered in [8,18]; the linear loss is used in [9,11]. It is reported that the linear loss generally outperforms the one-sided ℓ_1/ℓ_2 loss. Moreover, with proper regularization terms and constraints, the linear loss minimization can be solved analytically and enjoys great computational effectiveness.

In regular CS problems, non-convex penalties have been insightfully investigated and widely applied to enhance sparsity. Similarly, those non-convex techniques are applicable to 1bit-CS, and the recovery performance is expected to be improved. One obvious barrier is that nonconvex penalties lead to nonconvex problems, which are usually difficult to solve. An interesting result is recently represented in [12], which gives analytical solutions for two nonconvex penalties, namely the smoothly clipped absolute deviation (SCAD, [19]) and minimax concave penalty (MCP, [20]). Also Chen and Banerjee [13] propose an algorithm for 1bit-CS using the k-support norm. These nonconvex penalties are shown to obtain better results than convex ones in both theory and practice [12,13] and, therefore, have been extended to other applications including the multi-label learning task [21].

In this paper, we discuss more convex and nonconvex penalties, for which analytical solutions can be obtained, and we provide fast algorithms for finding these solutions. These penalties include SCAD, MCP, ℓ_p-norm ($0 \le p \le +\infty$, [22]), $\ell_1 - \ell_2$ norm [23], sorted ℓ_1 penalty [25,26], and so on. The contributions of this paper can be summarized as follows.

- We analyze a generic model for 1bit-CS and provide a sufficient condition for the global optimality.
- For positive homogeneous penalties, we show that an optimal solution can be obtained in two steps: a proximal operator and a normalization step. For general penalties, we provide a generic algorithm by solving the dual problem.
- We provide algorithms for finding analytical solutions for three nonconvex penalties: MCP, ℓ_0 norm, and the sorted ℓ_1 penalty. These algorithms are much faster than the existing 1bit-CS algorithms for nonconvex penalties and even comparable to that for the convex ℓ_1 minimization problem, e.g., our algorithm is averagely 200 times faster than the algorithm given in [12] for MCP. In addition, we compare these nonconvex penalties with the convex ℓ_1 penalty and show that the sorted ℓ_1 performs the best in both performance and computational time.

The rest of this paper is organized as follows. Section 2 briefly reviews the existing related 1bit-CS algorithms. The main contributions, i.e., analytical solutions for different penalties and corresponding algorithms, are presented in Section 3. The numerical experiments are reported in Section 4. We end this paper with a brief conclusion.

2. Related works

Model (2) for 1bit-CS has two main disadvantages: (i) it is difficult to solve because of the ℓ_0 norm in the objective and the constraint $\|\mathbf{x}\|_2 = 1$; (ii) the constraint $y_i = \mathrm{sgn}(\mathbf{u}_i^\top \mathbf{x})$ does not consider noisy sign measurements.

Several approaches are given to deal with both disadvantages. For the nonconvexity, the ℓ_0 norm is replaced by the ℓ_1 norm, and the constraint $\|\mathbf{x}\|_2 = 1$ is replaced by other convex constraints. The first convex model [27] for 1bit-CS is

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & \|\mathbf{x}\|_1, \\ \text{subject to} \quad & y_i(\mathbf{u}_i^\top \mathbf{x}) \ge 0, \ \ \forall i = 1, 2, \ldots, m, \\ & \|\mathbf{U}^\top \mathbf{x}\|_1 = r, \end{aligned} \tag{3}$$

where r is a given positive constant. In fact, the solutions for all positive r's have the same direction and the difference is only on the magnitudes of the reconstructed signals.

However, (3) still cannot be applied when there are noisy measurements, because, it, same as (2), requires the sign consistence in the measurements. Noisy measurements come from both the noise during the acquisition before the quantization and sign flips during the transmission. To deal with noisy measurements, Jacques et al. [18] introduces the following robust model using the one-sided ℓ_1 norm,

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & \frac{1}{m}\sum_{i=1}^{m} \max\left\{0, -y_i(\mathbf{u}_i^\top \mathbf{x})\right\}, \\ \text{subject to} \quad & \|\mathbf{x}\|_2 = 1, \\ & \|\mathbf{x}\|_0 = K. \end{aligned}$$

The robust model using the one-sided ℓ_2 norm is also introduced. Several modifications are designed by Yan et al. [8], Bahmani et al. [28], Dai et al. [29] to improve their robustness to sign flips and noise.

The linear loss for robust 1bit-CS attracts more attention because of its good performance and simplicity. Based on the linear loss, many results on sampling complexities are given recently [9,11–13]. In [9], the first model using the linear loss for 1bit-CS is proposed and takes the following formulation,

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & -\frac{1}{m}\sum_{i=1}^{m} y_i(\mathbf{u}_i^\top \mathbf{x}), \\ \text{subject to} \quad & \|\mathbf{x}\|_2 \le 1, \\ & \|\mathbf{x}\|_1 \le s, \end{aligned} \tag{4}$$

where s is a given positive constant. One can also put the ℓ_1-norm in the objective instead of in the constraint, resulting in the