A new hybrid conjugate gradient algorithm for unconstrained optimization with inexact line search

Fanar N. Jardow, Ghada M. Al-Naemi
Department of Mathematics, College of Computer Sciences and Mathematics, University of Mosul, Iraq

ABSTRACT
Many researchers are interested in developing and improving the conjugate gradient method for solving large-scale unconstrained optimization problems. In this work, a new parameter will be presented as a convex combination between RMIL and MMWU. The suggested method always produces a descent search direction at each iteration. Under Strong Wolfe Powell (SWP) line search conditions, the global convergence of the proposed method is established. The preliminary numerical comparisons with some other CG methods have shown that this new method is efficient and robust in solving all given problems.

Keywords:
Global convergence
Hybrid conjugate gradient method
Sufficient descent
SWP
Unconstrained optimization

1. INTRODUCTION
Conjugate gradient (CG) methods are among the most popular methods for solving optimization problems, especially for large-scale problems due to the simplicity and low storage of their iterative form [1]. The unconstrained optimization problem has the following general form:

\[ \min_{x \in \mathbb{R}^n} z(x) \]

where \( x \in \mathbb{R}^n \) is a real vector with \( n \geq 1 \) component and \( z: \mathbb{R}^n \rightarrow \mathbb{R} \) is smooth function and its gradient \( g \) is available. The nonlinear CG method that starts from an initial guess \( x_0 \in \mathbb{R}^n \) will be defined using the iterations of the sequence as in the following form:

\[ x_{n+1} = x_n + \rho_n d_n, \quad n = 0, 1, 2, 3, ... \]

where \( x_n \) is the \( n \)-th iterative point and \( \rho_n \) is the positive step size resulting from performing a one-dimensional search, known as the line searches [2]. The \( d_n \) is the direction of the search that is computed by

\[ d_{n+1} = \begin{cases} -g_{n+1} & n = 0 \\ -g_{n+1} + \beta_n d_n & n \geq 1 \end{cases} \]

where \( g_n \) record by \( \nabla z(x_n) \) is the gradient and the \( \beta_n > 0 \) is a scalar known as the CG-coefficient, the different choices for the parameter \( \beta_n \) correspond to different conjugate gradient methods. The step \( \rho_n \) length is very important for the global convergence of CG methods. It can either be exact or inexact. In the case of an exact steplength, one seeks \( \rho_n \) along the direction \( d_n \) such that
\[ f(x_n + \rho_n d_n) = \min_{\rho > 0} f(x_n + \rho d_n) \]

For inexact \( \rho_n \) a number of line search techniques can be used. For instance, the so-called SWP condition require that \[ z(x_n + \rho_n d_n) \leq z(x_n) + \epsilon_1 \rho_n d_n \] (4)

\[ |g(x_n + \rho_n d_n)| \leq \epsilon_2 |g_n| \] (5)

where \( 0 < \epsilon_1 < \frac{1}{2} < \epsilon_2 < 1 \), when \( x_n \) is far from the solution an approximation of \( \rho_n \) is found as the descending characteristic must be satisfied and the direction should not be searched. Thus by SWP we inherit the advantages of exact line search with inexpensive and low computational cost [5].

Different CG methods correspond to different choices of the parameter \( \beta_n \) [6]. The most popular formulas for parameters Hestenes Stiefel method (HS) [7], Fletcher-Reeves method (FR) [8], Polak-Ribiere – Polyak method (PR) [9, 10], conjugate – Descent method (CD) [11], Liu – Storey method (LS) [12], and Dai-Yuan method (DY) [13]. The parameters of these \( \beta_n \) are given as follows:

\[
\begin{align*}
\beta_{HS}^n &= \frac{\gamma_k + \eta_k}{\gamma_k} & \beta_{PR}^n &= \frac{\gamma_k + \eta_{k+1}}{\gamma_k} & \beta_{PRP}^n &= \frac{\gamma_k + \eta_{k+1}}{\gamma_k} \\
\beta_{CD}^n &= \frac{\gamma_k + \eta_{k+1}}{\gamma_k} & \beta_{LS}^n &= \frac{\gamma_k + \eta_{k+1}}{\gamma_k} & \beta_{DY}^n &= \frac{\gamma_k + \eta_{k+1}}{\gamma_k}
\end{align*}
\]

For a strictly convex quadratic function \( z(x) \), and the line search is exact, all these methods are identical, since the gradients are mutually orthogonal, so the parameters in these methods are equal. When implemented to general nonlinear function with inexact line searches, yet, the behavior of these methods is seeming different [14]. One of an important group of CG methods is the hybrid conjugate gradient algorithms, the hybrid computational schemes HCG work better than the classical CG methods because the HCG take the advantages of the two parameters \( \beta_n \) [15].

Many researchers devoted to the hybrid or mixed conjugate gradient methods which have better computational performances and strong convergence properties. Andrei [16] proposed the following hybrid method: \( \beta_n^\text{hyb} = (1 - \theta_n)\beta_n^\text{HS} + \theta_k \beta_n^\text{DY} \). Djordjevic’ [17] proposed the following HCG method \( \beta_n^\text{HCG} = (1 - \theta_n)\beta_n^\text{HS} + \theta_k \beta_n^\text{DY} \). Xiuyun, et al [18], proposed the following HCG method \( \beta_n^\text{HCG} = \lambda \beta_n^\text{HS} + (1 - \lambda) \beta_n^\text{PR} \). Livieris, et al [19], proposed the following HCG method \( \beta_n^\text{HCG} = \lambda \beta_n^\text{HS} + (1 - \lambda) \beta_n^\text{PR} \). Al-Namat et al [20], proposed the following HCG method \( \beta_n^\text{HCG} = (1 - \theta_n)\beta_n^\text{RMIL} + \theta_k \beta_n^\text{RMAR} \).

In this work we focus on hybrid conjugate gradient methods as a convex combination of RMI and MMWU [21, 22]. CG methods for solving unconstrained optimization method with suitable conditions. The corresponding conjugate gradient (CG) parameters are:

\[
\beta_n^\text{RMIL} = \frac{\gamma_k + \eta_n}{\|d_n\|^2}
\]

and

\[
\beta_n^\text{MMWU} = \frac{\|g_{n+1}\|^2}{\|d_n\|^2}
\]

The proposed method defined by set the parameter \( \beta_n \) by:

\[
\beta_n^\text{HA} = (1 - \theta_n)\beta_n^\text{RMIL} + \theta_k \beta_n^\text{MMWU}
\]

Choosing the appropriate value of the \( \theta_n \) in the convex combination, the search direction \( d_n \) of our algorithm not only is the Newton direction [23], so satisfies the famous DL conjugate condition proposed by Dai and Liao [24]. Under the SWP conditions, we prove the global convergence of the proposed algorithm, the numerical results also show the feasibility and activity of our algorithm. This study is organized as follows, Section 2 we introduce the new proposed hybrid CG method (HHA), and we got the parameter \( \theta_n \) using some approaches and give us specific algorithm. Section 3, we prove that it generates direction satisfying the sufficient descent condition under SWP condition. Section 4, The global convergence property of the proposed method is established. in Section 5, Some numerical results are reported.
2. A NEW HYBRID CONJUGATE GRADIENT METHOD

In this section, we will describe a new proposed HCG method, in order to get the sufficient descent direction, we will compute \( \theta_n \) as follows: we combine \( \beta_n^\text{RMIL} \) and \( \beta_n^\text{MMWU} \) in (8). The direction \( d_{n+1} \) are generated by:

\[
d_{n+1} = -g_{n+1} + \beta_n^\text{HHA} d_n
\]

(9)

The iterates \( x_1, x_2, x_3, \ldots \) of the proposed method are computed by means of the recurrence (2), where the step size \( \rho_n \) is definition according to the SWP conditions (4) and (5). The scale parameter \( \theta_n \) in (8) satisfying \( 0 \leq \theta_n \leq 1 \), which will be determined a specific way to be described later. If \( \theta_n \leq 0 \), then \( \beta_n^\text{HHA} = \beta_n^\text{RMIL} \), and if \( \theta_n \geq 1 \), then \( \beta_n^\text{HHA} = \beta_n^\text{MMWU} \). On the other hand, if \( 0 < \theta_n < 1 \), then \( \beta_n^\text{HHA} \) is a convex combination of \( \beta_n^\text{RMIL} \) and \( \beta_n^\text{MMWU} \). From (8) and (9) it is clear that:

\[
d_{n+1} = \begin{cases} 
-g_{n+1}, & n = 1 \\
-g_{n+1} + (1 - \theta_n) \frac{\partial^T \frac{\partial g_{n+1}}{\partial n} }{\Vert \partial n \Vert^2} d_n + \theta_n \frac{\Vert g_{n+1} \Vert^2}{\Vert \partial n \Vert^2} d_n, & n > 1
\end{cases}
\]

(10)

Our motivation to select the parameter \( \theta_n \) in such a manner that the deflection \( d_{n+1} \) given (10) is equal to the Newton direction \( d_n^N = -\nabla^2 f(x_{n+1})^{-1} g_{n+1} \). Therefore

\[
-\nabla^2 f(x_{n+1})^{-1} g_{n+1} = -g_{n+1} + (1 - \theta_n) \frac{\partial^T \frac{\partial g_{n+1}}{\partial n} }{\Vert \partial n \Vert^2} d_n + \theta_n \frac{\Vert g_{n+1} \Vert^2}{\Vert \partial n \Vert^2} d_n
\]

\[
-\nabla^2 f(x_{n+1})^{-1} g_{n+1} = -g_{n+1} + \frac{\partial^T \frac{\partial g_{n+1}}{\partial n} }{\Vert \partial n \Vert^2} d_n - \theta_n \frac{\Vert g_{n+1} \Vert^2}{\Vert \partial n \Vert^2} d_n
\]

(11)

Therefore, in order to have an algorithm for solving large scale problems we assume that pair \((s_n, y_n)\) satisfies the secant equation, \( y_n = \nabla^2 f(x_{n+1})s_n \) so,

\[
s_n^T \nabla^2 f(x_{n+1}) = y_n^T
\]

(12)

Multiplying the as shown in (11) by \( s_n^T \nabla^2 f(x_{n+1}) \) from the left and denoting \( \theta_n^\text{HHA} = \theta_n \), we get

\[
-s_n^T g_{n+1} = -s_n^T \nabla^2 f(x_{n+1}) g_{n+1} + \frac{\partial^T \frac{\partial g_{n+1}}{\partial n} }{\Vert \partial n \Vert^2} s_n^T \nabla^2 f(x_{n+1}) d_n - \theta_n^\text{HHA} \frac{\Vert g_{n+1} \Vert^2}{\Vert \partial n \Vert^2} s_n^T \nabla^2 f(x_{n+1}) d_n
\]

\[
-s_n^T g_{n+1} = -y_n^T g_{n+1} + \frac{\partial^T \frac{\partial g_{n+1}}{\partial n} }{\Vert \partial n \Vert^2} y_n^T d_n - \theta_n^\text{HHA} \frac{\Vert g_{n+1} \Vert^2}{\Vert \partial n \Vert^2} y_n^T d_n
\]

(11)

After some algebra, we get

\[
\theta_n^\text{HHA} = \frac{(s_n^T g_{n+1} - y_n^T g_{n+1}) \Vert \partial n \Vert^2 + (\partial^T \frac{\partial g_{n+1}}{\partial n} ) (y_n^T d_n)}{(g_n^T g_n)(y_n^T d_n)}
\]

(13)

We will specify a complete (HHA) which posses some nice properties of CG and Newton method.

**ALGORITHM HHA**

Step 1 : choose \( x_0 \in \mathbb{R}^n \), \( \varepsilon > 0 \), Calculate \( f(x_0) \) and \( g_0 = -\nabla f(x_0) \), set \( d_0 = -g_0 \), when \( n = 0 \).

Step 2 : The stopping criteria, i.e. if \( \Vert g_n \Vert \leq \varepsilon \), then stop.

Step 3 : Calculate \( \rho_n \) by SWP conditions in (3) & (4).

Step 4 : Calculate \( x_{n+1} = x_n + \rho_n d_n \), and \( g_{n+1} = g(x_{n+1}) \).

Step 5 : If \( \theta_n \geq 1 \) then put \( \theta_n = 1 \). If \( \theta_n \leq 0 \), then put \( \theta_n = 0 \), otherwise calculate \( \theta_n \) as (13).

Step 6 : Calculate \( \beta_n^\text{HHA} \) by (8).

Step 7 : Generate \( d_n = -g_n + \beta_n^\text{HHA} d_n \).

Step 8 : If the restart criteria of Powell \( |g_n^T g_n| \geq 0.2 |g_{n+1}|^2 \) is satisfied, then set \( d_{n+1} = -g_{n+1} \), otherwise put \( d_{n+1} = d \).

Step 9 : put \( n = n + 1 \), and go to step 2.
3. THE SUFFICIENT DESCENT CONDITION

In this section, we use to the following theorem to clear up that the search direction \(d_n\) obtained by HHA satisfies the sufficient descent condition which plays of role in analyzing the global convergence. For further considerations we need the assumptions below:

3.1. Assumption

The level sets \(Q = \{x : f(x) \leq f(x_1)\}\) at \(x_1\) is bounded where \(x_1\) is starting point, namely, that there exists \(M > 0\), such that \(|x| \leq M, \forall x \in Q\) [25].

3.2. Assumption

In a neighborhood \(N\) of \(Q\), the function \(z\) is continuously differentiable and its gradient is Lipschitz continuous, i.e, there exists a constant \(L > 0\), such that

\[ \|\nabla z(x) - \nabla z(y)\| \leq L\|x - y\|, \forall x, y \in N. \]

Under assumptions (3.1) and (3.2) , there exists positive constant \((u, \bar{u}, \varepsilon \text{ and } e^-)\), such that:

\[ \bar{u} \leq \|g_{n+1}\| \leq u, \text{and } \bar{e} \leq \|g_n\| \leq e \ \forall x \in Q \] [25].

3.3. Theorem

Let generated the sequences \(\{g_n\}\) and \(\{d_n\}\) by a HHA method.then \(d_n\) is the search direction satisfies the sufficient descent condition:

\[ g_{n+1}^T d_{n+1} \leq -\tau \|g_{n+1}\|^2, \forall \tau \geq 0 \]

with \(\tau = [\tau_3 \tau_2 + (1 - \tau_3)\tau_1] \)

3.4. Proof.

We show that search direction \(d_n\) shall satisfies the sufficient descent condition holds for \(n = 0\), the proof is a trivial one, i.e. \(d_0 = g_0\) and so \(g_0^T d_0 = -\|g_0\|^2\). Now we have

\[ d_{n+1} = -g_{n+1} + \beta_n^{HHA} d_n, \]

\[ d_{n+1} = -g_{n+1} + [(1 - \theta_n)\beta_n^{RMIL} + \theta_n \beta_n^{MMWU}] d_n. \]

We can rewrite the direction by the followin below:

\[ d_{n+1} = -((1 - \theta_n)g_{n+1}) + ((1 - \theta_n)\beta_n^{RMIL} + \theta_n \beta_n^{MMWU}) d_n. \]

The above equation can be written after arrange the terms as:

\[ d_{n+1} = \theta_n(-g_{n+1} + \beta_n^{MMWU} d_n) + (1 - \theta_n)(-g_{n+1} + \beta_n^{RMIL} d_n). \]

produces after some arrangement

\[ d_{n+1} = \theta_n d_n^{MMWU} + (1 - \theta_n) d_n^{RMIL}. \]

produces after multiplying the (15) from the left by \(g_{n+1}^T\), we get

\[ g_{n+1}^T d_{n+1} = \theta_n g_{n+1}^T d_n^{MMWU} + (1 - \theta_n) g_{n+1}^T d_n^{RMIL}. \]

Firstly, if \(\theta_n = 0\), then \(d_{n+1} = d_n^{RMIL}\), in [21] they proved that the sufficient descent condition holds with exact line search. We are going to prove that the sufficient descent condition holds for RMIL when inexact line search is used

\[ g_{n+1}^T d_{n+1}^{RMIL} \leq -\tau_1 \|g_{n+1}\|^2, \]

where \(\tau_1 = -(1 - 0.8 \rho_n L > 0, \text{ with } 0 < L < \frac{1}{0.8 \rho_n} \).
Now let $\theta_n = 1$ then $d_n = d_{n}^{MMWU}$, in [22] they proved that the sufficient descent condition holds with exact line search. In [20], they proved that the sufficient descent condition holds with exact line search.

$$ g_{n+1}^T d_{n+1}^{MMWU} \leq -\tau_2 ||g_{n+1}||^2, \tag{18} $$

where $\tau_2 = (1 - \rho_n L) > 0$, with $0 < L < \frac{1}{\rho_n}$. Now, we are going to prove the direction satisfy the sufficient descent condition when $0 < \theta_n < 1$, we have $g_{n+1}^T s_n \leq \gamma_n^T s_n \leq L ||s_n||^2$, and $y_n = (g_{n+1} - g_n)$, then (13) become

$$ \theta_n^{HA} \leq \frac{(L||v||^2 - (\gamma_n^T s_n)^2 - \gamma_n^T g_n) ||d_n||^2 + L ||g_{n+1}||^2 - \gamma_n^T g_n}{\theta_n^T g_n}, \tag{19} $$

we have $v(1 - \varepsilon) ||g_n|| \leq y_n^T d_n$, $y_n = (g_{n+1} - g_n)$. $||g_{n+1}|| \geq 0.2 ||g_n||^2$, and we know that

$$ s_n = \rho_n d_n \Rightarrow d_n = \frac{s_n}{\rho_n} \Rightarrow ||d_n|| \leq \frac{||s_n||}{||\rho_n||} \leq \frac{||x_{n+1} - x_n||}{||\rho_n||} \leq \frac{M}{||\rho_n||} = D, $$

put the above in (19) become

$$ \theta_n^{HA} \leq \frac{1}{0.2} \left( \frac{(LM^2 - 0.8u^2)\rho^2}{v(1 - \varepsilon) ||d_n||^2} + 0.8 \right) = \tau_3 \tag{20} $$

From (15), (17), (18), and (20) we get

$$ : g_{n+1}^T d_{n+1} \leq -[\tau_3 \tau_2 + (1 - \tau_3)\tau] ||g_{n+1}||^2 $$

$$ : g_{n+1}^T d_{n+1} \leq -\tau ||g_{n+1}||^2, \text{ with } \tau = [\tau_3 \tau_2 + (1 - \tau_3)\tau_1]. $$

So, it is proved that $d_{n+1}$ satisfied the sufficient descent condition

4. CONVERGE ANALYSIS

Let Assumption (3.1) and (3.2) hold. In [26] it is proved that for any conjugate gradient method with SWP conditions, it holds:

4.1. Lemma

Let Assumption (3.1) and (3.2) holds. Consider the method (2) and (5) where the $d_n$. Is a descent direction and $\rho_n$ is received from the SWP. If

$$ \sum_{n=1}^{\infty} \frac{1}{||d_n||^2} = \infty. $$

then

$$ \lim_{n \to \infty} \inf ||g_n|| = 0. $$

4.2. Theorem

Suppose that assumption (3.1) and (3.2) holds. Consider the algorithm HHA were $0 \leq \theta_n \leq 1$, and $\rho_n$ is obtained by the strong wolfe line search and $d_{n+1}$ is the descent direction. Then

$$ \lim_{n \to \infty} \inf ||g_n|| = 0 $$

4.3. Proof

Because the descent condition holds, we have $d_{n+1} \neq 0$. So using lemma 4.1, it is sufficient to prove that $||d_{n+1}||$ is bounded above. From (10).

$$ ||d_{n+1}|| = ||g_{n+1} + [(1 - \theta_n)\beta_n^RMIL + \theta_n \beta_n^{MMWU}]d_n|| $$

$$ \leq ||g_{n+1}|| + ||(1 - \theta_n)||\beta_n^RMIL|| + ||\theta_n||\beta_n^{MMWU}|| ||d_n|| $$

A new hybrid conjugate gradient algorithm for unconstrained optimization with .... (Fanar N. Jardow)
They proved that in [21] and [22], that

$$|\beta^{RMIL}_n| \leq \frac{\|g_n+1\|_2 \gamma_n}{\|d_n\|_2^2} \leq \frac{u_A}{\beta^2} = G_1,$$

And

$$|\beta^{MMWU}_n| = \frac{\|g_{n+1}\|_2^2}{\|d_n\|_2^2} \leq \frac{u^2}{\beta^2} = G_2.$$  

Now, we have

$$\| \theta_n \| = \left( \frac{g^T_{n+1} y_n - \gamma_n g_{n+1}}{(h_{n+1} y_n) (y_n^T d_n)} \right).$$

Using SWC, we get $-(1 - \varepsilon_1) \gamma_n \|d_n\| \leq \gamma_n^T d_n \leq (1 + \gamma_n) \|d_n\|$.  

$$\| \theta_n \| \leq \left( \frac{(g^T_{n+1} y_n - \gamma_n g_{n+1}) \|d_n\|^2 + \gamma_n^T \rho_n \|d_n\|^2}{0.2 \|d_n\|^2 (1 - \varepsilon_1) \gamma_n \|d_n\|^2} \right)$$

$$\leq \frac{(1 + L) \gamma_n \|d_n\|^2 + \gamma_n^2 \|d_n\|^2}{0.2 (1 - \varepsilon_1) \gamma_n \|d_n\|^2} = G_3$$

$\therefore |\theta_n| \leq G_3$

$\|d_{n+1}\| \leq \|g_{n+1}\| + [(1 - G_3) G_2 + G_3 G_2] \|d_n\|$

$\leq u + GB = \varphi$

$$\Rightarrow \sum_{n=1}^{\infty} \frac{1}{\|d_n\|^2} = \frac{1}{\varphi^2} \sum_{n=1}^{\infty} \frac{1}{\varphi^2} = \infty$$

$$\Rightarrow lim_{n \to \infty} \inf \|g_n\| = 0.$$  

5. NUMERICAL EXPERIMENTS

In this section we selected some of test functions from CUTE [27] library, along with other large scale optimization problems presented in Andrei [28] and Bongartz [29]. 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 $\rho_n$ is always compute by cubic fitting procedure.

We selecte (24) large scale unconstrained optimization problems in the extended or generalized form. Each problem was tested three times for a gradually increasing number of variables: $N = 1000, 5000$ and $10000$, all algorithms implemented the SWP (3) and (4) conditions with $\varepsilon_1 = 0.001$ and $\varepsilon_2 = 0.9$ and the stopping criterion $\|g_n\| \leq 10^{-6}$ 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 (ni), the number of function evaluations calls (nf), and the of test problems calls (N), for purpose of our comparisons. Table 1 gives the comparison depending in the ni and nf between $\beta^{RMIL}_n$, $\beta^{MMWU}_n$ and the proposed method $\beta^{HA}_n$.  

Table 2 gives the percentage performance of the proposed methods $\beta^{HA}_n$ against $\beta^{RMIL}_n$ and $\beta^{MMWU}_n$. We have seen that $\beta^{RMIL}_n$ method saves (ni 9.94%), (nf 17.11%), and $\beta^{HA}_n$ method saves (ni 53.42%), (nf 36.01%) compared with $\beta^{MMWU}_n$ method. While Figure 1 gives the comparison between $\beta^{RMIL}_n$, $\beta^{MMWU}_n$ and $\beta^{HA}_n$, using Well-known EX-Wood test function.
A new hybrid conjugate gradient algorithm for unconstrained optimization with .... (Fanar N. Jardow)


6. CONCLUSION

In this paper, a new parameter $\theta_n$ for a hybrid conjugate gradient is derived. The practical results indicated that the proposed hybrid method is faster and more efficient compared to the $\beta_n^{BMW}$ and $\beta_n^{MMWU}$ algorithms used.

ACKNOWLEDGEMENTS

The authors are very grateful to the University of Mosul, College of Computer Science and Mathematics for their provided facilities, which helped to improve the quality of this work.

REFERENCES

[1] J. K. Liu and S. J. Li, “New hybrid conjugate gradient method for unconstrained optimization,” Applied Mathematics and Computation, vol. 245, pp. 36-43, 2014.

[2] Z. Salleh and A. Alhawarat, “An efficient modification of the Hestenes-Siefel nonlinear conjugate gradient method with restart property,” Journal of Inequalities and Applications, vol. 110, 2016.

[3] P. Wolfe, “Convergence conditions for ascent methods,” SIAM review, vol. 11, no. 2, pp. 226-235, 1969.

[4] P. Wolfe, “Convergence Conditions for Ascent Methods. II: Some Corrections,” SIAM Review, vol. 13, no. 2, pp. 185-188, 1971.

[5] A. Alhawarat, et al., “An efficient hybrid conjugate gradient method with the strong Wolfe-Powell line search,” Mathematical Problems in Engineering, vol. 2015, no. 14, pp. 1-7, 2015.

[6] C. Li, “A Modified Conjugate Gradient Method for Unconstrained Optimization,” TELKOMNIKA, vol. 11, no. 11, pp. 6373-6380, 2013.

[7] M. R. Hestenes and E. Siefel, “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] R. Fletcher and C. M. Reeves, “Function minimization by conjugate gradients,” The计算机 journal, vol. 7, no. 2, pp. 149-154, 1964.

[9] 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. R1, pp. 35-43, 1969.

[10] B. T. Polyak, “The conjugate gradient method in extreme problems,” USSR Computational Mathematics and Mathematical Physics, vol. 9, no. 4, pp. 94-112, 1969.

[11] R. Fletcher, “Practical methods of optimization,” John Wiley & Sons, 1987.

[12] 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.

[13] 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.

[14] G. M. Al-Naemi and E. T. Hamed, “New conjugate gradient method with Wolfe type line searches for nonlinear programming,” Australian Journal of Basic and Applied Sciences, vol. 7, no. 14, pp. 622-632, 2013.

[15] N. S. Mohamed, et al., “A New Hybrid Coefficient of Conjugate Gradient method,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 18, no. 3, pp. 1454-1463, 2020.

[16] N. Andrei, “Another nonlinear conjugate gradient algorithm for unconstrained optimization,” Optimization methods & software, vol. 24, no. 1, pp. 89-104, 2009.
A new hybrid conjugate gradient algorithm for unconstrained optimization with ... (Fanar N. Jardow)