Filomat 36:6 (2022), 2011–2024
https://doi.org/10.2298/FIL2206011I

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Derivative-Free MLSCD Conjugate Gradient Method for Sparse Signal and Image Reconstruction in Compressive Sensing

Abdulkarim Hassan Ibrahim[a,e], Poom Kumam[a,b,c], Auwal Bala Abubakar[d,e], Jamilu Abubakar[a,f], Jewaidu Rilwan[d], Guash Haile Taddele[a,g]

[a]*KMUTTFixed Point Research Laboratory, Room SCL 802 Fixed Point Laboratory, Science Laboratory Building, Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), 126 Pracha-Uthit Road, Bang Mod, Thung Khru, Bangkok 10140, Thailand*
[b] *Center of Excellence in Theoretical and Computational Science (TaCS-CoE), Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), 126 Pracha Uthit Rd., Bang Mod, Thung Khru, Bangkok 10140, Thailand.*
[c]*Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan .*
[d]*Department of Mathematical Sciences, Faculty of Physical Sciences, Bayero University, Kano. Kano, Nigeria.*
[e]*Department of Mathematics and Applied Mathematics,Sefako Makgatho Health Sciences University, Ga-Rankuwa, Pretoria, Medunsa-0204, South Africa.*
[f]*Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria.*
[g]*Department of Mathematics, Debre Berhan University, Ethiopia.*

Abstract. Finding the sparse solution to under-determined or ill-condition equations is a fundamental problem encountered in most applications arising from a linear inverse problem, compressive sensing, machine learning and statistical inference. In this paper, inspired by the reformulation of the ℓ_1-norm regularized minimization problem into a convex quadratic program problem by Xiao et al. (Nonlinear Anal Theory Methods Appl, 74(11), 3570-3577), we propose, analyze, and test a derivative-free conjugate gradient method to solve the ℓ_1-norm problem arising from the reconstruction of sparse signal and image in compressive sensing. The method combines the MLSCD conjugate gradient method proposed for solving unconstrained minimization problem by Stanimirović et al. (J Optim Theory Appl, 178(3), 860-884) and a line search method. Under some mild assumptions, the global convergence of the proposed method is established using the backtracking line search. Computational experiments are carried out to reconstruct sparse signal and image in compressive sensing. The numerical results indicate that the proposed method is stable, accurate and robust.

1. Introduction

Let $w \in \mathbf{R}^n$ be a sparse or a nearly sparse original signal, $A \in \mathbf{R}^{k\times n}(k < n)$ be a linear map and b be an observed data. The relation between the signal w and the observed data b is given by:

$$b = Aw. \tag{1}$$

2020 *Mathematics Subject Classification*. Primary 90C30, 65K05;
Keywords. Compressive sensing, Nonlinear equations, Conjugate gradient method, Projection method, Global convergence.
Received: 10 February 2020; Revised: 28 February 2021; Accepted: 28 April 2021
Communicated by Predrag Stanimirović
Corresponding author: Poom Kumam
Email addresses: ibrahimkarym@gmail.com (Abdulkarim Hassan Ibrahim), poom.kum@kmutt.ac.th (Poom Kumam), ababubakar.mth@buk.edu.ng (Auwal Bala Abubakar), abubakar.jamilu@udusok.edu.ng (Jamilu Abubakar), jrilwan.mth@buk.edu.ng (Jewaidu Rilwan), guashhaile79@gmail.com (Guash Haile Taddele)

In most of the applications where sparsity constraint plays a significant role, we are dealing with an ill-conditioned or under-determined system of linear equations [1]. In this article, we focus our attention on finding sparse solutions to an under-determined linear system arising from compressive sensing (CS). In CS, it is possible to regain the sparse signal w from the linear system (1), by finding the solution of the ℓ_0-regularized problem:

$$\min_{w}\{\|w\|_0 \mid Aw = b\}, \tag{2}$$

where $\|w\|_0$ denotes the nonzero components in w. However, the ℓ_0-norm is not a proper norm and is not computationally implementable. Base on these, researchers developed alternative model by replacing the ℓ_0-norm with ℓ_1-norm. Thus, solving the basis Pursuit problem formulated as:

$$\min_{w}\{\|w\|_1 \mid Aw = b\}. \tag{3}$$

Here, $\|w\|_1 = \sum_{i=1}^{n} |w_i|$ is the ℓ_1-norm of w. Under some mild assumptions, Donoho [2] proved that solution(s) of problem (2) also solves (3). In most application, the observed value b usually contains some noise, thus the problem (3) can be relaxed to the penalized least squares problem

$$\min_{w} f(w) := \frac{1}{2}\|Aw - b\|_2^2 + \tau\|w\|_1, \tag{4}$$

where $\tau > 0$, balancing the tradeoff between sparsity and residual error. Problems of the form (4) have become familiar over the past three decades, particularly in *compressive sensing* context. Interested readers may refer to the recent papers (see, [3] and [4]) for more details.

Many approaches abound in the literature for solving (4): iterative shrinkage thresholding algorithm (IST) [5], fast iterative shrinkage thresholding algorithm (FISTA) [6], fixed-point continuous search method [7], gradient projection method [8]. Quite recent, Figueiredo, Nowak, Wright [8] first developed a gradient projection method to solve the penalized least squares problem (4). Thereafter, Xiao et al. [9, 10] proposed a conjugate gradient projection method and a spectral gradient method to solve problem (4), respectively. Unlike IST and FISTA, in order to solve problem (4), the problem was first transformed into a monotone system of equations.

Referring to [8], we briefly present a review on the reformulation procedure of (4) into a convex quadratic problem.

1.1. Reformulation of the Model

In the following, we give a short overview of the reformulation of (4) into a convex quadratic problem by Figuredo et al.[8].

Consider any vector $w \in \mathbf{R}^n$, w can be rewritten as follows

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

where $u \in \mathbf{R}^n, v \in \mathbf{R}^n$ and $u_i = (w_i)_+$, $v_i = (-w_i)_+$ for all $i \in [1, n]$ with $(\cdot)_+ = \max\{0, \cdot\}$. Therefore, the ℓ_1-norm could be represented as $\|w\|_1 = e_n^T u + e_n^T v$, where e_n is an n-dimensional vector with all element one. Thus, (4) can be written as

$$\min_{u,v}\{\frac{1}{2}\|b - A(u - v)\|^2 + \tau e_n^T u + \tau e_n^T v : u, v \geq 0\}. \tag{5}$$

Moreover, from [8], with no difficulty, (5) can be rewritten as the quadratic problem with box constraints. That is,

$$\min_{z} \frac{1}{2} z^T H z + c^T z, \quad z \geq 0, \tag{6}$$

where $z = \begin{bmatrix} u \\ v \end{bmatrix}$, $\quad c = \tau e_{2n} + \begin{bmatrix} -y \\ y \end{bmatrix}$, $\quad b = A^T y$, $\quad H = \begin{bmatrix} A^T A & -A^T A \\ -A^T A & A^T A \end{bmatrix}$.

Simple calculation shows that H is a semi-definite positive matrix. Hence, (6) is a convex quadratic problem, and equivalent to

$$G(z) = \min\{z, Hz + c \ : z \in E\} = 0, \tag{7}$$

where $E = \mathbf{R}_+^{2n}$ is a convex set. The function G is vector-valued and the "min" interpreted as componentwise minimum. From [[11], Lemma 3] and [[9], Lemma 2.2], we know that the mapping G is Lipschitz continuous and monotone. Hence, algorithms for solving (7) can be used to effectively solve (4).

The model (7) is a special class of optimization problem that has been discussed extensively by several authors with well known numerical methods developed. For instance Newton method, Quasi-Newton method, trust region method, Levenberg Marquardt method and projection method (see [12–16] and references therein). However, these methods need to compute and store the Jacobian matrix as well as solving linear equation at every iteration. These reasons make them unsuitable for large-scale problems. To overcome this drawback, several researchers have proposed derivative-free methods for solving (7). These methods incorporates conjugate gradient (CG) methods for solving unconstrained optimization problem with the projection technique of Solodov and Svaiter [17]; yielding efficient methods for solving (7) which do not need to compute and store the Jacobian matrix at every iteration. For more relevant contributions on CG methods and derivative free methods, interested readers can refer to [18–41] and the references therein.

Motivated by the approximate equivalence between problem (4) and a system of equations, we propose, analyze, and test a derivative-free conjugate gradient method to solve the ℓ_1-norm problem arising from the reconstruction of sparse signal and image in compressive sensing. The method combines the mixed LSCD conjugate gradient method (MLSCD) proposed for solving unconstrained minimization problem by Stanimirović et al. [42] and a line search method. Under some mild assumptions, the global convergence of the proposed method is established using the backtracking line search. Computational experiments are carried out to reconstruct sparse signal and image in compressive sensing. The numerical results indicate that the proposed method is stable, accurate and robust.

The paper is organised as follows: In Sections 2 and 3 of this paper, we focus on the motivation of the method, and prove that it converge globally. We perform some numerical experiments and analyze the experimental results in Section 4.

Notation. Unless stated otherwise, throughout this article, the symbol $\|\cdot\|$ denotes for Euclidean norm on $\mathbf{R}^n$. Furthermore, the projection map denoted as P_E, which is a mapping from $\mathbf{R}^n$ onto the nonempty convex set E, is defined as

$$P_E(w) = \arg\min\{\|w - y\| \ \ y \in E\}.$$

It has the well known nonexpansive property, that is,

$$\|P_E(w) - P_E(y)\| \leq \|w - y\|, \forall\ w, y \in \mathbf{R}^{2n}. \tag{8}$$

2. Algorithm

In this section, we present our framework after recalling the MLSCD conjugate gradient method by Stanimirović et al. [42]. Consider the following unconstrained optimization problem

$$\min\{f(w) | w \in \mathbf{R}^n\},$$

where the function f is assumed to be continuoisly differentiable from $\mathbf{R}^n$ into $\mathbf{R}$ and the gradient $\nabla f(w_k)$ is available. The iterative scheme of the conjugate gradient method by Stanimirović et al. [42] generates a sequence of iterate w_k using the following recursive relation:

$$w_{k+1} = w_k + \alpha_k d_k, \; k \geq 0,$$

where α_k is the step-length and the search direction d_k is updated by

$$d_k := \begin{cases} -\nabla f(w_k) + \delta_k \left(I - \frac{\nabla f(w_k) \nabla f(w_k)^T}{\|\nabla f(w_k)\|^2} \right) d_{k-1} & \text{if } k > 0, \\ -\nabla f(w_k) & \text{if } k = 0. \end{cases} \tag{9}$$

δ_k is a parameter defined as:

$$\begin{aligned} \delta_k := \delta_k^{MLSCD} &:= \max \left\{ 0, \; \min \left\{ \delta_k^{LS}, \; \delta_k^{CD} \right\} \right\} \\ &:= \max \left\{ 0, \; \min \left\{ \frac{y_{k-1}^T \nabla f(w_k)}{-\nabla f(w_{k-1})^T d_{k-1}}, \; \frac{\|\nabla f(w_k)\|^2}{-\nabla f(w_{k-1})^T d_{k-1}} \right\} \right\}, \end{aligned}$$

where $y_{k-1} := \nabla f(w_k) - \nabla f(w_{k-1})$. In what follows, we describe a derivative-free MLSCD conjugate gradient method (DF-MLSCD) for solving (7).

Algorithm 2.1. ***(DF-MLSCD)***
Input. *Choose any arbitrary initial point $w_0 \in E$, the positive constants: $Tol \in (0,1)$, $\xi \in (0,1)$, $\kappa > 0$, $\gamma > 0$. Set $k := 0$.*

Step 0. *Compute $G(w_k)$. If $\|G(w_k)\| \leq Tol$, stop. Otherwise, compute the search direction d_k by*

$$d_k := \begin{cases} -G(w_k) & \text{if } k = 0, \\ -G(w_k) + \delta_k \left(I - \frac{G(w_k) G(w_k)^T}{\|G(w_k)\|^2} \right) s_{k-1}, & \text{if } k > 0, \end{cases} \tag{10}$$

where $s_k = \alpha_k d_k$,

$$\begin{aligned} \delta_k := \delta_k^{EMLSCD} &:= \max \left\{ 0, \; \min \left\{ \delta_k^{LS}, \delta_k^{CD} \right\} \right\} \\ &:= \max \left\{ 0, \; \min \left\{ \frac{y_{k-1}^T G(w_k)}{-G(w_{k-1})^T d_{k-1}}, \frac{\|G(w_k)\|^2}{-G(w_{k-1})^T d_{k-1}} \right\} \right\}, \end{aligned}$$

and $y_{k-1} = G(w_k) - G(w_{k-1})$.

Step 1. *Determine the step-length $\alpha_k = \kappa \xi^i$ for $i = 0, 1, 2, \cdots$, satisfying the following line-search*

$$-G(w_k + \alpha_k d_k)^T d_k \geq \gamma \alpha_k \|d_k\|^2. \tag{11}$$

Step 2. *Compute*

$$z_k = w_k + \alpha_k d_k. \tag{12}$$

Step 3. *If $z_k \in E$ and $\|G(z_k)\| \leq Tol$, stop. Otherwise, compute the next iterate by*

$$w_{k+1} = P_E[w_k - \varrho_k G(z_k)], \tag{13}$$

where

$$\varrho_k = \frac{G(z_k)^T (w_k - z_k)}{\|G(z_k)\|^2}. \tag{14}$$

Step 4. *Finally we set $k := k + 1$ and return to step 0.*

Lemma 2.2. *Let δ_k be any CG parameter. Then, the search direction d_k defined by* (10) *satisfies*

$$G(w_k)^T d_k = -||G(w_k)||^2, \quad \forall k \geq 0. \tag{15}$$

Proof. For $k = 0$, multiplying both sides of (10) by $G(w_0)^T$, we have

$$G(w_0)^T d_0 = -||G(w_0)||^2.$$

Also for $k \geq 1$, multiplying both sides of (10) by $G(w_k)^T$, we get

$$\begin{aligned} G(w_k)^T d_k &= -||G(w_k)||^2 + \delta_k G(w_k)^T s_{k-1} - \frac{\delta_k ||G(w_k)||^2 G(w_k)^T s_{k-1}}{||G(w_k)||^2} \\ &= -||G(w_k)||^2 + \delta_k G(w_k)^T s_{k-1} - \delta_k G(w_k)^T s_{k-1} \\ &= -||G(w_k)||^2. \end{aligned}$$

□

3. Convergence Analysis

In order to establish the global convergence of the DF-MLSCD method for solving (7), we need the following assumption.

Assumption 3.1.

(A1) The mapping G is Lipschitz continuous, that is, there exists a constant $L > 0$ such that

$$||G(w) - G(y)|| \leq L||w - y|| \; \forall w, y \in \mathbf{R}^{2n}. \tag{16}$$

(A2) The mapping G is monotone. That is,

$$(G(w) - G(y))^T (w - y) \geq 0, \quad \forall w, y \in \mathbf{R}^{2n}. \tag{17}$$

Lemma 3.2. *Let $\{z_k\}$ and $\{w_k\}$ be sequences generated by (12) and (13) in Algorithm 2.1. Using (16) and (17), the following statements hold*

(i) *$\{w_k\}$ and $\{z_k\}$ are bounded.*
(ii) $\lim_{k\to\infty} ||z_k - w_k|| = 0$
(iii) $\lim_{k\to\infty} ||w_{k+1} - w_k|| = 0$

Proof. (i) Since G is monotone from (17), for any solution w_* of problem (7), we have

$$\begin{aligned} G(z_k)^T (w_k - w_*) &= G(z_k)^T (w_k - z_k) + G(z_k)^T (z_k - w_*) \\ &\geq G(z_k)^T (w_k - z_k) + G(w_*)^T (z_k - w_*) \\ &= G(z_k)^T (w_k - z_k) && (18) \\ &\geq \gamma \alpha_k^2 \cdot ||d_k||^2 && (19) \\ &= \gamma ||w_k - z_k||^2 \geq 0. && (20) \end{aligned}$$

Note that, inequality (19) is obtained from the line search.

From (8), it holds that ,

$$\begin{aligned}\|w_{k+1}-w_*\| &= \|P_E[w_k-\varrho_k G(z_k)]-w_*\|^2\\ &\le \|w_k-\varrho_k G(z_k)-w_*\|^2\\ &= \|w_k-w_*\|^2-2\varrho_k G(z_k)^T(w_k-w_*)+\varrho_k^2\|G(z_k)\|^2\\ &\le \|w_k-w_*\|^2-2\varrho_k G(z_k)^T(w_k-z_k)+\varrho_k^2\|G(z_k)\|^2 \quad (21)\\ &= \|w_k-w_*\|^2-\frac{(G(z_k)^T(w_k-z_k))^2}{\|G(z_k)\|^2}\\ &\le \|w_k-w_*\|^2-\frac{\gamma^2\|w_k-z_k\|^4}{\|G(z_k)\|^2} \quad (22)\\ &\le \|w_k-w_*\|^2, \quad (23)\end{aligned}$$

where (21) and (22) follows from (18) and (20), respectively.
Also from (23), we have

$$\|w_{k+1}-w_*\|^2 \le \|w_k-w_*\|^2,\ \forall k\ge 0,$$

which shows that the sequence $\{w_k\}$ is bounded. Furthermore, by (16), we have

$$\|G(w_k)\| = \|G(w_k)-G(w_*)\| \le L\|w_k-w_*\| \le L\|w_0-w_*\|.$$

Letting $M = L\|w_0-w_*\|$, we get

$$\|G(w_k)\| \le M. \quad (24)$$

By Assumption (A2) and Cauchy-Schwarz inequality, we have that

$$\begin{aligned}G(z_k)^T(w_k-z_k) &= (G(z_k)-G(w_k))^T(w_k-z_k)+G(w_k)^T(w_k-z_k)\\ &\le \|G(w_k)\|\|w_k-z_k\|\end{aligned}$$

Therefore,

$$\|G(w_k)\|\|w_k-z_k\| \ge G(z_k)^T(w_k-z_k) \ge \gamma\|w_k-z_k\|^2,$$

where the last inequality can be implied from (20). Thus, it is easy to obtain that

$$\gamma\|w_k-z_k\| \le \|G(w_k)\| \le M,$$

which implies that $\{z_k\}$ is bounded.

(ii) Using the continuity of G, we know that there exist a constant $M_1 > 0$ such that

$$\|G(z_k)\| \le M_1\ \forall k\ge 0.$$

It follows from (22) that

$$\frac{\gamma^2\|w_k-z_k\|^4}{\|G(z_k)\|^2} \le \|w_k-w_*\|^2-\|w_{k+1}-w_*\|^2. \quad (25)$$

Adding (25) for $k \ge 0$, we obtain

$$\frac{\gamma^2}{M_1^2}\sum_{k=0}^{\infty}\|w_k-z_k\|^4 \le \sum_{k=0}^{\infty}(\|w_k-w_*\|^2-\|w_{k+1}-w_*\|^2) \le \|w_0-w_*\|^2 < \infty. \quad (26)$$

Inequality (26) implies that

$$\lim_{k\to\infty}\|w_k-z_k\| = 0. \quad (27)$$

Hence, second assertion holds.

(iii) From (8) we have

$$\|w_{k+1} - w_k\| = \|P_E[w_k - \varrho_k G(z_k)] - w_k\| \leq \|\varrho_k G(z_k)\|.$$

Then by (14) and Cauchy-Schwartz inequality, we obtain

$$\|w_{k+1} - w_k\| \leq \|w_k - z_k\|,$$

which shows that the third assertion holds.
□

Lemma 3.3. *Let the search direction sequence $\{d_k\}$ be obtained by (10) in Algorithm 2.1. If there exist positive constant κ_0 such that*

$$\|G(w_k)\| \geq \kappa_0 \ \forall k \geq 0, \tag{28}$$

holds, then the sequence $\{d_k\}$ is bounded.

Proof. First, notice that,

$$\delta_k := \delta_k^{EMLSCD} := \max\left\{0, \quad \min\left\{\frac{y_{k-1}^T G(w_k)}{-G(w_{k-1})^T d_{k-1}}, \frac{\|G(w_k)\|^2}{-G(w_{k-1})^T d_{k-1}}\right\}\right\} \leq \frac{\|G(w_k)\|^2}{|-G(w_{k-1})^T d_{k-1}|}.$$

Thus,

$$|\delta_k| \leq \frac{\|G(w_k)\|^2}{|-G(w_{k-1})^T d_{k-1}|} \leq \frac{M^2}{\kappa_0}.$$

Then from (10), it holds that

$$\begin{aligned}\|d_k\| &\leq \|G(w_k)\| + 2|\delta_k| \cdot \|s_{k-1}\| \\ &\leq M + 2\frac{M^2}{\kappa_0}\alpha_{k-1}\|d_{k-1}\|,\end{aligned}$$

for all $k \in \mathbb{N}$. Having in view of (27), it follows that for every $\kappa_1 > 0$ there exist κ_0 such that $\alpha_{k-1}\|d_{k-1}\| < \kappa_1$ for every $k > \kappa_0$. Choosing $\kappa_1 = \kappa_0$ and $J = \max\{\|d_0\|, \|d_1\|, \cdots \|d_{k_0}\|, J_1\}$ where $J_1 = M(1 + 2M)$, it holds

$$\|d_k\| \leq J \ \forall k \in \mathbb{N}.$$

□

Theorem 3.4. *Let the sequence $\{z_k\}$ and $\{w_k\}$ be generated by (12) and (13) in Algorithm 2.1. From (16) and (17), we have*

$$\liminf_{k\to\infty} \|G(w_k)\| = 0. \tag{29}$$

Proof. Suppose the conclusion (29) does not hold, then (28) holds which implies that the sequence $\{d_k\}$ is bounded. That is, there exist a positive constant Λ such that

$$\|d_k\| \leq \Lambda, \ \forall k \geq 0.$$

From (21), $\frac{\alpha_k}{\xi}$ does not satisfy (11). Thus, we have

$$-G(w_k + \frac{\alpha_k}{\xi}d_k)^T d_k < \gamma\frac{\alpha_k}{\xi}\|d_k\|^2.$$

It follows from Lemma 2.2 that

$$\begin{aligned} \|G(w_k)\|^2 &\leq -G(w_k)^T d_k \\ &\leq (G(w_k + \frac{\alpha_k}{\xi} d_k) - G(w_k))^T d_k - G(w_k + \frac{\alpha_k}{\xi} d_k)^T d_k \\ &\leq L\frac{\alpha_k}{\xi}\|d_k\|^2 + \gamma\frac{\alpha_k}{\xi}\|d_k\|^2 \\ &\leq \frac{\alpha_k}{\xi}(L+\gamma)\|d_k\|^2. \end{aligned} \tag{30}$$

Inequality (30) is obtained by using (16) and the Cauchy-Schwartz inequality. Therefore, it holds that

$$\alpha_k\|d_k\|^2 \geq \frac{\xi\|G(w_k)\|^2}{(L+\gamma)} \geq \frac{\xi\kappa_0^2}{(L+\gamma)} > 0,$$

which contradicts (27). Thus, (29) holds. □

4. Numerical experiments

In this section, two types of experiments are carried out. Algorithm 2.1 is tested on signal and image recovery problems. All the algorithms are coded in MATLAB and run on an HP PC (CPU 2.4 GHz, RAM 8.0GB) with Windows 10 operating system.

- **Algo.1**: Algo.1, the new method (Algorithm 2.1).
- **Algo.2**: CGD, the method proposed by Xiao et al. [10].
- **Algo.3**: PCG, the method proposed in [43].
- **Algo.4 and Algo.5**: Algorithm 4.1a and Algorithm 4.1b proposed in [44].

4.1. Recovery of sparse signals

Here, our main goal is to reconstruct a length n sparse signal from k observation. To validate the effectiveness of Algo.1 in recovering sparse signal in compressive sensing, Algo.1 is compared with four different algorithms which include Algo.2, Algo.3, Algo.4 and Algo.5. The numerical results are reported in Table 1, where the quality of the restoration is assessed by the mean of squared error (MSE) calculated according to

$$MSE := \frac{1}{n}\|w - \bar{w}\|^2,$$

where w is the original signal and $\bar{w}$ is the restored signal. In this experiment, a random Gaussian matrix A is generated using the command rand(n,k) in MATLAB where the original signal contains 2^6 randomly non-zero elements and the selected size of the signal is chosen with $n = 2^{12}$ and $k = 2^{10}$. Furthermore, noise is appropriately added to the measurement, that is

$$b = Aw + \delta$$

where δ is the Gaussian noise distributed as $N(0, 10^{-4})$. The initial process starts at $w_0 = A^T b$ where the merit function used is given by $f(w) = \frac{1}{2}\|b - Aw\|^2 + \tau\|w\|$. The process terminates when

$$\frac{|f_k - f_{k-1}|}{|f_{k-1}|} < 10^{-5},$$

where f_k denotes the function value at w_k. Note, for this test we only observe the convergence behavior of each method to obtain a similar accuracy solution. The parameter τ in the merit function is selected as $\tau = 0.005\|A^T b\|_\infty$ which is inline with the suggestion given in [45].

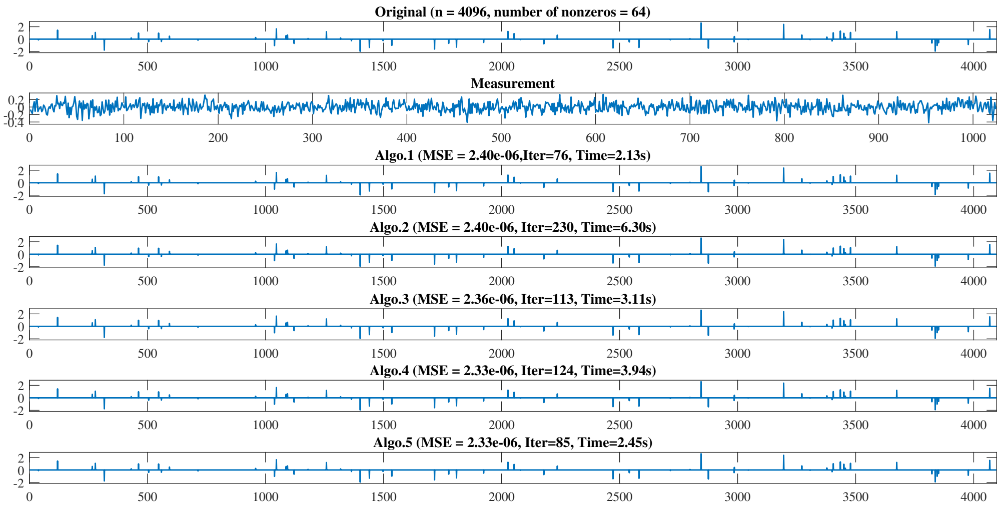

Figure 1: Reconstruction of the sparse signal by the various methods

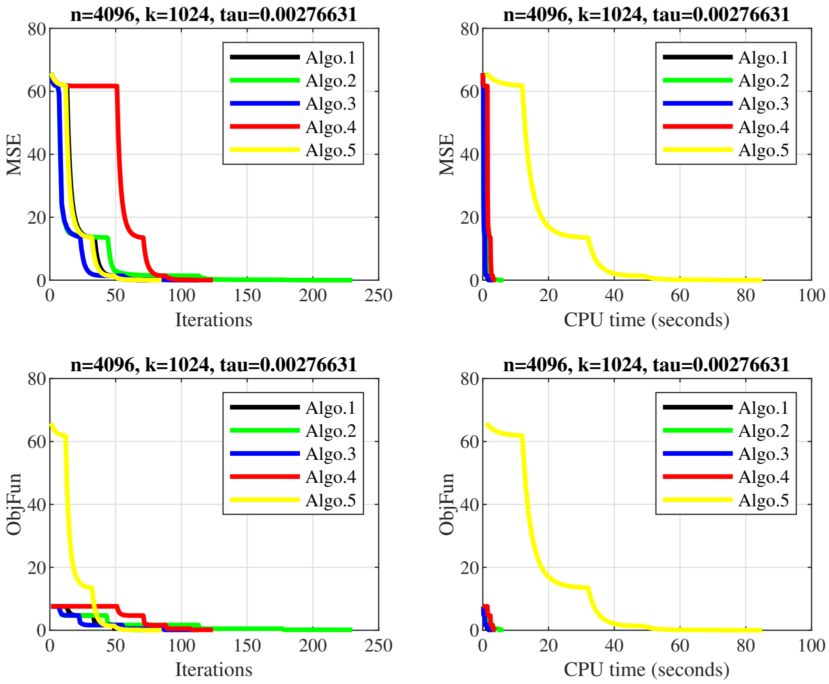

Figure 2: Comparison results of the various algorithms. From left to right: the changed trend of MSE goes along with the number of iterations or CPU time in seconds, and the changed trend of the objective function values accompany the number of iterations or CPU time in seconds.

In Figure 2, we give a visual illustration of the performance of each method relative to their convergence behavior from the view of merit function values and relative error as the iteration number and computing time increases. To demonstrate the effectiveness of the Algo.1, the experiment is carried out ten times using different noise samples. The detail of the test instances is reported in Table 1. Figure 1 is a visual illustration of the results of the reconstruction of the sparse signal. From Table 1, we can observe that the disturbed signal is restored almost exactly by the five methods, this is reflected by their MSE. However, Algo.1 performs better than the compared methods in terms of iterations and CPU time.

Note. The parameters used for the implementation of our algorithm for the signal recovery problem are as follows: $\gamma = 0.0001,\ \kappa = 1,\ \xi = 0.9$.

Table 1: Result of the sparse signal reconstruction by the various algorithms

	Algo.1			Algo.2			Algo.3			Algo.4			Algo.5		
SN	ITER	CPU	MSE	ITER	CPU	MSE	ITER	CPU	MSE	ITER	CPU	MSE	ITER	CPU	MSE
1	79	2.77	1.57E-06	129	3.72	1.54E-06	234	6.56	1.59E-06	119	3.27	1.56E-06	91	2.72	1.54E-06
2	75	2.13	1.47E-06	127	3.48	1.43E-06	181	5.25	8.40E-06	110	4	1.45E-06	89	2.33	1.43E-06
3	79	2.34	3.42E-06	125	3.42	3.37E-06	243	6.77	3.46E-06	128	3.52	3.41E-06	86	2.41	3.37E-06
4	73	1.97	3.32E-06	116	2.97	3.25E-06	224	5.42	3.30E-06	120	2.97	3.27E-06	77	1.88	3.25E-06
5	74	1.97	1.76E-06	125	3.41	1.73E-06	225	5.97	1.78E-06	106	3.17	1.75E-06	87	2.22	1.73E-06
6	75	2.14	2.20E-06	124	3.31	2.16E-06	221	5.67	2.21E-06	115	2.91	2.18E-06	85	2.14	2.16E-06
7	77	2.25	2.03E-06	129	3.97	2.01E-06	207	5.97	5.96E-06	120	3.23	2.03E-06	89	2.83	2.01E-06
8	78	1.95	3.27E-06	123	3.09	3.22E-06	232	5.89	3.32E-06	122	2.86	3.25E-06	85	2.11	3.22E-06
9	77	2.16	3.11E-06	119	3.34	3.04E-06	199	5.36	3.11E-06	117	3.17	3.07E-06	80	2.19	3.04E-06
10	69	2.06	2.10E-06	110	2.97	2.08E-06	224	6.09	2.12E-06	115	3.08	2.10E-06	71	1.83	2.08E-06
Average	75.6	2.174	2.43E-06	122.7	3.368	2.38E-06	219	5.895	3.52E-06	117.2	3.218	2.41E-06	84	2.266	2.38E-06

4.2. Image restoration

We present experimental results demonstrating the performance of the proposed algorithm and comparing it with some related methods (Algo.2, [10] and Algo.6, [46]). The test images for the experiments are; Tiffany (512×512), Lena (512×512), Barbara (720×576), Malamute (1616×1080), Mars (1280×1024), Abdul (800×800) and Poom (720×720) degraded by Gaussian blur and Gaussian noise.

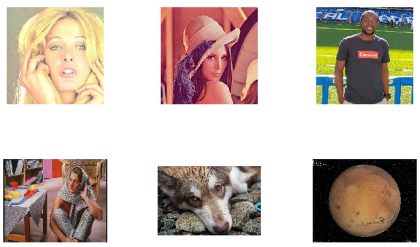

Figure 3: The original test images

All classical test images are obtained `http://hlevkin.com/06testimages.htm`. In this experiment, a matrix A (partial DWT matrix) whose k rows are randomly selected from the $m \times m$ DWT matrix. This type of matrix A requires no storage and helps in speeding up the matrix-vector multiplications involving A and A^T. Therefore, making it possible to test large-size images without storing any matrix. For fairness in comparing the algorithms, the iterative process of all algorithms start at $w_0 = A^T b$ and terminates when the relative change between successive iterates falls below 10^{-5}. The quality of the restored images are evaluated in terms of Signal-to-ratio (SNR), Peak Signal to noise ratio (PSNR) [47]) and Structural similarity index (SSIM [48]).

For comparison, we present restoration results obtained by the various methods in restoring the degraded images. Experimental results from Table 2 indicates that the quality of the restored images by Algo.1 is better than the restored image by Algo.2 and Algo.6. Larger PSNR, SNR and SSIM value indicate that the restored images by Algo.1 are closer to the original ones than those by Algo.2 and Algo.6 in almost all cases.

Note. The parameters used for the implementation of our algorithm for the image restoration problem are as follows: $\gamma = 0.0001,\ \kappa = 0.5,\ \xi = 0.05$.

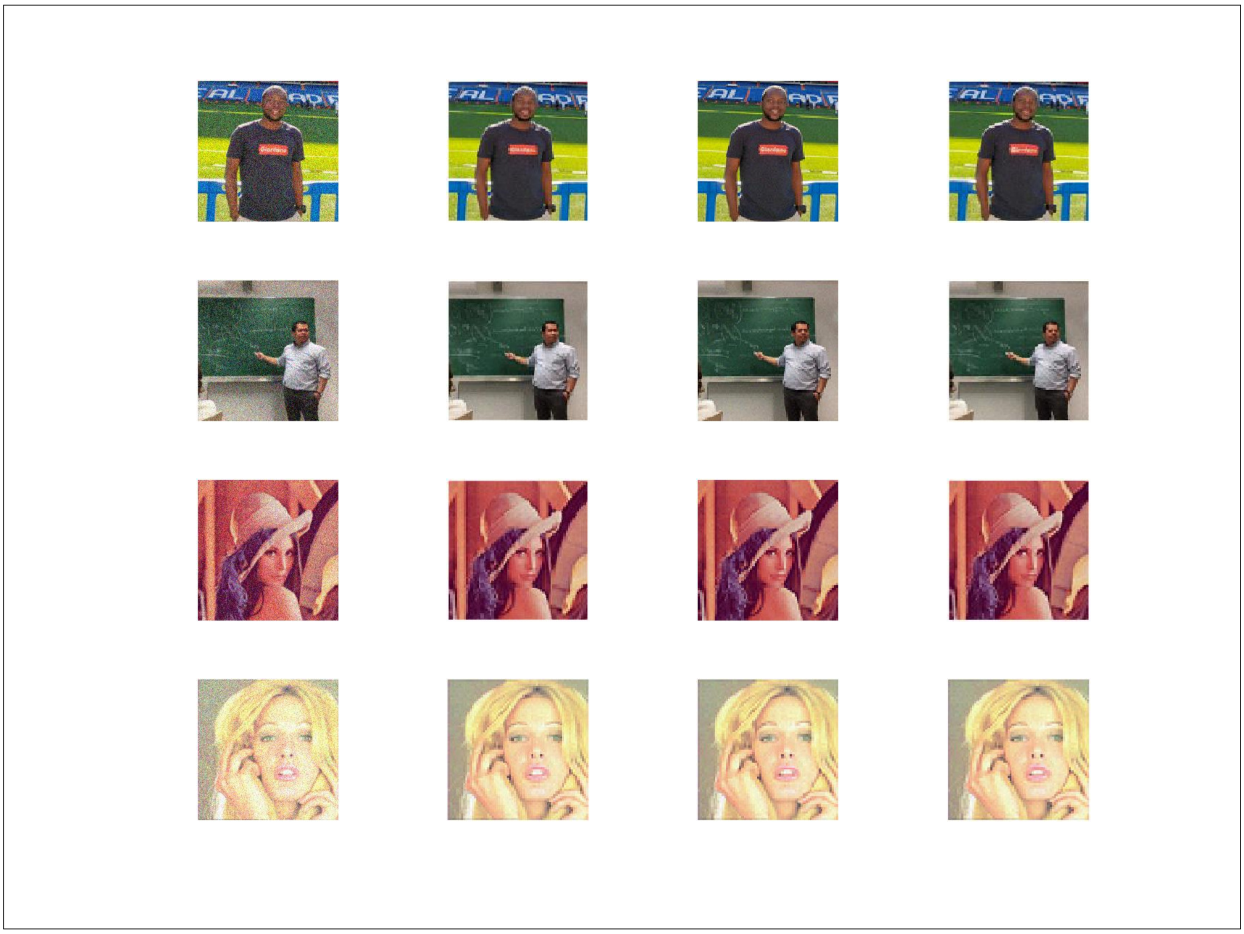

Figure 4: The image restoration of some of the test images: blurred and noisy image (10% noise) (left), image restored by Algo.1 (middle left), Algo.2 (middle right) and Algo.6 (right)

Table 2: Test results on image restoration

		Algo.1			Algo.2			Algo.6		
Noise	Images	SNR	PSNR	SSIM	SNR	PSNR	SSIM	SNR	PSNR	SSIM
10%	Tiffany	20.95	22.78	0.9134	20.93	22.76	0.9126	20.87	22.70	0.9114
	Lena	16.75	22.08	0.9128	16.70	22.04	0.9118	16.62	21.95	0.9101
	Barbara	13.64	20.06	0.6283	13.62	20.04	0.6269	13.56	19.98	0.6238
	Malute	15.32	21.74	0.5842	15.30	21.72	0.5823	15.25	21.67	0.5799
	Mars	14.69	24.58	0.7885	14.68	24.57	0.7883	14.65	24.54	0.7873
	Airoplane	18.41	21.10	0.6789	18.36	21.05	0.6738	18.23	20.92	0.6682
	Poom	16.73	22.86	0.7751	16.70	22.83	0.7725	16.64	22.77	0.7698
	Abdul	14.27	20.85	0.8159	14.22	20.80	0.8128	14.10	20.68	0.8077
	Average	16.35	22.01	0.7621	16.31	21.98	0.7601	16.24	21.90	0.7573
20%	Tiffany	20.34	22.17	0.8838	20.29	22.13	0.8817	20.21	22.05	0.8792
	Lena	16.25	21.58	0.8977	16.18	21.52	0.8961	16.04	21.38	0.8936
	Barbara	13.32	19.74	0.5990	13.28	19.70	0.5960	13.22	19.64	0.5920
	Malute	14.84	21.26	0.5191	14.80	21.22	0.5157	14.73	21.15	0.5103
	Mars	14.21	24.10	0.7729	14.18	24.07	0.7722	14.14	24.03	0.7710
	Airoplane	18.00	20.68	0.5598	17.92	20.61	0.5519	17.74	20.42	0.5422
	Poom	16.09	22.22	0.6682	16.04	22.17	0.6630	15.97	22.10	0.6574
	Abdul	13.88	20.46	0.7468	13.81	20.39	0.7415	13.68	20.26	0.7333
	Average	15.87	21.53	0.7059	15.81	21.48	0.7023	15.72	21.38	0.6974

Conclusion

In this paper, we have proposed a derivative-free gradient projection algorithm for solving the ℓ_1-norm regularized problems for reconstructing sparse signal and image restoration in compressive sensing. The method combines the line search method and the MLSCD conjugate gradient method. Furthermore, we have shown that the proposed derivative-free algorithm converges globally. We have presented numerical experiments on the recovery of sparse signal and image restoration. These experiments illustrate clearly the effectiveness of our approach in reconstructing sparse signal and image in compressive sensing compared to related methods.

Acknowledgments

We are grateful to the anonymous referees and editor for their useful comments, which have made the paper clearer and more comprehensive than the earlier version. The first author was supported by the "Petchra Pra Jom Klao Ph.D. Research Scholarship from King Mongkut's University of Technology Thonburi" (Grant no. 16/2561). The authors also acknowledge the financial support provided by the Center of Excellence in Theoretical and Computational Science (TaCS-CoE), KMUTT. Moreover, this research project is supported by Thailand Science Research and Innovation (TSRI) Basic Research Fund: Fiscal year 2021 under project number 64A306000005. The first and the third author acknowledge with thanks, the Department of Mathematics and Applied Mathematics at the Sefako Makgatho Health Sciences University.

References

[1] Marvasti F, Amini A, Haddadi F, Soltanolkotabi M, Khalaj BH, Aldroubi A, et al. A unified approach to sparse signal processing. EURASIP journal on advances in signal processing. 2012;2012(1):1–45.

[2] Donoho DL. For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution. Communications on Pure and Applied Mathematics. 2006;59(6):797–829. Available from: `https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.20132`.

[3] Donoho DL, et al. Compressed sensing. IEEE Transactions on information theory. 2006;52(4):1289–1306.

[4] Candes E, Romberg J. Sparsity and incoherence in compressive sampling. Inverse problems. 2007;23(3):969.
[5] Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences. 2004;57(11):1413–1457.
[6] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM journal on imaging sciences. 2009;2(1):183–202.
[7] Hale ET, Yin W, Zhang Y. A fixed-point continuation method for l1-regularized minimization with applications to compressed sensing. CAAM TR07-07, Rice University. 2007;43:44.
[8] Figueiredo MA, Nowak RD, Wright SJ. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. IEEE Journal of selected topics in signal processing. 2007;1(4):586–597.
[9] Xiao Y, Wang Q, Hu Q. Non-smooth equations based method for ℓ_1-norm problems with applications to compressed sensing. Nonlinear Analysis: Theory, Methods & Applications. 2011;74(11):3570–3577.
[10] Xiao Y, Zhu H. A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing. Journal of Mathematical Analysis and Applications. 2013;405(1):310–319.
[11] Pang JS. Inexact Newton methods for the nonlinear complementarity problem. Mathematical Programming. 1986;36(1):54–71.
[12] Tong X, Zhou S. A smoothing projected Newton-type method for semismooth equations with bound constraints. Journal of Industrial & Management Optimization. 2005;1(2):235–250.
[13] Wang C, Wang Y, Xu C. A projection method for a system of nonlinear monotone equations with convex constraints. Mathematical Methods of Operations Research. 2007 Aug;66(1):33–46. Available from: `https://doi.org/10.1007/s00186-006-0140-y`.
[14] Wang C, Wang Y. A superlinearly convergent projection method for constrained systems of nonlinear equations. Journal of Global Optimization. 2009 Jun;44(2):283–296.
[15] Fan J. On the Levenberg-Marquardt methods for convex constrained nonlinear equations. Journal of Industrial & Management Optimization. 2013;9(1):227–241.
[16] Jia Cx, Zhu Dt. Projected gradient trust-region method for solving nonlinear systems with convex constraints. Applied Mathematics-A Journal of Chinese Universities. 2011;26(1):57–69.
[17] Solodov MV, Svaiter BF. A globally convergent inexact Newton method for systems of monotone equations. In: Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods. Springer; 1998. p. 355–369.
[18] Yuan G, Lu J, Wang Z. The PRP conjugate gradient algorithm with a modified WWP line search and its application in the image restoration problems. Applied Numerical Mathematics. 2020;152:1–11.
[19] Yuan G, Wei Z, Yang Y. The global convergence of the Polak–Ribière–Polyak conjugate gradient algorithm under inexact line search for nonconvex functions. Journal of Computational and Applied Mathematics. 2019;362:262–275.
[20] Yuan G, Wang X, Sheng Z. The Projection Technique for Two Open Problems of Unconstrained Optimization Problems. Journal of Optimization Theory and Applications. 2020;186(2):590–619.
[21] Zhang M, Zhou Y, Wang S. A Modified Nonlinear Conjugate Gradient Method with the Armijo Line Search and Its Application. Mathematical Problems in Engineering. 2020;2020.
[22] Ibrahim AH, Garba AI, Usman H, Abubakar J, Abubakar AB. Derivative-free RMIL conjugate gradient method for convex constrained equations. Thai Journal of Mathematics. 2019;18(1):212–232.
[23] Abubakar AB, Rilwan J, Yimer SE, Ibrahim AH, Ahmed I. Spectral three-term conjugate descent method for solving nonlinear monotone equations with convex constraints. Thai Journal of Mathematics. 2020;18(1):501–517.
[24] Ibrahim AH, Kumam P, Abubakar AB, Jirakitpuwapat W, Abubakar J. A hybrid conjugate gradient algorithm for constrained monotone equations with application in compressive sensing. Heliyon. 2020;6(3):e03466.
[25] Ibrahim AH, Kumam P, Abubakar AB, Abubakar J, Muhammad AB. Least-square-based three-term conjugate gradient projection method for ℓ_1-norm problems with application to compressed sensing. Mathematics. 2020;8(4):602.
[26] Ibrahim AH, Kumam P, Abubakar AB, Yusuf UB, Rilwan J. Derivative-free conjugate residual algorithms for convex constraints nonlinear monotone equations and signal recovery. Journal of Nonlinear and Convex Analysis. 2020;21(9):1959–1972.
[27] Abubakar A, Ibrahim A, Muhammad A, Tammer C. A modified descent Dai-Yuan conjugate gradient method for constraint nonlinear monotone operator equations. Applied Analysis and Optimization. 2020;4:1–24.
[28] Abubakar AB, Kumam P, Ibrahim AH, Rilwan J. Derivative-free HS-DY-type method for solving nonlinear equations and image restoration. Heliyon. 2020;6(11):e05400.
[29] Ibrahim AH, Kumam P, Kumam W. A family of derivative-free conjugate gradient methods for constrained nonlinear equations and image restoration. IEEE Access. 2020;8:162714–162729.
[30] Ibrahim AH, Kumam P, Abubakar AB, Yusuf UB, Yimer SE, Aremu KO. An efficient gradient-free projection algorithm for constrained nonlinear equations and image restoration. AIMS Mathematics. 2020;6(1):235.
[31] Abubakar AB, Muangchoo K, Ibrahim AH, Muhammad AB, Jolaoso LO, Aremu KO. A New Three-Term Hestenes-Stiefel Type Method for Nonlinear Monotone Operator Equations and Image Restoration. IEEE Access. 2021;9:18262–18277.
[32] Ibrahima AH, Muangchoob K, Mohamedc NS, Abubakard AB. Derivative-free SMR conjugate gradient method for con-straint nonlinear equations. Journal of Mathematics and Computer Science. 2022;24(2):147–164.
[33] Abubakar AB, Muangchoo K, Ibrahim AH, Abubakar J, Rano SA. FR-type algorithm for finding approximate solutions to nonlinear monotone operator equations. Arabian Journal of Mathematics. 2021;p. 1–10.
[34] Abubakar AB, Kumam P, Mohammad H, Ibrahim AH. PRP-like algorithm for monotone operator equations. Japan Journal of Industrial and Applied Mathematics. 2021;p. 1–18.
[35] Ibrahim AH, Muangchoo K, Abubakar AB, Adedokun AD, Mohammed H. Spectral Conjugate Gradient Like Method for Signal Reconstruction. Thai Journal of Mathematics. 2020;18(4):2013–2022.
[36] Ibrahim AH, Kumam P. Re-modified derivative-free iterative method for nonlinear monotone equations with convex constraints. Ain Shams Engineering Journal. 2021;.

[37] Mohammad H. Barzilai-Borwein-like method for solving large-scale non-linear systems of equations. Journal of the Nigerian Mathematical Society. 2017;36(1):71–83.
[38] Abubakar AB, Kumam P. A descent Dai-Liao conjugate gradient method for nonlinear equations. Numerical Algorithms. 2019;81(1):197–210.
[39] Abubakar AB, Kumam P. An improved three-term derivative-free method for solving nonlinear equations. Computational and Applied Mathematics. 2018;37(5):6760–6773.
[40] Abubakar AB, Kumam P, Mohammad H. A note on the spectral gradient projection method for nonlinear monotone equations with applications. Computational and Applied Mathematics. 2020;39(2):1–35.
[41] Mohammad H, Abubakar AB. A descent derivative-free algorithm for nonlinear monotone equations with convex constraints. RAIRO-Operations Research. 2020;54(2):489–505.
[42] Stanimirović PS, Ivanov B, Djordjević S, Brajević I. New Hybrid Conjugate Gradient and Broyden–Fletcher–Goldfarb–Shanno Conjugate Gradient Methods. Journal of Optimization Theory and Applications. 2018;178(3):860–884.
[43] Liu J, Li S. A projection method for convex constrained monotone nonlinear equations with applications. Computers & Mathematics with Applications. 2015;70(10):2442–2453.
[44] Gao P, He C, Liu Y. An adaptive family of projection methods for constrained monotone nonlinear equations with applications. Applied Mathematics and Computation. 2019;359:1–16.
[45] Kim S, Koh K, Lustig M, Boyd S, Gorinevsky D. A method for large-scale. 1-regularized least squares problems with applications in signal processing and statistics. 2007;50.
[46] Abubakar AB, Kumam P, Mohammad H, Awwal AM, Sitthithakerngkiet K. A Modified Fletcher–Reeves Conjugate Gradient Method for Monotone Nonlinear Equations with Some Applications. Mathematics. 2019;7(8):745.
[47] Bovik AC. Handbook of image and video processing. Academic press; 2010.
[48] Lajevardi SM. Structural similarity classifier for facial expression recognition. Signal, Image and Video Processing. 2014;8(6):1103–1110.