Hindawi
Mathematical Problems in Engineering
Volume 2019, Article ID 5976595, 9 pages
https://doi.org/10.1155/2019/5976595

Research Article

Modified Three-Term Conjugate Gradient Method and Its Applications

Jiankun Liu and Shouqiang Du

School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China

Correspondence should be addressed to Shouqiang Du; sqdu@qdu.edu.cn

Received 16 January 2019; Revised 15 March 2019; Accepted 25 March 2019; Published 17 April 2019

Academic Editor: Yann Favennec

We propose a modified three-term conjugate gradient method with the Armijo line search for solving unconstrained optimization problems. The proposed method possesses the sufficient descent property. Under mild assumptions, the global convergence property of the proposed method with the Armijo line search is proved. Due to simplicity, low storage, and nice convergence properties, the proposed method is used to solve $\mathscr{M}$-tensor systems and a kind of nonsmooth optimization problems with l_1-norm. Finally, the given numerical experiments show the efficiency of the proposed method.

1. Introduction

We consider the following unconstrained optimization problem:

$$\min_{z\in R^n} \quad f(z), \tag{1}$$

where $f: R^n \longrightarrow R$ is a continuous function. It is well-known that the nonlinear conjugate gradient method is one of the most effective methods for solving large-scale unconstrained optimization problems due to its simplicity and low storage [1–8]. Let z_0 be the initial approximation of the solution to (1); the general format of the nonlinear conjugate gradient method is as follows:

$$z_{k+1} = z_k + \alpha_k d_k, \quad k = 0, 1, \cdots, \tag{2}$$

where α_k can be obtained by some linear searches, i.e., [6–8], and search direction d_k is computed by

$$d_k = \begin{cases} -g_k & if\ k = 0, \\ -g_k + \beta_k d_{k-1} & if\ k \geq 1, \end{cases} \tag{3}$$

where g_k is the gradient of f at point z_k and β_k is a parameter. Different choices for the parameter β_k correspond to different nonlinear conjugate gradient methods. The Fletcher-Reeves (FR) method, the Polak-Ribiere-Polyak (PRP) method, the Hestenes-Stiefel (HS) method, the Dai-Yuan (DY) method, and the Conjugate Descent (CD) method are some famous nonlinear conjugate gradient methods [1, 2, 9–12], and the parameters β_k of them are, respectively, defined by

$$\begin{aligned} \beta_k^{FR} &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \\ \beta_k^{PRP} &= \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \\ \beta_k^{HS} &= \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \\ \beta_k^{DY} &= \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \\ \beta_k^{CD} &= \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}}, \end{aligned} \tag{4}$$

where $y_{k-1} = g_k - g_{k-1}$ and $\|\cdot\|$ is the Euclidean norm. Because of the good numerical performance of the conjugate gradient method, in recent years, the nonlinear three-term conjugate gradient method has been paid much attention by researchers, such as the three-term conjugate gradient method [5], the three-term form of the L-BFGS method [13], the three-term PRP conjugate gradient method [14], and the new-type conjugate gradient update parameter similar with β_k^{PRP} [15]. On the other hand, we know that the Armijo line search is widely used in solving optimization problems; i.e., see [8]. So, in this paper, we propose a new modified three-term conjugate gradient method with the Armijo line search. The proposed method is used to solve $\mathscr{M}$-tensor systems [16, 17] and a kind of nonsmooth optimization problems with l_1-norm [18–22].

The remainder of this paper is organized as follows: In the next section, we give the new modified three-term conjugate gradient method. Firstly, we give the smooth case of the proposed method and prove the sufficient descent property and the global convergence property of it. Then, we give the nonsmooth case of the proposed method. In Section 3, we present $\mathscr{M}$-tensor systems and a kind of nonsmooth minimization problems with l_1-norm, which can be solved by the proposed method. And, we also give some numerical results to show the efficiency of the proposed method. In Section 4, we give the conclusion of this paper.

2. Modified Three-Term Conjugate Gradient Method

In this section, we consider the nonlinear conjugate gradient method for solving (1); we discuss the problem in two cases: (1) f is a smooth function; (2) f is a nonsmooth function.

2.1. Smooth Case. Based on nonlinear conjugate gradient methods in [5, 8], we propose a modified three-term conjugate gradient method with the Armijo line search. We consider the search direction

$$\begin{aligned} d_0 &= -g_0, \\ d_k &= -g_k + \beta_k d_{k-1} + \theta_k y_{k-1}, \end{aligned} \tag{5}$$

where $g_k = \nabla f(z_k)$ is the gradient of f at z_k, $y_{k-1} = g_k - g_{k-1}$, and

$$\beta_k = \frac{g_k^T y_{k-1}}{\|d_{k-1}\|^2}, \tag{6}$$

$$\theta_k = -\frac{g_k^T d_{k-1}}{\|d_{k-1}\|^2}. \tag{7}$$

From (5), (6), and (7), we can obtain that

$$d_k^T g_k = -\|g_k\|^2, \quad \forall k \geq 0. \tag{8}$$

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

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

Step 0. Choose $0 < \sigma < 1, 0 < \rho < 1, \varepsilon > 0$ and give an initial point $z_0 \in R^n$, let $k = 0$, compute $g_0 = \nabla f(z_0)$, and let $d_0 = -g_0$.

Step 1. If $\|\nabla f(z_k)\| \leq \varepsilon$, stop; otherwise, go to Step 2.

Step 2. Compute the search direction d_k by (5), where β_k and θ_k are defined by (6) and (7).

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

$$f(z_k + \alpha_k d_k) - f(z_k) \leq \sigma \alpha_k g_k^T d_k. \tag{9}$$

Step 4. Compute $z_{k+1} = z_k + \alpha_k d_k$, where d_k is given in Step 2 and α_k is given in Step 3.

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

Next, we will give the global convergence analysis of Algorithm 1. Firstly, we give the following assumptions.

Assumption 2. The level set $R_0 = \{z \in R^n \mid f(z) \leq f(z_0)\}$ is bounded; i.e., there exists a positive constant $B > 0$ such that $\|z\| \leq B$ for all $z \in R_0$.

Assumption 3. In the neighborhood N of R_0, f is continuously differentiable and its gradient g is Lipschitz continuous; that is, there exists a positive constant $L > 0, \forall x, y \in N$, such that

$$\|g(x) - g(y)\| \leq L\|x - y\|. \tag{10}$$

Remark 4. Because $\{f(z_k)\}$ is a decreasing sequence, so the sequence $\{z_k\}$ generated by Algorithm 1 is contained in R_0. And by Assumptions 2 and 3, we can easily obtain that there exists a positive constant γ such that

$$\|g_k\| \leq \gamma, \quad \forall z \in R_0. \tag{11}$$

Lemma 5. *Suppose $\{g_k\}$ and $\{d_k\}$ are generated by Algorithm 1, then*

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \tag{12}$$

Proof. Firstly, we prove that there exists a constant $c > 0$ such that, for sufficiently large k,

$$\alpha_k \geq c\frac{\|g_k\|^2}{\|d_k\|^2}. \tag{13}$$

The proof of (13) can be divided into two following cases.

Case 1 ($\alpha_k = 1$). By (8) and the Cauchy inequality, $\|g_k\|^2 = \|d_k^T g_k\| \leq \|d_k\|\|g_k\|$, then we have $\|g_k\| \leq \|d_k\|$. Let $c_1 \leq 1$, then we obtain (13).

Case 2 ($\alpha_k < 1$). Due to the linear search step, that is the Step 3 of Algorithm 1, $\rho^{-1}\alpha_k$ does not satisfy (9); i.e.,

$$f\left(z_k + \rho^{-1}\alpha_k d_k\right) - f\left(z_k\right) > \sigma\rho^{-1}\alpha_k g_k^T d_k. \tag{14}$$

By Assumption 3 and the mean value theorem, there exists $t_k \in (0,1)$ such that

$$\begin{aligned} &f\left(z_k + \rho^{-1}\alpha_k d_k\right) - f\left(z_k\right) \\ &\quad = \rho^{-1}\alpha_k g\left(z_k + t_k\rho^{-1}\alpha_k d_k\right)^T d_k \\ &\quad = \rho^{-1}\alpha_k g_k^T d_k \\ &\qquad + \rho^{-1}\alpha_k\left(g\left(z_k + t_k\rho^{-1}\alpha_k d_k\right) - g_k\right)^T d_k \\ &\quad \leq \rho^{-1}\alpha_k g_k^T d_k + L\rho^{-2}\alpha_k^2\|d_k\|^2. \end{aligned} \tag{15}$$

By the above formula, (8) and (14), we have

$$\alpha_k \geq \frac{(1-\sigma)\rho\|g_k\|^2}{L\|d_k\|^2}. \tag{16}$$

Let $c = \min\{c_1, (1-\sigma)\rho/L\}$, then we obtain (13).

By (9) and Assumption 2, we have

$$-\sum_{k=1}^{\infty}\alpha_k g_k^T d_k < +\infty. \tag{17}$$

From (8), (13), and (17), we have

$$-\sum_{k=1}^{\infty}\alpha_k g_k^T d_k = \sum_{k=1}^{\infty}\alpha_k\|g_k\|^2 \geq \sum_{k=1}^{\infty}c\frac{\|g_k\|^4}{\|d_k\|^2}, \tag{18}$$

then we get

$$\sum_{k=0}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \tag{19}$$

Hence, the result follows. □

Now we can get the global convergence of Algorithm 1.

Theorem 6. *Suppose $\{g_k\}$ and $\{d_k\}$ are generated by Algorithm 1, then*

$$\liminf_{k\longrightarrow\infty}\|g_k\| = 0. \tag{20}$$

Proof. Using the technique similar to Theorem 3.1 in [5], we can get this theorem. □

Remark 7. The Armijo type line search [7] is given as follows:

$$f\left(z_k + \alpha_k d_k\right) - f\left(z_k\right) \leq \sigma_1\alpha_k g_k^T d_k - \sigma_2\alpha_k^2\|d_k\|^2, \tag{21}$$

where $\alpha_k = \max\{\rho^j, j = 0,1,2,\ldots,\}$, $\rho \in (0,1)$, $\sigma_1 \in (0,1)$, and $\sigma_2 > 0$. The Wolfe type line search [6] is given as follows:

$$\begin{aligned} f\left(z_k + \alpha_k d_k\right) - f\left(z_k\right) &\leq -\rho\alpha_k^2\|d_k\|^2, \\ g\left(z_k + \alpha_k d_k\right)^T d_k &\geq -2\sigma\alpha_k\|d_k\|^2, \end{aligned} \tag{22}$$

where $0 < \sigma < 1, 0 < \rho < 1$. Obviously, Algorithm 1 is also true for the Armijo type line search and the Wolfe type line search.

2.2. Nonsmooth Case. In this subsection, by using smoothing function, we extend Algorithm 1 to the nonsmooth case. Firstly, we give the definition of smoothing function.

Definition 8. Let $f : R^n \longrightarrow R$ be a local Lipschitz continuous function. If $\forall z \in R^n$, $\forall\mu > 0$, and μ is fixed, $\widetilde{f}(\cdot,\mu)$ is continuously differentiable and satisfies

$$\lim_{\mu\longrightarrow 0}\widetilde{f}\left(z,\mu\right) = f\left(z\right), \tag{23}$$

then we call $\widetilde{f} : R^n \times R^+ \longrightarrow R$ is a smoothing function of f.

Denote $\widetilde{g}_k = \nabla\widetilde{f}(z_k,\mu_k)$. Now, we present the following smoothing modified three-term conjugate gradient method.

Algorithm 9 (smoothing modified three-term conjugate gradient method).

Step 0. Choose $0 < \sigma < 1, 0 < \rho < 1, \varepsilon > 0, r > 0, \mu_0 > 1, 0 < \sigma_1 < 1$ and give an initial point $z_0 \in R^n$, let $k = 0$, compute $\widetilde{g}_0 = \nabla\widetilde{f}(z_0,\mu_0)$, and let $d_0 = -\widetilde{g}_0$.

Step 1. If $\|\widetilde{g}_k\| \leq \varepsilon$, stop; otherwise, go to Step 2.

Step 2. Compute the search direction d_k by using β_k and θ_k, where

$$\beta_k = \frac{\widetilde{g}_k^T y_{k-1}}{\|d_{k-1}\|^2},$$

$$\theta_k = -\frac{\widetilde{g}_k^T d_{k-1}}{\|d_{k-1}\|^2},$$

$$d_k = \begin{cases} -\tilde{g}_k & if\ k = 0, \\ -\tilde{g}_k + \beta_k d_{k-1} + \theta_k y_{k-1} & if\ k \geq 1, \end{cases} \tag{24}$$

where $y_{k-1} = \tilde{g}_k - \tilde{g}_{k-1}$.

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

$$\tilde{f}(z_k + \alpha_k d_k, \mu_k) - \tilde{f}(z_k, \mu_k) \leq \sigma \alpha_k \tilde{g}_k^T d_k. \tag{25}$$

Step 4. Compute $z_{k+1} = z_k + \alpha_k d_k$, if $\|\nabla \tilde{f}(z_{k+1}, \mu_k)\| \geq r\mu_k$, set $\mu_{k+1} = \mu_k$; otherwise, let $\mu_{k+1} = \sigma_1 \mu_k$.

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

Next, we give the global convergence analysis of Algorithm 9.

Theorem 10. *Suppose that $\tilde{f}(\cdot, \mu)$ is a smoothing function of f. If for every fixed $\mu > 0$, $\tilde{f}(\cdot, \mu)$ satisfies Assumptions 2 and 3, then $\{z_k\}$ generated by Algorithm 9 satisfies*

$$\liminf_{k \longrightarrow \infty} \left\|\nabla \tilde{f}(z_k, \mu_{k-1})\right\| = 0. \tag{26}$$

Proof. Denote $K = \{k | \mu_{k+1} = \sigma_1 \mu_k\}$. If K is finite, then there exists an integer $\bar{k}$ such that, for all $k > \bar{k}$,

$$\left\|\nabla \tilde{f}(z_k, \mu_{k-1})\right\| \geq r\mu_{k-1} \tag{27}$$

and $\mu_k = \mu_{\bar{k}} = \bar{\mu}$. That is to solve

$$\min_{z \in R^n} \quad \tilde{f}(z, \bar{\mu}). \tag{28}$$

Hence, from Theorem 6, we get

$$\liminf_{k \longrightarrow \infty} \left\|\nabla \tilde{f}(z_k, \bar{\mu})\right\| = 0, \tag{29}$$

which contradicts with (27). This shows that K must be infinite and $\lim_{k \longrightarrow \infty} \mu_k = 0$. Since K is infinite, we can assume that $K = \{k_0, k_1, \cdots\}$ with $k_0 < k_1 < \cdots$. Then we have

$$\lim_{i \longrightarrow \infty} \left\|\nabla \tilde{f}\left(z_{k_i+1}, \mu_{k_i}\right)\right\| \leq r \lim_{i \longrightarrow \infty} \mu_{k_i} = 0. \tag{30}$$

□

3. Applications

In this section, the applications of the proposed modified three-term conjugate gradient method are given. The conjugate gradient method is suitable for solving unconstrained optimization problems. In the first subsection, we consider the $\mathscr{M}$-tensor systems, which can be transformed into the unconstrained minimization problem and solved by Algorithm 1. Then in the second subsection, we consider a kind of nonsmooth optimization problems with l_1-norm, which can be solved by Algorithm 9. And in each subsection, the numerical results are given to show the feasibility of the proposed method.

3.1. Applications in Solving $\mathscr{M}$-Tensor Systems. In this subsection, we consider the $\mathscr{M}$-tensor systems, which can be transformed into the general unconstrained minimization problem. We use Algorithm 1 to solve it. The problem of tensor systems [16, 17] is an important problem in tensor optimization [23–26]. We consider the tensor system

$$\mathscr{A}\mathbf{z}^{m-1} = \mathbf{b}, \tag{31}$$

where $\mathscr{A} \in \mathbb{C}^{[m,n]} := \mathbb{C}^{n \times n \times \cdots \times n}$ and $\mathbf{b} \in \mathbb{C}^n$. Then the ith element of (31) is defined as

$$\left(\mathscr{A}\mathbf{z}^{m-1}\right)_i = \sum_{i_2, \ldots, i_m = 1}^{n} a_{ii_2 \cdots i_m} z_{i_2} \cdots z_{i_m}, \quad 1 \leq i \leq n. \tag{32}$$

And if $\mathbf{z} \in \mathbb{C}^n \setminus \mathbf{0}$ and $\lambda \in \mathbb{C}$ satisfy

$$\mathscr{A}\mathbf{z}^{m-1} = \lambda \mathbf{z}^{[m-1]}, \tag{33}$$

where

$$\mathbf{z}^{[m-1]} = \left[z_1^{m-1}, z_2^{m-1}, \ldots, z_n^{m-1}\right]^T, \tag{34}$$

then we call λ is an eigenvalue of $\mathscr{A}$ and z is a corresponding eigenvector of λ [25]. The spectral radius [26] of a tensor $\mathscr{A}$ is defined as

$$\rho(\mathscr{A}) = \max\{|\lambda| : \lambda \textit{ is an eigenvalue of } \mathscr{A}\}. \tag{35}$$

Let $\mathscr{L} \in \mathbb{C}^{[m,n]}$ be the identity tensor [17], i.e.,

$$\mathscr{L}_{i_1 i_2 \ldots i_m} = \begin{cases} 1, & \text{if } i_1 = i_2 = \cdots = i_m, \\ 0, & \text{otherwise}, \end{cases} \tag{36}$$

for all $1 \leq i_1, i_2, \ldots, i_m \leq n$. If there exists a nonnegative tensor $\mathscr{B}$ and a positive real number $s \geq \rho(\mathscr{B})$ such that $\mathscr{A} = s\mathscr{L} - \mathscr{B}$, then the tensor $\mathscr{A}$ is called an $\mathscr{M}$-tensor [16]. And if $s > \rho(\mathscr{B})$, it is called a nonsingular $\mathscr{M}$-tensor. Suppose $\mathscr{A}$ is a nonsingular $\mathscr{M}$-tensor, then for every positive vector $\mathbf{b}$, (31) has a unique positive solution [16]. Then (31) can be transformed into the following unconstrained minimization problem

$$\min_{z \in R^n} \quad \frac{1}{2}\left\|\mathscr{A}z^{m-1} - b\right\|^2. \tag{37}$$

Now, we present numerical experiments for solving $\mathscr{M}$-tensor systems. Some examples are taken from [16]. We implement Algorithm 1 with the codes in Matlab Version R2014a and Tensor Toolbox Version 2.6 on a laptop with an Intel(R) Core(TM) i5-2520M CPU(2.50GHz) and RAM of 4.00GB. The parameters involved in the algorithm are taken as $\sigma = 0.2, \rho = 0.25, \epsilon = 10^{-6}, \mu = 0.6$.

Example 11. Consider (31) with a 3rd-order 2-dimensional $\mathscr{M}$-tensor, where $\mathscr{A} = s\mathscr{I} - \mathscr{B} \in R^{n \times n \times n}$. And $\mathscr{B}$ contains the entries $b_{ijk} = 1$ with $i = 1, 2$ and $j, k \geq i$, and other entries are zeros. Let $s = 10$, $s > \rho(\mathscr{B})$. Hence $\mathscr{A}$ is a upper triangular nonsingular $\mathscr{M}$-tensor. The starting point z_0 is set to be rand$(n, 1)$ and b is set to be *ones*$(n, 1)$.

Table 1: The numerical results of Example 11.

z_0	k	t	z^*
$(0.7577, 0.7431)^T$	7	0.6055	$(0.3902, 0.3332)^T$
$(0.3922, 0.6555)^T$	6	0.4911	$(0.3902, 0.3332)^T$
$(0.1712, 0.7060)^T$	6	0.4553	$(0.3905, 0.3335)^T$
$(0.8235, 0.6948)^T$	6	0.4623	$(0.3902, 0.3332)^T$
$(0.3171, 0.9502)^T$	10	0.7561	$(0.3902, 0.3335)^T$
$(0.3816, 0.7655)^T$	6	0.4374	$(0.3905, 0.3335)^T$
$(0.7952, 0.1869)^T$	7	0.4651	$(0.3904, 0.3334)^T$
$(0.4898, 0.4456)^T$	4	0.3351	$(0.3905, 0.3335)^T$

Table 2: The numerical results of Example 12.

z_0	k	t	z^*
$(0.8147, 0.9058)^T$	6	2.8056	$(0.8103, 0.7653)^T$
$(0.9575, 0.9649)^T$	7	0.4469	$(0.8108, 0.7658)^T$
$(0.9572, 0.4854)^T$	8	0.4957	$(0.8106, 0.7654)^T$
$(0.4218, 0.9157)^T$	9	0.4930	$(0.8103, 0.7652)^T$
$(0.7922, 0.9595)^T$	7	0.4878	$(0.8106, 0.7654)^T$
$(0.8491, 0.9340)^T$	5	0.3502	$(0.8105, 0.7654)^T$
$(0.6787, 0.7577)^T$	6	0.3900	$(0.8107, 0.7655)^T$
$(0.7431, 0.3922)^T$	7	0.4123	$(0.8103, 0.7656)^T$

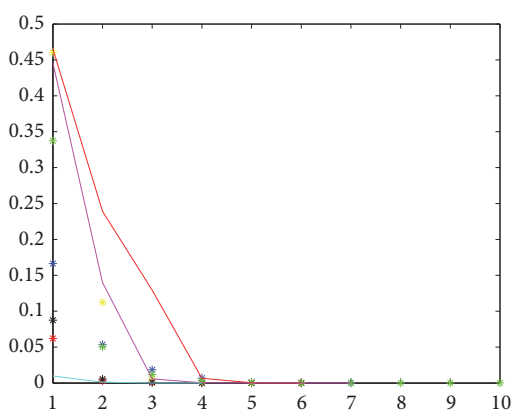

Figure 1: Numerical results for Example 11(n=2).

The numerical results are given in Table 1 and Figure 1.

Example 12. Consider (31) with a 3rd-order $\mathscr{M}$-tensor, where $\mathscr{A} = s\mathscr{I} - \mathscr{B} \in R^{n\times n\times n}$ with

$$b_{ijk} = \left|\cos\left(i + j + k\right)\right|. \tag{38}$$

By $\rho(\mathscr{B}) \leq \max_{1\leq i_1\leq n}\Sigma_{i_2,i_3=1}^{n} b_{i_1i_2i_3} < \max_{1\leq i_1\leq n}\Sigma_{i_2,i_3=1}^{n} 1 = n^2$, let $s = n^2$, $s > \rho(\mathscr{B})$. Hence $\mathscr{A}$ is a symmetric nonsingular $\mathscr{M}$-tensor. The starting point z_0 is set to be rand$(n, 1)$ and b is set to be *ones*$(n, 1)$.

When $n = 2$, the corresponding numerical results are given in Table 2 and Figure 2.

When $n = 5$, the starting points z_0 are set to be rand$(n, 1)$, then the corresponding numerical results are shown as follows:

$$z_0 = (0.6463, 0.7094, 0.7547, 0.2760, 0.6797)^T,$$
$$k = 15,$$
$$t = 4.1352,$$
$$z^* = (0.3227, 0.3169, 0.3212, 0.3195, 0.3132)^T,$$
$$z_0 = (0.3404, 0.5853, 0.2238, 0.7513, 0.2551)^T,$$
$$k = 19,$$
$$t = 2.0446,$$
$$z^* = (0.3226, 0.3168, 0.3211, 0.3194, 0.3131)^T,$$
$$z_0 = (0.1386, 0.1493, 0.2575, 0.8407, 0.2543)^T,$$
$$k = 18,$$
$$t = 2.7024,$$
$$z^* = (0.3226, 0.3168, 0.3210, 0.3193, 0.3131)^T,$$
$$z_0 = (0.5308, 0.7792, 0.9340, 0.1299, 0.5688)^T,$$
$$k = 37,$$
$$t = 3.4322,$$
$$z^* = (0.3227, 0.3169, 0.3212, 0.3195, 0.3132)^T,$$
$$z_0 = (0.3112, 0.5285, 0.1656, 0.6020, 0.2630)^T,$$
$$k = 15,$$
$$t = 1.2784,$$
$$z^* = (0.3226, 0.3168, 0.3210, 0.3193, 0.3131)^T,$$
$$z_0 = (0.1818, 0.2638, 0.1455, 0.1361, 0.8693)^T,$$
$$k = 17,$$
$$t = 1.4333,$$
$$z^* = (0.3226, 0.3168, 0.3210, 0.3193, 0.3131)^T,$$
$$z_0 = (0.3510, 0.5132, 0.4018, 0.0760, 0.2399)^T,$$
$$k = 15,$$

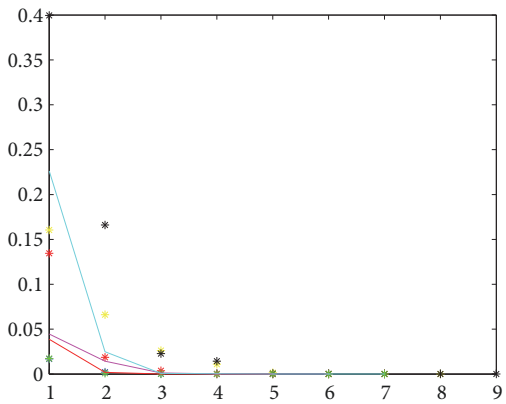

FIGURE 2: Numerical results for Example 12(n=2).

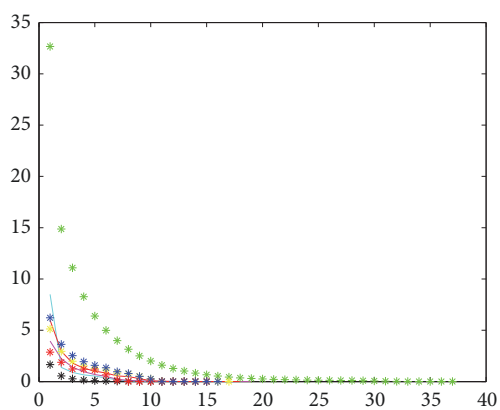

FIGURE 3: Numerical results for Example 12(n=5).

$$\begin{aligned} t &= 1.2209, \\ z^* &= (0.3228, 0.3169, 0.3212, 0.3195, 0.3133)^T, \\ z_0 &= (0.2417, 0.4039, 0.0965, 0.1320, 0.9421)^T, \\ k &= 16, \\ t &= 1.4087, \\ z^* &= (0.3226, 0.3168, 0.3210, 0.3193, 0.3131)^T, \end{aligned} \tag{39}$$

Figure 3 shows the numerical results of this example.

3.2. Applications in Solving l_1-Norm Problems. In this subsection, we consider a kind of nonsmooth optimization problems with l_1-norm. This kind of nonsmooth optimization problems can be solved by Algorithm 9. We consider

$$\min_{x\in R^n} \quad \frac{1}{2}\|Ax-b\|_2^2 + \tau\|x\|_1, \tag{40}$$

where $A \in R^{m\times n}(m \ll n)$, $b \in R^m$, and $\tau > 0$ is a parameter to trade off both terms for minimization. This problem is widely used in compressed sensing, signal reconstruction, and some related problems [18–22, 27–29]. In this subsection, we translate (40) into the absolute value equation problem based on the equivalence between the linear complementary problem and the absolute value equation problem [30] and then use Algorithm 9 to solve it.

We first give the transformation form of (40). As in [19, 21], let

$$\begin{aligned} u_i &= (x_i)_+, \\ v_i &= (-x_i)_+, \\ & i = 1, 2, \ldots, n, \end{aligned} \tag{41}$$

where $(x_i)_+ = \max\{x_i, 0\}$. $\forall x \in R^n$, we set

$$x = u - v, \quad u \geq 0, \ v \geq 0. \tag{42}$$

Due to the definition $\|x\|_1 = \sum_{i=1}^n |x_i|$, we get that

$$\|x\|_1 = e_n^T u + e_n^T v, \tag{43}$$

where $e_n = [1, 1, \ldots, 1]^T$ is an n-dimensional vector. Therefore, as in [19, 21], problem (40) can be rewritten as follows:

$$\min_{z=(u,v)^T\geq 0} \quad \frac{1}{2}\|b - Az\|_2^2 + \tau e_n^T u + \tau e_n^T v. \tag{44}$$

Then, the above problem can be transformed into

$$\min_{z\geq 0} \quad \frac{1}{2} z^T H z + c^T z, \tag{45}$$

where $z = (u, v)^T, c = \tau e_{2n} + \begin{pmatrix} -\bar{c} \\ \bar{c} \end{pmatrix}, c = A^T b, H = \begin{pmatrix} A^T A & -A^T A \\ -A^T A & A^T A \end{pmatrix}$, e_{2n} is a 2n-dimensional vector. Solving (45) is equivalent to solving the following linear complementary problem.

To find $z \in R^{2n}$, such that

$$\begin{aligned} z &\geq 0, \\ Hz + c &\geq 0, \\ z^T (Hz + c) &= 0, \end{aligned} \tag{46}$$

then (46) can be transformed into the following absolute value equation problem

$$(H + I) z + c = |(H - I) z + c|, \tag{47}$$

that is,

$$\min_{z\in R^{2n}} \quad f(z) = \frac{1}{2}\|(H + I) z + c - |(H - I) z| + c\|^2. \tag{48}$$

By the smoothing approximation function of $|(H-I)z+c|$, i.e.,

$$\begin{aligned} \Phi_i(z, \mu) &= \sqrt{((H - I) z + c)_i^2 + \mu^2}, \\ & \mu \in R^+, \ i = 1, 2, \ldots, 2n, \end{aligned} \tag{49}$$

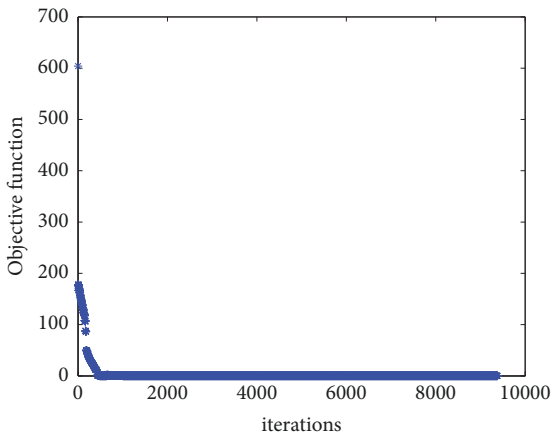

FIGURE 4: Numerical results for solving Example 13 with Algorithm 9.

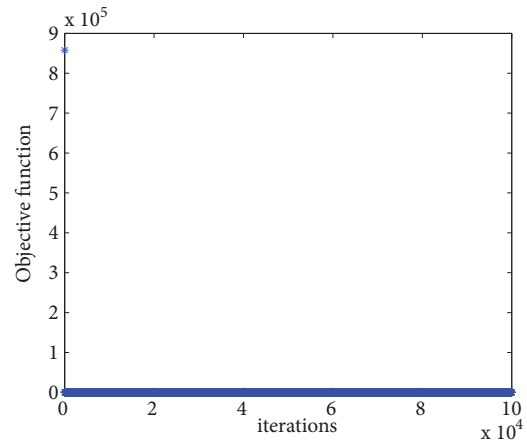

FIGURE 5: Numerical results for solving Example 14 with Algorithm 9.

then we get

$$\min_{z\in R^{2n}} \quad \tilde{f}(z,\mu) = \frac{1}{2}\sum_{i=1}^{2n} \tilde{f}_i^2(z,\mu), \tag{50}$$

where

$$\tilde{f}_i(z,\mu) = ((H+I)z+c)_i - \Phi_i(z,\mu), \quad i = 1,2,\ldots,2n. \tag{51}$$

Now, we give some numerical experiments of Algorithm 9, which are also considered in [19, 21, 22, 27, 28]. The numerical results of all examples indicate that the modified three-term conjugate gradient method is also effective for solving the l_1-norm minimization problem (40). In our numerical experiments, all codes run in Matlab R2014a. For Examples 13 and 14, the parameters used in Algorithm 9 are chosen as $\sigma = \sigma_1 = 0.2$, $r = 0.5$, $\varepsilon = 10^{-6}$, and $\rho = 0.4$.

Example 13. Consider (40) with

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & \cdots & 0 & 0 & 0 & 1 \\ . & . & . & . & . & . & \cdots & . & . & . & . \\ . & . & . & . & . & . & \cdots & . & . & . & . \\ . & . & . & . & . & . & \cdots & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 1 & 1 & 1 \end{pmatrix}_{m\times n}, \tag{52}$$

$$\tau = 3,$$

$$y = (1,1,\ldots,1)^T.$$

In this example, we choose $m = 30, n = 100$. The numerical results are given in Figure 4.

Example 14. Consider (40) with

$$A = \begin{pmatrix} 4 & -1 & 0 & \ldots & 0 & 0 & 0 & 1 & \cdots & 1 \\ -1 & 4 & -1 & \ldots & . & . & . & 1 & \cdots & 1 \\ 0 & -1 & 4 & \ldots & . & . & . & 1 & \cdots & 1 \\ . & . & . & . & . & . & . & . & \cdots & . \\ . & . & . & . & 4 & -1 & 0 & . & \cdots & . \\ . & . & . & . & -1 & 4 & -1 & . & \cdots & . \\ 0 & 0 & 0 & \ldots & 0 & -1 & 4 & 1 & \cdots & 1 \end{pmatrix}_{m\times n}, \tag{53}$$

$$\tau = 10,$$

$$y = (1,1,\ldots,1)^T.$$

In this example, we take $m = 200, n = 210$. The numerical results are given in Figure 5.

Example 15. Consider a typical compressed sensing problem with the form as (40), which is also considered in [21, 22, 27, 28]. We choose $m = 2^4$, $n = 2^6$, $\sigma = 0.5$, $\rho = 0.4$, $r = 0.5$, $\varepsilon = 10^{-6}$, $\mu = 5$, and $\eta = 2$. 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. We choose $\sigma^2 = 10^{-4}$ and $\tau = 0.1\|A^T y\|_\infty$. The numerical results are shown in Figure 6.

4. Conclusion

In this paper, we propose a modified three-term conjugate gradient method and give the applications in solving $\mathscr{M}$-tensor systems and a kind of nonsmooth optimization problems with l_1-norm. The global convergence of the proposed method is also given. Finally, we present some numerical experiments to demonstrate the efficiency of the proposed method.

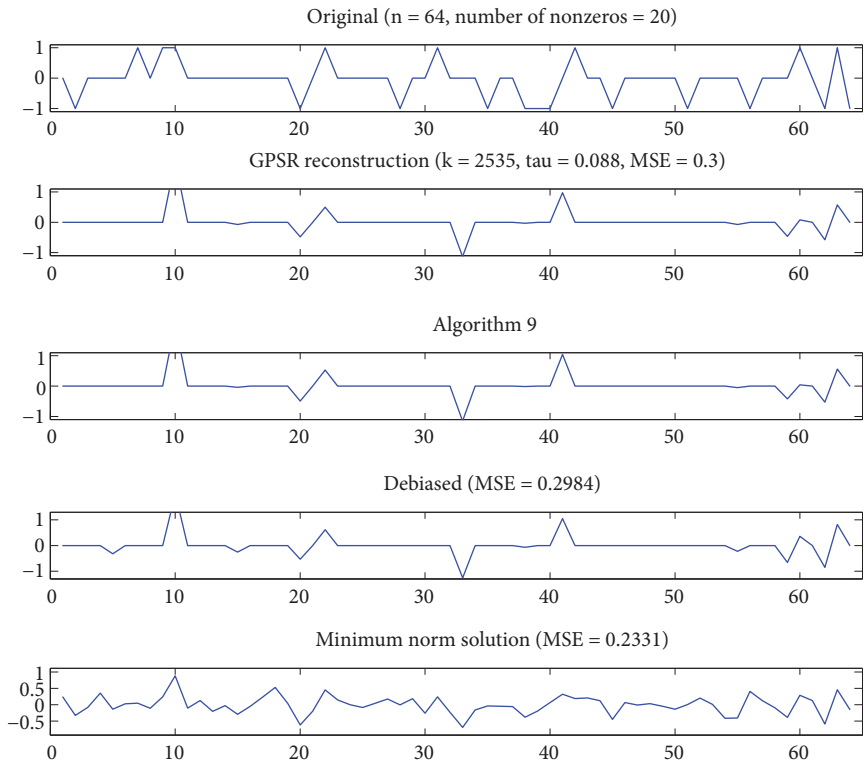

Figure 6: Numerical results for solving Example 15 with Algorithm 9.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Shandong Provincial Natural Science Foundation, China (no. ZR2016AM29), and National Natural Science Foundation of China (no. 11671220).

References

[1] E. Polak and G. Ribière, "Note sur la convergence de méthodes de directions conjuguées," *Revue Française d'Informatique et de Recherche Opérationnelle*, vol. 3, no. 16, pp. 35–43, 1969.

[2] B. T. Polyak, "The conjugate gradient method in extreme problems," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, pp. 94–112, 1969.

[3] W. W. Hager and H. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific Journal of Optimization. An International Journal*, vol. 2, no. 1, pp. 35–58, 2006.

[4] X. Chen and W. Zhou, "Smoothing nonlinear conjugate gradient method for image restoration using nonsmooth minimization," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 765–790, 2010.

[5] J. K. Liu, Y. M. Feng, and L. M. Zou, "Some three-term conjugate gradient methods with the inexact line search condition," *Calcolo*, vol. 55, no. 2, article 16, 2018.

[6] S. Du and Y. Chen, "Global convergence of a modified spectral FR conjugate gradient method," *Applied Mathematics and Computation*, vol. 202, no. 2, pp. 766–770, 2008.

[7] L. Zhang, W. J. Zhou, and D. H. Li, "Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search," *Numerische Mathematik*, vol. 104, no. 2, pp. 561–572, 2006.

[8] L. Zhang and W. Zhou, "On the global convergence of the Hager-Zhang conjugate radient method with Armijo line search," *Acta Mathematica Scientia*, vol. 28, no. 5, pp. 840–845, 2008.

[9] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, pp. 149–154, 1964.

[10] M. R. Hestenes and E. L. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, pp. 409–432, 1952.

[11] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.

[12] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1987.

[13] J. Nocedal, "Updating quasi-newton matrices with limited storage," *Mathematics of Computation*, vol. 35, no. 151, pp. 773–782, 1980.

[14] L. Zhang, W. Zhou, and D. H. Li, "A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence," *IMA Journal of Numerical Analysis (IMAJNA)*, vol. 26, no. 4, pp. 629–640, 2006.

[15] M. Rivaie, M. Mamat, L. W. June, and I. Mohd, "A new class of nonlinear conjugate gradient coefficients with global convergence properties," *Applied Mathematics and Computation*, vol. 218, no. 22, pp. 11323–11332, 2012.

[16] W. Ding and Y. Wei, "Solving multi-linear systems with M-tensors," *Journal of Scientific Computing*, vol. 68, no. 2, pp. 689–715, 2016.

[17] Z. Xie, X. Jin, and Y. Wei, "Tensor methods for solving symmetric M-tensor," *Journal of Scientific Computing*, vol. 74, no. 1, pp. 412–425, 2018.

[18] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.

[19] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction, application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, pp. 586–597, 2007.

[20] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for l_1-minimization: methodology and convergence," *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1107–1130, 2008.

[21] Y. Xiao, Q. Wang, and Q. Hu, "Non-smooth equations based method for l_1-norm problems with applications to compressed sensing," *Nonlinear Analysis*, vol. 74, no. 11, pp. 3570–3577, 2011.

[22] Y. Chen, Y. Gao, Z. Liu, and S. Du, "The smoothing gradient method for a kind of special optimization problem," *Operations Research Transactions*, vol. 21, pp. 119–125, 2017.

[23] X. Li and M. K. Ng, "Solving sparse non-negative tensor equations: algorithms and applications," *Frontiers of Mathematics in China*, vol. 10, no. 3, pp. 649–680, 2015.

[24] S. Du, L. Zhang, C. Chen, and L. Qi, "Tensor absolute value equations," *Science China Mathematics*, vol. 61, no. 9, pp. 1695–1710, 2018.

[25] L. Qi, "Eigenvalues of a real supersymmetric tensor," *Journal of Symbolic Computation*, vol. 40, no. 6, pp. 1302–1324, 2005.

[26] Y. Yang and Q. Yang, "Further results for Perron-Frobenius theorem for nonnegative tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 5, pp. 2517–2530, 2010.

[27] S. Du and M. Chen, "A new smoothing modified three-term conjugate gradient method for l_1-norm minimization problem," *Journal of Inequalities and Applications*, vol. 105, 2018.

[28] J. Yang and Y. Zhang, "Alternating direction algorithms for l_1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.

[29] 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*, vol. 1, no. 1, pp. 143–168, 2008.

[30] O. L. Mangasarian, "Absolute value equation solution via concave minimization," *Optimization Letters*, vol. 1, no. 1, pp. 3–8, 2007.