

Iterative thresholding algorithm based on non-convex method for modified l_p -norm regularization minimization

Angang Cui^a, Jigen Peng^{b,*}, Haiyang Li^c, Meng Wen^c, Junxiong Jia^a

^a*School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China.*

^b*School of Mathematics and Information Science, Guangzhou University, Guangzhou, 510006, China.*

^c*School of Science, Xi'an Polytechnic University, Xi'an, 710048, China.*

Abstract

Recently, the l_p -norm regularization minimization problem (P_p^λ) has attracted great attention in compressed sensing. However, the l_p -norm $\|x\|_p^p$ in problem (P_p^λ) is nonconvex and non-Lipschitz for all $p \in (0, 1)$, and there are not many optimization theories and methods are proposed to solve this problem. In fact, it is NP-hard for all $p \in (0, 1)$ and $\lambda > 0$. In this paper, we study two modified l_p regularization minimization problems to approximate the NP-hard problem (P_p^λ). Inspired by the good performance of Half algorithm and 2/3 algorithm in some sparse signal recovery problems, two iterative thresholding algorithms are proposed to solve the problems ($P_{p,1/2,\epsilon}^\lambda$) and ($P_{p,2/3,\epsilon}^\lambda$) respectively. Numerical results show that our algorithms perform effectively in finding the sparse signal in some sparse signal recovery problems for some proper $p \in (0, 1)$.

Keywords: Compressed sensing, Modified l_p regularization minimization problem, Thresholding representation theory, $1/2 - \epsilon$ algorithm, $2/3 - \epsilon$ algorithm

2000 MSC: 90C26, 65K10, 49M20

*Corresponding author

Email addresses: cuiangang@163.com (Angang Cui), jgpengxjtu@126.com (Jigen Peng), fplihaiyang@126.com (Haiyang Li), wen5495688@163.com (Meng Wen), jjx323@xjtu.edu.cn (Junxiong Jia)

1. Introduction

During the last decade, the l_p -norm regularization minimization problem has attracted great attention in compressed sensing [1, 2, 3, 4, 5]. In mathematics, it can be modeled into the following minimization problem

$$(P_p^\lambda) \quad \min_{z \in \mathbb{R}^n} \left\{ \|Az - b\|_2^2 + \lambda \|z\|_p^p \right\} \quad (1)$$

for some $\lambda \in (0, +\infty)$ and $p \in (0, 1)$, where $A \in \mathbb{R}^{m \times n}$ is a real matrix of full row rank with $m \ll n$, $b \in \mathbb{R}^m$ is a nonzero real column vector and $\|z\|_p^p = \sum_{i=1}^n |z_i|^p$ for any $z \in \mathbb{R}^n$. We can see that as $p \rightarrow 0$, the problem (P_p^λ) tends to the l_0 -norm regularization minimization problem

$$(P_0^\lambda) \quad \min_{z \in \mathbb{R}^n} \left\{ \|Az - b\|_2^2 + \lambda \|z\|_0 \right\}, \quad (2)$$

where $\|z\|_0$ denotes the number of nonzero components of z , and the iterative hard thresholding algorithm (Hard algorithm)[6] has been proposed for solving the l_0 regularization minimization problem (P_0^λ) . In addition, as $p \rightarrow 1$, the problem (P_p^λ) tends to the l_1 -norm regularization minimization problem

$$(P_1^\lambda) \quad \min_{z \in \mathbb{R}^n} \left\{ \|Az - b\|_2^2 + \lambda \|z\|_1 \right\}, \quad (3)$$

where $\|z\|_1 = \sum_{i=1}^n |z_i|$, and z_i represents the i -th component of vector z . As the compact convex relaxation of the NP-hard problem (P_0^λ) , there are many efficient methods (e.g., see [7, 8, 9, 10, 11]) being proposed for solving the l_1 -norm regularization minimization problem (P_1^λ) . It is obvious that the problem (P_p^λ) is intermediate between the l_0 -norm regularization minimization problem (P_0^λ) and the l_1 -norm regularization minimization problem (P_1^λ) for any $p \in (0, 1)$, because of the relationship

$$\lim_{p \rightarrow 0^+} \sum_{i=1, z_i \neq 0}^n |z_i|^p = \|z\|_0 \quad \text{and} \quad \lim_{p \rightarrow 1} \sum_{i=1}^n |z_i|^p = \|z\|_1. \quad (4)$$

Unfortunately, the problem (P_p^λ) is a nonconvex and non-Lipschitz minimization problem. There are not many optimization theories on analyzing this type of problems and it is NP-hard for all $p \in (0, 1)$ and $\lambda > 0$ (see [1]). At present, the most direct way to solve the problem (P_p^λ) is that the iterative thresholding algorithm only when $p = 1/2, 2/3$ (see [2, 3]). In fact,

the corresponding thresholding functions for the problem (P_p^λ) are in closed form only for $p = 1/2, 2/3$. Xu et al.[2] and Cao et al.[3] have shown that the problem (P_p^λ) could be fast solved by the iterative $l_{1/2}$ thresholding algorithm (Half algorithm) and iterative $l_{2/3}$ thresholding algorithm (2/3 algorithm), and the computational complexity of these two iterative algorithms are all $\mathcal{O}(mn)$. A major drawback of the iterative thresholding algorithm for the problem (P_p^λ) is that the closed form iterative thresholding for the problem (P_p^λ) available only at $p = 1/2, 2/3$.

In this paper, we propose a modified ℓ_p -norm to replace the nonconvex and non-Lipschitz l_p norm $\|x\|_p^p$ given by

$$\sum_{i=1}^n \frac{|z_i|^\theta}{(|z_i| + \epsilon_i)^{\theta-p}} \quad (5)$$

for all $p \in (0, 1)$, where $\theta > 0$ and $\epsilon_i > 0$. With the change of parameter $\epsilon_i > 0$, we have

$$\lim_{\epsilon_i \rightarrow 0^+} \frac{|z_i|^\theta}{(|z_i| + \epsilon_i)^{\theta-p}} = |z_i|^p, \quad (6)$$

and the modified ℓ_p -norm (6) approximates the l_p -norm of vector z :

$$\lim_{\epsilon_i \rightarrow 0^+} \sum_{i=1}^n \frac{|z_i|^\theta}{(|z_i| + \epsilon_i)^{\theta-p}} = \|z\|_p^p. \quad (7)$$

Therefore, the l_p -norm regularization minimization problem transformed by the modified ℓ_p -norm (6) could be written as the following minimization problem

$$(P_{p,\theta,\epsilon}^\lambda) \quad \min_{z \in \mathbb{R}^n} \left\{ \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^\theta}{(|z_i| + \epsilon_i)^{\theta-p}} \right\} \quad (8)$$

for all $p \in (0, 1)$. In particular, we do claim that the problem $(P_{p,\theta,\epsilon}^\lambda)$ matches the following special version

$$(P_{p,1/2,\epsilon}^\lambda) \quad \min_{z \in \mathbb{R}^n} \left\{ \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|z_i| + \epsilon_i)^{1/2-p}} \right\} \quad (9)$$

for $\theta = 1/2$, and

$$(P_{p,2/3,\epsilon}^\lambda) \quad \min_{z \in \mathbb{R}^n} \left\{ \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^{2/3}}{(|z_i| + \epsilon_i)^{2/3-p}} \right\} \quad (10)$$

for $\theta = 2/3$.

Throughout this paper, we just consider the special problems $(P_{p,1/2,\epsilon}^\lambda)$ and $(P_{p,2/3,\epsilon}^\lambda)$ for all $p \in (0, 1)$, and extend the aforementioned well-known Half algorithm and the 2/3 algorithm [2, 3] to solve these two problems. The outline of this paper is as follows. In Section 2, some preliminary results used in this paper are given. In Section 3, we propose two iterative thresholding algorithms to solve the problems $(P_{p,1/2,\epsilon}^\lambda)$ and $(P_{p,2/3,\epsilon}^\lambda)$ respectively. In Section 4, we conduct some numerical experiments to show the performance of our algorithm. Some conclusion remarks are presented in Section 5.

2. Preliminaries

In this section, we give some crucial preliminary results that are used in this paper.

Lemma 1. (see [2]) *For any fixed $\lambda > 0$ and $\beta, r \in \mathbb{R}$, suppose that*

$$h_{1/2,\lambda}(r) \triangleq \arg \min_{\beta \in \mathbb{R}} \left\{ (\beta - r)^2 + \lambda |\beta|^{1/2} \right\}, \quad (11)$$

then the operator $h_{1/2,\lambda}(r)$ can be expressed by

$$h_{1/2,\lambda}(r) = \begin{cases} f_{1/2,\lambda}(r), & \text{if } |r| > \frac{\sqrt[3]{54}}{4} \lambda^{2/3}, \\ 0, & \text{if } |r| \leq \frac{\sqrt[3]{54}}{4} \lambda^{2/3}. \end{cases} \quad (12)$$

where

$$f_{1/2,\lambda}(r) = \frac{2}{3} r \left(1 + \cos \left(\frac{2\pi}{3} - \frac{2}{3} \arccos \left(\frac{\lambda}{8} \left(\frac{|r|}{3} \right)^{-3/2} \right) \right) \right). \quad (13)$$

Lemma 2. (see [3]) *For any fixed $\lambda > 0$ and $\beta, r \in \mathbb{R}$, suppose that*

$$h_{2/3,\lambda}(r) \triangleq \arg \min_{\beta \in \mathbb{R}} \left\{ (\beta - r)^2 + \lambda |\beta|^{2/3} \right\}, \quad (14)$$

then the operator $h_{2/3,\lambda}(r)$ can be expressed by

$$h_{2/3,\lambda}(r) = \begin{cases} f_{2/3,\lambda}(r), & \text{if } |r| > \frac{\sqrt[4]{48}}{3} \lambda^{3/4}, \\ 0, & \text{if } |r| \leq \frac{\sqrt[4]{48}}{3} \lambda^{3/4}. \end{cases} \quad (15)$$

where

$$f_{2/3,\lambda}(r) = \frac{1}{8} \left(|\Phi_{2/3,\lambda}(r)| + \sqrt{\frac{2|r|}{|\Phi_{2/3,\lambda}(r)|} - |\Phi_{2/3,\lambda}(r)|^2} \right)^3 \text{sign}(r) \quad (16)$$

and

$$\Phi_{2/3,\lambda}(r) = \frac{2}{\sqrt{3}} \lambda^{1/4} \left(\cosh \left(\frac{1}{3} \operatorname{arccosh} \left(\frac{27}{16} \lambda^{-3/2} r^2 \right) \right) \right)^{1/2}. \quad (17)$$

Definition 1. ([12]) *The nonincreasing rearrangement of the vector $x \in \mathbb{R}^n$ is the vector $[x] \in \mathbb{R}^n$ for which*

$$[x]_1 \geq [x]_2 \geq \cdots \geq [x]_n \geq 0$$

and there is a permutation $\pi : [n] \rightarrow [n]$ with $[x]_i = |x_{\pi(i)}|$ for all $i \in [n]$.

3. Two iterative thresholding algorithms for solving problem $(P_{p,\theta,\epsilon}^\lambda)$

In this section, we propose two iterative thresholding algorithms, namely, $1/2 - \epsilon$ algorithm ($\theta = 1/2$) and the $2/3 - \epsilon$ algorithm ($\theta = 2/3$), to solve the problems $(P_{p,1/2,\epsilon}^\lambda)$ and $(P_{p,2/3,\epsilon}^\lambda)$ respectively. Moreover, we also provide some convergence analysis for our methods. We should declare that the study of the $1/2 - \epsilon$ algorithm and the $2/3 - \epsilon$ algorithm proposed below are motivated by the well-known Half algorithm and $2/3$ algorithm proposed in Xu et al. [2] and Cao et al. [3].

3.1. The $1/2 - \epsilon$ algorithm for solving the problem $(P_{p,1/2,\epsilon}^\lambda)$

In the subsection, we propose the $1/2 - \epsilon$ algorithm to solve the problem $(P_{p,1/2,\epsilon}^\lambda)$ for all $p \in (0, 1)$. Before the analytic expression of the $1/2 - \epsilon$ algorithm, we should derive the closed form representation of the optimal solution to the problem $(P_{p,1/2,\epsilon}^\lambda)$, which underlies the algorithm to be proposed.

For any $\lambda, \mu \in (0, \infty)$, $p \in (0, 1)$ and $z, y \in \mathbb{R}^n$, let

$$\mathcal{C}_\lambda^{1/2}(z) = \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|z_i| + \epsilon_i)^{1/2-p}}, \quad (18)$$

$$\mathcal{C}_{\lambda,\mu}^{1/2}(z, y) = \mu \|Az - b\|_2^2 + \lambda \mu \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|y_i| + \epsilon_i)^{1/2-p}} - \mu \|Az - Ay\|_2^2 + \|z - y\|_2^2 \quad (19)$$

and

$$B_\mu(z) = z + \mu A^\top (b - Az). \quad (20)$$

Lemma 3. For any $\lambda, \mu \in (0, \infty)$ and $p \in (0, 1)$, if $\tilde{z} = (\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_n)^\top$ is a local minimizer to $\mathcal{C}_{\lambda, \mu}^{1/2}(z, y)$, then

$$\tilde{z}_i = 0 \Leftrightarrow |[B_\mu(y)]_i| \leq t_{1/2, \lambda\mu/(|y_i| + \epsilon_i)^{1/2-p}} \quad (21)$$

and

$$\tilde{z}_i = f_{1/2, \lambda\mu/(|y_i| + \epsilon_i)^{1/2-p}}([B_\mu(y)]_i) \Leftrightarrow |[B_\mu(y)]_i| > t_{1/2, \lambda\mu/(|y_i| + \epsilon_i)^{1/2-p}}, \quad (22)$$

where $[B_\mu(y)]_i$ represents the i -th component of vector $B_\mu(y)$, and $t_{1/2, \lambda\mu/(|y_i| + \epsilon_i)^{1/2-p}}$ and $f_{1/2, \lambda\mu/(|y_i| + \epsilon_i)^{1/2-p}}$ are obtained by replacing λ with $\lambda\mu/(|y_i| + \epsilon_i)^{1/2-p}$ in $t_{1/2, \lambda}$ and $f_{1/2, \lambda}$, respectively.

proof. We notice that, $\mathcal{C}_{\lambda, \mu}^{1/2}(z, y)$ can be rewritten as

$$\begin{aligned} \mathcal{C}_{\lambda, \mu}^{1/2}(z, y) &= \|z - (y - \mu A^T A y + \mu A^T b)\|_2^2 + \lambda\mu \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|y_i| + \epsilon_i)^{1/2-p}} + \mu \|b\|_2^2 \\ &\quad + \|y\|_2^2 - \mu \|A y\|_2^2 - \|y - \mu A^T A y + \mu A^T b\|_2^2 \\ &= \|z - B_\mu(y)\|_2^2 + \lambda\mu \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|y_i| + \epsilon_i)^{1/2-p}} + \mu \|b\|_2^2 + \|y\|_2^2 - \mu \|A y\|_2^2 \\ &\quad - \|B_\mu(y)\|_2^2. \end{aligned}$$

This implies that minimizing $\mathcal{C}_{\lambda, \mu}^{1/2}(z, y)$ for any fixed $\lambda, \mu \in (0, \infty)$ and $y \in \mathbb{R}^n$ is equivalent to

$$\min_{z \in \mathbb{R}^n} \left\{ \|z - B_\mu(y)\|_2^2 + \lambda\mu \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|y_i| + \epsilon_i)^{1/2-p}} \right\}, \quad (23)$$

i.e.,

$$\min_{z \in \mathbb{R}^n} \left\{ \sum_{i=1}^n \left((z_i - [B_\mu(y)]_i)^2 + \lambda\mu \frac{|z_i|^{1/2}}{(|y_i| + \epsilon_i)^{1/2-p}} \right) \right\}. \quad (24)$$

Noting that the summation of equation (25) is separable; hence, solving equation (25) is equivalent to solving the following n subproblem, for $i \in [1, 2, \dots, n]$,

$$\min_{z \in \mathbb{R}^n} \left\{ (z_i - [B_\mu(y)]_i)^2 + \lambda\mu \frac{|z_i|^{1/2}}{(|y_i| + \epsilon_i)^{1/2-p}} \right\}. \quad (25)$$

Therefore, the proof is completed by Lemma 1. \square

Theorem 1. For any $\lambda, \mu \in (0, \infty)$, if $z^* = (z_1^*, z_2^*, \dots, z_n^*)^\top$ is an optimal solution to the problem $(P_{p,\theta,\epsilon}^\lambda)$ and μ satisfies $0 < \mu < \frac{1}{\|A\|_2^2}$, then

$$z_i^* = \begin{cases} f_{1/2,\lambda\mu/(|z_i^*|+\epsilon_i)^{1/2-p}}([B_\mu(z^*)]_i), & \text{if } |[B_\mu(z^*)]_i| > t_{1/2,\lambda\mu/(|z_i^*|+\epsilon_i)^{1/2-p}}; \\ 0, & \text{if } |[B_\mu(z^*)]_i| \leq t_{1/2,\lambda\mu/(|z_i^*|+\epsilon_i)^{1/2-p}}. \end{cases} \quad (26)$$

proof. By condition $0 < \mu < \frac{1}{\|A\|_2^2}$, we can get that

$$\begin{aligned} \mathcal{C}_{\lambda,\mu}^{1/2}(z, z^*) &= \mu \left\{ \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|z_i^*| + \epsilon_i)^{1/2-p}} \right\} \\ &\quad + \left\{ \|z - z^*\|_2^2 - \mu \|Az - Az^*\|_2^2 \right\} \\ &\geq \mu \left\{ \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^{1/2}}{(|z_i^*| + \epsilon_i)^{1/2-p}} \right\} \\ &\geq \mu \mathcal{C}_\lambda^{1/2}(z^*) \\ &= \mathcal{C}_{\lambda,\mu}^{1/2}(z^*, z^*) \end{aligned}$$

for any $z \in \mathbb{R}^n$. This implies that z^* is a local minimizer of $\mathcal{C}_{\lambda,\mu}^{1/2}(z, z^*)$ as long as z^* is an optimal solution to the problem $(P_{p,1/2,\epsilon}^\lambda)$. Combined with Lemma 3, we finish the proof. \square

Next, we present an iterative thresholding algorithm for solving the problem $(P_{p,1/2,\epsilon}^\lambda)$ for all $p \in (0, 1)$ based on the above theoretical analysis.

With the thresholding representation (26), the $1/2 - \epsilon$ algorithm for solving the regularization problem $(P_{p,1/2,\epsilon}^\lambda)$ can be naturally defined as

$$z_i^{k+1} = h_{1/2,\lambda\mu/(|z_i^k|+\epsilon_i)^{1/2-p}}([B_\mu(z^k)]_i), \quad k = 0, 1, 2, \dots, \quad (27)$$

where $B_\mu(z^k) = z^k + \mu A^\top(b - Az^k)$, and $h_{1/2,\lambda\mu/(|z_i^k|+\epsilon_i)^{1/2-p}}$ is obtained by replacing λ with $\lambda\mu/(|z_i^k| + \epsilon_i)^{1/2-p}$ in $h_{1/2,\lambda}$.

In general, the quality of the solution to a regularization problem depends seriously on the setting of the regularization parameter $\lambda > 0$. Suppose that the vector z^* of sparsity r is the optimal solution of the regularization problem $(P_{p,1/2,\epsilon}^\lambda)$. In $1/2 - \epsilon$ algorithm, we set

$$\lambda = \frac{8(\lceil B_\mu(z^k) \rceil_{r+1})^{3/2}(\lceil z^k \rceil_{r+1} + \lceil \epsilon \rceil_{r+1})^{1/2-p}}{\sqrt{54}\mu} \quad (28)$$

in each iteration, where $\lceil \cdot \rceil$ is defined in Definition 1, and $\lceil B_\mu(z^k) \rceil_i$ ($\lceil z^k \rceil_i$, $\lceil \epsilon \rceil_i$) represents the i -th component of vector $\lceil B_\mu(z^k) \rceil$ ($\lceil z^k \rceil$, $\lceil \epsilon \rceil$) for all $i \in [1, 2, \dots, n]$. When doing so, the $1/2 - \epsilon$ algorithm will be adaptive and free from the choice of regularization parameter.

Algorithm 1 : The $1/2 - \epsilon$ algorithm

Initialize: Choose $x^0 \in \mathbb{R}^n$, $\epsilon_i > 0$, $\lambda_0 > 0$, $\mu_0 = \frac{1-\eta}{\|A\|_2^2}(\eta \in (0, 1))$ and $p \in (0, 1)$;
 $k = 0$;
while not converged **do**
 $B_\mu(z^k) = z^k + \mu A^\top (b - Az^k)$;
 $\lambda = \frac{8(\lceil B_\mu(z^k) \rceil_{r+1})^{3/2}(\lceil z^k \rceil_{r+1} + \lceil \epsilon \rceil_{r+1})^{1/2-p}}{\sqrt{54}\mu}$, $\mu = \mu_0$;
 $t = t_{1/2, \lambda\mu / (|z_i^k| + \epsilon_i)^{1/2-p}} = \frac{\sqrt[3]{54}}{4}(\lambda\mu / (|z_i^k| + \epsilon_i)^{1/2-p})^{2/3}$;
for $i = 1 : n$
1. $|\lceil B_\mu(z^k) \rceil_i| > t$, then $z_i^{k+1} = f_{1/2, \lambda\mu / (|z_i^k| + \epsilon_i)^{1/2-p}}(\lceil B_\mu(z^k) \rceil_i)$
2. $|\lceil B_\mu(z^k) \rceil_i| \leq t$, then $z_i^{k+1} = 0$
 $k \rightarrow k + 1$;
end while
return: z^{k+1}

Remark 1. *It is worth emphasizing that the $1/2 - \epsilon$ algorithm reduces to the Half algorithm[2] when we set $p = 1/2$.*

In the following, we provide some convergence analysis for the $1/2 - \epsilon$ algorithm under some specific conditions.

Theorem 2. *Let $\{z^k\}$ be the sequence generated by the $1/2 - \epsilon$ algorithm with the step size μ satisfying $0 < \mu < \frac{1}{\|A\|_2^2}$. Then*

- 1) *The sequence $\{z^k\}$ is a minimization sequence, and the sequence $\{\mathcal{C}_\lambda^{1/2}(z^k)\}$ converging to $\mathcal{C}_\lambda^{1/2}(z^*)$, where z^* is a limit point of minimization sequence $\{z^k\}$;*
- 2) *The sequence $\{z^k\}$ is asymptotically regular, i.e., $\lim_{k \rightarrow \infty} \|z^{k+1} - z^k\|_2^2 = 0$;*
- 3) *Any accumulation point of the sequence $\{z^k\}$ is a stationary point of the problem $(P_{p, 1/2, \epsilon}^\lambda)$.*

proof. The proof of above theorem follows from the fact that the step size μ satisfying $0 < \mu < \frac{1}{\|A\|_2^2}$ and a similar argument as used in the proof of [2, Theorem 3]. \square

3.2. *The $2/3 - \epsilon$ algorithm for solving the problem $(P_{p,2/3,\epsilon}^\lambda)$*

In the subsection, we propose the $2/3 - \epsilon$ algorithm to solve the problem $(P_{p,2/3,\epsilon}^\lambda)$ for all $p \in (0, 1)$.

For any $\lambda, \mu \in (0, \infty)$, $p \in (0, 1)$ and $z, y \in \mathbb{R}^n$, let

$$\mathcal{C}_\lambda^{2/3}(z) = \|Az - b\|_2^2 + \lambda \sum_{i=1}^n \frac{|z_i|^{2/3}}{(|z_i| + \epsilon_i)^{2/3-p}} \quad (29)$$

and

$$\mathcal{C}_{\lambda,\mu}^{2/3}(z, y) = \mu \|Az - b\|_2^2 + \lambda \mu \sum_{i=1}^n \frac{|z_i|^{2/3}}{(|y_i| + \epsilon_i)^{2/3-p}} - \mu \|Az - Ay\|_2^2 + \|z - y\|_2^2. \quad (30)$$

Similar argument as the generation of $1/2 - \epsilon$ algorithm, the $2/3 - \epsilon$ algorithm for solving the problem $(P_{p,2/3,\epsilon}^\lambda)$ can be defined as

$$z_i^{k+1} = h_{2/3,\lambda\mu/(|z_i^k| + \epsilon_i)^{2/3-p}}([B_\mu(z^k)]_i), \quad k = 0, 1, 2, \dots, \quad (31)$$

where $B_\mu(z^k) = z^k + \mu A^\top(b - Az^k)$, and $h_{2/3,\lambda\mu/(|z_i^k| + \epsilon_i)^{2/3-p}}$ is obtained by replacing λ with $\lambda\mu/(|z_i^k| + \epsilon_i)^{2/3-p}$ in $h_{2/3,\lambda}$.

In $2/3 - \epsilon$ algorithm, we set the regularization parameter λ as

$$\lambda = \frac{\sqrt[3]{4^4}([\mathcal{B}_\mu(z^k)]_{r+1})^{4/3}([\mathcal{Z}^k]_{r+1} + [\epsilon]_{r+1})^{2/3-p}}{\sqrt[9]{48^4}\mu} \quad (32)$$

in each iteration.

Similar argument as the Theorem 2, the sequence $\{X^k\}$ generated by the $2/3 - \epsilon$ algorithm is a minimization sequence and asymptotically regular. Moreover, any accumulation point of $\{X^k\}$ is a stationary point of the problem $(P_{p,2/3,\epsilon}^\lambda)$. Its proof also follows from the fact that the step size μ satisfying $0 < \mu < \frac{1}{\|A\|_2^2}$ and a similar argument as used in the proof of [2, Theorem 3].

Remark 2. *If we set $p = 2/3$, the $2/3 - \epsilon$ algorithm reduces to the $2/3$ algorithm [3].*

Algorithm 2 : The $2/3 - \epsilon$ algorithm

Initialize: Choose $x^0 \in \mathbb{R}^n$, $\epsilon_i > 0$, $\mu_0 = \frac{1-\eta}{\|A\|_2^2}(\eta \in (0, 1))$ and $p \in (0, 1)$;
 $k = 0$;

while not converged **do**

$$B_\mu(z^k) = z^k + \mu A^\top (b - Az^k);$$

$$\lambda = \frac{\sqrt[3]{4^4}([B_\mu(z^k)]_{r+1})^{4/3}([z^k]_{r+1} + [\epsilon]_{r+1})^{2/3-p}}{\sqrt[9]{48^4 \mu}}, \mu = \mu_0;$$

$$t = t_{2/3, \lambda \mu / (|z_i^k| + \epsilon_i)^{2/3-p}} = \frac{\sqrt[4]{48}}{3} (\lambda \mu / (|z_i^k| + \epsilon_i)^{2/3-p})^{3/4};$$

for $i = 1 : n$

1. $|[B_\mu(z^k)]_i| > t$, then $z_i^{k+1} = f_{2/3, \lambda \mu / (|z_i^k| + \epsilon_i)^{2/3-p}}([B_\mu(z^k)]_i)$

2. $|[B_\mu(z^k)]_i| \leq t$, then $z_i^{k+1} = 0$

$k \rightarrow k + 1$;

end while

return: z^{k+1}

4. Numerical experiments

In this section, we first present numerical results of the $1/2 - \epsilon$ algorithm and $2/3 - \epsilon$ algorithm for some sparse signal recovery problems and then compare them with some state-of-art methods including iterative hard thresholding algorithm (Hard algorithm)[6], iterative soft thresholding algorithm (Soft algorithm)[7], Half algorithm [2] and $2/3$ algorithm [3] in some sparse recovery problems. We generate a measurement matrix $A \in \mathbb{R}^{m \times n}$ with entries independently drawn by random from a Gaussian distribution, $\mathcal{N}(0, 1)$. To show the success rate of these algorithms in the recovery of the sparse signals with the different sparsity for a given measurement matrix $A \in \mathbb{R}^{m \times n}$, we randomly generate sparse vectors $z_0 \in \mathbb{R}^n$ and generate vectors b by $b = Az_0$. Therefore, we know the sparsest solution to the linear system $b = Az_0$. In our experiments, the stopping criterion is defined as

$$\frac{\|z^{k+1} - z^k\|_2}{\|z^k\|_2} \leq \text{Tol}$$

where z^{k+1} and z^k are numerical results from two continuous iterative steps and Tol is a small given number (we set Tol = 10^{-8}). The success is measured by computing the relative error (RE):

$$\text{RE} = \frac{\|z^* - z_0\|_2}{\|z_0\|_2}$$

to indicate a perfect recovery of the original sparse vector z_0 , and the success is declared when $\text{RE} \leq 10^{-4}$. Moreover, we adapt ϵ at each iteration as a function of the current guess z^k , and set

$$\epsilon_i = \max\{\gamma|[\mu A^\top(b - Az^k)]_i|, 10^{-3}\}, \quad \gamma = 0.7.$$

In all of our experiments, we repeat 20 tests and present average results. The experiments are all performed on a Lenovo-PC with an Intel(R) Core(TM) i7-6700 CPU @3.40GHZ with 16GB of RAM running Microsoft Windows 7.

4.1. Performance of $1/2 - \epsilon$ algorithm and $2/3 - \epsilon$ algorithm

In this subsection, we carry out a series of experiments to demonstrate the performance of the $1/2 - \epsilon$ algorithm and the $2/3 - \epsilon$ algorithm for some sparse signal recovery problems. In our experiments, we set $m = 128$, $n = 512$.

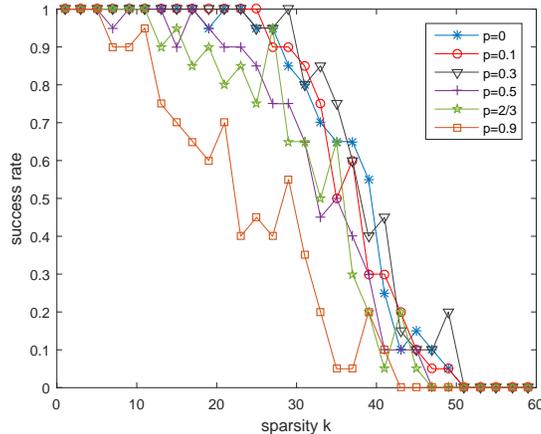


Figure 1: The comparison of the success rate for the $1/2 - \epsilon$ algorithm in the recovery of a sparse signal with different parameter values p .

The graphs presented in Figure 1 and Figure 2 show the success rate of $1/2 - \epsilon$ algorithm and $2/3 - \epsilon$ algorithm in recovering the true (sparsest) solution. From Figures 1 and 2, we can see that $1/2 - \epsilon$ algorithm can exactly recover the ideal signal until k is around 25 when $p = 0.1$, and $2/3 - \epsilon$ algorithm's counterpart is around 17 when $p = 0$. As we can see, the parameter $p = 0.1$ is the best strategy for the $1/2 - \epsilon$ algorithm, and the the parameter $p = 0$ is the best strategy for the $2/3 - \epsilon$ algorithm.

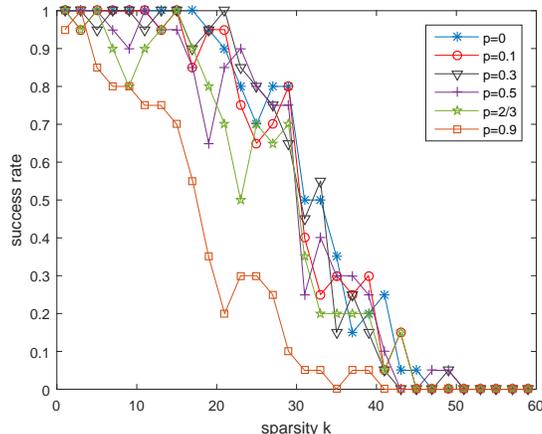


Figure 2: The comparison of the success rate for the $2/3 - \epsilon$ algorithm in the recovery of a sparse signal with different parameter values p .

4.2. Compared with some state-of-art methods

In this subsection, we compare our algorithms (the $1/2 - \epsilon$ algorithm and $2/3 - \epsilon$ algorithm) with some state-of-art methods including Hard algorithm[6], Soft algorithm[7], Half algorithm[2] and $2/3$ algorithm[3] in some sparse recovery problems. In our experiments, we set $p = 0.001$ in the $1/2 - \epsilon$ algorithm and the $2/3 - \epsilon$ algorithm, and set $m = 256$, $n = 1024$ to size the dimension of the matrix $A \in \mathbb{R}^{m \times n}$ and the length of the vector $z_0 \in \mathbb{R}^n$. Two different cases in sparse recovery problems will be considered: exactly sparse signals recovery in the noiseless case and exactly sparse signals recovery in the noise case. In noiseless case, we generate vectors b by $b = Az_0$, where $A \in \mathbb{R}^{256 \times 1024}$ is a measurement matrix with entries independently drawn by random from a Gaussian distribution, $\mathcal{N}(0, 1)$, and $z_0 \in \mathbb{R}^{1024}$ is a randomly sparse vector $z_0 \in \mathbb{R}^{1024}$. Turn to the noise case, we use the same matrix A , and generate a random vector z_0 with a prespecified cardinality of nonzeros. We compute $b = Az_0 + e$, where $e \in \mathcal{N}(0, \sigma)$ ($\sigma = 10^{-5}$). Thus, the original vector z_0 is a feasible solution and close to the optimal solution. Due to the presence of noise, it becomes harder to accurately recover the original signal z_0 . The graphs presented in Figure 3 and Figure 4 show the success rate of $1/2 - \epsilon$ algorithm ($p = 0.1$), $2/3 - \epsilon$ algorithm ($p = 0$), Half algorithm, $2/3$ algorithm, Soft algorithm and Hard algorithm in recovering the true (sparsest) solution. From Figure 3, we can see that the $1/2 - \epsilon$ algorithm ($p = 0.1$) can exactly recover the ideal signal until k is around 78, and the $2/3 - \epsilon$

algorithm ($p = 0$) is around 70. The results in noise case are consistent with the noiseless case. We can see that the $1/2 - \epsilon$ algorithm ($p = 0.1$) again has the best performance in recovering the sparse signals in the six algorithms with noise or not, and the $2/3 - \epsilon$ algorithm ($p = 0$) performs the second best.

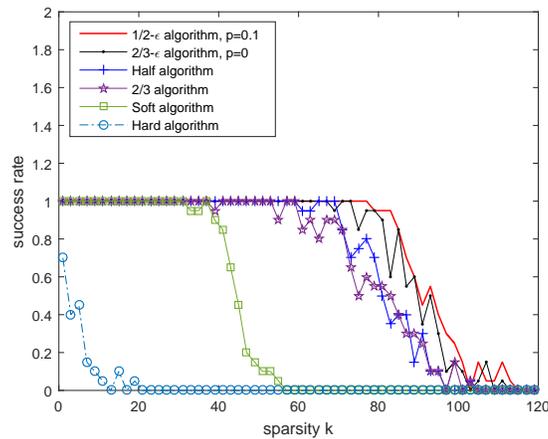


Figure 3: The success rate of six algorithms in the recovery of a sparse signal with different sparsity (noiseless case).

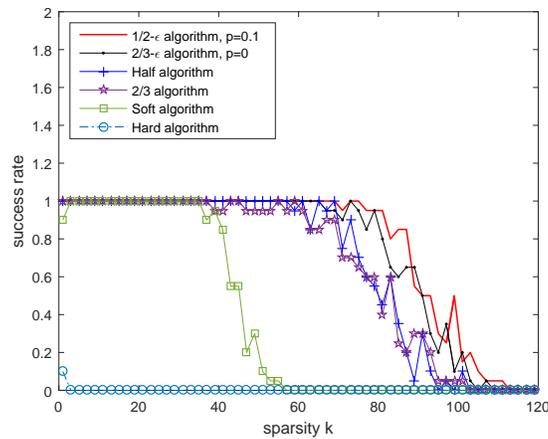


Figure 4: The success rate of six algorithms in the recovery of a sparse signal with different sparsity (noise case).

5. Conclusions

In this paper, we studied two modified l_p -norm regularization minimization problems $(P_{p,1/2,\epsilon}^\lambda)$ and $(P_{p,2/3,\epsilon}^\lambda)$ to approximate the NP-hard problem (P_p^λ) for all $p \in (0, 1)$. Inspired by the good performances of iterative $l_{1/2}$ thresholding algorithm and iterative $l_{2/3}$ algorithm in some sparse signal recovery problems, the $1/2 - \epsilon$ algorithm and $2/3 - \epsilon$ algorithm are generated to solve the problems $(P_{p,1/2,\epsilon}^\lambda)$ and $(P_{p,2/3,\epsilon}^\lambda)$ for all $p \in (0, 1)$. Numerical results show that our algorithms perform effectively in finding the sparse signals in some sparse signal recovery problems for some proper $p \in (0, 1)$. Moreover, the numerical results also show that the $1/2 - \epsilon$ algorithm performs the best in some sparse signal recovery problems compared with some state-of-art methods, and the $2/3 - \epsilon$ algorithm performs the second best for some proper p .

Acknowledgments

We would like to thank editors and reviewers for their comments which help us to enrich the content and improve the presentation of the results in this paper. The work was supported by the National Natural Science Foundations of China (11771347, 91730306, 41390454, 11271297) and the Science Foundations of Shaanxi Province of China (2016JQ1029, 2015JM1012).

References

References

- [1] X. Chen, D. Ge, Z. Wang, and Y. Ye, Complexity of unconstrained $L_2 - L_p$ minimization. *Mathematical Programming*, 143 (2014) 371–383.
- [2] Z. Xu, X. Chang, F. Xu, and H. Zhang. $L_{1/2}$ Regularization: A thresholding representation theory and a fast solver. *IEEE Transactions on Neural Networks and Learning Systems*, 24(7) (2012) 1013–1027.
- [3] W. Cao, J. Sun, Z. Xu, Fast image deconvolution using closed-form thresholding formulas of $L_q(q = \frac{1}{2}, \frac{2}{3})$ regularization. *Journal of Visual Communication and Image Representation*, 24(1) (2013) 31–41.
- [4] X. Chen, F. Xu, Y. Ye, Lower bound theory of nonzero entries in solutions of $l_2 - l_p$ minimization. *SIAM Journal on Scientific Computing*, 32(5) (2010) 2832–2852.

- [5] X. Chen, L. Niu, Y. Yuan, Optimality conditions and a smoothing trust region newton method for nonlipschitz optimization. *SIAM Journal on Optimization*, 23(3) (2013) 1528–1552.
- [6] T. Blumensath, M. E. Davies, Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14(5-6) (2008) 629–654.
- [7] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57 (11) (2004) 1413–1457.
- [8] D. L. Donoho, Denoising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3) (1995) 613–627.
- [9] T. Goldstein, S. Osher, The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2) (2009) 323–343.
- [10] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1) (2008) 143–168.
- [11] J. Yang, Y. Zhang, Alternating direction algorithms for ℓ_1 problems in compressive sensing. *SIAM Journal on Scientific Computing*, 33(1) (2011) 250–278.
- [12] S. Foucart, H. Rauhut, A mathematical introduction to compressive sensing. Springer, New York (2010).