RESEARCH

Open Access

CrossMark

Sparse signals recovered by non-convex penalty in quasi-linear systems

Angang Cui[1], Haiyang Li[2], Meng Wen[2] and Jigen Peng[1*]

*Correspondence:
jgpengxjtu@126.com
[1] School of Mathematics and
Statistics, Xi'an Jiaotong University,
Xi'an, China
Full list of author information is
available at the end of the article

Abstract

The goal of compressed sensing is to reconstruct a sparse signal under a few linear measurements far less than the dimension of the ambient space of the signal. However, many real-life applications in physics and biomedical sciences carry some strongly nonlinear structures, and the linear model is no longer suitable. Compared with the compressed sensing under the linear circumstance, this nonlinear compressed sensing is much more difficult, in fact also NP-hard, combinatorial problem, because of the discrete and discontinuous nature of the ℓ_0-norm and the nonlinearity. In order to get a convenience for sparse signal recovery, we set the nonlinear models have a smooth quasi-linear nature in this paper, and study a non-convex fraction function ρ_a in this quasi-linear compressed sensing. We propose an iterative fraction thresholding algorithm to solve the regularization problem (QP_a^λ) for all $a > 0$. With the change of parameter $a > 0$, our algorithm could get a promising result, which is one of the advantages for our algorithm compared with some state-of-art algorithms. Numerical experiments show that our method performs much better than some state-of-the-art methods.

MSC: 34A34; 78M50; 93C10

Keywords: Compressed sensing; Quasi-linear; Non-convex fraction function; Iterative thresholding algorithm

1 Introduction

In compressed sensing (see, e.g., [1, 2]), the problem of reconstructing a sparse signal under a few linear measurements which are far fewer than the dimension of the ambient space of the signal can be modeled into the following ℓ_0-minimization:

$$(P_0) \quad \min_{x \in \mathbb{R}^n} \|x\|_0 \quad \text{subject to } Ax = b, \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$ is an $m \times n$ real matrix of full row rank with $m < n$, and $b \in \mathbb{R}^m$ is a nonzero real vector of m-dimension, and $\|x\|_0$ is the ℓ_0-norm of the real vector x, which counts the number of the nonzero entries in x (see, e.g., [3–5]). In general, the problem (P_0) is computational and NP-hard because of the discrete and discontinuous nature of the ℓ_0-norm. However, many real-life applications in physics and biomedical sciences carry some strongly nonlinear structures [6], so that the linear model in problem (P_0) is no longer suitable. In this nonlinear case, we consider a map $A : \mathbb{R}^n \to \mathbb{R}^m$, which is no longer necessarily

linear, and reconstruct a sparse vector $x \in \mathbb{R}^n$ from the measurements $b \in \mathbb{R}^m$ given by

$$A(x) = b. \tag{2}$$

Compared with the ℓ_0-minimization under the linear circumstance, this nonlinear minimization is much more difficult, in fact also NP-hard, combinatorial problem, because of the discrete and discontinuous nature of the ℓ_0-norm and the nonlinearity. In order to get a convenience for sparse signal recovery, in this paper, we set the nonlinear models have a smooth quasi-linear nature. By this means, there exists a Lipschitz map

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n} \tag{3}$$

such that

$$A(x) = F(x)x \tag{4}$$

for all $x \in \mathbb{R}^n$. So, the ℓ_0-minimization under the quasi-linear case can be mathematically viewed as the following form:

$$(QP_0) \quad \min_{x \in \mathbb{R}^n} \|x\|_0 \quad \text{subject to } F(x)x = b. \tag{5}$$

In fact, the minimization (QP_0) under the quasi-linear case is also combinatorial and NP-hard [6, 7]. To overcome this problem, the authors in [6, 7] proposed the ℓ_1-minimization

$$(QP_1) \quad \min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to } F(x)x = b \tag{6}$$

for the constrained problem and

$$(QP_1^\lambda) \quad \min_{x \in \mathbb{R}^n} \left\{ \|F(x)x - b\|_2^2 + \lambda \|x\|_1 \right\} \tag{7}$$

for the regularization problem, where $\|x\|_1 = \sum_{i=1}^n |x_i|$ is the ℓ_1-norm of vector x.
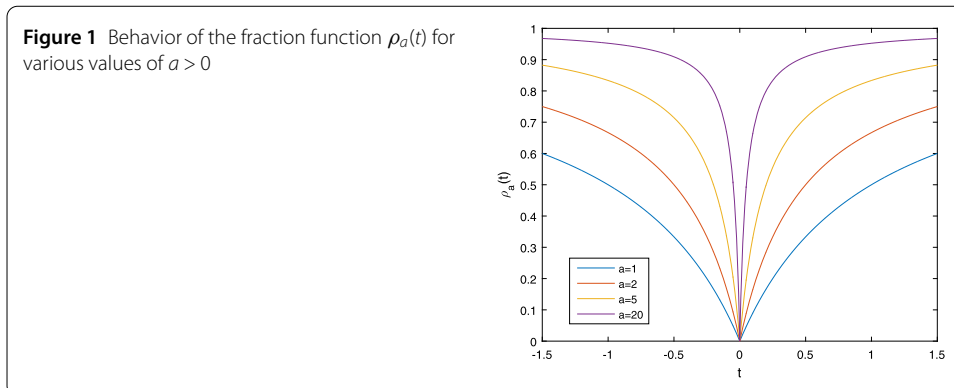
In [6, 7], the authors have shown that the ℓ_1-norm minimization can really make an exact recovery in some specific conditions. In general, however, these conditions are always hard to satisfied in practice. Moreover, the regularization problem (QP_1^λ) always leads to a biased estimation by shrinking all the components of the vector toward zero simultaneously, and sometimes results in over-penalization in the regularization model (QP_1^λ) as the ℓ_1-norm in linear compressed sensing.

In pursuit of better reconstruction results, in this paper, we propose the following fraction minimization:

$$(QP_a^\lambda) \quad \min_{x \in \mathbb{R}^n} \left\{ \|F(x)x - b\|_2^2 + \lambda P_a(x) \right\}, \tag{8}$$

where

$$P_a(x) = \sum_{i=1}^n \rho_a(x_i), \quad a > 0 \tag{9}$$

Figure 1 Behavior of the fraction function $\rho_a(t)$ for various values of $a > 0$

and

$$\rho_a(t) = \frac{a|t|}{a|t| + 1} \tag{10}$$

is the fraction function which performs outstanding in image restoration [8], linear compressed sensing [9] and matrix rank minimization problem [10]. Clearly, with the change of parameter $a > 0$, the non-convex function $P_a(x)$ could approximately interpolate the ℓ_0-norm

$$\lim_{a \to +\infty} \rho_a(x_i) = \begin{cases} 0 & \text{if } x_i = 0; \\ 1 & \text{if } x_i \neq 0. \end{cases} \tag{11}$$

Figure 1 shows the behavior of the fraction function $\rho_a(t)$ for various values of $a > 0$.

The rest of this paper is organized as follows. Some preliminary results that are used in this paper are given in Sect. 2. In Sect. 3, we propose an iterative fraction thresholding algorithm to solve the regularization problem (QP_a^λ) for all $a > 0$. In Sect. 3, we present some numerical experiments to demonstrate the effectiveness of our algorithm. The concluding remarks are presented in Sect. 4.

2 Preliminaries

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

Define a function of $\beta \in \mathbb{R}$ as

$$f_\lambda(\beta) = (\beta - \gamma)^2 + \lambda \rho_a(\beta) \tag{12}$$

and let

$$\text{prox}_{a,\lambda}^\beta(\gamma) \triangleq \arg \min_{\beta \in \mathbb{R}} f_\lambda(\beta). \tag{13}$$

Lemma 1 (see [9–11]) *The operator* $\text{prox}_{a,\lambda}^\beta$ *defined in* (13) *can be expressed as*

$$\text{prox}_{a,\lambda}^\beta(\gamma) = \begin{cases} g_{a,\lambda}(\gamma) & \text{if } |\gamma| > t_{a,\lambda}^*; \\ 0 & \text{if } |\gamma| \leq t_{a,\lambda}^*, \end{cases} \tag{14}$$

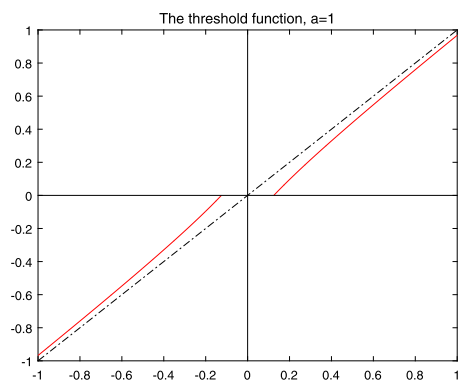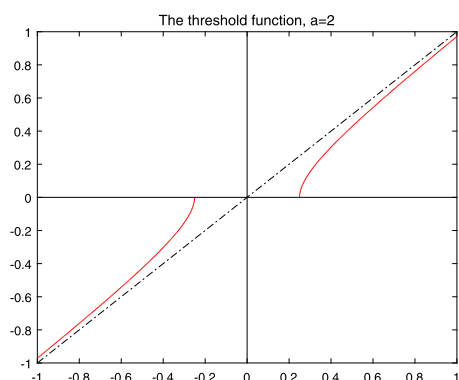**Figure 2** Plot of the threshold function $g_{a,\lambda}$ for $a = 1$ and $\lambda = 0.25$

The threshold function, a=1

Figure 3 Plot of the threshold function $g_{a,\lambda}$ for $a = 2$ and $\lambda = 0.25$

The threshold function, a=2

where $g_{a,\lambda}(\gamma)$ is defined as

$$g_{a,\lambda}(\gamma) = \mathrm{sign}(\gamma)\left(\frac{\frac{1+a|\gamma|}{3}(1 + 2\cos(\frac{\phi(\gamma)}{3} - \frac{\pi}{3})) - 1}{a}\right), \tag{15}$$

$$\phi(\gamma) = \arccos\left(\frac{27\lambda a^2}{4(1 + a|\gamma|)^3} - 1\right)$$

and the threshold value satisfies

$$t_{a,\lambda}^* = \begin{cases} t_{a,\lambda}^1 & \textit{if } \lambda \leq \frac{1}{a^2}; \\ t_{a,\lambda}^2 & \textit{if } \lambda > \frac{1}{a^2}, \end{cases} \tag{16}$$

where

$$t_{a,\lambda}^1 = \frac{\lambda}{2}a, \qquad t_{a,\lambda}^2 = \sqrt{\lambda} - \frac{1}{2a}. \tag{17}$$

Figures 2, 3, 4, and 5 show the plots of the threshold function $g_{a,\lambda}$ for $a = 1, 2, 3, 5$, and $\lambda = 0.25$.

Figures 6 and 7 show the plots of the hard/soft threshold functions with $\lambda = 0.25$.

Figure 4 Plot of the threshold function $g_{a,\lambda}$ for $a = 3$ and $\lambda = 0.25$
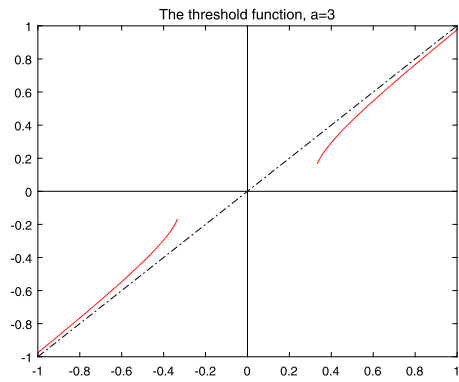

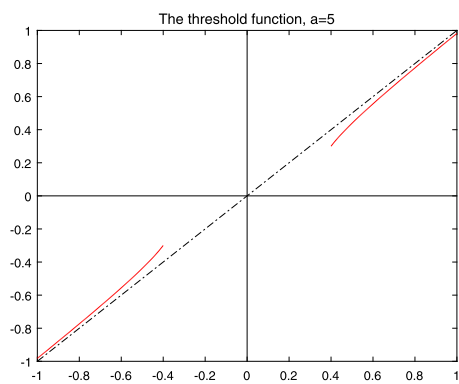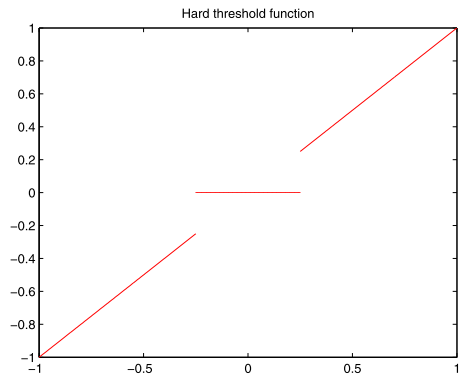**Figure 5** Plot of the threshold function $g_{a,\lambda}$ for $a = 5$, and $\lambda = 0.25$

Figure 6 Plot of the hard threshold function with $\lambda = 0.25$

**Definition 1** The iterative thresholding operator $G_{a,\lambda}$ can be defined by

$$G_{a,\lambda}(x) = \left(\text{prox}_{a,\lambda}^{\beta}(x_1), \dots, \text{prox}_{a,\lambda}^{\beta}(x_n) \right)^{\top}, \tag{18}$$

where $\text{prox}_{a,\lambda}^{\beta}$ is defined in Lemma 1.

3 Thresholding representation theory and algorithm for problem (QP_a^{λ})

In this section, we establish a thresholding representation theory of the problem (QP_a^{λ}), which underlies the algorithm to be proposed. Then an iterative fraction thresholding algorithm (IFTA) is proposed to solve the problem (QP_a^{λ}) for all $a > 0$.

Figure 7 Plot of the soft threshold function with $\lambda = 0.25$
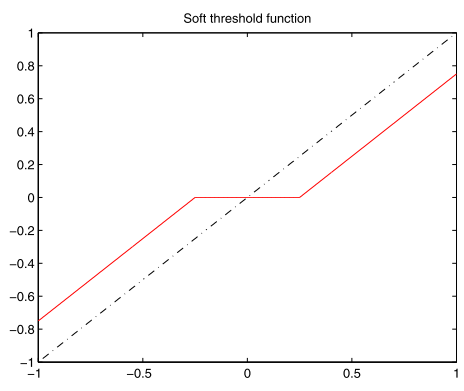
3.1 Thresholding representation theory

For any fixed positive parameters $\lambda > 0$, $\mu > 0$, $a > 0$ and $x \in \mathbb{R}^n$, let

$$C_1(x) = \left\| F(x)x - b \right\|_2^2 + \lambda P_a(x) \tag{19}$$

and

$$C_2(x, y) = \mu \left\| F(y)x - b \right\|_2^2 + \lambda \mu P_a(x) - \mu \left\| F(y)x - F(y)y \right\|_2^2 + \|x - y\|_2^2. \tag{20}$$

It is clear that $C_2(x, x) = \mu C_1(x)$ for all $\mu > 0$.

Theorem 1 *For any $\lambda > 0$ and $0 < \mu < L_*^{-1}$ with $\|F(x^*)x - F(x^*)x^*\|_2^2 \leq L_* \|x - x^*\|_2^2$. If x^* is the optimal solution of $\min_{x \in \mathbb{R}^n} C_1(x)$, then x^* is also the optimal solution of $\min_{x \in \mathbb{R}^n} C_2(x, x^*)$, that is,*

$$C_2\left(x^*, x^*\right) \leq C_2\left(x, x^*\right)$$

for any $x \in \mathbb{R}^n$.

Proof By the definition of $C_2(x, y)$, we have

$$
\begin{aligned}
C_2\left(x, x^*\right) &= \mu \left\| F\left(x^*\right)x - b \right\|_2^2 + \lambda \mu P_a(x) - \mu \left\| F\left(x^*\right)x - F\left(x^*\right)x^* \right\|_2^2 + \left\| x - x^* \right\|_2^2 \\
&\geq \mu \left\| F\left(x^*\right)x - b \right\|_2^2 + \lambda \mu P_a(x) \\
&\geq \mu C_1\left(x^*\right) \\
&= C_2\left(x^*, x^*\right).
\end{aligned}
$$
$\square$

Theorem 2 *For any $\lambda > 0$, $\mu > 0$ and solution x^* of $\min_{x \in \mathbb{R}^n} C_1(x)$, $\min_{x \in \mathbb{R}^n} C_2(x, x^*)$ is equivalent to*

$$\min_{x \in \mathbb{R}^n} \left\{ \left\| x - B_\mu\left(x^*\right) \right\|_2^2 + \lambda \mu P_a(x) \right\}, \tag{21}$$

where $B_\mu(x^*) = x^* + \mu F(x^*)^\top (b - F(x^*)x^*)$.

Proof By the definition, $C_2(x, y)$ can be rewritten as

$$C_2\left(x, x^*\right) = \left\| x - \left(x^* - \mu F\left(x^*\right)^\top F\left(x^*\right)x^* + \mu F\left(x^*\right)^\top b\right) \right\|_2^2 + \lambda\mu P_a(x) + \mu\|b\|_2^2 + \left\|x^*\right\|_2^2$$
$$- \mu\left\|F\left(x^*\right)x^*\right\|_2^2 - \left\|x^* - \mu F\left(x^*\right)^\top F\left(x^*\right)x^* + \mu F\left(x^*\right)^\top b\right\|_2^2$$
$$= \left\|x - B_\mu\left(x^*\right)\right\|_2^2 + \lambda\mu P_a(x) + \mu\|b\|_2^2 + \left\|x^*\right\|_2^2 - \mu\left\|F\left(x^*\right)x^*\right\|_2^2 - \left\|B_\mu\left(x^*\right)\right\|_2^2,$$

which implies that $\min_{x\in\mathbb{R}^n} C_2(x, x^*)$ for any $\lambda > 0$, $\mu > 0$ is equivalent to

$$\min_{x\in\mathbb{R}^n}\left\{ \left\|x - B_\mu\left(x^*\right)\right\|_2^2 + \lambda\mu P_a(x)\right\}. \qquad \square$$

Combining Theorem 2, Theorem 1 and Lemma 1, the thresholding representation of (QP_a^λ) can be concluded by

$$x^* = G_{a,\lambda\mu}\left(B_\mu\left(x^*\right)\right), \tag{22}$$

where the operator $G_{a,\lambda\mu}$ is defined in Definition 1 and obtained by replacing λ with $\lambda\mu$. With the thresholding representations (22), the IFTA for solving the regularization problem (QP_a^λ) can be naturally defined as

$$x^{k+1} = G_{a,\lambda\mu}\left(B_\mu\left(x^k\right)\right), \quad k = 0, 1, 2, \dots, \tag{23}$$

where $B_\mu(x^k) = x^k + \mu F(x^k)^\top (b - F(x^k)x^k)$.

3.2 Adjusting the values for the regularization parameter $\lambda > 0$

In this subsection, the cross-validation method (see [9, 10, 12]) is accepted to automatically adjust the value for the regularization parameter $\lambda > 0$. In other words, when some prior information is known for a regularization problem, this selection is more reasonable and intelligent. Suppose that the vector x^* of sparsity r is the optimal solution of the regularization problem (QP_a^λ), and without loss of generality, set

$$\left|B_\mu\left(x^*\right)\right|_1 \geq \left|B_\mu\left(x^*\right)\right|_2 \geq \cdots \geq \left|(B_\mu\left(x^*\right)\right|_r \geq \left|(B_\mu\left(x^*\right)\right|_{r+1} \geq \cdots \geq \left|(B_\mu\left(x^*\right)\right|_n \geq 0.$$

Then it follows from (14) that

$$\left|B_\mu\left(x^*\right)\right|_i > t_{a,\lambda\mu}^* \quad \Leftrightarrow \quad i \in \{1, 2, \dots, r\},$$
$$\left|B_\mu\left(x^*\right)\right|_i \leq t_{a,\lambda\mu}^* \quad \Leftrightarrow \quad i \in \{r+1, r+2, \dots, n\},$$

where $t_{a,\lambda\mu}^*$ is obtained by replacing λ with $\lambda\mu$ in $t_{a,\lambda}^*$.

By $t_{a,\lambda\mu}^2 \leq t_{a,\lambda\mu}^1$, we have

$$\begin{cases} |B_\mu(x^*)|_r \geq t_{a,\lambda\mu}^* \geq t_{a,\lambda\mu}^2 = \sqrt{\lambda\mu} - \frac{1}{2a}; \\ |B_\mu(x^*)|_{r+1} < t_{a,\lambda\mu}^* \leq t_{a,\lambda\mu}^1 = \frac{\lambda\mu}{2}a. \end{cases} \tag{24}$$

It follows that

$$\frac{2|B_\mu(x^*)|_{r+1}}{a\mu} \leq \lambda \leq \frac{(2a|B_\mu(x^*)|_r + 1)^2}{4a^2\mu}. \tag{25}$$

From (25), we obtain

$$\lambda \in \left[\frac{2|B_\mu(x^*)|_{r+1}}{a\mu}, \frac{(2a|B_\mu(x^*)|_r + 1)^2}{4a^2\mu} \right].$$

We denote by λ_1 and λ_2 the left and the right of the above interval, respectively:

$$\lambda_1 = \frac{2|B_\mu(x^*)|_{r+1}}{a\mu} \quad \text{and} \quad \lambda_2 = \frac{(2a|B_\mu(x^*)|_r + 1)^2}{4a^2\mu}.$$

A choice of λ is

$$\lambda = \begin{cases} \lambda_1 & \text{if } \lambda_1 \leq \frac{1}{a^2\mu}; \\ \lambda_2 & \text{if } \lambda_1 > \frac{1}{a^2\mu}. \end{cases}$$

Since x^* is unknown, and x^k is the best available approximation to x^*, so we can take

$$\lambda = \begin{cases} \lambda_{1,k} = \frac{2|B_\mu(x^k)|_{r+1}}{a\mu} & \text{if } \lambda_{1,k} \leq \frac{1}{a^2\mu}; \\ \lambda_{2,k} = \frac{(2a|B_\mu(x^k)|_r + 1)^2}{4a^2\mu} & \text{if } \lambda_{1,k} > \frac{1}{a^2\mu}, \end{cases} \tag{26}$$

in the kth iteration. That is, (26) can be used to automatically adjust the value of the regularization parameter $\lambda > 0$ during iteration.

Remark 1 Notice that (26) is valid for any $\mu > 0$ satisfying $0 < \mu \leq \|F(x_k)\|_2^{-2}$. In general, we can take $\mu = \mu_k = \frac{1-\epsilon}{\|F(x_k)\|_2^2}$ with any small $\epsilon \in (0, 1)$ below. Especially, the threshold value is $t^*_{a,\lambda\mu} = \frac{\lambda\mu}{2}a$ when $\lambda = \lambda_{1,k}$, and $t^*_{a,\lambda\mu} = \sqrt{\lambda\mu} - \frac{1}{2a}$ when $\lambda = \lambda_{2,k}$.

3.3 Iterative fraction thresholding algorithm (IFTA)
Based on the thresholding representation (23) and the analyses given in Sect. 3.2, the proposed iterative fraction thresholding algorithm (IFTA) for regularization problem (QP_a^λ) can be naturally described in Algorithm 1.

Remark 2 The convergence of IFTA is not proved theoretically in this paper, and this is our future work.

4 Numerical experiments
In the section, we carry out a series of simulations to demonstrate the performance of IFTA, and compare them with those obtained with some state-of-art methods (iterative soft thresholding algorithm (ISTA) [6, 7]), iterative hard thresholding algorithm (IHTA) [6, 7]. In our numerical experiments, we set

$$F(x) = A_1 + \eta f(\|x - x_0\|_2)A_2, \tag{27}$$

where $A_1 \in \mathbb{R}^{100 \times 400}$ is a fixed Gaussian random matrix, $x_0 \in \mathbb{R}^{400}$ is a reference vector, $f : [0, \infty) \to \mathbb{R}$ is a positive and smooth Lipschitz continuous function with $f(t) = \ln(t + 1)$, η is a sufficiently small scaling factor (we set $\eta = 0.003$), and $A_2 \in \mathbb{R}^{30 \times 100}$ is a fixed matrix with every entry equals 1. Then the form of nonlinearity considered in (27) is a quasi-

---

Algorithm 1 Iterative fraction thresholding algorithm (IFTA)

Initialize: Given $x^0 \in \mathbb{R}^n$, $\mu_0 = \frac{1-\epsilon}{\|F(x^0)\|_2^2}$ ($0 < \epsilon < 1$) and $a > 0$;

while not converged **do**

$\quad z^k := B_{\mu_k}(x^k) = x^k + \mu_k F(x^k)^\top (y - F(x^k)x^k)$, $\lambda_{1,k} = \frac{2|B_{\mu_k}(x^k)|_{r+1}}{a\mu_k}$, $\lambda_{2,k} = \frac{(2a|B_{\mu_k}(x^k)|_r + 1)^2}{4a^2\mu_k}$,

$\quad \mu_k = \frac{1-\epsilon}{\|F(x^k)\|_2^2}$;

$\quad$ if $\lambda_{1,k} \leq \frac{1}{a^2\mu_k}$ then

$\quad\quad \lambda = \lambda_{1,k}$; $t^*_{a,\lambda\mu_k} = \frac{\lambda\mu_k a}{2}$

$\quad\quad$ for $i = 1 : \text{length}(x)$

$\quad\quad$ 1. $|z_i^k| > t^*_{a,\lambda\mu_k}$, then $x_i^{k+1} = g_{a,\lambda_k\mu_k}(z_i^k)$;

$\quad\quad$ 2. $|z_i^k| \leq t^*_{a,\lambda\mu_k}$, then $x_i^{k+1} = 0$;

$\quad$ else

$\quad\quad \lambda = \lambda_{2,k}$; $t^* = \sqrt{\lambda\mu_k} - \frac{1}{2a}$;

$\quad\quad$ for $i = 1 : \text{length}(x)$

$\quad\quad$ 1. $|z_i^k| > t^*_{a,\lambda\mu_k}$, then $x_i^{k+1} = g_{a,\lambda_k\mu_k}(z_i^k)$;

$\quad\quad$ 2. $|z_i^k| \leq t^*_{a,\lambda\mu_k}$, then $x_i^{k+1} = 0$;

$\quad$ end

$\quad\quad k \to k + 1$;

end while

return: x^{k+1}.

---

linear, and the more detailed accounts of the setting in the form of (27) can be found in [6, 7]. By randomly generating such sufficiently sparse vectors x_0 (choosing the nonzero locations uniformly over the support in random, and their values from $N(0, 1)$), we generate vectors b. In this way, we know the sparsest solution to $F(x_0)x_0 = b$, and we are able to compare this with algorithmic results. The stopping criterion is usually as follows:

$$\frac{\|x_k - x_{k-1}\|_2}{\|x_k\|_2} \leq \text{Tol},$$

where x_k and x_{k-1} are numerical results from two continuous iterative steps and Tol is a given small number. The success is measured by computing

$$\text{relative error} = \frac{\|x^* - x_0\|_2}{\|x_0\|_2} \leq \text{Re},$$

where x^* is the numerical results generated by IFTA, and Re is also a given small number. In all of our experiments, we set Tol $= 10^{-8}$ to indicate the stopping criterion, and set Re $= 10^{-4}$ to indicate a perfect recovery of the original sparse vector x_0.

Figure 8 shows the success rate of three algorithms in the recovery of a sparse signal with different cardinality. In this experiment, we repeatedly perform 30 tests and present average results and take $a = 2.5$.
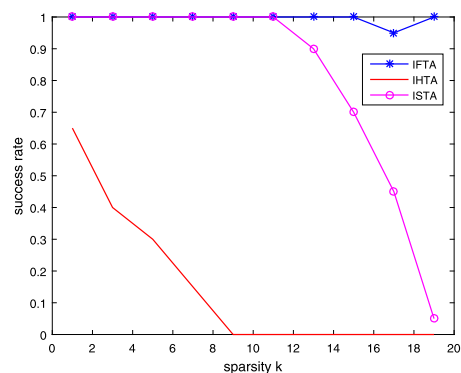
Figure 9 shows the relative error between the solution x^* and the given signal x_0. In this experiment, we repeatedly perform 30 tests and present average results and take $a = 2.5$.

The graphs presented in Fig. 8 and Fig. 9 show the performance of the ISTA, IHTA and IFTA in recovering the true (sparsest) signals. From Fig. 8, we can see that IFTA performs best, and IST algorithm the second. From Fig. 9, we see that the IFTA has the smallest relative error value with sparsity growing.

Figure 8 Success rate of three algorithms in the recovery of a sparse signal with different cardinality. In this experiment, we repeatedly perform 30 tests and present average results and take $a = 2.5$

Figure 9 Relative error between the solution x^* and the given signal x_0. In this experiment, we repeatedly perform 30 tests and present average results and take $a = 2.5$
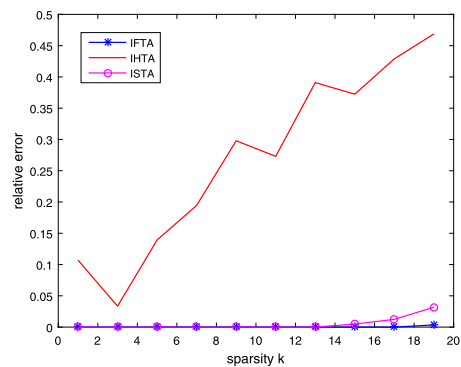
5 Conclusion

In this paper, we take the fraction function as the substitution for ℓ_0-norm in quasi-linear compressed sensing. An iterative fraction thresholding algorithm is proposed to solve the regularization problem (QP_a^λ) for all $a > 0$. With the change of parameter $a > 0$, our algorithm could get a promising result, which is one of the advantages for our algorithm compared with some state-of-art algorithms. We also provide a series of experiments to assess performance of our algorithm and the experiment results have illustrated that our algorithms is able to address the sparse signal recovery problems in nonlinear systems. Compared with ISTA and IHTA, IFTA performs best in sparse signal recovery and has the smallest relative error value with sparsity growing. However, the convergence of our algorithm is not proved theoretically in this paper, and it is our future work.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to this work. All authors read and approved the final manuscript.

Author details
[1]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China. [2]School of Science, Xi'an Polytechnic University, Xi'an, China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Candes, E., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math. **59**(8), 1207–1223 (2006)
2. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
3. Bruckstein, A.M., Donoho, D.L., Elad, M.: From sparse solutions of systems of equations to sparse modelling of signals and images. SIAM Rev. **51**(1), 34–81 (2009)
4. Elad, M.: Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springe, New York (2010)
5. Theodoridis, S., Kopsinis, Y., Slavakis, K.: Sparsity-aware learning and compressed sensing: an overview. https://arxiv.org/pdf/1211.5231
6. Ehler, M., Fornasier, M., Sigl, J.: Quasi-linear compressed sensing. Multiscale Model. Simul. **12**(2), 725–754 (2014)
7. Sigl, J.: Quasilinear compressed sensing. Master's thesis, Technische University München, Munich, Germany (2013)
8. Geman, D., Reynolds, G.: Constrained restoration and recovery of discontinuities. IEEE Trans. Pattern Anal. Mach. Intell. **14**(3), 367–383 (1992)
9. Li, H., Zhang, Q., Cui, A., Peng, J.: Minimization of fraction function penalty in compressed sensing. https://arxiv.org/pdf/1705.06048
10. Cui, A., Peng, J., Li, H., Zhang, C., Yu, Y.: Affine matrix rank minimization problem via non-convex fraction function penalty. J. Comput. Appl. Math. **336**, 353–374 (2018)
11. Xing, F.: Investigation on solutions of cubic equations with one unknown. J. Cent. Univ. Natl. (Nat. Sci. Ed.) **12**(3), 207–218 (2003)
12. Xu, Z., Chang, X., Xu, F., Zhang, H.: L1/2 regularization: a thresholding representation theory and a fast solver. IEEE Trans. Neural Netw. Learn. Syst. **23**(7), 1013–1027 (2012)