

RESEARCH

Open Access



A new smoothing modified three-term conjugate gradient method for l_1 -norm minimization problem

Shouqiang Du^{1*}  and Miao Chen¹

*Correspondence:
sqdu@qdu.edu.cn

¹School of Mathematics and
Stochastic, Qingdao University,
Qingdao, China

Abstract

We consider a kind of nonsmooth optimization problems with l_1 -norm minimization, which has many applications in compressed sensing, signal reconstruction, and the related engineering problems. Using smoothing approximate techniques, this kind of nonsmooth optimization problem can be transformed into a general unconstrained optimization problem, which can be solved by the proposed smoothing modified three-term conjugate gradient method. The smoothing modified three-term conjugate gradient method is based on Polak–Ribière–Polyak conjugate gradient method. For the Polak–Ribière–Polyak conjugate gradient method has good numerical properties, the proposed method possesses the sufficient descent property without any line searches, and it is also proved to be globally convergent. Finally, the numerical experiments show the efficiency of the proposed method.

MSC: 90C30; 90C33

Keywords: Nonsmooth optimization problem; Smoothing modified three-term conjugate gradient method; Global convergence

1 Introduction

In this paper, we consider the following nonsmooth optimization problems with l_1 -norm minimization problem

$$\min_{x \in R^n} \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1, \quad (1)$$

where $x \in R^n$, $A \in R^{m \times n}$ ($m \ll n$), $b \in R^m$, $\tau > 0$, $\|v\|_2$ denotes the Euclidean norm of v and $\|v\|_1 = \sum_{i=1}^n |v_i|$ is the l_1 -norm of v . This problem is widely used in compressed sensing, signal reconstruction, analog-to-information conversion and related to many mathematical problems [1–16]. Problem (1) is also a typical compressed sensing scenario, which can reconstruct a length- n sparse signal from m observations. From the Bayesian perspective, problem (1) can also be seen as a maximum a posteriori criterion for estimating x from observations $b = Ax + \omega$, where ω is the Gaussian noise of variance σ^2 . Many researchers have studied the numerical algorithms, which can be used to solve problem (1) with large-scale data such as fixed point method [1], gradient projection method for sparse reconstruction [2], interior-point continuation method [3, 4], iterative shrinkage thresholds algorithms in

[5, 6], linearized Bregman method [7, 8], alternating direction algorithms [9], nonsmooth equations-based method [10] and some related methods [11, 12]. Just recently, a smoothing gradient method has been given for solving problem (1) based on the new transformed absolute value equations in [14, 15]. The transformation is based on the equivalence between a linear complementarity problem and an absolute value equation problem [17, 18]. The complementarity problem, the absolute value equation problem, and the related constrained optimization problem are three kinds of important optimization problems [19–23]. On the other hand, the nonlinear conjugate gradient methods and smoothing methods are used widely to solve large-scale optimization problems [24, 25], total variation image restoration [26], monotone nonlinear equations with convex constraints [27], and nonsmooth optimization problems, such as nonsmooth nonconvex problems [28], minimax problem [29], P0 nonlinear complementarity problems [30]. Specially, the effectiveness of widely used and attained different numerical outcomes three-term conjugate gradient method, which is based on Hang–Zhang conjugate gradient method and Polak–Ribière–Polyak conjugate gradient method [31–33], has been widely studied. Based on the above analysis, in this paper, we propose a new smoothing modified three-term conjugate gradient method to solve problem (1). The global convergence analysis of the proposed method is also presented.

The remainder of this paper is organized as follows. In Sect. 2, we give the transformation of problem (1), which includes the transformation of a linear complementarity problem transformed into an absolute value equation problem. In Sect. 3, we present the smoothing modified three-term conjugate gradient method and give the convergence analysis of it. Finally, we give some numerical results of the given method which show the effectiveness of it.

2 Results: the transformation of the problem

In this section, as in [9, 10, 14, 15], we set

$$x = u - v, \quad u \geq 0, v \geq 0,$$

where $u_i = (x_i)_+$ and $v_i = (-x_i)_+$ for all $i = 1, 2, \dots, n$ with $(x_i)_+ = \max\{x_i, 0\}$. And we also have $\|x\|_1 = 1_n^T u + 1_n^T v$, where $1_n = [1, 1, \dots, 1]^T$ is an n -dimensional vector with n ones. Thus, problem (1) can be transformed into the following problem:

$$\min_{z=(u,v)^T \geq 0} \frac{1}{2} \|b - Az\|_2^2 + \tau 1_n^T u + \tau 1_n^T v,$$

i.e.,

$$\min_{z \geq 0} \frac{1}{2} z^T H z + c^T z, \tag{2}$$

where

$$z = (u, v)^T, \quad c = \tau 1_{2n} + \begin{pmatrix} -c^- \\ c^- \end{pmatrix}, \quad c = A^T b, \quad H = \begin{pmatrix} A^T A & -A^T A \\ -A^T A & A^T A \end{pmatrix}.$$

Since H is a positive semi-definite matrix, problem (2) can be translated into a linear variable inequality problem, which is to find $z \in R^{2n}$ such that

$$\langle Hz + c, \tilde{z} - z \rangle \geq 0, \quad \forall \tilde{z} \geq 0. \quad (3)$$

By the feasible structure of the feasible region of z , problem (3) is equivalent to the linear complementary problem, which to find $z \in R^{2n}$ such that

$$z \geq 0, \quad Hz + c \geq 0, \quad z^T(Hz + c) = 0. \quad (4)$$

Due to the equivalence of linear complementarity problems and absolute value equation problems, problem (4) can be transformed into the following absolute value equation problem, which is defined by

$$(H + I)z + c = |(H - I)z + c|.$$

Then problem (1) can be transformed into the following problem:

$$\min_{z \in R^{2n}} f(z) = \frac{1}{2} \|(H + I)z + c - |(H - I)z + c|\|^2. \quad (5)$$

3 Main results and discussions

In this section, we present the smoothing modified three-term conjugate gradient method to solve problem (1). Firstly, we give the definition of smoothing function and smoothing approximation function of the absolute value function [14, 15, 29].

Definition 1 Let $f : R^n \rightarrow R$ be a local Lipschitz continuous function. We call $\tilde{f} : R^n \times R_+ \rightarrow R$ a smoothing function of f , if

$$\lim_{\mu \rightarrow 0} \tilde{f}(x) = f(x),$$

where $f_\mu(\cdot)$ is continuously differentiable in R^n for any fixed $\mu > 0$.

The smoothing function of the absolute value function is defined by

$$\Phi_{i\mu}(z) = \sqrt{((H - I)z + c)_i^2 + \mu^2}, \quad \mu \in R_+, i = 1, 2, \dots, 2n, \quad (6)$$

and satisfies

$$\lim_{\mu \rightarrow 0} \Phi_{i\mu}(z) = |(H - I)z + c|_i, \quad i = 1, 2, \dots, 2n.$$

Based on (6), we obtain the following unconstrained optimization problem:

$$\min_{z \in R^{2n}} \tilde{f}_\mu(z) = \frac{1}{2} \sum_{i=1}^{2n} \tilde{f}_{i\mu}^2(z),$$

where $\tilde{f}_{i\mu}(z) = ((H + I)z + c)_i - \Phi_{i\mu}(z)$ is a smoothing function of $f(z)$ in (5) for $i = 1, 2, \dots, 2n$.

Now, we give the smoothing modified three-term conjugate gradient method.

Algorithm 1 (Smoothing modified three-term conjugate gradient method)

- Step 0. Choose $0 < \sigma < 1$, $0 < \rho < 1$, $r > 0$, $\mu = 2$, $\eta = 1$, $\varepsilon > 0$, $\mu_0 > 1$ and, given an initial point $z_0 \in R^n$, let $d_0 = -\tilde{g}_0$, where $\tilde{g}_0 = \nabla_z \tilde{f}(z_0, \mu_0)$.
- Step 1. If $\|\nabla_z \tilde{f}\| \leq \varepsilon$, stop; otherwise, go to Step 2.
- Step 2. Compute search direction by using $\tilde{\beta}_k^{\text{BZAU}}$ and $\tilde{\theta}_k^{\text{BZAU}}$, which are defined by

$$\tilde{\beta}_k^{\text{BZAU}} = \frac{\nabla_z \tilde{f}_\mu(z_k)^T (\nabla_z \tilde{f}_\mu(z_k) - \nabla_z \tilde{f}_\mu(z_{k-1}))}{-\eta \nabla_z \tilde{f}_\mu(z_{k-1})^T d_{k-1} + \mu |\nabla_z \tilde{f}_\mu(z_k)^T d_{k-1}|}, \quad (7)$$

$$\tilde{\theta}_k^{\text{BZAU}} = \frac{\nabla_z \tilde{f}_\mu(z_k)^T d_{k-1}}{-\eta \nabla_z \tilde{f}_\mu(z_{k-1})^T d_{k-1} + \mu |\nabla_z \tilde{f}_\mu(z_k)^T d_{k-1}|}, \quad (8)$$

$$d_k = \begin{cases} -\nabla_z \tilde{f}_\mu(z_k) & \text{if } k = 0, \\ -\nabla_z \tilde{f}_\mu(z_k) + \tilde{\beta}_k^{\text{BZAU}} d_{k-1} - \tilde{\theta}_k^{\text{BZAU}} y_{k-1} & \text{if } k \geq 1, \end{cases}$$

where $y_{k-1} = \nabla_z \tilde{f}_\mu(z_k) - \nabla_z \tilde{f}_\mu(z_{k-1})$.

- Step 3. Compute α_k by the Armijo line search, where $\alpha_k = \max\{\rho^0, \rho^1, \rho^2, \dots\}$ and ρ^i satisfies

$$\tilde{f}(z_k + \rho^i d_k, \mu_k) \leq \tilde{f}(z_k, \mu_k) + \sigma \rho^i \nabla_z \tilde{f}_\mu(z_k)^T d_k. \quad (9)$$

- Step 4. Compute $z_{k+1} = z_k + \alpha_k d_k$, if $\|\nabla_z \tilde{f}(z_{k+1}, \mu_k)\| \geq r \mu_k$, set $\mu_{k+1} = \mu_k$. Otherwise, let $\mu_{k+1} = \sigma \mu_k$.

- Step 5. Set $k := k + 1$ and go to Step 1.

Now, we give convergence analysis of Algorithm 1. In order to get the global convergence of Algorithm 1, we give the following assumptions.

Assumption 1

- (i) The level set $\Omega = \{z \in R^{2n} | \tilde{f}_\mu(z) \leq \tilde{f}_\mu(z_0)\}$ is bounded.
- (ii) There exists a positive constant $L > 0$ such that $\nabla_z \tilde{f}_\mu(z_k)$ is Lipschitz continuous on an open convex set $B \subseteq \Omega$ and for any $z_1, z_2 \in B$, i.e.,

$$\|\nabla_z \tilde{f}_\mu(z_1) - \nabla_z \tilde{f}_\mu(z_2)\| \leq L \|z_1 - z_2\|.$$

- (iii) There exists a positive constant m such that

$$m \|d_x\|^2 \leq d^T \nabla_z^2 \tilde{f}_\mu(z_k) d, \quad \forall x, d \in R^n,$$

where $\nabla_z^2 \tilde{f}_\mu(z_k)$ is the Hessian matrix of \tilde{f} .

By Assumption 1, we can see that there exist positive constants $\gamma > 0$ and b such that

$$\|\nabla_z \tilde{f}_\mu(z_k)\| \leq \gamma, \quad \forall z_k \in \Omega$$

and

$$\|z_1 - z_2\| \leq b, \quad \forall z_1, z_2 \in \Omega.$$

Lemma 1 Suppose $\{z_k\}$ and $\{d_k\}$ are generated by Algorithm 1, then

$$\nabla_{Z\tilde{f}_\mu}(z_k)^T d_k = -\|\nabla_{Z\tilde{f}_\mu}(z_k)\|^2$$

and

$$\|\nabla_{Z\tilde{f}_\mu}(z_k)\| \leq \|d_k\|.$$

Proof By Algorithm 1, we have

$$d_k = -\nabla_{Z\tilde{f}_\mu}(z_k) + \tilde{\beta}_k^{\text{BZAU}} d_{k-1} - \tilde{\theta}_k^{\text{BZAU}} y_{k-1}.$$

Multiplying both sides of the above equation by $\nabla_{Z\tilde{f}_\mu}(z_k)^T$, we obtain

$$\begin{aligned} \nabla_{Z\tilde{f}_\mu}(z_k)^T d_k &= -\|\nabla_{Z\tilde{f}_\mu}(z_k)\|^2 + \frac{\nabla_{Z\tilde{f}_\mu}(z_k)^T (\nabla_{Z\tilde{f}_\mu}(z_k) - \nabla_{Z\tilde{f}_\mu}(z_{k-1})) (\nabla_{Z\tilde{f}_\mu}(z_k)^T d_{k-1})}{-\eta \nabla_{Z\tilde{f}_\mu}(z_{k-1})^T d_{k-1} + \mu |\nabla_{Z\tilde{f}_\mu}(z_k)^T d_{k-1}|} \\ &\quad - \frac{(\nabla_{Z\tilde{f}_\mu}(z_k)^T d_{k-1} \nabla_{Z\tilde{f}_\mu}(z_k)^T (\nabla_{Z\tilde{f}_\mu}(z_k) - \nabla_{Z\tilde{f}_\mu}(z_{k-1})))}{-\eta \nabla_{Z\tilde{f}_\mu}(z_{k-1})^T d_{k-1} + \mu |\nabla_{Z\tilde{f}_\mu}(z_k)^T d_{k-1}|}, \end{aligned}$$

i.e.,

$$\nabla_{Z\tilde{f}_\mu}(z_k)^T d_k = -\|\nabla_{Z\tilde{f}_\mu}(z_k)\|^2.$$

Now, we have

$$|\nabla_{Z\tilde{f}_\mu}(z_k)^T d_k| = \|\nabla_{Z\tilde{f}_\mu}(z_k)\|^2$$

and

$$|\nabla_{Z\tilde{f}_\mu}(z_k)^T d_k| \leq \|\nabla_{Z\tilde{f}_\mu}(z_k)\| \|d_k\|.$$

By

$$\|\nabla_{Z\tilde{f}_\mu}(z_k)\|^2 \leq \|\nabla_{Z\tilde{f}_\mu}(z_k)\| \|d_k\|$$

we have

$$\|\nabla_{Z\tilde{f}_\mu}(z_k)\| \leq \|d_k\|.$$

Hence, the proof is complete. \square

Lemma 2 Suppose Assumption 1 holds and $\{z_k\}$ and $\{d_k\}$ are generated by Algorithm 1, then

$$\sum_{k=0}^{\infty} \frac{(\nabla_{Z\tilde{f}_\mu}(z_k)^T d_k)^2}{\|d_k\|^2} < +\infty$$

and

$$\sum_{k=0}^{\infty} \frac{\|\nabla_Z \tilde{f}_{\mu}(z_k)\|^4}{\|d_k\|^2} < +\infty.$$

Proof Using the techniques similar to lemmas in [31–33], we can get this lemma. The description will not be repeated again. \square

Lemma 3 Suppose Assumption 1 holds and x_k and d_k are generated by Algorithm 1, then

$$a_1 \alpha_k \|d_k\|^2 \leq -\nabla_Z \tilde{f}_{\mu}(z_k)^T d_k, \quad (10)$$

where $a_1 = (1 - \sigma)^{-1}(m/2)$, m is a positive constant and $0 < \sigma < 1$.

Proof By using Taylor's expansion, we have

$$\tilde{f}(z_{k+1}) = \tilde{f}(z_k) + \nabla_Z \tilde{f}_{\mu}(z_k)^T s_k + \frac{1}{2} s_k^T G_k s_k, \quad (11)$$

where $s_k = z_{k+1} - z_k = \alpha_k d_k$ and

$$G_k = \int_0^1 \nabla_Z^2 \tilde{f}_{\mu}(z_k + \tau s_k) d\tau s_k.$$

By Armijo line search, we know that

$$\tilde{f}(z_{k+1}) \leq \tilde{f}(z_k) + \sigma \nabla_Z \tilde{f}_{\mu}(z_k)^T s_k. \quad (12)$$

By (11) and (12), we have

$$\frac{1}{2} s_k^T G_k s_k \leq (1 - \sigma) (-\nabla_Z \tilde{f}_{\mu}(z_k)^T s_k),$$

i.e.,

$$\frac{1}{2} (1 - \sigma)^{-1} m \alpha_k \|d_k\|^2 \leq -\nabla_Z \tilde{f}_{\mu}(z_k)^T d_k.$$

Denote $a_1 = (1 - \sigma)^{-1}(m/2)$, we get (10). Thus, we complete the proof. \square

By Lemmas 1, 2, and 3, we can get global convergence of the given method, i.e., the following theorem.

Theorem 1 Suppose Assumption 1 holds, then

$$\lim_{k \rightarrow \infty} \|\nabla_Z \tilde{f}_{\mu}(z_k)\| = 0.$$

Proof From Assumption 1, (7), and (10), we have

$$|\tilde{\beta}_k^{\text{BZAU}}| \leq \left| \frac{\nabla_Z \tilde{f}_{\mu}(z_k)^T (\nabla_Z \tilde{f}_{\mu}(z_k) - \nabla_Z \tilde{f}_{\mu}(z_{k-1}))}{\eta (-\nabla_Z \tilde{f}_{\mu}(z_{k-1})^T d_{k-1})} \right| \leq \frac{\|\nabla_Z \tilde{f}_{\mu}(z_k)\| L \alpha_{k-1} \|d_{k-1}\|}{\eta (a_1 \alpha_{k-1} \|d_{k-1}\|^2)}, \quad (13)$$

$$|\tilde{\beta}_k^{\text{BZAU}}| \|d_{k-1}\| \leq \left(\frac{L \|\nabla_Z \tilde{f}_\mu(z_k)\|}{\eta(a_1 \|d_{k-1}\|)} \right) \|d_{k-1}\| = \frac{L \|\nabla_Z \tilde{f}_\mu(z_k)\|}{\eta a_1},$$

i.e.,

$$|\tilde{\theta}_k^{\text{BZAU}}| \|y_{k-1}\| \leq \left| \frac{\nabla_Z \tilde{f}_\mu(z_k)^T d_{k-1}}{\eta(-\nabla_Z \tilde{f}_\mu(z_{k-1})^T d_{k-1})} \right| \|y_{k-1}\|.$$

From Assumption 1, (8), and (10), we have

$$|\tilde{\theta}_k^{\text{BZAU}}| \|y_{k-1}\| \leq \left(\frac{\|\nabla_Z \tilde{f}_\mu(z_k)\| L \|x_k - x_{k-1}\|}{\eta(a_1 \alpha_{k-1} \|d_{k-1}\|^2)} \right) \|d_{k-1}\| = \frac{L \|\nabla_Z \tilde{f}_\mu(z_k)\|}{\eta a_1}. \quad (14)$$

Combining (13), (14), and d_k generated in Algorithm 1, we obtain

$$\begin{aligned} \|d_k\| &\leq \|\nabla_Z \tilde{f}_\mu(z_k)\| + |\tilde{\beta}_k^{\text{BZAU}}| \|d_{k-1}\| + |\tilde{\theta}_k^{\text{BZAU}}| \|y_{k-1}\| \\ &\leq \|\nabla_Z \tilde{f}_\mu(z_k)\| + \frac{L \|\nabla_Z \tilde{f}_\mu(z_k)\|}{\eta a_1} + \frac{L \|\nabla_Z \tilde{f}_\mu(z_k)\|}{\eta a_1} \\ &= \left(1 + \frac{2L}{\eta a_1} \right) \|\nabla_Z \tilde{f}_\mu(z_k)\|. \end{aligned}$$

Denote $\sqrt{B} = (1 + \frac{2L}{\eta a_1})$, we have $\|d_k\|^2 \leq B \|\nabla_Z \tilde{f}_\mu(z_k)\|^2$, i.e.,

$$\frac{1}{\|d_k\|^2} \geq \frac{1}{B \|\nabla_Z \tilde{f}_\mu(z_k)\|^2}$$

and

$$\frac{B \|\nabla_Z \tilde{f}_\mu(z_k)\|^4}{\|d_k\|^2} \geq \frac{\|\nabla_Z \tilde{f}_\mu(z_k)\|^4}{\|\tilde{g}_k\|^2} = \|\nabla_Z \tilde{f}_\mu(z_k)\|^2$$

By Lemma 2, we have

$$\sum_{k=0}^{\infty} \|\nabla_Z \tilde{f}_\mu(z_k)\|^2 < +\infty.$$

This completes the proof. \square

4 Numerical experiments

In this section, we give some numerical experiments of Algorithm 1, which are also considered in [2, 9, 10, 14, 15]. We compare Algorithm 1 with smoothing gradient method, GPSR method, debiased and minimum norm methods proposed in [2, 9, 10, 14] respectively. The numerical results of all the examples show that Algorithm 1 is effective. All codes run in MATLAB 8.0. For Examples 1 and 2, the parameters used in Algorithm 1 are chosen as $\sigma = 0.2$, $\mu = 5$, $\eta = 2$, $\gamma = 0.5$, $\varepsilon = 10^{-6}$, $\rho = 0.4$.

Example 1 Consider

$$A = \begin{pmatrix} 3 & 5 & 8 & 4 & 1 & 5 \\ 2 & 9 & 6 & 5 & 7 & 4 \\ 3 & 4 & 7 & 2 & 1 & 6 \\ 8 & 9 & 6 & 5 & 7 & 4 \end{pmatrix}, \quad b = (2 \quad 4 \quad 1 \quad 7)^T,$$

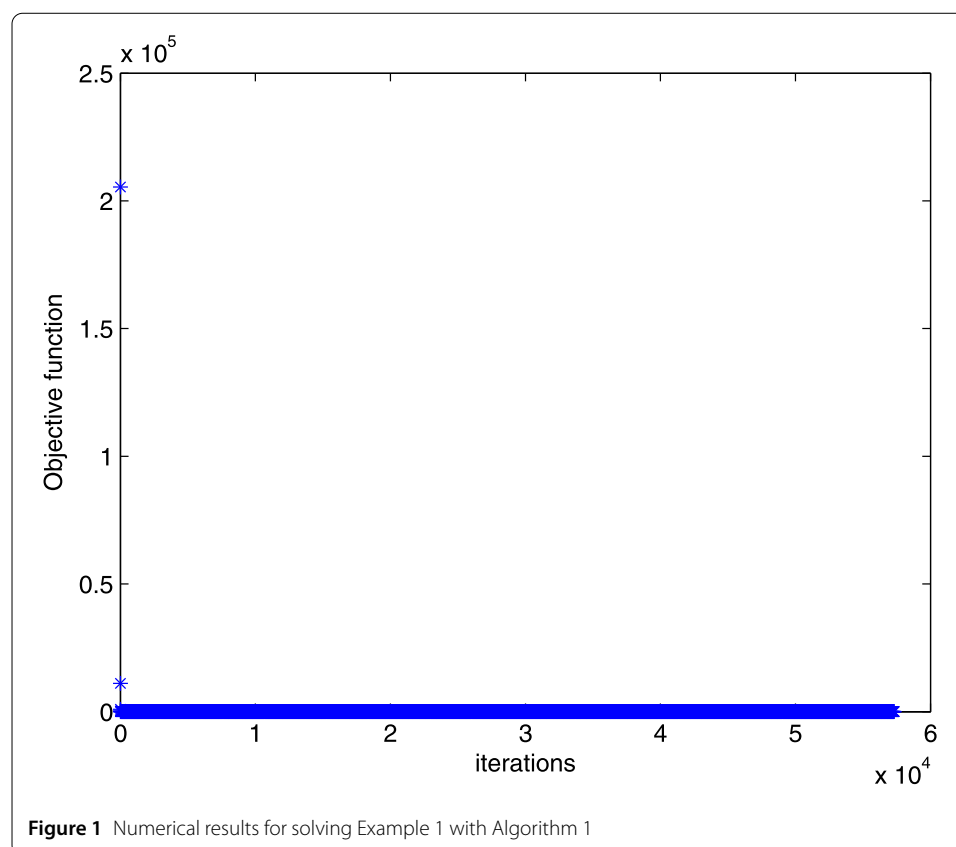
and $\tau = 5$.

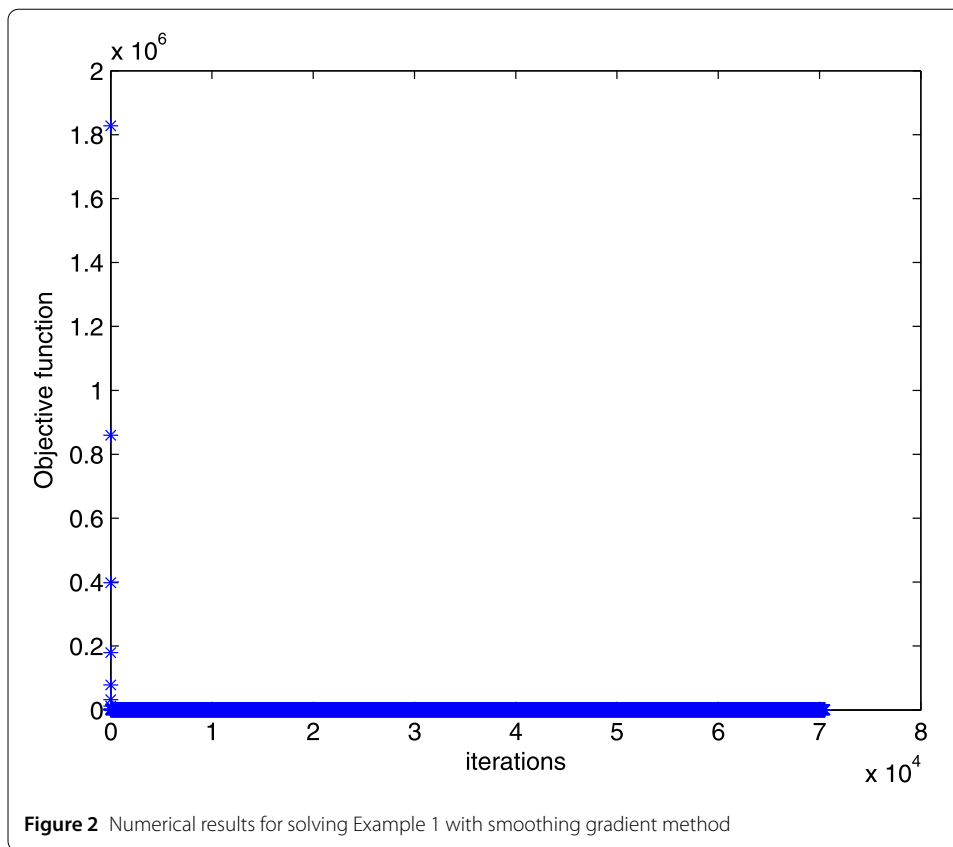
From [14], we know that this example has a solution $x^* = (0.3461, 0.0850, 0, 0, 0.3719, 0)^T$. The optimal solution of Algorithm 1 is $x^* = (0.3459, 0.0850, 0.0001, 0.0009, 0.3717, -0.0001)^T$. In Figs. 1 and 2, we plot the evolution of the objective function versus the number of iterations when solving Example 1 with Algorithm 1 and the smoothing gradient method respectively.

Example 2 Consider

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 1 & 1 & 1 & \cdots & 1 \end{pmatrix}_{m \times n},$$

$$b = (1 \quad 1 \quad \cdots \quad 1)^T,$$





and $\tau = 2$. In this example, we choose $m = 30$, $n = 100$. The numerical results are given in Figs. 3 and 4.

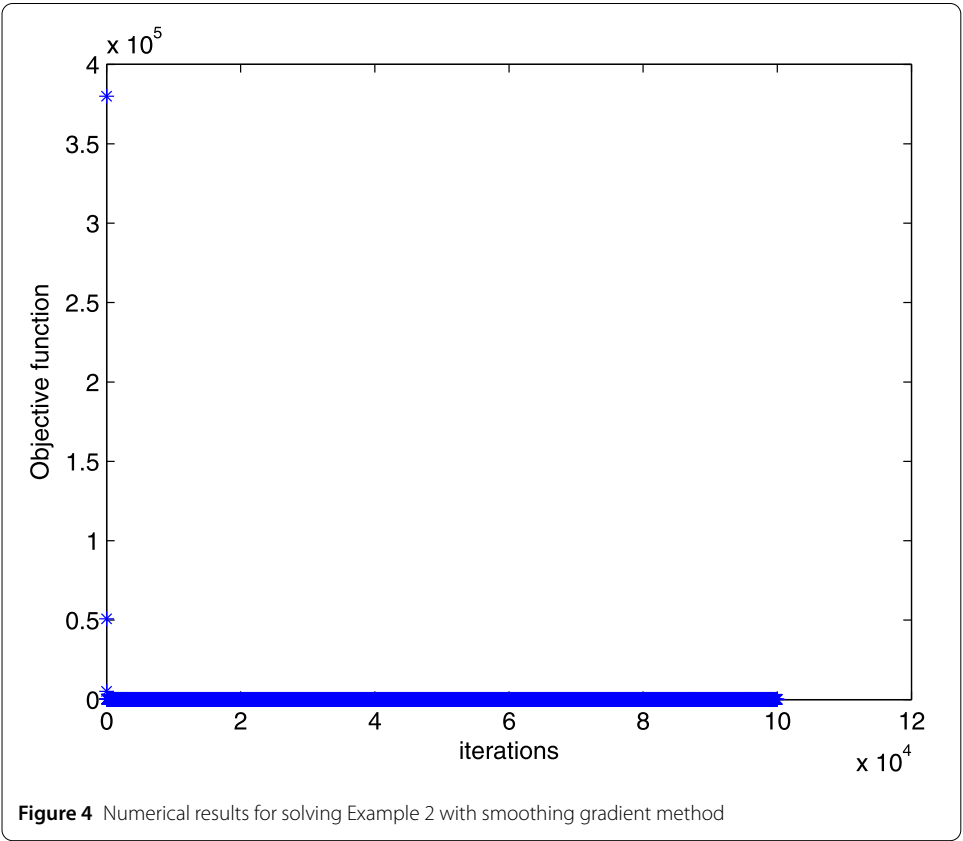
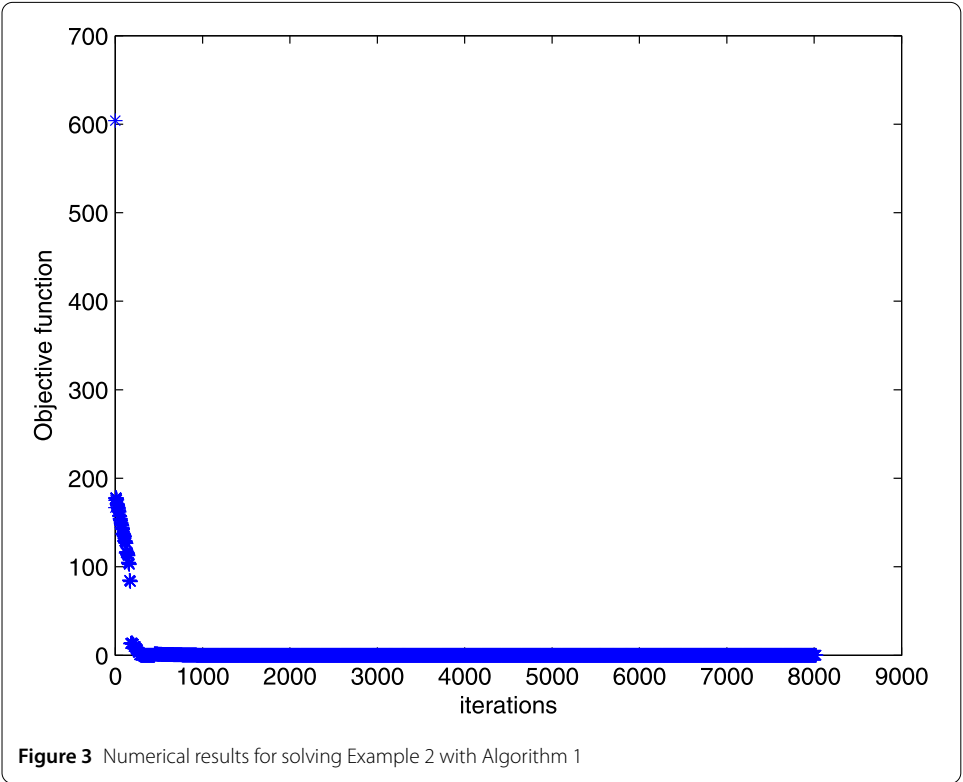
Example 3 Consider

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 1 & 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \end{pmatrix}_{m \times n}, \quad b = (1 \ 1 \ \cdots \ 1)^T,$$

and $\tau = 6$. In this example, we choose $m = 100$, $n = 110$. The numerical results are given in Figs. 5 and 6.

Example 4 Consider

$$A = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ -1 & 4 & -1 & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & -1 & 4 & \ddots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & -1 & 4 & 1 & \cdots & 1 \end{pmatrix}_{m \times n}, \quad b = (1 \ 1 \ \cdots \ 1)^T,$$



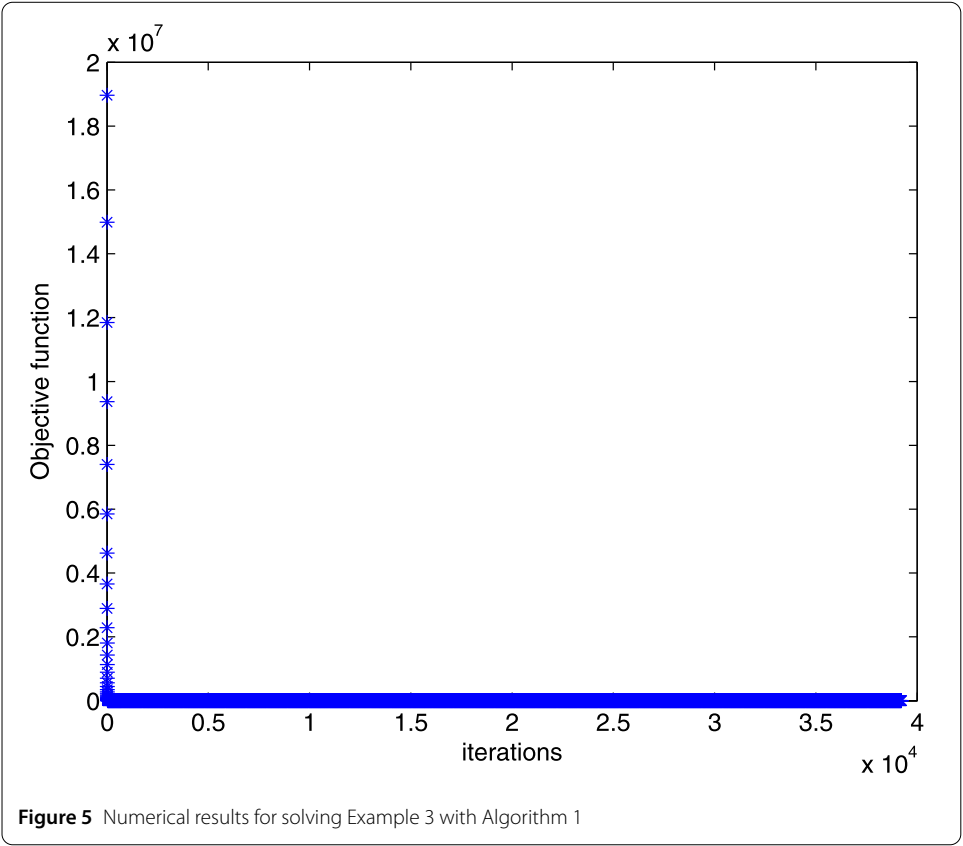


Figure 5 Numerical results for solving Example 3 with Algorithm 1

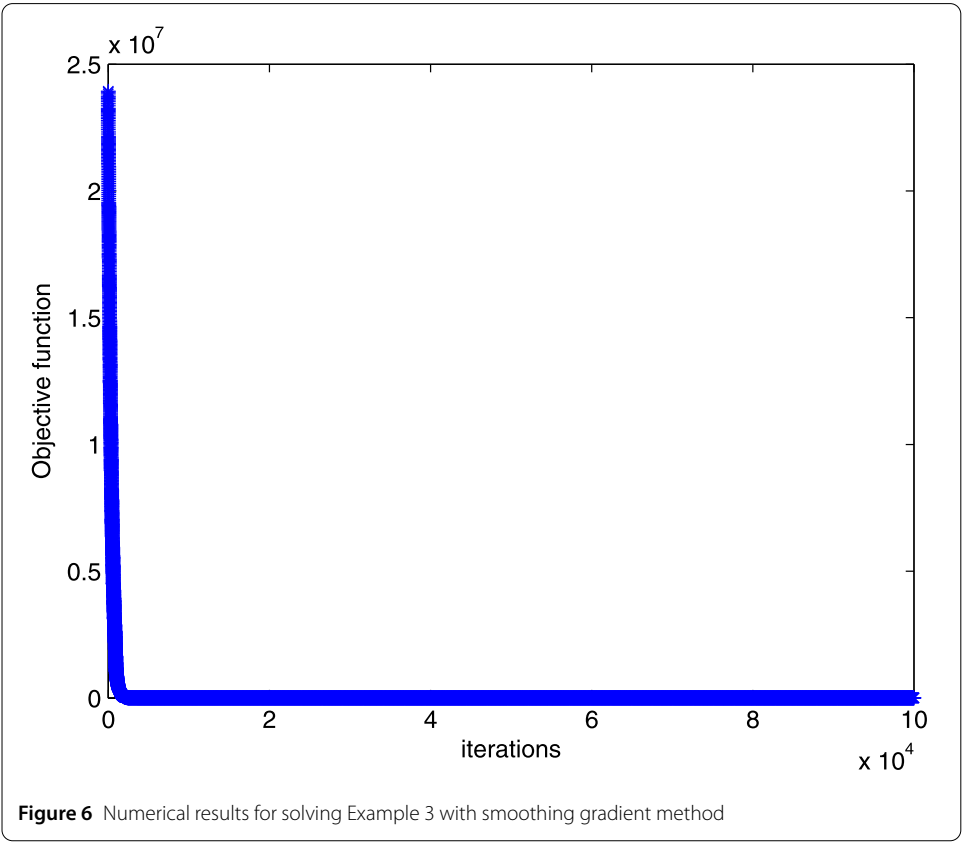
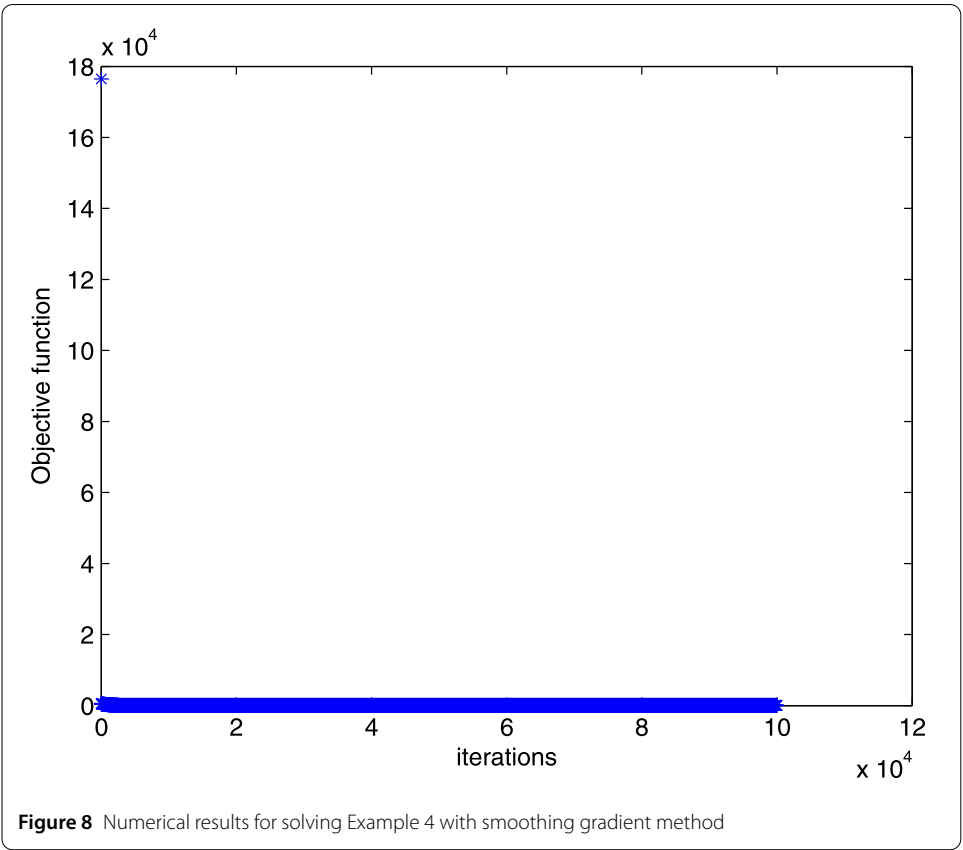
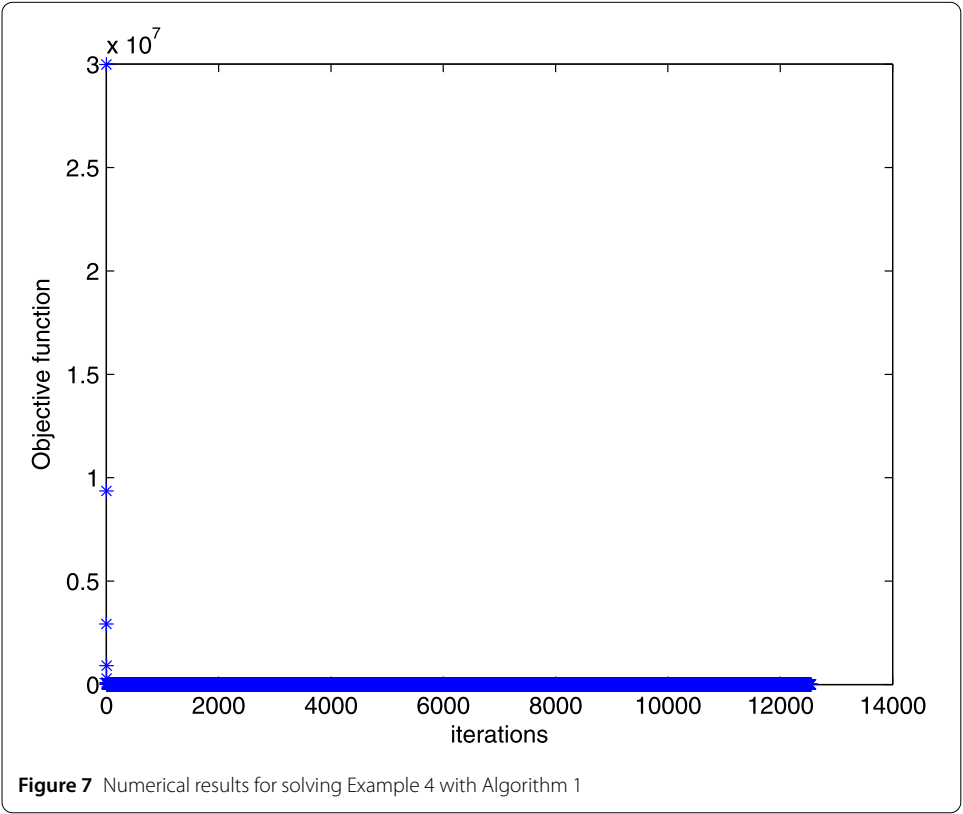


Figure 6 Numerical results for solving Example 3 with smoothing gradient method



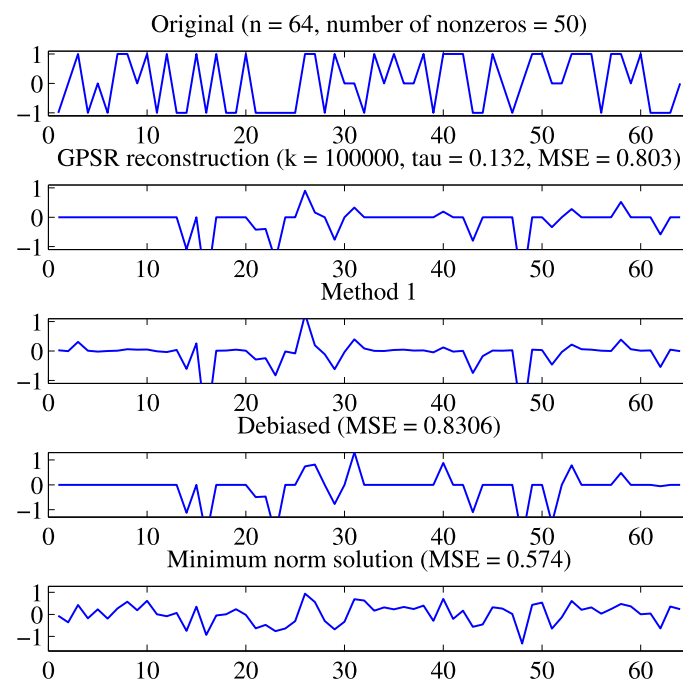


Figure 9 Numerical results for solving Example 5 with Algorithm 1

and $\tau = 10$. In this example, we choose $m = 200$, $n = 210$. The numerical results are given in Figs. 7 and 8.

Example 5 In this example, we consider a typical compressed sensing problem, which is also considered in [9, 10, 14, 15]. In this example, we choose $m = 2^4$, $n = 2^6$, $\sigma = 0.5$, $\rho = 0.4$, $\gamma = 0.5$, $\varepsilon = 10^{-6}$, $\mu = 5$, $\eta = 2$. The original signal contains 520 randomly generated ± 1 spikes. Further, the $m \times n$ matrix A is obtained by first filling it with independent samples of a standard Gaussian distribution and then orthogonalization of its rows. In this example, we choose $\sigma^2 = 10^{-4}$ and $\tau = 0.1 \|A^T y\|_\infty$ the same as suggested in [14]. The numerical results are shown in Fig. 9.

5 Conclusion

In this paper, we have proposed a new smoothing modified three-term conjugate gradient method for solving l_1 -norm nonsmooth problems. Comparing with the smoothing gradient method, GPSR method, and other methods proposed in [2, 9, 10, 14], we can see that the smoothing modified three-term conjugate gradient method is simple and needs small storage. Comparing with the smoothing gradient method proposed in [14], the smoothing modified three-term conjugate gradient method is significantly faster especially in solving large-scale problems.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 11671220) and the Natural Science Foundation of Shandong Province (No. ZR2016AM29).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally. All authors read and approved the final manuscript.

Publisher's Note

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

Received: 26 February 2018 Accepted: 21 April 2018 Published online: 03 May 2018

References

- Hale, E.T., Yin, W., Zhang, Y.: Fixed-point continuation for l_1 -minimization: methodology and convergence. *SIAM J. Optim.* **19**(3), 1107–1130 (2008)
- Figueiredo, M.A.T., Nowak, R.D., Wright, S.J.: Gradient projection for sparse reconstruction, application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.* **1**(4), 586–597 (2007)
- Kim, S.J., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D.: An interior-point method for large-scale l_1 -regularized least squares. *IEEE J. Sel. Top. Signal Process.* **4**, 606–617 (2007)
- Turlach, B.A., Venables, W.N., Wright, S.J.: Simultaneous variable selection. *Technometrics* **47**(3), 349–363 (2005)
- Daubechies, I., Defrise, M., Mol, C.D.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **57**(11), 1413–1457 (2004)
- Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2**, 183–202 (2009)
- Osher, S., Mao, Y., Dong, B., Yin, W.: Fast linearized Bregman iteration for compressive sensing and sparse denoising. *Math. Comput.* **8**(1), 93–111 (2011)
- Yin, W., Osher, S., Goldfarb, D., Darbon, J.: Bregman iterative algorithm for l_1 minimization with applications to compressed sensing. *SIAM J. Imaging Sci.* **1**(1), 143–168 (2008)
- Yang, J., Zhang, Y.: Alternating direction algorithms for l_1 -problems in compressive sensing. *SIAM J. Sci. Comput.* **33**, 250–278 (2011)
- Xiao, Y.H., Wang, Q.Y., Hu, Q.J.: Non-smooth equations based method for l_1 -norm problems with applications to compressed sensing. *Nonlinear Anal.* **74**(11), 3570–3577 (2011)
- Wakin, M.B., Laska, J.N., Duarte, M.F., Baron, D., Sarvotham, S., Takhar, D., Kelly, K.F., Baraniuk, R.G.: An architecture for compressive imaging. In: 2006 International Conference on Image Processing, pp. 1273–1276 (2006)
- Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* **58**(6), 1182–1195 (2010)
- Farajzadeh, A.P., Plubtieng, S., Ungchittrakool, K.: On best proximity point theorems without ordering. *Abstr. Appl. Anal.* **2014**, Article ID 130439 (2014)
- Chen, Y.Y., Gao, Y., Liu, Z.M., Du, S.Q.: The smoothing gradient method for a kind of special optimization problem. *Oper. Res. Trans.* **21**(2), 119–125 (2017)
- Chen, M., Du, S.Q.: The smoothing FR conjugate gradient method for solving a kind of nonsmooth optimization problem with l_1 -norm. *Math. Probl. Eng.* **2018**, Article ID 5817931 (2018)
- Wang, Y.J., Zhou, G.L., Caccetta, L., Liu, W.Q.: An alternative Lagrange-dual based algorithm for sparse signal reconstruction. *IEEE Trans. Signal Process.* **59**(4), 1895–1901 (2011)
- Mangasarian, O.L.: Absolute value programming. *Comput. Optim. Appl.* **36**(1), 43–53 (2007)
- Mangasarian, O.L.: A generalized Newton method for absolute value equations. *Optim. Lett.* **3**(1), 101–108 (2009)
- Caccetta, L., Qu, B., Zhou, G.L.: A globally and quadratically convergent method for absolute value equations. *Comput. Optim. Appl.* **48**(1), 45–58 (2011)
- Mangasarian, O.L.: Absolute value equation solution via concave minimization. *Optim. Lett.* **1**(1), 3–8 (2007)
- Chen, H.B., Wang, Y.J., Zhao, H.G.: Finite convergence of a projected proximal point algorithm for generalized variational inequalities. *Oper. Res. Lett.* **40**(4), 303–305 (2012)
- Chen, Y.Y., Gao, Y.: Two kinds of new Levenberg–Marquardt method for nonsmooth nonlinear complementarity problem. *ScienceAsia* **40**, 89–93 (2014)
- Qu, B., Xiu, N.H.: A relaxed extragradient-like method for a class of constrained optimization problem. *J. Ind. Manag. Optim.* **3**(4), 645–654 (2007)
- Fletcher, R., Reeves, C.M.: Function minimization by conjugate gradients. *Comput. J.* **7**(2), 149–154 (1964)
- Dai, Y.H., Yuan, Y.: A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.* **10**(1), 177–182 (1999)
- Yu, G.H., Qi, L.Q., Dai, Y.H.: On nonmonotone chambolle gradient projection algorithms for total variation image restoration. *J. Math. Imaging Vis.* **35**(2), 143–154 (2009)
- Yu, Z.S., Lin, J., Sun, J., Xiao, Y.H., Lin, L.Y., Li, Z.H.: Spectral gradient projection method for monotone nonlinear equations with convex constraints. *Appl. Numer. Math.* **59**(10), 2416–2423 (2009)
- Chen, X.J.: Smoothing methods for nonsmooth, nonconvex minimization. *Math. Program.* **134**, 71–99 (2012)
- Pang, D.Y., Du, S.Q., Ju, J.J.: The smoothing Fletcher–Reeves conjugate gradient method for solving finite minimax problems. *ScienceAsia* **42**, 40–45 (2016)
- Zhang, L., Wu, S.Y., Gao, T.: Improved smoothing Newton methods for P_0 nonlinear complementarity problems. *Appl. Math. Comput.* **215**(1), 324–332 (2009)
- Zhang, L., Zhou, W.: On the global convergence of the Hang–Zhang conjugate gradient method with Armijo line search. *Acta Math. Sci.* **28**(5), 840–845 (2008)
- Sun, M., Liu, J.: Three modified Polak–Ribière–Polyak conjugate gradient methods with sufficient descent property. *J. Inequal. Appl.* **2015**, Article ID 125 (2015)
- Baluch, B., Salleh, Z., Alhawarat, A., Roslan, U.: A new modified three-term conjugate gradient method with sufficient descent property and its global convergence. *J. Math.* **2017**, Article ID 2715854 (2017)