Journal of Computational and Applied Mathematics 311 (2017) 115–129

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

The proximal alternating iterative hard thresholding method for l_0 minimization, with complexity $\mathcal{O}(1/\sqrt{k})$ [☆]

Fan Yang [a], Yi Shen [a,*], Zhi Song Liu [b]

[a] Department of Mathematics, Zhejiang Sci-Tech University, Hangzhou 310028, China

[b] School of Mathematics, Physics and Information Science, Zhejiang Ocean University, Zhoushan, 316004, China

ARTICLE INFO

Article history:
Received 6 August 2015
Received in revised form 8 July 2016

Keywords:
Sparse approximation
Alternating minimization
Hard thresholding
Tight wavelet frame
Kurdyka–Łojasiewicz property

ABSTRACT

Since digital images are usually sparse in the wavelet frame domain, some nonconvex minimization models based on wavelet frame have been proposed and sparse approximations have been widely used in image restoration in recent years. Among them, the proximal alternating iterative hard thresholding method is proposed in this paper to solve the nonconvex model based on wavelet frame. Through combining the proposed algorithm with the iterative hard thresholding algorithm which is well studied in compressed sensing theory, this paper proves that the complexity of the proposed method is $\mathcal{O}(1/\sqrt{k})$. On the other hand, a more general nonconvex–nonsmooth model is adopted and the pseudo proximal alternating linearized minimization method is developed to solve the above problem. With the Kurdyka–Łojasiewicz (KL) property, it is proved that the sequence generated by the proposed algorithm converges to some critical points of the corresponding model. Finally, the proposed method is applied to restore the blurred noisy gray images. As the numerical results reveal, the performance of the proposed method is comparable or better than some well-known convex image restoration methods.

1. Introduction

This paper will study how to restore the digital images corrupted by the given blur operators and white Gaussian noise. Mathematically speaking, if the columns of a digital image are stacked one by one, the digital image can be regarded as a column vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^n$. The generic image restoration problem can be formulated as a linear inverse problem accordingly. Namely, an observed degraded digital image $\boldsymbol{b}$ in $\mathbb{R}^m$ is given by

$$\boldsymbol{b} = \mathcal{A}\hat{\boldsymbol{x}} + \boldsymbol{w},$$

where $\mathcal{A} \in \mathbb{R}^{m\times n}$ is a known linear blur operator; $\boldsymbol{w}$ denotes white Gaussian noise with variance σ^2 and $\hat{\boldsymbol{x}}$ is the unknown true image. For the given vector $\boldsymbol{x}$, $\|\boldsymbol{x}\|_0$ denotes the number of nonzero entries. Although $\|\cdot\|_0$ is not a norm, it is still called

☆ This work is supported by the Zhejiang Provincial Natural Science Foundation of China under Grant No. LY15A010020, the Key project of NSF of China under number 11531013, the Science Foundation of Zhejiang Sci-Tech University under Grant 1113834–Y, the Key Laboratory of Oceanographic Big Data Mining & Application of Zhejiang Province under Grants No. OBDMA201505 and the Open Foundation from Marine Sciences in the Most Important Subjects of Zhejiang under Grant No. 111040602136.

* Corresponding author.
E-mail addresses: kiraofsin@yahoo.co.jp (F. Yang), yshen@zstu.edu.cn (Y. Shen), liuzhisong@zjou.edu.cn (Z.S. Liu).

http://dx.doi.org/10.1016/j.cam.2016.07.013
0377-0427/

as l_0-norm for convenience. As usual, $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$ denotes the l_1-norm and $\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$ denotes the l_2-norm. This paper considers recovering the unknown image by solving the following l_0 minimization model:

$$\min_{\boldsymbol{x},\boldsymbol{y}} \Psi_0(\boldsymbol{x},\boldsymbol{y}) := \frac{1}{2}\|\mathcal{A}\boldsymbol{x}-\boldsymbol{b}\|^2 + \lambda\|\boldsymbol{y}\|_0 + \frac{\eta}{2}\|\mathcal{D}^\top\boldsymbol{x}-\boldsymbol{y}\|^2, \tag{1.1}$$

where $\mathcal{D}$ is a given matrix and $\mathcal{D}^\top$ denotes its transpose, while $\mathcal{I} \in \mathbb{R}^{n\times n}$ is assumed to be the identity matrix. This paper mainly regards $\mathcal{D} \in \mathbb{R}^{n\times d}$ as a tight frame, which means $\mathcal{D}\mathcal{D}^\top = \mathcal{I}$. Usually, the number of the columns d is larger than that of the rows n, indicating $\mathcal{D}$ is a redundant system in $\mathbb{R}^n$. Natural images are often sparse with respect to tight wavelet frames. Hence, the regularization term used for models based on wavelet frame can be the l_0-norm or l_1-norm of the wavelet frame coefficients. Other nonconvex image restoration models based on wavelet frame can be found in [1–4] and referred therein. The minimization model (1.1) is solved in [4,5] via the block coordinate descent (BCD) method. Roughly speaking, the first step of the BCD method is deconvoluting in the pixel domain and the second step is denosing in the frame domain. It is proved in [5] that the sequence generated by the BCD method is bounded, which motivates the authors to propose a new method to generate a convergent sequence as below,

$$\begin{cases} \boldsymbol{x}^{k+1} = \operatorname*{argmin}_{\boldsymbol{x}} \frac{1}{2}\|\mathcal{A}\boldsymbol{x}-\boldsymbol{b}\|^2 + \frac{\eta}{2}\|\mathcal{D}^\top\boldsymbol{x}-\boldsymbol{y}^k\|^2, \\ \boldsymbol{y}^{k+1} \in \operatorname*{argmin}_{\boldsymbol{y}} \lambda\|\boldsymbol{y}\|_0 + \frac{\eta}{2}\|\boldsymbol{y}-\mathcal{D}^\top\boldsymbol{x}^{k+1}\|^2 + \frac{d_k}{2}\|\boldsymbol{y}-\boldsymbol{y}^k\|^2, \end{cases} \quad k = 0, 1, 2, \ldots, \tag{1.2}$$

where d_k refer to positive real numbers. The initial value $\boldsymbol{y}^0$ is set 0. If all d_k are set 0, the iterative scheme (1.2) will turn to be the BCD method. The second step of the iterative scheme (1.2) can be connected with the iterative hard thresholding method in [6]. Therefore, the proposed method in this study is called as *Proximal Alternating Iterative Hard Thresholding* (PAIHT) method. To study the convergence properties of the PAIHT method, more general nonconvex–nonsmooth problems of the form will be considered

$$\min_{\boldsymbol{x},\boldsymbol{y}} \Psi(\boldsymbol{x},\boldsymbol{y}) := f(\boldsymbol{x}) + g(\boldsymbol{y}) + H(\boldsymbol{x},\boldsymbol{y}). \tag{1.3}$$

If

$$f(\boldsymbol{x}) = \frac{1}{2}\|\mathcal{A}\boldsymbol{x}-\boldsymbol{b}\|^2, \qquad g(\boldsymbol{y}) = \lambda\|\boldsymbol{y}\|_0 \quad \text{and} \quad H(\boldsymbol{x},\boldsymbol{y}) = \frac{\eta}{2}\|\mathcal{D}^\top\boldsymbol{x}-\boldsymbol{y}\|^2,$$

then it is easy to observe that the minimization model (1.1) is a special case of the minimization model (1.3). If the functions $f(\boldsymbol{x})$ and $g(\boldsymbol{y})$ are both semi-continuous, and the function $H(\boldsymbol{x},\boldsymbol{y})$ is smooth, then the sequence generated by the proximal alternating linearized minimization algorithm [7] is convergent with the *Kurdyka–Łojasiewicz* (KL) property. Motivated by these works on the KL property, and based on the image restoration model (1.1), this paper explores the nonconvex–nonsmooth minimization problem (1.3) under different situations. For instance, the function $f(\boldsymbol{x})$ is convex; the function $g(\boldsymbol{y})$ is semi-continuous and the function $H(\boldsymbol{x},\boldsymbol{y})$ is strongly convex. Subsequently, a new iterative scheme is established for solving the problem (1.3) approximately under the new settings. More specifically, the iteration scheme for solving (1.3) is given by

$$\begin{cases} \boldsymbol{x}^{k+1} = \operatorname*{argmin}_{\boldsymbol{x}} H(\boldsymbol{x},\boldsymbol{y}^k) + f(\boldsymbol{x}), \\ \boldsymbol{y}^{k+1} \in \operatorname*{argmin}_{\boldsymbol{y}} \langle \nabla_{\boldsymbol{y}} H(\boldsymbol{x}^{k+1},\boldsymbol{y}^k), \boldsymbol{y}-\boldsymbol{y}^k\rangle + g(\boldsymbol{y}) + \frac{d_k}{2}\|\boldsymbol{y}-\boldsymbol{y}^k\|^2, \end{cases} \tag{1.4}$$

where d_k refer to positive real numbers. Since the direct convex method is applied to solve the first subproblem of the iterative scheme (1.4), it is called the *Pseudo Proximal Alternating Linearized Minimization* (PPALM) method.

The rest of this paper is organized as follows. In Section 2, some notations and propositions are introduced and some technical results on nonconvex programming are also included. In Section 3.1, the PAIHT method is proved to be sublinearly convergent. In Section 3.2, the PAIHT method is connected with the PPALM method. Then the convergence analysis of the PPALM method is applied to discuss the convergence properties of the PAIHT method. In Section 4, the comparison results of the proposed iterative algorithms with the BCD method in [5] are reported, and then the image restoration model (1.1) is compared with two convex wavelet frame image restoration models. Some conclusions are given in Section 5.

2. Some preliminaries

This section reviews some notations and propositions. The support of a given vector $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ is defined by

$$S(\boldsymbol{x}) := \{i : x_i \neq 0\}.$$

Let $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function, then the domain of σ is defined by

$$\operatorname{dom} \sigma := \{\boldsymbol{x} \in \mathbb{R}^n : \sigma(\boldsymbol{x}) < +\infty\}.$$