The convergence properties of a new hybrid conjugate gradient parameter for unconstrained optimization models

I M Sulaiman¹, M Mamat¹, *, M Y Waziri², U A Yakubu³, M Malik¹

¹Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.
²Department of Mathematics, Faculty of Physical Sciences, Bayero University, Kano, Nigeria.
³Department of Mathematics, Faculty of Sciences, Maitama Sule University, Kano, Nigeria

Corresponding author: must@unisza.edu.my

Abstract. The hybrid conjugate gradient (CG) algorithms are among the efficient modifications of the conjugate gradient methods. Some interesting features of the hybrid modifications include inherenting the nice convergence properties and efficient numerical performance of the existing CG methods. In this paper, we proposed a new hybrid CG algorithm that inherits the features of the Rivaie et al. (RMIL*) and Dai (RMIL+) conjugate gradient methods. The proposed algorithm generates a descent direction under the strong Wolfe line search conditions. Preliminary results on some benchmark problems reveal that the proposed method efficient and promising.

1. Introduction

Given an unconstrained optimization model

$$\min f(x)$$

where \( f: \mathbb{R}^n \rightarrow \mathbb{R} \) is a smooth function and \( g(x) = \nabla f(x) \) is its gradient, \( x \in \mathbb{R}^n \) is an \( n \)-dimensional real vector. The CG algorithms are amongst the efficient optimization algorithms for obtaining the solution of problem (1), mainly when the dimension \( n \) is large [1]. The solution of the unconstrained optimization problem (1) is often in form of a local minimum or global minimal point. In practice, most optimization algorithms achieved only the local minima, because the global minimum is sometimes very difficult to attain as per knowledge of the function is commonly local. Beginning with a starting guess \( x_0 \in \mathbb{R}^n \), the CG algorithm would generates a sequence of points \( \{x_k\}_{k=0}^{\infty} \) using the iterative formula defined as follows

$$x_{k+1} = x_k - \alpha_k d_k$$

\( \alpha_k \) is a step length calculated via suitable line search method along the direction of search \( d_k \). For the first iteration, the \( d_k \) is generally the negative of the gradient and known as the direction of steepest descent, i.e. \( d_0 = -g_0 \). However, resulting \( d_k \)’s are computed using

$$d_k = -g_k + \beta_k d_{k-1}$$

where the scalar \( \beta_k \) is referred to as conjugate gradient update coefficient [2]. In computing the step size \( \alpha_k \), some of the inexpensive commonly used line searches algorithms include the weak Wolfe conditions

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

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k$$

1. Introduction

Given an unconstrained optimization model

$$\min f(x)$$

where \( f: \mathbb{R}^n \rightarrow \mathbb{R} \) is a smooth function and \( g(x) = \nabla f(x) \) is its gradient, \( x \in \mathbb{R}^n \) is an \( n \)-dimensional real vector. The CG algorithms are amongst the efficient optimization algorithms for obtaining the solution of problem (1), mainly when the dimension \( n \) is large [1]. The solution of the unconstrained optimization problem (1) is often in form of a local minimum or global minimal point. In practice, most optimization algorithms achieved only the local minima, because the global minimum is sometimes very difficult to attain as per knowledge of the function is commonly local. Beginning with a starting guess \( x_0 \in \mathbb{R}^n \), the CG algorithm would generates a sequence of points \( \{x_k\}_{k=0}^{\infty} \) using the iterative formula defined as follows

$$x_{k+1} = x_k - \alpha_k d_k$$

\( \alpha_k \) is a step length calculated via suitable line search method along the direction of search \( d_k \). For the first iteration, the \( d_k \) is generally the negative of the gradient and known as the direction of steepest descent, i.e. \( d_0 = -g_0 \). However, resulting \( d_k \)’s are computed using

$$d_k = -g_k + \beta_k d_{k-1}$$

where the scalar \( \beta_k \) is referred to as conjugate gradient update coefficient [2]. In computing the step size \( \alpha_k \), some of the inexpensive commonly used line searches algorithms include the weak Wolfe conditions

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

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k$$
the strong Wolfe (SWP) condition (4) and
\[ |g_{k+1}^T d_k| \leq |\sigma g_k^T d_k| \]  \hspace{1cm} (6)
and the generalized Wolfe line search (4) and
\[ \sigma g_k^T d_k \leq g_{k+1}^T d_k \leq \sigma_1 g_k^T d_k \]  \hspace{1cm} (7)
with $0 < \delta < \sigma < 1$, and $\sigma_1 \geq 0$ being constants that are frequently used. The efficient and effective line search of a CG method gives an approximate value of the step length by guaranteeing that the stages are precisely either too long or too short.

Generally, CG algorithms are characterized by the choice of coefficient $\beta_k$. Some of the classical formulas for $\beta_k$ are

\[ \beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \]  \hspace{1cm} Fletcher - Reeves (FR) [3]

\[ \beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}} \]  \hspace{1cm} Polak - Ribiere – Polyak (PRP) [4,5]

\[ \beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \]  \hspace{1cm} Hestenes - Stiefel (HS) [6]

\[ \beta_k^{LS} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}} \]  \hspace{1cm} Liu – storey (LS) [7]

\[ \beta_k^{CD} = \frac{g_k^T g_k}{d_{k-1}^T g_{k-1}} \]  \hspace{1cm} Conjugate Descent (CD) [8]

\[ \beta_k^{DY} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \]  \hspace{1cm} Dai - Yuan (DY) [9]

The above formulas can be characterized into two groupings. One group includes PRP [4,5], HS [6], and LS [7]. These formulas are regarded as amongst the greatest efficient CG algorithms for obtaining the solution of large-scale functions. This is as a result of an inbuilt automatic restart feature that helps prevent them from jamming during the computation. Yet, these algorithms may fail to converge to the solution for some problems, and their convergence are until now not established under certain inexact line search conditions. The other group includes FR [3], CD [8], and DY [9] methods. Though, these algorithms possess strong convergence properties, their numerical performance is often very poor due to the jamming phenomena [10,11].

The above stated drawbacks motivated researchers to study and proposed numerous modifications of the CG method with the aim of overcoming the lapses encountered by the classical methods. Some of these modifications include the three-term CG algorithms, spectral CG algorithms, the hybrid CG method, and many more. However, most researcher focussed on the hybrid CG algorithms that combine various CG parameters $\beta_k$ so as to overcome the weaknesses and exploit the advantages of the parent CG algorithm [12]. Among the earliest hybrid CG algorithms is $\beta_k^{TS}$ developed by Touati-Ahmed and Storey [13] with formula defined as follows

\[ \beta_k^{TS} = \begin{cases} 
\beta_k^{PRP} & \text{if } 0 \leq \beta_k^{PRP} \leq \beta_k^{FR}, \\
\beta_k^{FR} & \text{otherwise}.
\end{cases} \]  \hspace{1cm} (8)

This method numerically outperformed both methods of PRP and FR in addition to its nice convergence properties. Based on the above discussion, numerous studies and effort have been done
focusing on finding new CG methods with not only good efficient numerical performance, but also, good convergence properties, (see Refs [14 – 30,42,44,45,46]). For the application of the CG method to real-world situation, (see Refs 31-35,43,47). Motivated by the recent trend on conjugate gradient method, this research aim to study a new hybrid CG method that would possess nice convergence features in addition to efficient numerical performance.

The remaining part of the study is planned as follows. In the subsequent section, we derive the proposed CG parameter $\beta_k$ using the idea of existing CG algorithm and further present the specific algorithm. Section 3 presents the global convergence analysis with strong Wolfe conditions. Preliminary numerical results are analysed via the performance profiles introduced by Dolan-Morè [36] in Section 4. In conclusion, we presented the summary of the research in Section 5.

2. A new hybrid CG method and algorithm

In an attempt to overcome some of the drawbacks discussed above, Rivaie et al. [37] developed a variant of PRP formula by replacing $\|g_k\|^2$ in the denominator by $\|d_k\|^2$. The coefficient $\beta_k$ of RMIL is computed as

$$\beta_k^{RMIL} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{\|d_k\|^2}$$

This method possesses an efficient numerical performance because of an in-built restart feature that prevent it from jamming. The authors show that the method converge globally under exact minimization condition. Rivaie [38] further defined a variant of RMIL method by retain the denominator $\beta_k^{RMIL}$ while adding a negative previous $d_k$ to the numerator. This method is defined as follows

$$\beta_k^{RMIL^+} = \frac{g_{k+1}^T(g_{k+1} - g_k - d_{k-1})}{\|d_k\|^2}$$

This method inherited the nice convergence properties of RMIL algorithm and would reduce to $\beta_k^{RMIL}$ under exact minimization conditions. Rivaie et al. [38] further show that the formula is globally convergent provided the SWP condition is satisfied.

Recently, Dai [39] pointed the convergence prove of RMIL [37] is not correct and pointed out that a wrong inequality known to play a vital part in the convergence study of the proposed algorithm. The author further presented as modification of RMIL method as follows:

$$\beta_k^{RMIL^+} = \begin{cases} \frac{g_{k+1}^T(g_{k+1} - g_k)}{\|d_k\|^2}, & \text{if } 0 \leq \|g_{k+1}^T g_k\| \leq \|g_{k+1}\|^2 \\ 0, & \text{otherwise,} \end{cases}$$

Though, the performance of both methods are similar, Dai [38] show that RMIL+ converges globally using exact minimization condition. The convergence RMIL+ was further studied under the strong Wolfe conditions by Yousif [40].

Motivated by the nice convergence properties and efficient numerical performance of [37–40] and taking into account the ideas of the hybrid methods of [13], [21], we suggest a new hybrid CG coefficient as follows:

$$\beta_k^{hRMIL} = \begin{cases} \beta_k^{RMIL^+}, & \text{if } 0 \leq \beta_k^{RMIL^+} \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}, \\ \frac{g_{k+1}^T(g_{k+1} - g_k - d_{k-1})}{\|d_k\|^2}, & \text{otherwise}. \end{cases}$$
Next, we give the algorithm of (9) as follows.

**Algorithm 1.**

**Stage 1.** Starting: Assumed $x_0 \in \mathbb{R}^n$, $d_0 = -g_0$, fixe $k := 0$.

**Stage 2.** Solve for $\alpha_k$ using (4) and (6).

**Stage 3.** Update $x_k$ via (2).

**Stage 4.** Calculate $\beta_k$ by (9) and update $d_k$ by (3).

**Stage 5.** If $\|g_k\| \leq 10^{-6}$, terminate. Else, go to stage 2 with $k := k + 1$.

The assumptions given below are very vital in the study and analysis of various CG algorithm convergence properties.

**Assumption A.** $f(x)$ is bounded from below on the level set $\Omega = \{x \in \mathbb{R}^n / f(x) \leq f(x_0)\}$.

**Assumption B.** In some neighborhood $N$ of $\Omega$, the function $f$ is smooth and $g(x)$ is Lipchitz continuous in $N$, such that, $\exists L > 0$ (constant) satisfying:

$$\|g(x) - g(y)\| \leq L\|x - y\| \quad \forall x, y \in N. \quad (10)$$

### 3. Global convergence analysis

This section discusses the convergence of $\beta_k^{\text{HAMILT}}$. One of the general requirements that any CG algorithm should have is that the descent property is specified as $g_k^T d_k \leq -c\|g_k\|^2, c > 0.$

(11)

To ease the theoretical proof, we need to simplify $\beta_k^{\text{HAMILT}}$ as follows

$$\beta_k^{\text{HAMILT}} = \begin{cases} \beta_k^{\text{HAMILT}} + \frac{g_{k+1}^T (g_{k+1} - g_k - d_{k-1})}{\|d_k\|^2} & \text{if } 0 \leq \beta_k^{\text{HAMILT}} + \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}, \\ \beta_k^{\text{HAMILT}} & \text{otherwise.} \end{cases} \quad (12)$$

If $0 \leq \beta_k^{\text{HAMILT}} + \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}$, then, it is obvious from [39] that

$$\beta_k^{\text{HAMILT}} < \beta_k^{\text{HAMILT}} + \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}$$

Otherwise,

$$\beta_k^{\text{HAMILT}} = \frac{g_{k+1}^T (g_{k+1} - g_k - d_{k-1})}{\|d_k\|^2} = \frac{\|g_{k+1}\|^2 - g_{k+1}^T g_k - g_{k+1}^T d_{k-1}}{\|d_k\|^2}$$

From [38], it follows that

$$0 \leq \beta_k^{\text{HAMILT}} \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}$$

Hence, for both cases, we have

$$0 \leq \beta_k^{\text{HAMILT}} \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}. \quad (12)$$

The convergence prove of the proposed CG formula is built on (12) and the theorems that would be discussed below.
3.1 Sufficient descent condition

The following theorems would be applied to proof that $\beta^\text{hRMIL}_k$ possess (11) under inexact line search.

**Theorem 1**: For any CG scheme or algorithm or scheme defined by (2) and (3), with $\beta_k$ defined by (9), and $\alpha_k$ is calculated using the SWP conditions (4) and (6) with $0 < \sigma < \frac{1}{4}$. Then,

$$\frac{\|z_k\|}{\|d_k\|^2} < 2, \quad \forall k \geq 0. \tag{13}$$

**Proof**: The prove of this theorem follows from Osman [40]. □

**Theorem 2**: For any CG scheme or algorithm defined by (2) and (3), with $\beta_k$ defined by (9), and $\alpha_k$ calculated using the SWP conditions (4) and (6) with $0 < \sigma < \frac{1}{4}$. Then,

$$-1 - 2\sigma \leq \frac{g^T_k d_k}{\|g_k\|^2} \leq -1 + 2\sigma, \quad \forall k \geq 0. \tag{14}$$

Hence, the condition (11) holds.

**Proof**: For $k = 0$, it is obvious that $\frac{\|g_0\|^2}{\|d_0\|^2} = 1 < 2$ which follows from (13). Hence Theorem 1 is true for $k = 0$. Next, we need to also show that Theorem 1 is true for $k > 0$.

**Case 1**: If $0 \leq \beta^\text{hRMIL}_k \leq \frac{\|g_{k+1}\|^2}{\|d_k\|^2}$, then for $0 \leq |g^T_{k+1} g_k| \leq \|g_{k+1}\|$ from [39], we have

$$\beta^\text{hRMIL}_k = \frac{g^T_{k+1}(g_{k+1} - g_k)}{\|d_k\|^2}$$

Multiplying both side of (3) by $g^T_k$ gives

$$g^T_k d_k = -\|g_k\|^2 + \beta^\text{hRMIL}_k g^T_k d_{k-1}. \tag{15}$$

From (12) and SWP condition (6), we have

$$-\sigma \beta^\text{hRMIL}_k |g^T_{k-1} g_{k-1}| \leq \beta^\text{hRMIL}_k g^T_{k-1} d_{k-1} \leq \sigma \beta^\text{hRMIL}_k |g^T_{k-1} g_{k-1}| \tag{16}$$

Considering (15) and (16) and applying Cauchy Schwartz inequality, we get

$$-\|g_k\|^2 - \sigma \beta^\text{hRMIL}_k \|g_{k-1}\| \|d_{k-1}\| \leq g^T_k d_k \leq -\|g_k\|^2 + \sigma \beta^\text{hRMIL}_k \|g_{k-1}\| \|d_{k-1}\| \tag{17}$$

Substituting $\beta^\text{hRMIL}_k$ in (17) gives

$$-\|g_k\|^2 - \sigma \frac{\|g_k\|^2}{\|d_{k-1}\|^2} \|g_{k-1}\| \|d_{k-1}\| \leq g^T_k d_k \leq -\|g_k\|^2 + \sigma \frac{\|g_k\|^2}{\|d_{k-1}\|^2} \|g_{k-1}\| \|d_{k-1}\| \tag{18}$$

Dividing (18) by $\|g_k\|^2$

$$-1 - \sigma \frac{\|g_{k-1}\|}{\|d_{k-1}\|} \leq \frac{g^T_k d_k}{\|g_k\|^2} \leq -1 + \sigma \frac{\|g_{k-1}\|}{\|d_{k-1}\|} \tag{19}$$

From (19) and Theorem 1, it follows

$$-1 - 2\sigma \leq \frac{g^T_k d_k}{\|g_k\|^2} \leq -1 + 2\sigma. \tag{20}$$

and thus, completes the prove. □
Case 2: Otherwise, that is, when

\[ \beta_k^{hRMIL} = \frac{g_{k+1}^T(g_{k+1} - g_k - d_{k-1})}{\|d_k\|^2} \]

**Proof:** The proof of case 2 follows from Rivaie et al. [38].

4. Numerical results

Here, we studied the performance of defined \( hRMIL \) formula by equating with the performance of \( RMIL + [39] \) and \( RMIL * [38] \) algorithms using 23 benchmark functions taken from Andrei [41]. This comparison was done based on number of iterations and CPU time using the strong Wolfe line search procedures. All problems and formulas are coded and run on the same Matlab programs using 23 benchmark functions taken from Andrei [41]. This comparison was done based on number of iterations and CPU time using the strong Wolfe line search procedures. All problems and formulas are coded and run on the same Matlab programs with termination criterion set as \( \|g_k\| \leq 10^{-6} \). All problems have been implement using dimensions ranging from \( 2 \leq n \leq 10,000 \) to illustrates the robustness of \( hRMIL \) as shown in Table 1. The performance was also analysed via the performance profile software developed by Dolan and More [36] as can be seen in Figure 1 and 2.

**Table 1:** List of Unconstrained Optimization Test Functions

| No | Functions           | N     | Initial points                      |
|----|---------------------|-------|-------------------------------------|
| 1  | Booth               | 2     | (-8,-8)(49,49)(80,80)               |
| 2  | TRECCANI            | 2     | (-2.1,2)(20,20)(79,79)              |
| 3  | Zettl               | 2     | (6.6)(20,20)(-100,-100)             |
| 4  | Raydan 2            | 2,4   | (1.3)(-17,16)(2.24)                 |
| 5  | Dixon and Price     | 2,4   | (-55,-55)(85,85)(101,...,101)       |
| 6  | Hager               | 2,4   | (6,...,6)(-17,...,-17)(-78,...,-78) |
| 7  | Ext. Friedenstein and Roth | 2,4,10 | (2,...,2)(19.2,19.2)(0.5,30) |
| 8  | Raydan 1            | 2,4,10| (1,...,1)(-10,...,-10)(-20,...,-20)|
| 9  | Ext. penalty        | 2,4,10| (2,...,2)(19,...,19)(59,...,59)     |
| 10 | Extended Maratos    | 2,4,10,100| (18,...,18)(-4.5,...,-4.5)(-84,...,-106)|
| 11 | Generalized Tridiagonal 1 | 2,4,10,100| (1.1)(20,20)(40,40) |
| 12 | Extended Beale      | 2,4,10,100| (-1.3,...,-1.3)(5,...,5)(11.3,...,11.3)|
| 13 | Extended Denschnb   | 2,4,10,100| (3,...,3)(23,...,23)(200,...,200) |
| 14 | Extended Tridiagonal 1 | 2,4,10,100,500| (3.3)(8.8)(24.6, 24.7)|
| 15 | Generalized Quartic 1 | 2,4,10,100,500| (10,...,10)(20,...,20)(80,...,80)|
| 16 | Extended Shallow    | 2,4,10,100,500| (11,...,11)(-1,...,-1)(-50,110) |
| 17 | Extended Himmelblau | 2,4,10,100,500| (1.5)(10,...,10)(41,...,41) |
| 18 | Sum Squares         | 2,4,10,100,500| (3.7,...,3.7)(15,...,15)(35,...,35)|
| 19 | Quadratic 2         | 2,4,10,100,500,1000| (0.5,...,0.5)(20,...,20)(80,...,80)|
| 20 | Diagonal 2          | 2,4,10,100,500,1000| (1,...,1)(5,...,5)(15,...,15)|
| 21 | Extended White and Holst | 2,4,10,100,500,1000| (-1.3,...,-1.3)(10,...,100)(11,...,11)|
| 22 | Ext. Quadratic penalty 2 | 2,4,10,100,500,1000| (0.5,...,0.5)(21,...,21)(50,...,50)|
| 23 | Extended Rosenbrock  | 2,4,10,100,500,1000| (2,...,2)(20,...,20)(80,...,80)|

The performance profile is employed to evaluates and compares the performance for the classes of involved solvers \( S \) on a whole set of test problems \( P \). Assume that \( n_p \) problems and \( n_s \) solvers exists, for every solver \( s \) and problem \( p \), Dolan and More defined

\[ \tau_{p,s} = \text{Computation time (CPU time or NO.IT.) needed by solver } s \text{ to solve problem } p. \]

The output program graphs the portion of each given problem for each algorithm, such that the process is a factor of the fastest period in the neighborhood. The uppermost curve shows the highest
functioning algorithm. That is, the algorithm that obtained the solutions of nearly all the given function within the shortest time.

**Figure 1:** Performance profile with regards to iteration number

**Figure 2:** Performance profile with regards to CPU time

Considering both figures, we can see that the derived hRMIL method behaves more like the RMIL+ method. However, hRMIL has the least iterations and CPU time under SWP as can be observe on the left side of both figures. Also, we can notice that hRMIL lies above RMIL+ and RMIL* methods, both under CPU time and iteration number. These show that the hybrid hRMIL algorithm is efficient and promising. Hence, can be considered as a substitute for solution of optimization models.

5. Conclusion

In this paper, based on the nice convergence analysis and efficient numerical performance of previous hybrid CG algorithms, we present an alternative hybrid CG algorithm that inherits the features of the known Riviae et al. (RMIL*), Dai (RMIL+) and Osman (RMIL+) CG algorithms. The convergence analysis of the defined hRMIL was studied under SWP. The performance of $\beta_k^{RMIL}$ was compared with that of the existing $\beta_k^{RMIL*}$ and $\beta_k^{RMIL+}$ methods on several unconstrained optimization benchmark functions using the performance software by Dolan and More [36]. The computational results show that the proposed hRMIL algorithm is both efficient and promising.

References

[1] Sulaimana, I M., Mamat, M., Owoyemi, A E., Ghazali, P L., Rivaie, M., Malik, M. The convergence properties of some descent conjugate gradient algorithms for optimization models. *Journal of Mathematics and Computer Science*, 22, 204-215, (2021).

[2] Sulaiman, I M., Mamat, M., Kamfa, K., Muktar D. A Descent Modification of Conjugate Gradient Method for Optimization Models. *Iraqi Journal of Science*. 61(7), 1745-1750, (2020). DOI: 10.24996/ijs.2020.61.7.23.

[3] Fletcher R., Reeves C. Function minimization by conjugate gradients, *Comput. J.*, 7, 149–154, (1964).

[4] Polak. E., Ribiere, G. Note sur la convergence de directions conjuguées, *Rev. Francaise Informat Recherche Opertionelle*, 3e ann ee, 16, 35–43, (1969).

[5] Polyak B. T. The conjugate gradient method in extremem problems, *USSR Comp. Math. Math. Phys.*, 9, pp. 94–112, (1969).

[6] Hestenes, M. R., Stiefel E. L. Methods of conjugate gradients for solving linear systems, *J. Research Nat. Bur. Standards*, 49, 409–436, (1952).

[7] Liu, Y., Storey, C. Efficient generalized conjugate gradient algorithms, part 1: theory, *J.optim.theory Appl*, 69, 129-137, (1992).

[8] Fletcher R.. *Practical Methods of Optimization*, 2nd ed., John Wiley, New York, (1989).

[9] Dai, Y., Yuan, Y. A nonlinear conjugate gradient with strong global convergence properties, *SIAM J.optim*, 10, 177-182, (2000).
[10] Hager, W. W., Zhang, H. A survey of nonlinear conjugate gradient methods. Pac. J. Optim., 2, 35-58, (2006).
[11] Gilbert J. C and Nocedal J. 1992. Global convergence properties of conjugate gradient methods for optimization. SIAM. J. Optim., vol 2, pp. 21–42.
[12] Kaelo, P., Mtaligulwa, P., Thuto, M.V. A globally convergent hybrid conjugate gradient method with strong Wolfe conditions for unconstrained optimization. Math Sci 14, 1–9, (2020).
[13] Touati-Ahmed, D., Stokey, C. Efficient hybrid conjugate gradient techniques. J. Optim. Theory Appl. 64(2), 379–397 (1990).
[14] Liu, G. H., Han J. Y., Yin H. X. Global convergence of the Fletcher-Reeves algorithm with an inexact line search. Report, Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, (1993)
[15] Hager, W.W., Zhang, H. A new conjugate gradient method with guaranteed descent and an efficient line search, SIAM J. Optim., 16, 170-192, (2005).
[16] Aini, N, Rivaie, M, Mamat, M., Sulaiman, I.M. A Hybrid of Quasi-Newton Method with CG Method for Unconstrained Optimization. J. Phys.: Conf. Ser. 1366(012079), (2019).
[17] Yu, G. H., L. Guan, L.T., Li, G.Y. Global convergence of modified Polak–Ribière–Polyak conjugate gradient methods with sufficient descent property. J. Ind. Manag. Optim., 4, 565-579, (2008).
[18] Sulaiman, I.M., Sukono, Sudrajat, S., Mamat, M. New class of hybrid conjugate gradient coefficients with guaranteed descent and efficient line search. IOP Conf. Ser.: Mater. Sci. Eng. 621(012021). (2019).
[19] Alhawarat, A., Salleh, Z., Mamat, M., Rivaie, M. An efficient modified Polak-Ribiere-Polyak conjugate gradient method with global convergence properties. Optim. Methods Softw. 32(6), 1299–1312 (2017).
[20] Babaie-Kafaki, S. A quadratic hybridization of Polak-Ribie`re–Polyak and Fletcher-Reeves conjugate gradient methods. J. Optim. Theory Appl. 154(3), 916–932 (2012).
[21] Sallih, Y., Hamoda, M.A., Rivaie, M. New hybrid conjugate gradient method with global convergence properties for unconstrained optimization. Malays. J. Comput. Appl. Math. 1(1), 29–38 (2018).
[22] Dai, Y.H., Yuan, Y.X. An efficient hybrid conjugate gradient method for unconstrained optimization. Ann. Oper. Res. 103, 33-47, (2001).
[23] Talat, A., Sulaiman, I.M., Mamat, M., Rivaie, M. An Efficient Hybrid Conjugate Gradient Coefficient under Inexact Line Search. International Journal of Advanced Trends in Computer Science and Engineering, 9(1), 784-788 (2020).
[24] Andrei, N. Hybrid conjugate gradient algorithm for unconstrained optimization. J. Optim. Theory Appl., 141, 249-264, (2009).
[25] Kamfa, K., Waziri, M.Y., Sulaiman, I.M., Ibrahim, M.H., Mamat, M., S.S. Abas. An Efficient Hybrid BFGS-CG Search Direction for Solving Unconstrained Optimization Problems. Journal of Advanced Research in Dynamical and Control Systems 12(2):1035-1041 (2020).
[26] Touati-Ahmed, D., Stokey. C Efficient hybrid conjugate gradient techniques. J. Optim. Theory Appl., 64, 379-397, (1990).
[27] Malik, M., S.S. Abas., Mamat, M., Sukono, Sulaiman, I.M.. A New Hybrid Conjugate Gradient Method with Global Convergence Properties. International Journal of Advanced Science and Technology, 29(05), 199-210 (2020).
[28] Xu, X., Kong, F.: New hybrid conjugate gradient methods with the generalized Wolfe line search. SpringerPlus 5(1), 881 (2016).
[29] Mtaligulwa, P., Kaelo, P.: An efficient modified PRP-FR hybrid conjugate gradient method for solving unconstrained optimization problems. Appl. Numer. Math. 145, 111–120 (2019)
[30] Jiang, X.Z., Han, L., Jian, J.B. A globally convergent mixed conjugate gradient method with Wolfe line search. Math. Numer. Sin., 34, 103-112 (2012).
[31] Abdulkarim, H.B., Kumam, p., Auwal, B.A., Jirakitpuwapat, W., Abubakar, J. A hybrid conjugate gradient algorithm for constrained monotone equations with application in compressive sensing. Helinyon. 6(3): e03466 (2020).
[32] Sulaiman, I.M., Usman, A.Y., Mamat, M. Application of Spectral Conjugate Gradient Methods for Solving Unconstrained Optimization Problems. *An International Journal of Optimization and Control: Theories & Applications*, 10(2), 2146-5703 (2020).

[33] Auwal, B.A., Kumam, p., Aliyu, M. A. Global convergence via descent modified three-term conjugate gradient projection algorithm with applications to signal recovery. Results in Applied Mathematics 4:100069, (2019).

[34] Umar, A.O., Sulaiman, I.M., Mamat, M., Waziri, M.Y. Foziah, H.M., Altien, J.R., Deiby, T.S. New Hybrid Conjugate Gradient Method for Solving Fuzzy Nonlinear Equations. *Journal of Advanced Research in Dynamical and Control Systems* 12(2):585-590 (2020).

[35] Mamat, M., Rokhayati, Y., Noor, N.M.M., Mohd, I. Optimizing human diet problem with fuzzy price using fuzzy linear programming approach. *Pakistan Journal of Nutrition*, 10(6), 594-598 (2011).

[36] Dolan, E.D., More, J.J. Benchmarking optimization software with performance profiles, *Mathematical Programming*, 91, 201-213 (2002).

[37] Rivaie, M., Mamat, M., June, L.W., Ismail, M. A new class of nonlinear conjugate gradient coefficients with global convergence properties. *Applied Mathematics and Computation*, 218(22), 11323-11332 (2012).

[38] Rivaie, M., Mamat, M., Abashar, A. A new class of nonlinear conjugate gradient coefficients with exact and inexact line search. *Applied Mathematics and Computation*, 268, 1152-1163 (2015).

[39] Dai, Z. Comments on a new class of nonlinear conjugate gradient coefficients with global convergence properties. *Applied Mathematics and Computation*, 276, 297-300 (2016).

[40] Osman, O.O.Y. The convergence properties of RMIL+ conjugate gradient method under the strong Wolfe line search. *Applied Mathematics and Computation*, 367, 124777 (2020).

[41] Andrei, N. An unconstrained optimization test functions collection. *Adv. Model. Optim.*, 10(1), 147-161 (2008).

[42] Sulaiman, I.M., Mamat, M., Owoyemi, A.E., Ghazali, P.L., Rivaie, M., Malik, M. "The convergence properties of some descent conjugate gradient algorithms for optimization models" *Journal of Mathematics and Computer Science*, 22(3), 204–215, (2020).

[43] Mamat, M., Rokhayati, Y., Noor, N.M.M., Mohd, I. Optimizing human diet problem with fuzzy price using fuzzy linear programming approach. *Pakistan Journal of Nutrition*, 10(6), 594-598, (2011).

[44] Abashar, A., Mamat, M., Rivaie, M., Mohd, I. Global convergence properties of a new class of conjugate gradient method for unconstrained optimization. *Applied Mathematical Sciences* Issue 65-68, 3307-3319, (2014).

[45] Sulaiman, I.M., Mamat, M., Kamfa, K., Muktar, D. A Descent Modification of Conjugate Gradient Method for Optimization Models. *Iraqi Journal of Science*. 61(7):1745-1750, (2020).

[46] Sulaiman, I.M., Supian, S., Sukono, F., and Mamat, M. New class of hybrid conjugate gradient coefficients with guaranteed descent and efficient line search. *In IOP Conference Series: Materials Science and Engineering*, vol. 621, no. 1, (2019). art. Id. 012021.

[47] Sulaiman, I.M. and Mamat, M. A New Conjugate Gradient Methods with Descent Properties and its Application to Regression analysis. *Journal of Numerical Analysis, Industrial and Applied Mathematics*. 14(3-4), 1-15, (2020).