Signal Processing 145 (2018) 167–174

Contents lists available at ScienceDirect

Signal Processing

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

A new accelerated alternating minimization method for analysis sparse recovery

Jiaxin Xie [a,b,*], Anping Liao [a], Yuan Lei [a]

[a] College of Mathematics and Econometrics, Hunan University, Changsha 410082, China
[b] LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 14 July 2017
Revised 14 November 2017
Accepted 6 December 2017
Available online 7 December 2017

MSC:
94A12
90C25
68U10

Keywords:
Compressed sensing
Moreau proximal smoothing
Accelerated alternating minimization
Sparse recovery
Total variation

ABSTRACT

This paper proposes a new method based on accelerated alternating minimization (AAM) for analysis sparse recovery. This method is extremely attractive as (1) it is very simple and computationally efficient, (2) it exhibits a fast convergence rate, (3) it is flexible and amenable to many kinds of reconstruction problems. We establish the connection between the classical alternating minimization (AM) method and the well-known proximal gradient (PG) method. Thus combining the accelerated proximal gradient (APG) method with the Moreau proximal smoothing technique, a new smoothing-based AAM (SAAM) method, which can obtain an ϵ-optimal solution within $O(1/\epsilon)$ iterations, is designed. Numerical experiments on randomly generated data and real image reconstruction show that this method compares favorably with several state-of-the-art methods.

1. Introduction

Compressed sensing [14] is a very active field of recent research which covers a wide range of applications, including signal processing, medical imaging, seismology, and statistics. It predicts that sparse signals can be reconstructed from incomplete measurements using efficient reconstruction methods. Formally, in compressed sensing, one considers the following linear form:

$$b = Ax + e, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ is a known measurement matrix, $b \in \mathbb{R}^m$ is the observation vector, and $e \in \mathbb{R}^m$ represents the noise term. The goal is to reconstruct the unknown signal x based on A and b.

Recently, the analysis sparsity model (or cosparsity model) [13,17,20,24] has attracted significant interest as well, where it is assumed that the signal x possesses a structure with respect to the given matrix $D \in \mathbb{R}^{p\times n}$ in the sense that the application of D to x produces a sparse vector, i.e., Dx is sparse. We refer to D as a analysis operator. If the ℓ_2 norm of the noise e is bounded by ϵ, then the recovery problem can be formulated as the following constrained analysis based approach, referred to as analysis basis pursuit (ABP):

$$\min_{x\in\mathbb{R}^n} \|Dx\|_1 \quad \text{subject to} \quad \|Ax-b\|_2 \le \epsilon. \tag{2}$$

Alternatively, the following unconstrained analysis based approach is more widely studied in many applications:

$$\min_{x\in\mathbb{R}^n} \|Dx\|_1 + \frac{\mu}{2}\|Ax-b\|_2^2, \tag{3}$$

which is called analysis LASSO (ALASSO). Here $\mu > 0$ denotes the regularization parameter, and $\|\cdot\|_1$ and $\|\cdot\|_2$ stand for the ℓ_1 and ℓ_2 norm, respectively. In fact, ABP is equivalent to ALASSO in the sense that for any $\epsilon > 0$ there exists a μ for which the optimal solutions of ABP and ALASSO are identical. In addition to the above ℓ_1-convex relaxation methods, there are some other alternative approaches including greedy-type algorithms such as Greedy Analysis Pursuit (GAP) [20,24], thresholding-based methods [15,27] or reweighted ℓ_1-minimization [10]. In this paper, we focus on the ALASSO formulation (3).

In the last few years, several efficient algorithms have been proposed and studied for solving the optimization problems ABP and ALASSO [7,18,21,22,33]. The algorithm YALL1 developed by Yang

* Corresponding author.
E-mail addresses: xiejiaxin@hnu.edu.cn (J. Xie), liaoap@hnu.edu.cn (A. Liao), yleimath@hnu.edu.cn (Y. Lei).

https://doi.org/10.1016/j.sigpro.2017.12.010
0165-1684/

and Zhang [33], which is an alternating direction algorithm, is able to solve the constrained problem (2) and the unconstrained formulation (3). In [7], based on a smoothing technique studied by Nesterov in [25], a first-order algorithm called NESTA was proposed. However, such two algorithms require the measurement matrix A to have a nice structure. YALL1 could be prone to a slow convergence if the rows (or columns) of A are neither orthonormalized nor normalized and NESTA requires the matrix A^TA to be an orthogonal projector. Recently, two fast algorithms called SFISTA and DFISTA were proposed in [31] by Tan et al. where the authors used the smoothing and decomposition transformations to relax the original sparse recovery problem (3) and then a monotone version of the fast iterative shrinkage-thresholding algorithm (MFISTA) [4] is implemented on the relaxed formulation. Experimental results provided in [31] show SFISTA converges faster than DFISTA. Although SFISTA can avoid imposing conditions on the measurement A, it is very slow as will be shown later in the numerical results section.

Alternating minimization (AM) method, which is a rather old and fundamental algorithm, is attractive due to its simplicity and efficiency. However, it has also been recognized as a slow method. Very recently, Beck [3, Theorem 5.4] studied the convergence properties of the AM method with decomposition-based approach designed to solve a composite convex model, and shown that $O(1/\epsilon^2)$ iterations are needed to obtain an ϵ-optimal solution.

In this paper, we aim to design a faster method than the classical AM method for analysis sparse recovery, in the sense that the computational effort of the new method will keep the simplicity of AM, while its convergence rate will be significantly better. The main contributions of the work are as follows.

- The Moreau proximal smoothing technique is adopted to transform the nondifferentiable and nonseparable term $\|Dx\|_1$ in (3) into a smooth counterpart, and then we establish the equivalence between the classical AM method and the well-known proximal gradient (PG) method.
- Based on the techniques and results of the accelerated proximal gradient (APG) method, a new smoothing-based accelerated alternating minimization (SAAM) that keeps the simplicity and efficiently of AM but shares the fast convergence rate is constructed for solving the analysis sparse recovery problem. In addition, we also show that the proposed SAAM method can obtain an ϵ-optimal solution within $O(1/\epsilon)$ iterations, which, of course, improves the convergence rate established in [3].
- To verify the efficiency of the proposed method, we compare SAAM with several state-of-the-art solvers on randomly generated data, such as YALL1 [33], NESTA [7] and SFISTA [31], which had been shown to be favorable among other algorithms including the interior point method (e.g., $\ell_1-\ell_s$) [18], the nonlinear conjugate gradient descend (CGD) [22] algorithm, the generalized iterative soft-thresholding (GIST) [21] algorithm, etc. Furthermore, we also show that our results are flexible and useful for many types of reconstruction problems such as the TV-based image restoration problem [12,19,28,32]. Numerical experiments indicate that SAAM is more efficient than several state-of-the-art TV solvers for image restoration, such as FTVd [32] and ADMM-based methods [1,2,12,19].

The remaining parts of this paper are organized as follows. In Section 2, we give a quick review of the proximal map and the PG method. In Section 3, we propose the SAAM method. In Section 4, we will establish the convergence properties of the algorithm. In Section 5, we provide the numerical results. Finally, we conclude this paper in Section 6.

2. Proximal map and proximal gradient method

2.1. Proximal map

In this subsection, we will recall the definition and some fundamental properties of Moreau's proximal map, which is essential for us to establish the smoothing-based algorithm for the analysis sparse recovery problem (3).

Given a proper closed convex function $g:\mathbb{R}^p\to(-\infty,+\infty]$ and any $t>0$, the proximal map associated to g is defined by

$$\mathrm{prox}_{tg}(x) := \arg\min_u \left\{ g(u) + \frac{1}{2t}\|u-x\|_2^2 \right\}, \tag{4}$$

and define

$$g_t(x) := \inf_u \left\{ g(u) + \frac{1}{2t}\|u-x\|_2^2 \right\}. \tag{5}$$

The function g_t is called the Moreau envelope of g and enjoys several important properties. The next proposition records three such important properties, for a proof see [25, Theorem 1], [6, Theorem 4.1], [30, Proposition 2.2].

Proposition 1. *Let $g:\mathbb{R}^p\to(-\infty,+\infty]$ be a closed proper convex function and let $D\in\mathbb{R}^{p\times n}$ be a given matrix. For any $y\in\mathbb{R}^n$ and $t>0$, the following results hold:*

(1) The function $g_t(Dy)$ is continuously differentiable with the gradient given by

$$\nabla g_t(Dy) = \frac{1}{t}D^T(Dy - \mathrm{prox}_{tg}(Dy)).$$

(2) Let D^T denote the transpose of D, then for every $y, y'\in\mathbb{R}^n$,

$$g_t(Dy') \le g_t(Dy) + \langle \nabla g_t(Dy), y'-y\rangle + \frac{1}{2t}\langle y'-y, D^TD(y'-y)\rangle.$$

(3) Suppose the subgradients of g over $\mathbb{R}^p$ are bounded by L_g, then

$$g(Dy) - \frac{t}{2}L_g^2 \le g_t(Dy) \le g(Dy).$$

2.2. Proximal gradient method

For the purpose of our analysis, we consider the following general convex optimization model:

$$\min_{u\in\mathbb{R}^p} \left\{ \Phi(u) := \phi(u) + \varphi(u) \right\}. \tag{6}$$

Here, $\phi:\mathbb{R}^p\to(-\infty,+\infty]$ is an extended-valued, proper, closed and convex function (possible nonsmooth); $\varphi:\mathbb{R}^p\to\mathbb{R}$ is convex and continuously differentiable with Lipschitz continuous gradient. Given any symmetric positive semidefinite matrix H, define $\omega(\cdot,\cdot):\mathbb{R}^p\times\mathbb{R}^p\to\mathbb{R}$ by

$$\omega_H(u,w) := \varphi(w) + \langle \nabla\varphi(w), u-w\rangle + \frac{1}{2}\langle u-w, H(u-w)\rangle.$$

Then for any $u, w\in\mathbb{R}^p$, there exists a symmetric positive semidefinite matrix H such that

$$\varphi(u) \le \omega_H(u,w). \tag{7}$$

For any $u^0\in\mathbb{R}^p$, the kth iteration of proximal gradient (PG) for solving (6) takes the following form (see [5]):

$$u^{k+1} := \arg\min_{u\in\mathbb{R}^p} \left\{ \phi(u) + \omega_H(u,u^k) \right\}, \tag{8}$$

where H is a symmetric positive definite matrix such that (7) holds. The main disadvantage of the PG method is that it suffers from a relatively slow $O(1/k)$ rate of convergence of the function values, i.e., $\Phi(u^k)-\Phi^*\simeq O(1/k)$. Here, Φ^* stands for the optimal value for the problem (6). An accelerated version is the accelerated proximal gradient (APG) method [5,16], which shares the