Open Access Library Journal
2020, Volume 7, e6048
ISSN Online: 2333-9721
ISSN Print: 2333-9705

Global Convergence Property with Inexact Line Search for a New Hybrid Conjugate Gradient Method

Fanar N. Al-Namat, Ghada M. Al-Naemi*

College of Computer Sciences and Mathematics, Department of Mathematics, University of Mosul, Mosul, Iraq
Email: *gh66alnaemi@gmail.com

How to cite this paper: Al-Namat, F.N. and Al-Naemi, G.M. (2020) Global Convergence Property with Inexact Line Search for a New Hybrid Conjugate Gradient Method. *Open Access Library Journal*, **7**: e6048.
https://doi.org/10.4236/oalib.1106048

Received: January 3, 2020
Accepted: February 8, 2020
Published: February 11, 2020

Abstract

In this study, we derive a new scale parameter φ for the CG method, for solving large scale unconstrained optimization algorithms. The new scale parameter φ satisfies the sufficient descent condition, global convergence analysis proved under Strong Wolfe line search conditions. Our numerical results show that the proposed method is effective and robust against some known algorithms.

Subject Areas

Numerical Optimization

Keywords

Unconstrained Optimization, Hybrid, Conjugate Gradient

1. Introduction

In unconstrained optimization, we minimize an objective function that depends on real variables with no restrictions at all on the value of these variables. The unconstrained optimization problem is stated by:

$$\min_{x \in R^n} f(x) \tag{1}$$

where $x \in R^n$ is a real vector with $n \geq 1$ component and $f: R^n \to R$ is a smooth function and its gradient g is available [1]. A nonlinear conjugate gradient method generates a sequence x_k Starting from an initial guess $x_0 \in R^n$ Using the recurrence

$$x_{k+1} = x_k + \alpha_k d_k,\ k = 0,1,2,\cdots \tag{2}$$

where α_k is the positive step size obtained by carrying out a one dimensional search, known as the line searches [2]. Among them, the so-called strong wolf line search conditions require that [3] [4].

$$f(x_k+\alpha_k d_k) \le f(x_k)+\sigma\alpha_k d_k, \tag{3}$$

$$\left|g(x_k+\alpha_k d_k)\right| \le \delta\left|g_k^{\mathrm{T}} d_k\right| \tag{4}$$

where $0<\sigma<\delta<1$, is to find an approximation of α_k where the descent property must be satisfied and no longer searching in the direction when x_k is far from the solution. Thus by strong Wolfe line search conditions we in herit the advantages of exact line search with inexpensive and low computational cost [5].

The search direction d_k is generated by:

$$d_k = \begin{cases} -g_k, & k=1 \\ -g_k+\beta_k d_k, & k>1 \end{cases} \tag{5}$$

where g_k and β_k is the gradient and conjugate gradient coefficient of $f(x)$ respectively at the point x_k. The different choices for the parameter β_k correspond to different conjugate gradient methods. The most popular formulas for β_k is Hestenes Stiefel method (HS), Fletcher-Reeves method (FR), Polak-Ribiere-Polyak method (PR), conjugate—Descent method (CD), Liu—Storey method (LS), and Dai-Yuan method (DY), etc

These methods are identical when f is a strongly convex quadratic function and the line search is exact, since the gradient are mutually orthogonal, and the parameters β_k in these methods are equal. When applied to general nonlinear function with inexact line searches, however, the behavior of these methods is marked different [1]. We are going to summarize some well known conjugate gradient method in **Table 1**.

An important class of conjugate gradient methods is the hybrid conjugate gradient algorithms. The hybrid computational schemes perform better than the classical conjugate gradient methods. They are defined by (2) and (5) where the parameter β_k is computed as projections or as convex combinations of different conjugate gradient methods [14].

We are going to summarize some well known hybrid conjugate gradient method in **Table 2**.

We propose a new hybrid CG method based on combination of MMWU [24] and RMAR [25] conjugate gradient methods for solving unconstrained optimization method with suitable conditions. The corresponding conjugate gradient parameters are

$$B_k^{MMWU} = \frac{\left\|g_{k+1}\right\|^2}{\left\|d_k\right\|^2} \tag{6}$$

and

$$\beta_k^{RMAR} = \frac{\left\|g_{k+1}\right\|^2 - \dfrac{\left\|g_{k+1}\right\|}{\left\|d_k\right\|} g_{k+1}^{\mathrm{T}} d_k}{\left\|d_k\right\|^2} \tag{7}$$

Table 1. Some well known conjugate gradient coefficients.

NO	Formula	Authors
1	$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T s_k}$	Hestenes and Stiefel (HS) [6]
2	$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$	Fletcher and Reeves (FR) [7]
3	$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{g_k^T g_k}$	Polak-Ribiere (PRP) [8] [9]
4	$\beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k}$	Conjugate Descent (CD) [10]
5	$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-g_k^T s_k}$	Liu and Storey (LS) [11]
6	$\beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k}$	Dai-Yuan method (DY) [12]
7	$\beta_k^{new} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - \alpha_k^2 \frac{d_k^T g_k}{y_k^T y_k}$	Al-Naemi and Hamed [13]

Table 2. Hybrid conjugate gradient methods.

NO	Formula	Authors
1	$\beta_k^c = (1-\theta_k)\beta_k^{HS} + \theta_k \beta_k^{DY}$	Andrei [15]
2	$\beta_k^{Ac} = (1-\theta_k)\beta_k^{PRP} + \theta_k \beta_k^{DY}$	Yan [16]
3	$\beta_k^{N} = (1-\theta_k)\beta_k^{FR} + \theta_k \beta_k^{MMWU}$	Li and Sun [17]
4	$\beta_k^{hyb} = (1-\theta_k)\beta_k^{LS} + \theta_k \beta_k^{DY}$	Liu, J.K. and Li, Sij [1]
5	$\beta_k^{hyb} = (1-\theta_k)\beta_k^{LS} + \theta_k \beta_k^{FR}$	Djordjevic' [18]
6	$\beta_k^{hyb} = (1-\theta_k)\beta_k^{HS} + \theta_k \beta_k^{FR}$	Djordjevic' [19]
7	$\beta_k^{hyb} = (1-\theta_k)\beta_k^{LS} + \theta_k \beta_k^{FR}$	Djordjevic' [20]
8	$\beta_k^c = (1-\theta_k)\beta_k^{HS} + \theta_k \beta_k^{CD}$	Xiuyun, *et al.* [21]
9	$\beta_k^{LSDY} = (1-\gamma_k)\beta_k^{LS} + \gamma_k \beta_k^{DY}$	Abdullahi and Ahmad [22]
10	$\beta_k^{HCG} = \lambda_k \beta_k^{DY} + (1-\lambda_k)\beta_k^{HS}$	Livieris, Tampakas, and Pintelas [23]

We defined the parameter β_k in the proposed method by:

$$\beta_k^{FG} = (1-\varphi_k)\beta_k^{MMWU} + \varphi_k \beta_k^{RMAR} \tag{8}$$

Observe that if $\varphi_k = 0$, then $\beta_k^{FG} = \beta_k^{MMWU}$, and if $\varphi_k = 1$, then $\beta_k^{FG} = \beta_k^{RMAR}$.

By choosing the appropriate value of the parameter φ_k In the convex combination, the search direction d_k of our algorithm not only is the Newton direction, but also satisfies the famous DL conjugate condition proposed by Dai and Liao [26]. Under the strong Wolfe line search conditions, we prove the global convergence of our algorithm. The numerical results also show the feasibility and effectiveness of our algorithm.

This paper is organized as follows. Section 2 we introduce our new hybrid conjugate gradient method (HFG), and we obtain the parameter φ_k using some approaches and give us a specific algorithm. Section 3, we prove that it generates direction satisfying the sufficient descent condition under strong Wolfe line search conditions. The global convergence property of the proposed method is established in Section 4. Some numerical results are reported in Section 5.

2. A New Hybrid Conjugate Gradient Method

In this section, we will describe a new proposed hybrid conjugate gradient method. In order to obtain the sufficient descent direction, we will compute φ_k as follows. We combine β_k^{MMWU} and β_k^{RMAR} in a convex combination in order to have a good algorithm for unconstrained optimization.

The direction d_{k+1} is generated by the rule

$$d_{k+1} = -g_{k+1} + \beta_k^{HFG} d_k \tag{9}$$

where β_k^{HFG} defined in (8), the iterates $x_1, x_2, x_3, \cdots$ of our method are computed by means of the recurrence (2), where the step size α_k Is determined according to the strong Wolf conditions (3) and (4).

The scale parameter φ_k satisfying $0 \le \varphi_k \le 1$, which will be determined in a specific way to be described later. Observe that if $\varphi_k = 0$, then $\beta_k^{HFG} = \beta_k^{MMWU}$, and

If $\varphi_k = 1$, then $\beta_k^{HFG} = \beta_k^{RMAR}$. On the other hand, if $0 < \varphi_k < 1$, then β_k^{HFG} is a convex combination of β_k^{MMWU} and β_k^{RMAR}.

From (8) and (9) it is obvious that:

$$d_{k+1} = \begin{cases} -g_{k+1}, & k = 1 \\ -g_{k+1} + (1-\varphi_k)\dfrac{\|g_{k+1}\|^2}{\|d_k\|^2} d_k + \varphi_k \dfrac{\|g_{k+1}\|^2 - \dfrac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^{\mathrm{T}} d_k}{\|d_k\|^2} d_k, & k > 1 \end{cases} \tag{10}$$

Our motivation to select the parameter φ_k in such a manner that the defection d_{k+1} given in (10) is equal to the Newton direction $d_{k+1}^N = -\nabla^2 f(x_{k+1})^{-1} g_{k+1}$. There for

$$\begin{aligned} &-\nabla^2 f(x_{k+1})^{-1} g_{k+1} \\ &= -g_{k+1} + (1-\varphi_k)\frac{\|g_{k+1}\|^2}{\|d_k\|^2} d_k + \varphi_k \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^{\mathrm{T}} d_k}{\|d_k\|^2} d_k \end{aligned} \tag{11}$$

Now multiplying (11) by $s_k^{\mathrm{T}} \nabla^2 f(x_{k+1})$ from the left, we get

$$\begin{aligned} -s_k^{\mathrm{T}} g_{k+1} &= -s_K^{\mathrm{T}} \nabla^2 f(x_{k+1}) g_{k+1} + (1-\varphi_k)\frac{\|g_{k+1}\|^2}{\|d_k\|^2} s_K^{\mathrm{T}} \nabla^2 f(x_{k+1}) d_k \\ &\quad + \varphi_k \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^{\mathrm{T}} d_k}{\|d_k\|^2} s_k^{\mathrm{T}} \nabla^2 f(x_{k+1}) d_k \end{aligned}$$

Therefore, in order to have an algorithm for solving large scale problems we assume that pair (s_k, y_k) satisfies the secant equation

$$\nabla^2 f(x_{k+1}) s_k = y_k. \tag{12}$$

From (12), we get

$$s_k^{\mathrm{T}} \nabla^2 f(x_{k+1}) = y_k^{\mathrm{T}}.$$

Denoting $\varphi_k^{FG} = \varphi_k$ we get

$$-s_k^{\mathrm{T}} g_{k+1} = -y_K^{\mathrm{T}} g_{k+1} + \frac{\|g_{k+1}\|^2}{\|d_k\|^2} y_k^{\mathrm{T}} d_k + \varphi_k^{FG} \left(\frac{\|g_{k+1}\| (g_{k+1}^{\mathrm{T}} d_k)}{\|d_k\|^2} \right) (y_k^{\mathrm{T}} d_k)$$

after some algebra, we get

$$\varphi_k^{FG} = \frac{(s_k^{\mathrm{T}} g_{k+1} - y_k^{\mathrm{T}} g_{k+1}) \cdot \|d_k\|^3 + \|g_{k+1}\|^2 \cdot \|d_k\| (y_k^{\mathrm{T}} d_k)}{\|g_{k+1}\| \cdot (g_{k+1}^{\mathrm{T}} d_k) \cdot (y_k^{\mathrm{T}} d_k)} \tag{13}$$

Now, we specify a complete hybrid conjugate gradient method (HFG) which posses some nice properties of conjugate gradient and Newton method.

Algorithm HFG

Step 1: Select $x_0 \in R^n$, $\in > 0$, set $k = 0$. Compute $f(x_0)$ and $g_0 = -\nabla f(x_0)$, set $d_0 = -g_0$.

Step 2: Test the stopping criteria, *i.e.* if $\|g_k\| \le \in$, then stop.

Step 3: Compute α_k by strong Wolfe line search conditions in (3) & (4).

Step 4: Compute $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = g(x_{k+1})$. Compute $s_k = x_{k+1} - x_k$ And $y_k = g_{k+1} - g_k$

Step 5: If $\varphi_k \ge 1$ then set $\varphi_k = 1$. If $\varphi_k \le 0$, then set $\varphi_k = 0$, otherwise compute φ_k as (13).

Step 6: Compute β_k^{FG} by (8).

Step 7: Generate $d = -g_{k+1} + \beta_k^{FG} d_k$

Step 8: If the restart criteria of Powell $|g_k^{\mathrm{T}} g_k| \ge 0.2 \|g_{k+1}\|^2$, is satisfied, then set $d_k = -g_{k+1}$, otherwise define $d_{k+1} = d$

Step 9: Set $k = k + 1$, and continue with step 2.

3. The Sufficient Descent Condition

In this section, we are going to apply the following theorem to illustrate that the search direction d_k Obtained by hybrid FG satisfies the sufficient descent condition which plays Avit of role in analyzing the global convergence.

For further considerations we need the following assumptions

3.1. Assumption

The level sets $S = \{x \in R^n, f(x_n)\}$ are bounded.

3.2. Assumption

In a neighborhood *N* of *S*, the function *f* is continuously differentiable and its gradient is Lipschitz continuous, *i.e.*, there exists a constant $L > 0$, such that

$$\|\nabla f(x) - \nabla f(y)\| \le L \|x - y\|, \forall x, y \in N$$

Under these assumptions of if there exists a positive constant ($\gamma, \bar{\gamma}, \omega$ & $\bar{\omega}$) & such that

$$\bar{\gamma} \le \|g_{k+1}\| \le \gamma \text{ and } \bar{\varpi} \le \|g_k\| \le \omega, \ \forall x \in S$$ [27].

Theorem.

Let the sequences $\{g_k\}$ and $\{d_k\}$ be generated by a hybrid FG method. Then the search direction d_k satisfies the sufficient descent condition:

$$g_{k+1}^{\mathrm{T}} d_{k+1} \le -\mu \|g_{k+1}\|^2, \ \forall \mu \ge 0 \tag{14}$$

where $\mu = 1-(E_4 - E_3)$, with $0 < (E_4 - E_3) < 1$.

Proof. We shall show that d_k satisfies the sufficient descent condition holds for $k=0$, the proof is a trivial one, *i.e.* $d_0 = -g_0$ and so $g_0^{\mathrm{T}} d_0 = -\|g_0\|^2$. Now we have

$$d_{k+1} = -g_{k+1} + \beta_k^{FG} d_k,$$

i.e.

$$d_{k+1} = -g_{k+1} + \left[(1-\varphi_k)\beta_k^{MMWU} + \varphi_k \beta_k^{RMAR}\right] d_k$$

We can rewrite the above direction by the following manner:

$$d_{k+1} = -\left(\varphi_k g_{k+1} + (1-\varphi_k) g_{k+1}\right) + \left((1-\varphi_k)\beta_k^{MMWU} + \varphi_k \beta_k^{RMAR}\right) d_k .$$

So,

$$d_{k+1} = \varphi_k\left(-g_{k+1} + \beta_k^{RMAR} d_k\right) + (1-\varphi_k)\left(-g_k + \beta_k^{MMWU} d_k\right),$$

After some arrangement, we get

$$d_{k+1} = \varphi_k d_{k+1}^{RMAR} + (1-\varphi_k) d_{k+1}^{MMWU} \tag{15}$$

Multiplying (15) by g_{k+1}^{T} from the left, we get

$$g_{k+1}^{\mathrm{T}} d_{k+1} = \varphi_k g_{k+1}^{\mathrm{T}} d_{k+1}^{RMAR} + (1-\varphi_k) g_{k+1}^{\mathrm{T}} d_k^{MMWU}$$

Firstly, if $\varphi_k = 0$, then $d_{k+1} = d_{k+1}^{MMWU}$, we are going to prove that the sufficient descent condition holds for MMWU method in the presence of the strong Wolfe line search condition, because in [24] they proved this method satisfied the sufficient descent condition with exact line search.

i.e.

$$g_{k+1}^{\mathrm{T}} d_{k+1}^{MMWU} = -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|d_k\|^2} g_{k+1}^{\mathrm{T}} d_k \tag{16}$$

Since,

$$g_{k+1}^{\mathrm{T}} d_k \le y_k^{\mathrm{T}} d_k \text{ and } y_k^{\mathrm{T}} d_k \le \alpha_k L \|d_k\|^2 \tag{17}$$

Applications (17) in (16), we get

$$\begin{aligned} g_{k+1}^{\mathrm{T}} d_{k+1}^{MMWU} &\le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|d_k\|^2} \alpha_k L \|d_k\|^2 \\ &= -(1-\alpha_k L)\|g_{k+1}\|^2 \\ &= -E_1 \|g_{k+1}\|^2 \end{aligned} \tag{18}$$

where $E_1 = (1-\alpha_k L) > 0$, with $0 < \alpha_k L < 1$.

So, it is proved that d_{k+1}^{MMWU} satisfies the sufficient descent condition.

Now let $\varphi_k = 1$ then $d_k = d_k^{RMAR}$, we are going to prove that the sufficient descent condition holds for RMAR method in the presence of the strong Wolfe line search condition because in [25] they proved this method satisfied the sufficient descent condition with exact line search.

$$d_{k+1}^{RMAR} = -g_{k+1} + \beta_k^{RMAR} d_k$$

Multiplying the above equation from left by g_{k+1}^{T} we get

$$g_{k+1}^{\mathrm{T}} d_{k+1}^{RMAR} = -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^{\mathrm{T}} d_k}{\|d_k\|^2} g_{k+1}^{\mathrm{T}} d_k .$$

In [25], they proved that

$$0 \le \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^{\mathrm{T}} d_k}{\|d_k\|^2} \le 2\frac{\|g_{k+1}\|^2}{\|d_k\|^2} \tag{19}$$

Used (17), and (19) the direction become

$$\begin{aligned} g_{k+1}^{\mathrm{T}} d_{k+1} &\le -\|g_{k+1}\|^2 + 2\alpha_k L \|g_{k+1}\|^2 \\ &= -(1-2\alpha_k L)\cdot\|g_{k+1}\|^2 \\ &= -E_2 \cdot \|g_{k+1}\|^2 \end{aligned} \tag{20}$$

where $E_2 = (1-2\alpha_k L) > 0$ with $0 < 2\alpha_k L < 1$ and $0 < L < \frac{1}{2}$.

So, it is proved that d_{k+1}^{RMAR} satisfied the sufficient descent condition.

Now, we are going to prove the direction satisfy the sufficient descent condition when $0 < \varphi_k < 1$, firstly for

$$(1-\varphi_k)\beta_K^{MMWU} g_{k+1}^{\mathrm{T}} d_k$$
$$= \frac{\|g_{k+1}\|^2}{\|d_k\|^2} g_k^{\mathrm{T}} d_k - \left[\frac{s_k^{\mathrm{T}} g_{k+1}\|d_k\|^3 - y_k^{\mathrm{T}} g_{k+1}\|d_k\|^3 + \|g_{k+1}\|^2\|d_k\| y_k^{\mathrm{T}} d_k}{\|g_{k+1}\|\left(g_{k+1}^{\mathrm{T}} d_k\right) y_k^{\mathrm{T}} d_k}\right] * \frac{\|g_{k+1}\|^2}{\|d_k\|^2} g_{k+1}^{\mathrm{T}} d_k$$

We have from Lipschitz condition $g_{k+1}^{\mathrm{T}} d_k < y_k^{\mathrm{T}} d_k$ and

$$-(1-\sigma)\|g_k\| \le y_k^{\mathrm{T}} d_k \le \alpha_k L \|d_k\|^2$$

with a mathematical calculation, we get

$$\begin{aligned} &(1-\varphi_k)\beta_k^{MMWU} g_{k+1}^{\mathrm{T}} d_k \\ &\le \left[\frac{\alpha_k L\|d_k\|^2}{\|d_k\|^2} - \frac{\|s_k\|\|g_{k+1}\|\|d_k\| - \alpha_k L\|d_k\|^2 + \|g_{k+1}\|^2 \alpha_k L \|d_k\|}{\|g_{k+1}\|\cdot\left(-(1-\sigma)\|g_k\|\right)}\right]\|g_{k+1}\|^2 \\ &\le \left[\alpha_k L + \frac{L}{(1-\sigma)} \frac{\|s_k\|^2\|d_k\| - \alpha_k L\|d_k\|^3 + \|g_{k+1}\|^2\|d_k\|}{\|g_{k+1}\|\|g_k\|^2}\right]\|g_{k+1}\|^2 \\ &\le \left[\alpha_k L + \frac{A\gamma B - \alpha_k L B^2 + Y^2\alpha_k L B}{(1-\sigma)\bar{Y}\bar{W}^2}\right]\|g_{k+1}\|^2 \end{aligned}$$

Let $E_1 = \alpha_k L + \frac{LAB - \alpha_k LB^3 + \gamma^2 \alpha_k LB}{(1-\sigma)\bar{\gamma}\bar{\varpi}^2}$

$$\therefore (1-\varphi_k)\beta^{MMWU} g_{k+1}^{\mathrm{T}} d_k \le E_3 \|g_{k+1}\|^2 \tag{21}$$

Now, secondly for

$$\varphi_k \beta_k^{RMAR} g_{k+1}^{\mathrm{T}} d_k = \left[\frac{s_k^{\mathrm{T}} g_{k+1} \|d_k\|^3 - y_k^{\mathrm{T}} g_{k+1} \|d_k\|^3 + \|g_{k+1}\|^2 \|d_k\| y_k^{\mathrm{T}} \|d_k\|}{\|g_{k+1}\| \left(g_{k+1}^{\mathrm{T}} d_k\right)\left(y_k^{\mathrm{T}} d_k\right)}\right] \cdot \left[\frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|d_k\|} g_{k+1}^{\mathrm{T}} d_k}{\|d_k\|^2}\right] g_{k+1}^{\mathrm{T}} d_k$$

From (19), Lipschitz condition $s_k^{\mathrm{T}} g_{k+1} \le y_k^{\mathrm{T}} s_k \le L\|s_k\|^2$ and $s_k = \alpha_k d_k$, we get

$$\varphi_k \beta_k^{RMAR} g_{k+1}^{\mathrm{T}} d_k = 2\left[\frac{L\|s_k\|^2 \|d_k\| - \|y_k\| \|g_{k+1}\| \|d_k\|^3 + \alpha_k L \|g_{k+1}\|^2 \|d_k\|^2}{\|g_{k+1}\|^2 \left(-(1-\sigma)\right) \|g_k\|^2 \|d_k\|^2}\right] \cdot \|g_{k+1}\|^2$$

Since $\|y_k\| \le \|g_{k+1}\| + \|g_k\|$, so

$$\begin{aligned} &\varphi_k \beta_k^{RMAR} g_{k+1}^{\mathrm{T}} d_k \\ &\le \frac{-2}{1-\sigma}\left[\frac{L\|s_k\|^2 \|d_k\| - 0.8\|g_{k+1}\|^2 \|d_k\| + \alpha_k L \|g_{k+1}\|^2 \|d_k\|}{\|g_{k+1}\|^2 \|g_k\|^2}\right] \cdot \|g_{k+1}\|^2 \\ &\le \frac{-2B}{1-\sigma}\left[\frac{LA - 0.8\gamma^2 + \alpha_k L\omega^2}{\bar{\gamma}\bar{\varpi}^2}\right] \cdot \|g_{k+1}\|^2 \end{aligned}$$

where $E_4 = \frac{2B}{1-\sigma}\left[\frac{LA - 0.8\gamma^2 + \alpha_k L\omega^2}{\bar{\gamma}\bar{\varpi}^2}\right]$

$$\therefore \varphi_k \beta_k^{RMAR} g_{k+1}^{\mathrm{T}} d_k \le -E_4 \|g_{k+1}\|^2 \tag{22}$$

From (18), (20), (21) and (22) we get

$$\begin{aligned} g_{k+1}^{\mathrm{T}} d_{k+1} &\le -\|g_{k+1}\|^2 + E_3 \|g_{k+1}\|^2 - E_4 \|g_{k+1}\|^2 \\ &= -\left[1 - \left(E_4 - E_3\right)\right] \|g_{k+1}\|^2 \\ &= -E \|g_{k+1}\|^2 \end{aligned}$$

with $E = 1 - (E_4 - E_3)$ and $0 < E_4 - E_3 < 1$.

So, it is proved that d_{k+1} Satisfied the sufficient descent condition.

4. Converge Analysis

Let Assumption 2.1 and 2.2 hold. In [26] it is proved that for any conjugate gradient method with strong Wolfe line search conditions, it holds:

4.1. Lemma

Let Assumption 2.1 and 2.2 holds. Consider the method (2) and (5) where the d_k Is a descent direction and α_k is received from the strong wolf line search. If

$$\sum_{k\geq 1}\frac{1}{\|d_k\|^2}=\infty .$$

Then

$$\lim_{k\to\infty}\inf\|g_k\|=0 .$$

4.2. Theorem

Suppose that assumption 2.1 and 2.2 holds. Consider the algorithm HFG were $0\leq\varphi_k\leq 1$ and α_k is obtained by the strong Wolfe line search and d_{k+1} is the descent direction. Then

$$\lim_{k\to\infty}\inf\|g_k\|=0 .$$

Proof. Because the descent condition holds, we have $d_{k+1}\neq 0$. So using lemma 3.1, it is sufficient to prove that $\|d_{k+1}\|$ is bounded above. From (10).

$$\begin{aligned}\|d_{k+1}\| &= -\left\|g_{k+1}+\left[(1-\varphi_k)\beta_k^{MMWU}+\varphi_k\beta_k^{RMAR}\right]d_k\right\| \\ &\leq \|g_{k+1}\|+\left[|1-\varphi_k|\cdot\left|\beta_k^{MMWU}\right|+|\varphi_k|\left|\beta_k^{RMAR}\right|\right]\cdot\|d_k\|\end{aligned}$$

They proved that in [24] and [25].

$$\left|\beta_k^{MMWU}\right|=\frac{\|g_{k+1}\|^2}{\|d_k\|^2}\leq\frac{\gamma^2}{\bar{B}^2},$$

And

$$\left|\beta_k^{RMAR}\right|\leq\frac{2\|g_{k+1}\|^2}{\|d_k\|^2}\leq\frac{2\gamma^2}{\bar{B}^2}.$$

Now for

$$|\varphi_k|=\left|\frac{\left(s_k^{\mathrm{T}}g_{k+1}-y_k^{\mathrm{T}}g_{k+1}\right)\|d_k\|^3-\|g_{k+1}\|^2\|d_k\|y_k^{\mathrm{T}}d_k}{\|g_{k+1}\|\left(g_{k+1}^{\mathrm{T}}d_k\right)\left(y_k^{\mathrm{T}}d_k\right)}\right|,$$

By (4), we have $-(1-\sigma)\leq y_k^{\mathrm{T}}d_k\leq L\alpha_k\|d_k\|^2$.
Since $\left|s_k^{\mathrm{T}}g_{k+1}\right|\leq\left|g_k^{\mathrm{T}}s_k\right|+L\|s_k\|^2,\left|g_k^{\mathrm{T}}s_k\right|\leq\|g_k\|\cdot\|s_k\|$ and

$$\left|y_k^{\mathrm{T}}g_{k+1}\right|\leq(1-0.2)\|g_{k+1}\|^2,$$

with some mathematical calculation, we get

$$\begin{aligned}|\varphi_k| &\leq \frac{\left[\|g_k\|\cdot\|s_k\|+L\|s_k\|^2+(1-0.2)\|g_{k+1}\|^2\right]\cdot\|d_k\|^3+\|g_{k+1}\|^2\cdot\|d_k\|\cdot\sigma\|g_k\|\cdot\|s_k\|}{\sigma(1-\sigma)\|g_{k+1}\|\cdot\|g_k\|^4} \\ &\leq \frac{\left[\left(\gamma\cdot A+L\cdot A^2\right)+(1-0.2)\omega^2\right]\cdot B^3+\omega^2 B\sigma\gamma A}{\sigma(1-\sigma)\bar{\gamma}\bar{\omega}^4} \\ &\leq D\end{aligned}$$

$$\therefore d_{k+1}\leq\gamma\left[1+(1-D)\gamma B+2D\gamma B^2\right]\cdot B=\gamma F$$

$$\Rightarrow\sum_{k\geq 1}\frac{1}{\|d_k\|^2}\geq\frac{1}{\gamma^2F^2}\sum_{k\geq 1}1=\infty$$

$$\therefore \lim_{k\to\infty} \inf \|g_k\| = 0$$

5. Numerical Experiments

In this section we selected some of test functions in Table 3 from CUTE library, along with other large scale optimization problems presented in Andrei [28] [29] and Bongartz *et al.* [30].

All codes are written in double precision FORTRAN Language and compiled Visual F90 (default compiler settings) on a Workstation Intel Pentium 4. The value of α_k is always computed by cubic fitting procedure.

We selected 26 large scale unconstrained optimization problems in the extended

Table 3. It gives the comparison depending in the NOI and NOF between β_k^{MMWU}, β_k^{RMAR} and the proposed method β_k^{FG}.

n	Test Function	Dimension (N)	β_k^{MMWU}		β_k^{RMAR}		β_k^{FG}	
			Total NOI	Total NOF	Total NOI	Total NOF	Total NOI	Total NOF
1	Beal	1000	12	29	12	29	12	29
		5000	12	29	12	29	12	29
		10,000	12	29	12	29	12	29
2	Biggsb1	1000	F	F	F	F	F	F
		5000	32	71	32	71	32	71
		10,000	241	511	241	511	240	506
3	Cosine	1000	10	22	10	22	10	22
		5000	11	27	11	27	11	27
		10,000	11	28	11	28	11	28
4	Cubic	1000	16	45	16	45	16	45
		5000	16	45	16	45	16	45
		10,000	16	45	16	45	16	45
5	Denschnb	1000	6	15	6	15	6	15
		5000	6	15	6	15	6	15
		10,000	6	15	6	15	6	15
6	Denschnf	1000	12	26	12	26	12	26
		5000	13	28	13	28	13	28
		10,000	15	31	15	31	15	31
7	Diagonal1	1000	32	71	32	71	32	71
		5000	F	F	52	123	51	121
		10,000	93	242	F	F	92	236
8	DiagonalI3	1000	24	49	24	49	24	49
		5000	54	110	54	110	53	108
		10,000	84	184	84	175	83	170
9	Diagonal4	1000	2	6	2	6	2	6
		5000	2	6	2	6	2	6
		10,000	2	6	2	6	2	6
10	Dixmaan A	1000	6	15	6	15	6	15
		5000	6	15	6	15	6	15
		10,000	5	13	5	13	5	13

Continued

11	Dixmaan E	1000	43	112	43	112	43	112
		5000	68	193	68	193	68	193
		10,000	115	335	116	338	111	305
12	Dixmaan I	1000	43	117	43	117	43	117
		5000	68	191	F	F	65	173
		10,000	111	327	F	F	110	322
13	Dqdrtic	1000	32	65	32	65	32	65
		5000	32	65	32	65	32	65
		10,000	32	65	32	65	32	65
14	Extended EP1function	1000	4	10	4	10	4	10
		5000	4	10	4	10	4	10
		10,000	4	10	4	10	4	10
15	Extended cliff	1000	6	29	6	29	6	29
		5000	6	29	6	29	6	29
		10,000	6	29	6	29	6	29
16	Exhimmelbau	1000	26	276	26	276	24	268
		5000	8	1138	8	416	7	382
		10,000	8	390	8	400	7	278
17	Ex tri2	1000	49	150	46	129	45	103
		5000	57	1372	50	314	46	339
		10,000	44	235	58	935	41	340
18	Ex Wood	1000	248	503	220	447	161	329
		5000	210	427	200	407	166	339
		10,000	207	421	204	416	171	349
19	Hager	1000	26	54	26	53	26	54
		5000	29	59	29	62	29	59
		10,000	77	5360	F	F	70	263
20	Helical	1000	65	134	58	121	43	90
		5000	68	140	58	121	43	90
		10,000	68	140	58	121	43	90
21	Miele	1000	134	510	146	521	108	368
		5000	141	549	150	543	120	419
		10,000	145	569	160	593	108	369
22	Nond	1000	30	78	30	78	30	78
		5000	30	78	30	78	30	78
		10,000	30	78	30	78	30	78
23	OSP	1000	197	758	195	714	149	540
		5000	329	1159	298	1041	297	1011
		10,000	401	1353	386	1342	383	1318
24	Powell 3	1000	31	66	27	58	26	56
		5000	32	68	28	61	27	58
		10,000	32	68	28	61	27	58
25	Powell4	1000	F	F	212	485	197	483
		5000	F	F	293	660	230	530
		10,000	F	F	293	660	230	530
26	Wood	1000	204	415	266	539	175	357
		5000	266	539	237	481	177	361
		10,000	246	499	243	493	191	389

Table 4. The percentage performance of the proposed methods.

Measures	β_k^{MMWU}	β_k^{RMAR}	β_k^{FG}
NOI	100%	99.2%	71.3%
NOF	100%	92.4%	60.0%

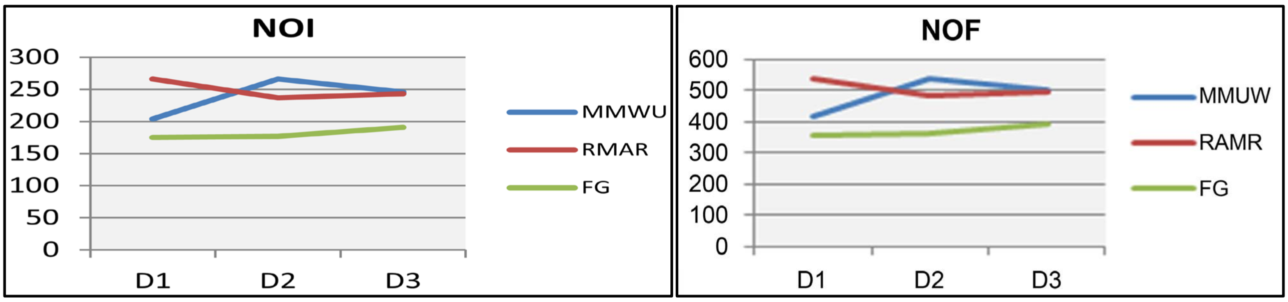

Figure 1. The comparison between the three methods.

or generalized form. Each problem was tested three times for a gradually increasing number of variables: $N =$ 1000, 5000 and 10,000, all algorithms implemented the strong Wolfe line search (3) and (4) conditions with $\sigma = 0.001$ and $\delta = 0.9$ and the same stopping criterion $\|g_k\| \leq 10^{-6}$ is used.

In some cases, the computation stopped due to the failure of the line search to find the positive step size, and thus it was considered as a failure denoted by (F).

We record the number of iteration calls (NOI), the number of function evaluations calls (NOF), and the dimensions of test problems calls (N), for the purpose of our comparisons.

Table 3 gives the comparison depending in the NOI and NOF between β_k^{MMWU}, β_k^{RMAR} and the proposed method β_k^{FG}.

Table 4 gives the percentage performance of the proposed methods β_k^{FG} against β_k^{MMWU} and β_k^{RMAR}. We have seen that β_k^{RMAR}. Method saves (NOI 0.8%), (NOF 7.6%), and β_k^{FG} method saves (NOI 28.7%), (NOF 40.0%) compared with β_k^{MMWU} method.

While Figure 1 gives the comparison between β_k^{MMWU}, β_k^{RMAR} and β_k^{FG}, using a well-known Wood test function.

Conflicts of Interest

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

References

[1] Liu, A.J. and Li, S. (2014) New Hybrid Conjugate Gradient Method for Unconstrained Optimization. *Applied Mathematics and Computation*, **245**, 36-43. https://www.sciencedirect.com/science/article/abs/pii/S0096300314010637 https://doi.org/10.1016/j.amc.2014.07.096

[2] Salleh, Z. and Alhawarat, A. (2016) An Efficient Modification of the Hestenes-Stiefel Nonlinear Conjugate Gradient Method with Restart Property. *Journal*

of Inequalities and Applications, **2016**, Article No. 110.
https://link.springer.com/article/10.1186/s13660-016-1049-5
https://doi.org/10.1186/s13660-016-1049-5

[3] Wolfe, P. (1969) Convergence Condition for Ascent Methods. *SIAM Review*, **11**, 226-235. https://doi.org/10.1137/1011036

[4] Wolfe, P. (1971) Convergence Conditions for Ascent Methods II: Some Corrections. *SIAM Review*, **13**, 185-188. https://doi.org/10.1137/1013035

[5] Alhawarat, A., Mamat, M., Rivaie, M. and Salleh, Z. (2015) An Efficient Hybrid Conjugate Gradient Method with the Strong Wolfe-Powell Line Search. *Mathematical Problems in Engineering*, **2015**, Article ID: 103517.
https://www.hindawi.com/journals/mpe/2015/103517
https://doi.org/10.1155/2015/103517

[6] Hestenes, M.R. and Stiefel, E. (1952) Methods of Conjugate Gradients for Solving Linear Systems. *Journal of Research of the National Bureau of Standards*, **49**, 409-436. https://nvlpubs.nist.gov/nistpubs/jres/049/jresv49n6p409_A1b.pdf
https://doi.org/10.6028/jres.049.044

[7] Fletcher, R. and Reeves, C. (1964) Function Minimization by Conjugate Gradients. *The Computer Journal*, **7**, 149-154. https://doi.org/10.1093/comjnl/7.2.149
https://academic.oup.com/comjnl/article/7/2/149/335311

[8] Polak, E. and Ribie're, G. (1969) Note sur la convergence de me'thodes de directions conjugue's. *Revue Francsis d'Infermatique et de recherché Operationnelle*, **3**, 35-43.
http://www.numdam.org/item?id=M2AN_1969__3_1_35_0
https://doi.org/10.1051/m2an/196903R100351

[9] Polyak, B.T. (1969) The Conjugate Gradient Method in Extreme Problems. *USSR Computational Mathematics and Mathematical Physics*, **9**, 94-112.
https://www.researchgate.net/publication/222365587
https://doi.org/10.1016/0041-5553(69)90035-4

[10] Fletcher, R. (1987) Practical Methods of Optimization. 2nd Edition, John Wiley & Sons, Inc., Hoboken. https://www.wiley.com/en-ba

[11] Liu, D. and Story, C. (1991) Efficient Generalized Conjugate Gradient Algorithms. Part 1: Theory. *Journal of Optimization Theory and Applications*, **69**, 129-137.
https://link.springer.com/article/10.1007/BF00940464
https://doi.org/10.1007/BF00940464

[12] Dai, Y.H. and Yuan, Y. (1999) A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. *SIAM Journal on Optimization*, **10**, 177-182.
https://doi.org/10.1137/S1052623497318992

[13] Al-Naemi, Gh.M. and Hamed, E.T. (2013) New Conjugate Method with Wolfe Type Line Searches for Nonlinear Programming. *Australian Journal of Basic and Applied Sciences*, **7**, 622-632. https://www.researchgate.net/publication/330686295

[14] Andrei, N. (2010) Acceleration Hybrid Conjugate Gradient Algorithm with Modified Secant Condition for Unconstrained Optimization. *Numerical Algorithms*, **54**, 23-46. https://link.springer.com/article/10.1007/s11075-009-9321-0
https://doi.org/10.1007/s11075-009-9321-0

[15] Andrei, N. (2008) Another Nonlinear Conjugate Gradient Algorithm for Unconstrained Optimization. *Optimization Methods and Software*, **24**, 89-104.
https://www.tandfonline.com/doi/abs/10.1080/10556780802393326
https://doi.org/10.1007/s11075-007-9152-9

[16] Yan, H., Chen, L. and Jiao, B. (2009) HS-LS-CD Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization. 2*nd International Workshop on Computer*

Science and Engineering, Vol. 1, Qingdao, 28-30 October 2009, 264-268. https://ieeexplore.ieee.org/document/5403315 https://doi.org/10.1109/WCSE.2009.667

[17] Li, S. and Sun, Z.B. (2010) A New Hybrid Conjugate Gradient Method and Its Global Convergence for Unconstrained Optimization. *International Journal of Pure and Applied Mathematics*, **63**, 84-93. https://ijpam.eu/contents/2010-63-3/4/index.html

[18] Djordjević, S.S. (2019) New Hybrid Conjugate Gradient Method as a Convex Combination of LS and CD Methods. *Filomat*, **31**, 1813-1825. https://doi.org/10.2298/FIL1706813D

[19] Djordjevic, S.S. (2018) New Hybrid Conjugate Gradient Method as a Convex Combination of HS and FR Conjugate Gradient Methods. *Journal of Applied Mathematics and Computation*, **2**, 366-378. https://doi.org/10.26855/jamc.2018.09.002 https://www.researchgate.net/publication/328125868

[20] Djordjević, S.S. (2019) New Hybrid Conjugate Gradient Method as a Convex Combination of LS and FR Conjugate Gradient Methods. *Acta Mathematica Scientia*, **39**, 214-228.

[21] Zheng, X.Y., Dong, X.L., Shi, J.R. and Yang, W. (2019) Further Comment on Another Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization by Andrei. *Numerical Algorithm*, 1-6. https://doi.org/10.1007/s11075-019-00771-1

[22] Abdullahi, I. and Ahmad, R. (2016) Global Convergence Analysis of a Nonlinear Conjugate Gradient Method for Unconstrained Optimization Problems. *Indian Journal of Science and Technology*, **9**, 1-9. http://www.indjst.org/index.php/indjst/article/view/90175/74779 https://doi.org/10.17485/ijst/2016/v9i41/90175

[23] Livieris, I.E., Tampakas, V. and Pintelas, P. (2018) A Descent Hybrid Conjugate Gradient Method Based on the Memoryless BFGS Update. *Numerical Algorithms*, **79**, 1169-1185. https://link.springer.com/article/10.1007/s11075-018-0479-1 https://doi.org/10.1007/s11075-018-0479-1

[24] Mandara, A.V., Mamat, M., Waziri, M.Y., Mohamed, M.A. and Yakubu, U.A. (2018) A New Conjugate Gradient Coefficient with Exact Line Search for Unconstrained Optimization. *Far East Journal of Mathematical Sciences*, **105**, 193-206. https://www.researchgate.net/publication/329522075 https://doi.org/10.17654/MS105020193

[25] Liu, J.K. and Li, S.J. (2014) New Hybrid Conjugate Gradient Method for Unconstrained Optimization. *Applied Mathematics and Computation*, **245**, 36-43. https://doi.org/10.1016/j.amc.2014.07.096 https://www.sciencedirect.com/science/article/abs/pii/S0096300314010637

[26] Yunus, R.B., Mamat, M. and Abashar, A. (2018) Comparative Study of Some New Conjugate Gradient Methods. *UniSZA Research Conference* (*URC* 2015), Kuala Terengganu, 14-16 April 2015, 616-621. https://www.researchgate.net/publication/326301618

[27] Al-Naemi, Gh.M. and Ahmed, H.I. (2013) Modified Nonlinear Conjugate Gradient Algorithms with Application in Neural Networks. LAP Lambert Academic Publishing, Saarbrucken.

[28] Andrei, N. (2008) An Unconstrained Optimization Test Functions Collection. *Advanced Modeling and Optimization*, **10**, 147-161. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.665.3152&rep=rep1&type=pdf

[29] Andrei, N. (2014) Test Functions for Unconstrained Global Optimization. 3-5.

[30] Bongartz, I., Conn, A.R., Gould, N. and Toint, P.L. (1995) CUTE: Constrained and Unconstrained Testing Environment. *ACM Transactions on Mathematical Software*, **21**, 123-160. https://doi.org/10.1145/200979.201043