Research Article

A Modified Nonlinear Conjugate Gradient Method with the Armijo Line Search and Its Application

Mengxiang Zhang,¹ Yingjie Zhou,¹ and Songhua Wang²

¹College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, China

²School of Mathematics and Statistics, Baise University, 533000 Baise, Guangxi, China

Correspondence should be addressed to Songhua Wang; wsh6600789@126.com

Received 19 July 2020; Accepted 10 August 2020; Published 28 August 2020

Guest Editor: Maojun Zhang

Copyright © 2020 Mengxiang Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this article, a modified Polak-Ribière-Polyak (PRP) conjugate gradient method is proposed for image restoration. The presented method can generate sufficient descent directions without any line search conditions. Under some mild conditions, this method is globally convergent with the Armijo line search. Moreover, the linear convergence rate of the modified PRP method is established. The experimental results of unconstrained optimization, image restoration, and compressive sensing show that the proposed method is promising and competitive with other conjugate gradient methods.

1. Introduction

Consider the following unconstrained optimization problem:

$$\min_{x \in \mathbb{R}^n} f(x), \quad (1)$$

where $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously differentiable function. The iteration formula is determined by

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \dots, \quad (2)$$

where α_k is stepsize obtained by some kind of line search and d_k is search direction.

It is well known that there are many methods ([1–4] etc.) for solving optimization problem (1), where the conjugate gradient method is a powerful method because of its simplicity, and this method can avoid the computation and storage of some matrices associated with the Hessian of objective functions; then its memory requirements are very low. Large-scale calculations are usually required in the process of image processing; therefore, the conjugate gradient algorithm has excellent application prospects in this aspect. The search direction d_k of the conjugate gradient methods is given by

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

where $\beta_k \in \mathbb{R}$ is a scalar and $g_k = g(x_k)$ is the gradient of objective function $f(x)$ at the point x_k . The choice of β_k determines different conjugate gradient methods. Classical conjugate gradient methods include PRP conjugate gradient method [5, 6], HS conjugate gradient method [7], LS conjugate gradient method [8], DY conjugate gradient method [9], FR conjugate gradient method [10], and CD conjugate gradient method [11]. The parameters β_k of these methods are specified as follows:

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

where $y_{k-1} = g_k - g_{k-1}$, and $\|\cdot\|$ stands the Euclidean norm. There has been much research on convergence properties of these methods ([12–15] etc.) and applications ([16–23] etc.).

In this paper, the focus is mainly on the PRP method. The PRP method is generally believed to have the best numerical performance in classical conjugate gradient methods, but there is much room for discussion in terms of convergence. When the exact line search was used, the global convergence of the PRP conjugate gradient method has been proved by Polak and Ribière [5] for convex objective functions. However, Powell [24] proposed a counter example that the PRP conjugate gradient method may fail for nonconvex functions even if the exact line search is used. Gilbert and Nocedal [14] showed a modified PRP conjugate gradient method that the modified PRP method is globally convergent if β_k^{PRP} is restricted to be not less than zero and α_k is determined by a line search step satisfying the sufficient descent condition:

$$-g_k^T d_k \geq c \|g_k\|^2, \quad (5)$$

in addition to the following weak Wolfe-Powell conditions, i.e.,

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f_k + \sigma \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k &\geq \varphi g_k^T d_k, \end{aligned} \quad (6)$$

where $0 < \sigma < \varphi < 1$. Meanwhile, Gilbert and Nocedal [14] pointed out that even if the objective function is guaranteed to be uniformly convex, β_k may be negative. With the strong Wolfe-Powell line search, Dai [25] gave an example showing that even though the objective function is uniformly convex, the PRP method cannot guarantee that the search direction is always the descent direction.

Through the above observations and [14, 24, 26–28], the sufficient descent condition (5) and the condition $\beta_k \geq 0$ play key roles in establishing the global convergence of the conjugate gradient methods. However, in the case where Armijo line search or Wolfe line search is used, the descent property of d_k determined by (3) is in general not guaranteed.

In order to obtain the generated sufficient descent direction, Hager and Zhang [29] showed a new conjugate gradient method (CG-C) obtained by modifying the HS method, which generates sufficient descent directions at each step, without relying on any line search. The scalar β_k in CG-C method is determined by

$$\beta_k^{\text{NHS}} = \max\{\beta_k^N, \mu_k\}, \quad (7)$$

where

$$\begin{aligned} \beta_k^N &= \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} - 2 \frac{\|y_{k-1}\|^2}{(d_{k-1}^T y_{k-1})^2} g_{k-1}^T d_{k-1}, \\ \mu_k &= \frac{1}{\|d_{k-1}\| \min\{\|g_{k-1}\|, \mu\}}, \end{aligned} \quad (8)$$

and $\mu > 0$ is a constant. With the weak Wolfe-Powell line search, Hager and Zhang [29] established a global convergence result for (7) when the objective function $f(x)$ is general nonlinear function.

Along this line, Yu and Guan [30, 31] extended the CG-C method to the PRP, LS, DY, FR, and CD methods. The parameter β_k in the modified PRP method (DPRP) is as follows:

$$\beta_k^{\text{DPRP}} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2} + \eta \frac{\|y_{k-1}\|^2}{\|g_{k-1}\|^4} g_k^T d_{k-1}, \quad (9)$$

where $\eta > (1/4)$ is a constant. The performance of the above methods is better than other conjugate gradient in practice. In addition, because Armijo line search is very simple, compared with Wolfe line search, it is more widely used in image restoration and compressed sensing problems. Therefore, it is necessary to explore the performance of the above methods under Armijo line search. However, the above methods cannot obtain global convergence with Armijo line search, i.e., to find a stepsize $\alpha_k = \max(\rho^i | i = 0, 1, 2, \dots)$ satisfying

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (10)$$

where $0 < \delta, \rho < 1$ are given constants. In order to overcome this defect, scholars have done a lot of related work ([32–34] etc). In order to overcome the defect of PRP, HS and LS conjugate gradient methods cannot globally converge with the Armijo line search, Li and Qu [33] proposed a series of modified conjugate gradient methods based on (3) and (4) as follows:

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k - \beta_k \frac{g_k^T d_{k-1}}{\|g_k\|^2} g_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \end{cases} \quad (11)$$

$$\beta_k = \frac{g_k^T y_{k-1}}{z_{k-1}}, \quad (12)$$

where z_{k-1} is a scalar to be specified. The parameter $z_{k-1} = \max\{t\|d_{k-1}\|, \|g_{k-1}\|^2\}$ in the modified PRP method (APRP). And the results of global convergence are established when the Armijo line search is used.

Motivated by the idea of [31, 33] and taking into account the excellent numerical performance, we propose a new modified PRP method. The parameter β_k is computed as

$$\beta_k^{\text{NPRP}} = \frac{g_k^T y_{k-1}}{z_{k-1}} + \eta \frac{\|y_{k-1}\|^2}{z_{k-1}^2} g_k^T d_{k-1}, \quad (13)$$

where $z_{k-1} = \max\{t\|d_{k-1}\|, \|g_{k-1}\|^2\}$, $t > 0$ and $\eta > (1/4)$ are two constants. And the search direction d_k is determined by the classical two-term conjugate gradient method (3).

In the next section, the proposed algorithm is stated. In Section 3, the properties and the convergent results of the new method are given. This method enjoys linear convergence under suitable conditions in Section 4. Numerical results and conclusion are presented in Section 5 and in Section 6, respectively.

2. Algorithm

This section will give algorithm of the new modified PRP method in association with the Armijo line search technique (10) for (1). The algorithm is stated as follows.

Remark 1. Without any line search method, the NPRP method can get sufficient descent properties.

Lemma 1. Let x_k and d_k be generated by the NPRP method. If $z_{k-1} \neq 0$, for all k , then

$$g_k^T d_k \leq -\left(1 - \frac{1}{4\eta}\right) \|g_k\|^2, \quad \eta > \frac{1}{4}. \quad (14)$$

Proof. Case (i) $k = 0$. From (3), we get

$$g_0^T d_0 = -\|g_0\|^2 \leq -\left(1 - \frac{1}{4\eta}\right) \|g_k\|^2, \quad \eta > \frac{1}{4}. \quad (15)$$

Hence (14) is satisfied.

Case (ii) $k \geq 1$. From (3) and (13), we have

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{\text{NPRP}} d_{k-1}), \\ &= -\|g_k\|^2 + \left(\frac{g_k^T y_{k-1}}{z_{k-1}} + \eta \frac{\|y_{k-1}\|^2}{z_{k-1}^2} g_k^T d_{k-1} \right) g_k^T d_{k-1}, \\ &= -\|g_k\|^2 + \left(\frac{z_{k-1} g_k^T d_{k-1} g_k^T y_{k-1} + \eta \|y_{k-1}\|^2 g_k^T d_{k-1} g_k^T d_{k-1}}{z_{k-1}^2} \right). \end{aligned} \quad (16)$$

Define

$$\begin{aligned} u_k &= \frac{1}{\sqrt{2\eta}} z_{k-1} g_k, \\ v_k &= \sqrt{2\eta} g_k^T d_{k-1} y_{k-1}. \end{aligned} \quad (17)$$

By the above equation, we get

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \left(\frac{u_k^T v_k - (1/2) \|v_k\|^2}{\|z_{k-1}\|^2} \right), \\ &= -\left(1 - \frac{1}{4\eta}\right) \|g_k\|^2 + \left(\frac{u_k^T v_k - (1/2) (\|u_k\|^2 + \|v_k\|^2)}{\|z_{k-1}\|^2} \right), \\ &\leq -\left(1 - \frac{1}{4\eta}\right) \|g_k\|^2, \end{aligned} \quad (18)$$

where $\eta > (1/4)$. Thus, (14) is satisfied. $\square$

3. Properties and Convergence Analysis

The following few suitable assumptions are often utilized in global convergence analysis for conjugate gradient algorithms.

Assumption 1

- (i) The level set $\Omega = \{x \in \mathfrak{R}^n | f(x) \leq f(x_0)\}$ is bounded.
- (ii) Function $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$ is continuously differentiable and its gradient is Lipschitz continuous on an open convex set Ω_1 containing Ω , i.e., there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L \|x - y\|, \quad \forall x, y \in \Omega_1. \quad (19)$$

It follows directly from Assumption 1 that there is a positive constant A , such that

$$\|x\| \leq A, \quad \forall x \in \Omega_1. \quad (20)$$

Lemma 2. If Assumptions 1 holds, $\{x_k\}$, $\{\alpha_k\}$, $\{d_k\}$, and $\{g_k\}$ be generated by Algorithm 1. There exists a constant $B > 1$ such that the following inequality holds for all k , then

$$\|d_k\| \leq B \|g_k\|. \quad (21)$$

Proof. When $k = 0$, from (3), we can obtain

$$\|d_k\| = \|g_k\|, \quad (22)$$

then (21) holds.

When $k \geq 1$, from (3) and (20), we have

$$\begin{aligned}
\|\beta_k^{\text{NPRP}}\| &= \left\| \frac{g_k^T y_{k-1}}{z_{k-1}} + \eta \frac{\|y_{k-1}\|^2}{z_{k-1}^2} g_k^T d_{k-1} \right\|, \\
&\leq \frac{\|g_k\| \|y_{k-1}\|}{t \|d_{k-1}\|} + \eta \frac{\|y_{k-1}\|^2}{t^2 \|d_{k-1}\|^2} \|g_k\| \|d_{k-1}\|, \\
&\leq \frac{\|g_k\| \|y_{k-1}\|}{t \|d_{k-1}\|} + \eta \frac{\|y_{k-1}\|^2}{t^2 \|d_{k-1}\|^2} \|g_k\| \|d_{k-1}\|, \\
&\leq \frac{L\alpha_k \|d_{k-1}\| \|g_k\|}{t \|d_{k-1}\|} + \eta \frac{(L\alpha_k \|d_{k-1}\|)^2}{t^2 \|d_{k-1}\|^2} \|g_k\| \|d_{k-1}\|, \\
&\leq \frac{2LB\|g_k\|}{t \|d_{k-1}\|} + \frac{\eta(2LA)^2 \|g_k\|}{t^2 \|d_{k-1}\|^2}, \\
&= \left(\frac{2LA}{t} + \frac{4\eta(LA)^2}{t^2} \right) \frac{\|g_k\|}{\|d_{k-1}\|}.
\end{aligned} \tag{23}$$

Together with (3) and the above formula implies

$$\|d_k\| \leq \|g_k\| + \|\beta_k^{\text{NPRP}}\| \|d_{k-1}\| \leq B\|g_k\|, \tag{24}$$

letting

$$B = 1 + \frac{2LA}{t} + \frac{4\eta(LA)^2}{t^2}. \tag{25}$$

Thus (21) is satisfied. $\square$

Lemma 3. *If the stepsize α_k is obtained by Armijo line search, there exists a constant $C > 0$ such that the following inequality holds for all k , then*

$$\alpha_k \geq C \frac{(-g_k^T d_k)}{\|d_k\|^2}. \tag{26}$$

Proof. Case (i) $\alpha_k = 1$. From (14) and (26), we get

$$\begin{cases} -g_k^T d_k \geq \left(1 - \frac{1}{4\eta}\right) \|g_k\|^2, \\ -g_k^T d_k \leq \frac{\|d_k\|^2}{C}. \end{cases} \tag{27}$$

Then we need to satisfy

$$\frac{\|d_k\|^2}{C} \geq \left(1 - \frac{1}{4\eta}\right) \|g_k\|^2. \tag{28}$$

From Lemma 2, we have

$$C \leq \frac{4\eta}{4\eta - 1} \frac{\|d_k\|^2}{\|g_k\|^2} \leq \frac{4\eta}{4\eta - 1} N^2, \quad N > 1. \tag{29}$$

Case (ii) $\alpha_k < 1$. By the Armijo line search condition, $\rho^{-1}\alpha_k$ does not satisfy inequality (10). This means

$$f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) > \delta \rho^{-1}\alpha_k g_k^T d_k. \tag{30}$$

By the mean-value theorem and inequality (19), there is a $p_k \in (0, 1)$ such as $x_k + p_k \rho^{-1}\alpha_k d_k \in \Omega_1$ and

$$\begin{aligned}
f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) &= \rho^{-1}\alpha_k g(x_k + p_k \rho^{-1}\alpha_k d_k)^T d_k, \\
&= \rho^{-1}\alpha_k g(x_k)^T d_k + \rho^{-1}\alpha_k (g(x_k) \\
&\quad + p_k \rho^{-1}\alpha_k d_k - g(x_k))^T d_k, \\
&\leq \rho^{-1}\alpha_k g(x_k)^T d_k + L p_k \rho^{-2}\alpha_k^2 \|d_k\|^2, \\
&\leq \rho^{-1}\alpha_k g(x_k)^T d_k + L \rho^{-2}\alpha_k^2 \|d_k\|^2.
\end{aligned} \tag{31}$$

Observing the last inequality and (30), we have

$$\alpha_k \geq \frac{(1 - \delta)\rho(-g_k^T d_k)}{L\|d_k\|^2}, \tag{32}$$

letting

$$C = \min \left\{ \frac{4\eta}{4\eta - 1}, \frac{(1 - \delta)\rho}{L} \right\}. \tag{33}$$

Thus (26) is satisfied. $\square$

Lemma 4. *Let the sequence $\{x_k\}$, $\{d_k\}$ and $\{g_k\}$ be generated by Algorithm 1, and suppose that Assumption 1 holds, then,*

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{34}$$

Proof. We have from (10) and Assumption 1 that

$$\sum_{k=0}^{\infty} -\delta \alpha_k g_k^T d_k < \infty. \tag{35}$$

Substituting (26) into the above formula, we have

$$\sum_{k=0}^{\infty} C \frac{(-g_k^T d_k)^2}{\|d_k\|^2} (\delta g_k^T d_k) \leq \sum_{k=0}^{\infty} -\delta \alpha_k g_k^T d_k < \infty, \tag{36}$$

where C and δ are two constants. Therefore, (34) holds.

Equation (34) is usually called the Zoutendijk condition [35], and it is very important for establishing global convergence. $\square$

Theorem 1. *Suppose the conditions in Assumption 1 hold. Let $\{x_k\}$ be generated by the NPRP method with Armijo line search, then either $g_k = 0$ for some k or $\lim_{k \rightarrow \infty} \|g_k\| = 0$.*

Proof. By Lemma 2 and Lemma 4, we have

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (37)$$

Together with (21), the above formula implies

$$\sum_{k=0}^{\infty} \|g_k\|^2 \leq B^2 \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty, \quad (38)$$

where $B = 1 + (2LA/t) + (4\eta(LA)^2/t^2)$.

From the above inequality, we can obtain $\lim_{k \rightarrow \infty} \|g_k\| = 0$. $\square$

Remark 2. From the proof of Theorem 1, with the Wolfe line search, the NPRP method can also establish global convergence without requirement of truncation and convexity of the objective function.

4. Linear Convergence Rate

When discussing the linear convergence rate for conjugate gradient methods, the following assumptions are often established on the basis of Assumption 1.

Assumption 2

- (i) Objective function f is twice continuously differentiable in Ω_1
- (ii) $\nabla^2 f(x)$ is a symmetric positive definite matrix in Ω_1 , i.e., there exist positive constants $M \geq m > 0$ such that

$$m^2 \|d\|^2 \leq d^T \nabla^2 f(x) d \leq M^2 \|d\|^2, \quad \forall x \in \Omega_1, d \in \mathbb{R}^n. \quad (39)$$

Lemma 5. Suppose that Assumption 2 holds and the sequence $\{x_k\}$ generated by the NPRP method converges with the Armijo line search. Then the sequence $\{x_k\}$ converges to the unique minimal point x^* .

Proof. If Assumption 2 holds, then Assumption 1 holds, and objective function f is uniformly convex. By the Armijo line search (10) and Lemma 1, it can be deduced that $\{f(x_k)\}$ is a decreasing sequence.

Therefore, by the conclusion of Theorem 1, the sequence $\{x_k\}$ converges to the unique minimal point x^* . $\square$

Lemma 6. Suppose that Assumption 2 holds. The problem (1) has a unique solution x^* and

$$\frac{1}{2} m \|x - x^*\|^2 \leq f(x) - f(x^*) \leq \frac{1}{2} M \|x - x^*\|^2, \quad \forall x \in \Omega_1, \quad (40)$$

$$m \|x - x^*\| \leq \|g(x)\| \leq M \|x - x^*\|, \quad \forall x \in \Omega_1, \quad (41)$$

$$m \|x - y\| \leq (g(x) - g(y))^T (x - y) \leq M \|x - y\|, \quad \forall x, y \in \Omega_1, \quad (42)$$

$$\|g(x) - g(y)\| \leq M \|x - y\|, \quad \forall x, y \in \Omega_1. \quad (43)$$

Theorem 2. Suppose Assumption 2 holds. Let x^* be the unique solution of problem (1) and the sequence $\{x_k\}$ be generated by the NPRP method with the Armijo line search. Then there are constants $a > 0$ and $r \in (0, 1)$ such that for all k , then

$$\|x_k - x^*\| \leq ar^k, \quad (44)$$

that is, $\{x_k\}$ converges to x^* at least R -linearly.

Proof. If Assumption 2 holds, then Assumption 1 holds. By (14), (21), and (26), it is easy to prove that

$$\alpha_k \geq C \frac{(-g_k^T d_k)}{\|d_k\|^2} \geq \frac{C}{B^2} \frac{(-g_k^T d_k)}{\|g_k\|^2} \geq \frac{C}{B^2} \left(1 - \frac{1}{4\eta}\right). \quad (45)$$

We get from (14), (40), (41), (45), and the Armijo condition (10)

$$\begin{aligned} f(x_{k+1}) - f(x^*) &\leq f(x_k) - f(x^*) + \delta \alpha_k g_k^T d_k, \\ &\leq f(x_k) - f(x^*) - \delta \left(1 - \frac{1}{4\eta}\right) \alpha_k \|g_k\|^2, \\ &\leq f(x_k) - f(x^*) - \delta \frac{C}{B^2} \left(1 - \frac{1}{4\eta}\right) m^2 \|x_k - x^*\|^2, \\ &\leq f(x_k) - f(x^*) \\ &\quad - \delta \frac{Cm^2}{B^2} \left(1 - \frac{1}{4\eta}\right)^2 \frac{2}{M} (f(x_k) - f(x^*)), \\ &= \left(1 - \delta \frac{2Cm^2}{MB^2} \left(1 - \frac{1}{4\eta}\right)^2\right) (f(x_k) - f(x^*)). \end{aligned} \quad (46)$$

$$\begin{aligned} \text{Letting } r &= \sqrt{1 - \delta(2Cm^2/MB^2)(1 - (1/4\eta))^2}, \text{ we have} \\ f(x_k) - f(x^*) &\leq r^2 (f(x_{k-1}) - f(x^*)) \leq \dots \leq r^{2k} (f(x_0) - f(x^*)). \end{aligned} \quad (47)$$

Combing (40), we obtain

Step 0: Choose an initial point $x_0 \in \mathfrak{R}^n$, constants $\delta, \rho \in (0, 1)$. Compute. $g_0 = \nabla f(x_0)$. Set $k = 0$.
Step 1: If $\|g_k\| \leq \epsilon$, then stop the iterations. Otherwise, continue with step 2.
Step 2: Compute the system of equations (3) to get d_k , where β_k is obtained by (13).
Step 3: Determine a step size α_k by the Armijo line search (10).
Step 4: Compute $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = \nabla f(x_{k+1})$.
Step 5: Let $k = k + 1$ and go to Step 1.

ALGORITHM 1: NPRP method with Armijo line search.

TABLE 1: Test problem.

No.	Test problem
1	Extended Freudenstein and roth function
2	Extended trigonometric function
3	Extended Rosenbrock function
4	Extended White and Holst function
5	Extended Beale function
6	Extended penalty function
7	Perturbed quadratic function
8	Raydan 1 function
9	Raydan 2 function
10	Diagonal 1 function
11	Diagonal 2 function
12	Diagonal 3 function
13	Hager function
14	Generalized tridiagonal-1 function
15	Extended tridiagonal-1 function
16	Extended three exponential terms function
17	Generalized tridiagonal-2 function
18	Diagonal 4 function
19	Diagonal 5 function
20	Extended Himmelblau function
21	Generalized PSC1 function
22	Extended PSC1 function
23	Extended Powell function
24	Extended block diagonal BD1 function
25	Extended Maratos function
26	Extended Cliff function
27	Quadratic diagonal perturbed function
28	Extended Wood function
29	Extended Hiebert function
30	Quadratic function QF1 function
31	Extended quadratic penalty QP1 function
32	Extended quadratic penalty QP2 function
33	A quadratic function QF2 function
34	Extended EP1 function
35	Extended Tridiagonal-2 function
36	BDQRTIC function (CUTE)
37	TRIDIA function (CUTE)
38	ARWHEAD function (CUTE)
39	NONDIA function (CUTE)
40	NONDQUAR function (CUTE)
41	DQDRTIC function (CUTE)
42	EG2 function (CUTE)
43	DIXMAANA function (CUTE)
44	DIXMAANB function (CUTE)
45	DIXMAANC function (CUTE)
46	DIXMAANE function (CUTE)
47	Partial perturbed quadratic function
48	Broyden tridiagonal function

TABLE 1: Continued.

No.	Test problem
49	Almost perturbed quadratic function
50	Tridiagonal perturbed quadratic function
51	EDENSCH function (CUTE)
52	VARDIM function (CUTE)
53	STAIRCASE S1 function
54	LIARWHD function (CUTE)
55	DIAGONAL 6 function
56	DIXON3DQ function (CUTE)
57	DIXMAANF function (CUTE)
58	DIXMAANG function (CUTE)
59	DIXMAANH function (CUTE)
60	DIXMAANI function (CUTE)
61	DIXMAANJ function (CUTE)
62	DIXMAANK function (CUTE)
63	DIXMAANL function (CUTE)
64	DIXMAAND function (CUTE)
65	ENGVAL1 function (CUTE)
66	FLETCHCR function (CUTE)
67	COSINE function (CUTE)
68	Extended DENSCHNB function (CUTE)
69	Extended DENSCHNF function (CUTE)
70	SINQUAD function (CUTE)
71	BIGGSB1 function (CUTE)
72	Partial perturbed quadratic PPQ2 function
73	Scaled quadratic SQ1 function
74	Scaled quadratic SQ2 function

$$\|x_k - x^*\| \leq \sqrt{\frac{2}{m} (f(x_k) - f(x^*))} \leq \sqrt{\frac{2}{m} (f(x_0) - f(x^*))} r^k. \quad (48)$$

Hence (44) holds, where $a = \sqrt{(2/m)(f(x_0) - f(x^*))}$.
The proof is completed. $\square$

Remark 3. From the proof of Theorem 2, with the Wolfe line search, the NPRP method can also establish that $\{x_k\}$ converges to x^* at least R -linearly.

5. Numerical Experiments

In this section, we report some numerical results of Algorithm 1 and compare the performance of some methods previously detailed in this paper. The following methods were compared:

FIGURE 1: Performance profiles of these methods (CPU Time).

NPRP: β_k^{NPRP} is computed by (13), d_k is determined by (3), the Armijo line search (10)).

DPRP [30, 31]: β_k^{DPRP} is computed by (9), d_k is determined by (3), the weak Wolfe-Powell line search (6).

APRP [33]: β_k^{APRP} is computed by (12), d_k is determined by (11) the Armijo line search (10).

All programs are written in MATLAB R2019a and run on a PC with an AMD Ryzen 5 3550 H with Radeon Vega Mobile Gfx CPU @2.10 GHz, 16 GB (14.9 GB available) of RAM and the Windows 10 operating system.

5.1. General Unconstrained Optimization Problems

5.1.1. Contrast Algorithm. This experiment tests three algorithms, NPRP, DPRP and APRP.

5.1.2. Tested Problems. A number of 74 unconstrained optimization test problems are described in Table 1 [36, 37].

5.1.3. Dimensionality. Problem instances with 3000, 9000, and 27000 variables are considered.

5.1.4. Parameters. All the algorithms run with $t = 1$, $\eta = 1$, $\delta = 0.1$, $\rho = 0.9$, $\sigma = 0.2$, and $\varphi = 0.85$.

5.1.5. Termination Rule. If $g_k < 10^{-6}$ or if the number of iterations exceeds 10^3 .

5.1.6. Symbol Representation. No: the test problem number, CPUTime: the CPU time in seconds, NI: the number of iterations, NFG: the total number of function and gradient evaluations.

5.1.7. Image Description. Figures 1–3 show the performance profiles of CPU Time, NI, and NFG of the three methods.

From Table 2 and Figures 1–3, we can see the performance of NPRP method and APRP method is significantly better than that of DPRP method. The reason is that the Armijo line search is more convenient than the Wolfe line search to obtain stepsize α_k . Meanwhile, we can get that the performance of NPRP method is slightly better than that of APRP method.

5.2. Image Restoration Problems. This subsection is done to recover the original image from an image corrupted by impulse noise.

5.2.1. Contrast Algorithm. This experiment tests two algorithms, NPRP and APRP.

5.2.2. Tested Problems. Restore the original image from the image destroyed by impulsive noise. The experiments chose Boat (256×256), Lena (512×512), and Man (1024×1024) are the test images.

5.2.3. Noise Level. Different algorithms perform for 25%, 50%, and 75% noise problems.

5.2.4. Parameters. All the algorithms run with $t = 1$, $\eta = 1$, $\delta = 0.1$, and $\rho = 0.9$.

5.2.5. Termination Rule. If $(\|g_{k+1} - g_k\|/\|g_k\|) < 10^{-6}$ or $(\|x_{k+1} - x_k\|/\|x_k\|) < 10^{-6}$, the algorithm is terminated.

FIGURE 2: Performance profiles of these methods (NI).

FIGURE 3: Performance profiles of these methods (NFG).

TABLE 2: Numerical results for the total CPU Time, NI and NFG.

Algorithm	CPU time	NI	NFG
NPRP	165.55	10199	16383
DPRP	463.53125	21724	48264
APRP	556.65625	10399	39369

5.2.6. *Symbol Representation.* CPU Time: the CPU time in seconds.

5.2.7. *Image Description.* Figures 4 and 5 show the recovery of different algorithms with different noise levels. Restoration of the Boat, Lena, and Man images is done by the NPRP

method and APRP method. From left to right: a noise image with different percentages of salt-and-pepper noise, restorations with NPRP method, and restorations with APRP method.

The results in Table 3 and Figures 4–6 show us at least two conclusions: (i) NPRP method and APRP method are successful for restoring these images with suitable CPU time; (ii) NPRP method is promising and competitive to the APRP method for 25% noise problem, 50% noise problem, and 75% noise problem.

5.3. *Compressive Sensing Problems.* The purpose of this section is to accurately recover the image from a few random projections by compressive sensing (CG), based on the

FIGURE 4: From left to right: 25% noise, NPRP method, APRP method.

gradient recovery algorithms. The experimental method comes from the model proposed by Dai and Sha [38].

5.3.1. Contrast Algorithm. This experiment tests two algorithms, NPRP and APRP.

5.3.2. Tested Problems. Recover the image accurately from a few of random projections. The experiments chose Phantom (256×256), Fruits (256×256), and Peppers (256×256) as the test images.

5.3.3. Parameters. All the algorithms run with $t = 0.3$, $\eta = 1/3$, $\delta = 0.1$, and $\rho = 0.4$.

5.3.4. Termination Rule. If the square root of the sum of the diagonal elements of $g_k^T g_k$ is less than 10^{-3} or if the number of iterations exceeds 500.

5.3.5. Symbol Representation. PSNR: peak signal to noise ratio. The greater the PSNR, the better the effect.

FIGURE 5: From left to right: 50% noise, NPRP method, APRP method.

TABLE 3: CPU times of the NPRP method and the APRP method in seconds.

	Boat	Lena	Man	Total
<i>25% noise</i>				
NPRP	6.6406	13.8594	78.2813	98.7813
APRP	7.1563	15.1563	80.3281	102.6407
<i>50% noise</i>				
NPRP	11.6094	45.8281	204.1406	261.5781
APRP	12.8281	49.3125	212.4844	274.6250
<i>75% noise</i>				
NPRP	20.9844	88.2188	414.4375	523.6407
APRP	21.6563	95.7344	435.8438	553.2345

FIGURE 6: From left to right: 75% noise, NPRP method, APRP method.

5.3.6. Image Description. Figure 7 shows the recovery of images after random measurement in Fourier domain. Restoration of the Phantom, Fruits, and Peppers images by NPRP method and APRP method. From left to right: original image, restorations with NPRP method, restorations with APRP method.

The results in Table 4 and Figure 7 show us at least two conclusions: (i) NPRP method and APRP method are successful for restoring these images with suitable PSNR; (ii)

NPRP method is promising and competitive to APRP method for compressive sensing problems.

From the above numerical experiment results, we can see that the performance of the NPRP method is more competitive than the APRP method and the DPRP method. The reasons are as follows. (i) Hager and Zhang pointed out in the survey paper [39] that the common molecule $g_k^T y_{k-1}$ of the PRP, HS, and LS methods possess a built-in restart feature that addresses the jamming problem: the factor

TABLE 4: Peak signal to noise ratio (PNSR).

Algorithm	Phantom	Fruits	Peppers	Total
NPRP	42.5962	34.3038	33.7987	110.6987
APRP	42.4038	31.9646	33.3072	107.6746

FIGURE 7: From left to right: original image, NPRP method, APRP method.

$y_{k-1} = g_k - g_{k-1}$ in the numerator of β_k tends to zero when the step $s_{k-1} = x_k - x_{k-1}$ is too small. In addition, β_k becomes smaller, and the new search direction d_k is closer to the steepest descent direction g_k . Meanwhile, NPRP method

is similar in form to CG-C and PRP methods and therefore inherits their good performance. (ii) The Armijo line search is more convenient than the Wolfe line search to obtain stepsize α_k .

6. Conclusion

In this paper, a modified PRP conjugate gradient method was proposed to solve image restoration and compression sensing problems, based on the well-known CG-C method [29]. With the Armijo line search, the global convergence and liner convergence rate of the algorithm is established under some suitable conditions. The sufficient descent property of the algorithm has been proved without the use of any line search method. The numerical results indicate that the algorithm is effective and competitive for solving unconstrained optimization problems, image restoration problems, and compressive sensing problems.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

There are no potential conflicts of interest.

Acknowledgments

The authors want to thank the financial support by the High Level Innovation Teams and Excellent Scholars Program in Guangxi Institutions of Higher Education (Grant No. [2019] 52), the National Natural Science Foundation of China (Grant no. 11661009), the Guangxi Natural Science Key Foundation (No. 2017GXNSFDA198046), and the Guangxi Natural Science Foundation (No. 2020GXNSFAA159069).

References

- [1] Y. H. Dai, "A nonmonotone conjugate gradient algorithm for unconstrained optimization," *Journal of System Science and Complexity*, vol. 15, no. 2, pp. 139–145, 2002.
- [2] J. Nocedal and Y. X. Yuan, "Analysis of a self-scaling quasi-newton method," *Mathematical Programming*, vol. 61, no. 1–3, pp. 19–37, 1993.
- [3] Z. Wei, G. Li, and L. Qi, "New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems," *Applied Mathematics and Computation*, vol. 179, no. 2, pp. 407–430, 2006.
- [4] Z. Wei, S. Yao, and L. Liu, "The convergence properties of some new conjugate gradient methods," *Applied Mathematics and Computation*, vol. 183, no. 2, pp. 1341–1350, 2006.
- [5] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, vol. 3, no. 16, pp. 35–43, 1969.
- [6] B. T. Polyak, "The conjugate gradient method in extremal problems," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94–112, 1969.
- [7] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.
- [8] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: theory," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.
- [9] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.
- [10] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.
- [11] R. Fletcher, "Practical methods of optimization," in *Unconstrained Optimization*, vol. 1, John Wiley & Sons, Hoboken, NJ, USA, 1980.
- [12] W. Y. Cheng, "A two-term prp-based descent method," *Numerical Functional Analysis and Optimization*, vol. 28, no. 11–12, pp. 1217–1230, 2007.
- [13] Z.-f. Dai and B.-S. Tian, "Global convergence of some modified prp nonlinear conjugate gradient methods," *Optimization Letters*, vol. 5, no. 4, pp. 615–630, 2011.
- [14] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21–42, 1992.
- [15] G. Yu, Y. Zhao, and Z. Wei, "A descent nonlinear conjugate gradient method for large-scale unconstrained optimization," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 636–643, 2007.
- [16] Z. Dai, X. Dong, J. Kang, and L. Hong, "Forecasting stock market returns: new technical indicators and two-step economic constraint method," *The North American Journal of Economics and Finance*, vol. 53, Article ID 101216, 2020.
- [17] Z. F. Dai and H. Zhu, "A modified hestenes-stiefel-type derivative-free method for large-scale nonlinear monotone equations," *Mathematics*, vol. 8, no. 2, p. 168, 2020.
- [18] F. Wen and X. Yang, "Skewness of return distribution and coefficient of risk premium," *Journal of Systems Science and Complexity*, vol. 22, no. 3, pp. 360–371, 2009.
- [19] G. Yu, J. Huang, and Y. Zhou, "A descent spectral conjugate gradient method for impulse noise removal," *Applied Mathematics Letters*, vol. 23, no. 5, pp. 555–560, 2010.
- [20] G. Yuan, T. Li, and W. Hu, "A conjugate gradient algorithm for large-scale nonlinear equations and image restoration problems," *Applied Numerical Mathematics*, vol. 147, pp. 129–141, 2020.
- [21] G. L. Yuan, J. Y. Lu, and Z. Wang, "The prp conjugate gradient algorithm with a modified wwp line search and its application in the image restoration problems," *Applied Numerical Mathematics*, vol. 152, pp. 1–11, 2020.
- [22] G. Yuan, Z. Wei, and G. Li, "A modified Polak-Ribière-Polyak conjugate gradient algorithm for nonsmooth convex programs," *Journal of Computational and Applied Mathematics*, vol. 255, pp. 86–96, 2014.
- [23] G. Yuan, Z. Wei, and Y. Yang, "The global convergence of the Polak-Ribière-Polyak conjugate gradient algorithm under inexact line search for nonconvex functions," *Journal of Computational and Applied Mathematics*, vol. 362, pp. 262–275, 2019.
- [24] M. J. D. Powell, "Nonconvex minimization calculations and the conjugate gradient method," in *Numerical Analysis*, pp. 122–141, Springer, Berlin, Germany, 1984.
- [25] Y. H. Dai, *Analyses of conjugate gradient methods*, PhD. thesis, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing, China, 1997.
- [26] M. Al-Baali, "Descent property and global convergence of the fletcher-reeves method with inexact line search," *IMA Journal of Numerical Analysis*, vol. 5, no. 1, pp. 121–124, 1985.

- [27] Y. F. Hu and C. Storey, "Global convergence result for conjugate gradient methods," *Journal of Optimization Theory and Applications*, vol. 71, no. 2, pp. 399–405, 1991.
- [28] D. Touati-Ahmed and C. Storey, "Efficient hybrid conjugate gradient techniques," *Journal of Optimization Theory and Applications*, vol. 64, no. 2, pp. 379–397, 1990.
- [29] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.
- [30] G. H. Yu and L. T. Guan, "New descent nonlinear conjugate gradient methods for large-scale optimization," *Department of Scientific Computation and Computer*, 2005.
- [31] G. Yu, L. Guan, and W. Chen, "Spectral conjugate gradient methods with sufficient descent property for large-scale unconstrained optimization," *Optimization Methods and Software*, vol. 23, no. 2, pp. 275–293, 2008.
- [32] Z. Dai and F. Wen, "Global convergence of a modified hestenes-stiefel nonlinear conjugate gradient method with armijo line search," *Numerical Algorithms*, vol. 59, no. 1, pp. 79–93, 2012.
- [33] M. Li and A. Qu, "Some sufficient descent conjugate gradient methods and their global convergence," *Computational and Applied Mathematics*, vol. 33, no. 2, pp. 333–347, 2014.
- [34] L. Zhang, W. Zhou, and D. Li, "Global convergence of the dy conjugate gradient method with armijo line search for unconstrained optimization problems," *Optimization Methods and Software*, vol. 22, no. 3, pp. 511–517, 2007.
- [35] G. Zoutendijk, "Nonlinear programming, computational methods," *Integer and nonlinear programming*, pp. 37–86, 1970.
- [36] I. Bongartz, A. R. Conn, N. Gould, and P. L. Toint, "Cute: constrained and unconstrained testing environment," *ACM Transactions on Mathematical Software*, vol. 21, no. 1, pp. 123–160, 1995.
- [37] J. J. Moré, B. S. Garbow, and K. E. Hillstom, "Testing unconstrained optimization software," *ACM Transactions on Mathematical Software (TOMS)*, vol. 7, no. 1, pp. 17–41, 1981.
- [38] Q. Dai and W. Sha, "The physics of compressive sensing and the gradient-based recovery algorithms," 2009, <https://arxiv.org/abs/0906.1487>.
- [39] W. W. Hager and H. C. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific Journal of Optimization*, vol. 2, no. 1, pp. 35–58, 2006.