EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 15, No. 3, 2022, 1254-1264
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

A Sufficient Descent Property to Improving a Three-Term Conjugate Gradient Algorithm

Ghada M. Al-Naemi[1,*], Firas Mahmood Saeed[2]

[1] *Department of Mathematics, Faculty of Computer Science and Mathematics, University of Mosul, Mosul, Nineveh, Iraq*
[2] *Department of Family Medicine, College of Medicine, University of Mosul, Mosul, Nineveh, Iraq*

Abstract. The nonlinear conjugate gradient (NLCGM) methods have received attention because due to their simplicity, low memory requirements, and global convergent property, which allows them to be used directly to solve large-scale nonlinear unconstrained optimization problems. We suggested a modification to the β_k^{KMAR} formula, applied with three-term conjugate gradient method that is both simple and effective, denoted by (TTKMAR), which has a sufficient descent property (SDP) and ensures global convergence (GCP) when we use any line search. The numerical efficiency of TTKMAR was assessed using a variety of standard test functions. TTCGM has been demonstrated to be more numerically efficient than two-term CG methods. This paper also quantifies the difference between TTCGM and two-term methods of performance. As a result, we compare our new modification to an efficient two-term and TTCGM in the numerical results. Finally, we conclude that our proposed modification.

2020 Mathematics Subject Classifications: 90C90, 90C30, 90C06, 65K05, 49M37
Key Words and Phrases: TTCGM, Nonlinear unconstrained optimization, SWPL search, SDP

1. Introduction

The NLCGM, an iterative method for solving the unconstrained optimization problem of the form: was used in this study.

$$\min u(x), \ x \in R^n \tag{1}$$

is investigated. $u : R^n \to R$ is smooth and $g(x) = \nabla u(x)$ is accessible, because it does not require any matrices, the NLCGM is one option for obtaining the bare minimum (1) [5]. CG methods are iterative methods in the form of:

$$x_{k+1} = x_k + \gamma_k d_k \ , \quad k = 0, 1, 2, 3, \ldots \tag{2}$$

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v15i3.4437

Email addresses: `drghadaalnaemi@uomosul.edu.iq` (GH. M. Al-Naemi), `fms@uomosul.edu.iq` (F. M. Saeed)

 © 2022 EJPAM All rights reserved.

where γ_k denotes a positive step size using by cubic interpolation and d_k denotes a search direction. Typically, the search direction is defined as:

$$d_k = \begin{cases} -g_k, & k=0 \\ -g_k + \beta_k d_{k-1}, & k \geq 1 \end{cases} \tag{3}$$

where $\beta_k \in R$ is a scalar parameter that defines CG-method. Typically, the parameter β_k is typically chosen in such a way (2)-(3) reduces the linear CG-method, if $f(x)$ is a strictly convex quadratic function and is calculated using the exact line search (ELS), the parameter γ_k is usually chosen so that (2)-(3) reduces to the linear CG-method [1]. In [8, 10–12, 14–16], six pioneering forms of β_k are defined.

For some TTKMAR methods, the line search is frequently used to achieve GCP and improve computational performance [17]

$$f\left(x_k + \gamma_k d_k\right) \leq f\left(x_k\right) + \rho \gamma_k d_k \tag{4}$$

$$\left| g(x_k + \gamma_k d_k)^T d_k \right| \leq \sigma \left| g_k^T d_k \right| \tag{5}$$

where d_k is the direction of descent and $0 < \rho < 0.5 < \sigma < 1$ is a very efficient value. In general, the different conjugate gradient parameter choices in (3) correspond to different TTCG methods.

2. Motivation and Algorithm

Several TTCG methods for unconstrained optimization problems have recently been proposed. This section begins with an explanation of our motivation before moving on to a detailed explanation of our method. Recently, Zhang et al. [23] presented a threeâ€“term MPRP method and used the Armijo line search to demonstrate that the direction meets GCP.

$$d_k = \begin{cases} -g_k, & k=0 \\ -g_k + \frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}} d_{k-1} - \frac{g_k^T d_{k-1}}{g_{k-1}^T g_{k-1}} y_{k-1}, & k \geq 1 \end{cases} \tag{6}$$

where $y_{k-1} = g_k - g_{k-1}$, in the same content, Zhang et al. [22] developed the three-term HS method. Which is expressed as:

$$d_k = \begin{cases} -g_k, & k=0 \\ -g_k + \frac{g_k^T y_{k-1}}{y_{k-1}^T d_{k-1}} d_{k-1} - \frac{g_k^T d_{k-1}}{y_{k-1}^T d_{k-1}} y_{k-1}, & k \geq 1 \end{cases} \tag{7}$$

The SDP is satisfied by the three-term HS method; if an ELS is used, it reduces to the original HS method. Furthermore, a modified three-term HS algorithm on the search direction is used to ensure the GCP of the search direction specified in (7):

$$d_k = \begin{cases} -g_k, & k=0 \\ -g_k + \frac{g_k^T \mu_{k-1}}{z_{k-1}^T d_{k-1}} d_{k-1} - \frac{g_k^T d_{k-1}}{z_{k-1}^T d_{k-1}} y_{k-1}, & k \geq 1 \end{cases} \tag{8}$$

with $\mu_{k-1} = y_{k-1} + t||g(x_k)||^T s_{k-1}$. Given that modified three-term HS were introduced in (7) to demonstrate the GCP of the search direction, one might wonder why (7) is not used to demonstrate the GCP of the search direction. Rather than disregarding (7), it should be made efficient and globally convergent. As a result, (7) can be modified to meet the GCP. In terms of numerical performance, such a modification is expected to outperform the MTTHS algorithm.

Recently, Kamilu and colleagues [13] proposed a new CG formula with the same numerator as the PRP, HS, and RMIL formulas. The numerator was kept to order to give the formula the ability to restart.

$$\beta_k^{KMAR} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T (g_k + g_{k-1})} \tag{9}$$

This formula can be reduced to the FR or PRP parameters under certain conditions. The denominator is geometrically similar to the well-known CG denominator. Because $(g_k + g_{k-1})$ forms a vector scaled by g_{k-1}. The new vector would be slightly different from the normal vector produced by the PRP parameter $d_k^T (g_k - g_{k-1})$. This minor modification would boost the performance while simplifying the convergence proof.

To increase the efficiency of the one-term and two-term CG methods. The TTCG has been extensively studied and emphasized since its exception. The performance of this class of CG methods is heavily dependent on how the scalar parameter is chosen. Yanlin [18] recently improved the Zhang [23] method by constructing a new modified PRP parameter, MPRP. The performance is quite good.

At this point, we should recall that the approximation matrix B_k must satisfy the secant condition for quasi-Newton methods:

$$B_k s_{k-1} = y_{k-1} \tag{10}$$

Zhang et al. [21] and Zhang and Xu [20] using Taylorâ€™s series proposed the following modified secant condition by expanding condition (10).

$$B_k s_{k-1} = \mu_{k-1}, \qquad \mu_{k-1} = y_{k-1} + \frac{\varphi_{k-1}}{||s_{k-1}||^2}.s_{k-1} \tag{11}$$

$$\varphi_{k-1} = 6 (u_{k-1} - u_k) + 3 \ (g_{k-1} + g_k) \tag{12}$$

where $u_k = u(x_k)$. Yabe and Takano [19] extended the modified secant relation (11) by multiplying a fixed parameter $\varepsilon \geq 0$, as expressed by the expression:

$$B_k s_{k-1} = \mu_{k-1}, \ \mu_{k-1} = y_{k-1} + \varepsilon \frac{\varphi_{k-1}}{||s_{k-1}||^2} s_{k-1} \tag{13}$$

Babaie-Kafaki et al. [6] presented the following modified version of Equation (12):

$$B_k s_{k-1} = \mu_{k-1}, \ B_k s_{k-1} = \mu_{k-1}, \ \mu_{k-1} = y_{k-1} + \varepsilon \frac{max\{0, \varphi_{k-1}\}}{||s_{k-1}||^2} s_{k-1} \tag{14}$$

In this paper, we are motivated to develop TTCGM that meets the SDP and achieves GCP under SWPL.

3. Modified TTKMAR CG-Method

To develop a new method that is globally convergent in the face of inexact line search. In the following section, we will create our new TTKMAR method using the modified β_k^{KMAR} by:

$$\beta_k^{MKMAR} = \frac{g_k^T \mu_{k-1}}{g_{k-1}^T (g_k + g_{k-1})} \tag{15}$$

with $\mu_{k-1} = y_{k-1} + \frac{\varphi_{k-1}}{||s_{k-1}||^2}.s_{k-1}$ and $\varphi_{k-1} = \frac{g_k^T d_{k-1}}{g_{k-1}^T (g_k + g_{k-1})}$. Then the direction will be defined by:

$$d_k = \begin{cases} -g_k, & k = 0 \\ -g_k + \beta_k^{MKMAR} d_{k-1} - \varphi_{k-1} y_{k-1}, & k \geq 1 \end{cases} \tag{16}$$

3.1. Algorithm (TTKMAR)

1. Choose $x_0 \in \mathbb{R}^n, \in > 0, d_k = -g_k$, set $k = 0$.
2. If $||g_k|| \leq 1 \times 10^{-5}$, then stop; otherwise, go to the next step.
3. Calculate γ_{k-1} by using SWPL defined in (4) and (5).
4. Calculate x_k by (2), and compute $g_k,\ u_k$.
5. Compute the direction d_k by (16).
6. If $||g_k|| \leq \in$, stop; otherwise go to the next step.
7. If $k = n$ or $\left|g_k^T g_{k-1}\right| \geq (0.2) . \left(||g_k||^2\right)$ is hold, proceed step 1; Otherwise, proceed to next step.
8. Put $k = k + 1$, and go to (2).

The TTKMAR method meets the SDP, as demonstrated by the following lemma.

Theorem 1. *Suppose that the sequences $\{g_n\}$ and $\{d_n\}$ are generated by the TTKMAR method and the step size γ_k by using SWPL defined in (4) and (5). Then*

$$d_k^T g_k \leq -\tau.||g_k||^2, \quad with\ \tau = \left[1 - \frac{\delta\omega}{a.\delta^2}\right] \forall k \geq 0 \tag{17}$$

Proof. We arrive at the conclusion (16) through mathematical induction. For $k = 0$, $d_0^T g_0 = -||g_0||^2$, holds. If we assume that conclusion (17) holds true for $k-1$, we have $d_{k-1}^T g_{k-1} \leq -\tau.||g_{k-1}||^2$ We have from (16)

$$d_k = -g_k + \frac{g_k^T \mu_{k-1}}{g_{k-1}^T (g_k + g_{k-1})} d_{k-1} - \frac{g_k^T d_{k-1}}{g_{k-1}^T (g_k + g_{k-1})} y_{k-1}$$

Multiply both sides of the above equation by g_k^T

$$g_k^T d_k \leq -||g_k||^2 + \frac{g_k^T \left(y_{k-1} + \frac{\varphi_{k-1}}{||s_{k-1}||^2}.s_{k-1} \right)}{g_{k-1}^T (g_k + g_{k-1})} g_k^T d_{k-1} - \frac{g_k^T d_{k-1}}{g_{k-1}^T (g_k + g_{k-1})} g_k^T y_{k-1}$$

$$\leq -||g_k||^2 + \frac{\left(y_{k-1}^T g_k\right)\left(g_k^T d_{k-1}\right)}{g_{k-1}^T (g_k + g_{k-1})} + \frac{\varphi_{k-1}\left(s_{k-1}^T g_k\right)\left(g_k^T d_{k-1}\right)}{||s_{k-1}||^2 \left[g_{k-1}^T (g_k + g_{k-1})\right]} - \frac{\left(y_{k-1}^T g_k\right)\left(g_k^T d_{k-1}\right)}{g_{k-1}^T (g_k + g_{k-1})}$$

$$\leq -||g_k||^2 + \frac{\left(s_{k-1}^T g_k\right)\left(g_k^T d_{k-1}\right)^2}{||s_{k-1}||^2 \left(g_{k-1}^T (g_k + g_{k-1})\right)^2}$$

Since, $g_k^T d_{k-1} \leq ||g_k|| ||d_{k-1}||$, we have $s_{k-1} = \gamma_{k-1} d_{k-1}$, and they proved in [13] that:

$$0 \leq \beta_k^{KMAR} \leq \frac{||g_k||^2}{||g_{k-1}||^2} \tag{18}$$

we get

$$d_k^T g_k \leq -||g_k||^2 + \frac{\gamma_{k-1} ||g_k||^3 . ||d_{k-1}||^3}{\gamma_{k-1}{}^2 ||d_{k-1}||^2 . ||g_{k-1}||^4}$$

$$d_k^T g_k \leq -\left[1 - \frac{||g_k||.||d_{k-1}||}{\gamma_{k-1} ||g_{k-1}||^4}\right] ||g_k||,$$

$$d_k^T g_k \leq -\tau.||g_k||^2 \quad \text{where} \quad \tau = \left[1 - \frac{\delta\omega}{a.\delta^2}\right]$$

The proof has been completed. The following assumptions are always useful for CG approach convergence analysis.

4. Global Convergence of TTKMAR

Assumption (1)

1. $u(x)$ is restricted from below to the level set $\Lambda = \{x \in \mathcal{R}^n, u(x) \leq u(x_0)$, x_0 is the starting point. i.e., there is a constant $\omega > 0$, which means $||x_k|| \leq \omega \ \forall x \in \Lambda$.
2. $u(x)$ is continuously differentiable in a specific neighborhood N of Λ, and its gradient is Lipschitz continuous, which means that there is a constant $\mathcal{M} > 0$, s.t.

$$||g(x) - g(y)|| \leq \mathcal{M} ||x - y||, \forall x, y \in N \tag{19}$$

It is worth noting that Assumption (1) implies existence of a positive constant δ, such that [2]:

$$||g_k|| \leq \delta, \ \forall k \tag{20}$$

Theorem 2. *Assume Assumption (1) is true. Consider methods (2) and (16), where d_k is a descent direction and γ_k is provided by SWPL. If*

$$\sum_{k\geq 1} \frac{1}{||d_k||^2} = \infty \tag{21}$$

Then

$$\liminf_{k\to\infty} ||g_k|| = 0 \tag{22}$$

Theorem 3. *Assuming Assumption (1) is correct, consider Algorithm (TTKMAR), where d_k and γ_k satisfy the sufficient descent condition (17) as well as (4) and (5), respectively. Then* $\liminf_{k\to\infty} ||g_k|| = 0$

Proof. Because the descending property holds, we have $d_k \neq 0$. As a result, lemma (2) suffices to demonstrate that $||d_k||$ is bounded above. Derived from (15), and (16)

$$||d_k|| = \left|\left| -g_k + \frac{g_k^T \mu_{k-1}}{\left|g_{k-1}^T g_k\right| + ||g_{k-1}||^2} d_{k-1} - \frac{g_k^T d_{k-1}}{\left|g_{k-1}^T g_k\right| + ||g_{k-1}||^2} y_{k-1} \right|\right|$$

$$= \left|\left| -g_k + \frac{g_k^T \left(y_{k-1} + \frac{\varphi_{k-1}}{||s_{k-1}||^2}.s_{k-1}\right)}{\left|g_{k-1}^T g_k\right| + ||g_{k-1}||^2} d_{k-1} - \frac{g_k^T d_{k-1}}{\left|g_{k-1}^T g_k\right| + ||g_{k-1}||^2} y_{k-1} \right|\right|$$

As a result of $0 \leq \beta_k^{KMAR} \leq \frac{||g_k||^2}{||g_{k-1}||^2}$, $s_{k-1} = \gamma_{k-1} d_{k-1}$, (17), and (19) we obtain

$$\begin{aligned} ||d_k|| &\leq ||g_k|| + \frac{||g_k||^2}{||g_{k-1}||^2}||d_{k-1}|| + \frac{\gamma_{k-1}||g_k||^2.||d_{k-1}||^2}{\gamma_{k-1}{}^2||d_{k-1}||^2||g_{k-1}||^2}||d_{k-1}|| + \frac{||g_k||||d_{k-1}||}{||g_{k-1}||^2}\mathcal{M}\gamma_{k-1}||d_{k-1}|| \\ &= \left[1 + \frac{\gamma_{k-1}||g_k|| + ||g_k||||d_{k-1}|| + \mathcal{M}\gamma_{k-1}{}^2||d_{k-1}||^2}{\gamma_{k-1}||g_{k-1}||^2}\right]||g_k|| \\ &\leq \left[1 + \frac{a\delta + \delta\omega + \mathcal{M}a^2\omega^2}{a.\bar{\delta}^2}\right]||g_k|| \leq A.\delta \end{aligned}$$

$$||d_k|| \leq V \ \Rightarrow ||d_k||^2 \leq V^2 \tag{23}$$

By taking the summation on both sides of (23), we get

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

Inconsistency with the Zountendijk theorem [24], so

$$\liminf_{k\to\infty} ||g_k|| = 0$$

CGP has been achieved by the new proposed algorithm.

5. Mathematical Experiments

In this part, we chose some of the test functions from CUTE [4] library, which, as well as Andrei [3], Bongartz [7], and others have solved large-scale optimization problems. All codes are written in FORTRAN with double precision, and compiled Visual F90 (using the default compiler settings). To compute the value of γ_k, the cubic fitting procedure is always used.

We chose twenty-four extended or generalized large-scale unconstrained optimization problems in an extended or generalized form. Every problem was tested three times for a progressively increasing number of dimensions: $N = 1000, 10000$, and 100000, respectively, and all algorithms implemented under the SWPL (4) and (5) conditions with $\rho = 0.01$ and $\sigma = 0.85$, respectively, and the stopping criterion $||g_n|| \leq 1 \times 10^{-5}$ is used.

The execution was also analyzed using the performance profile software developed by Dolan and Mor'e [9], as shown in Figures 1 and 2.

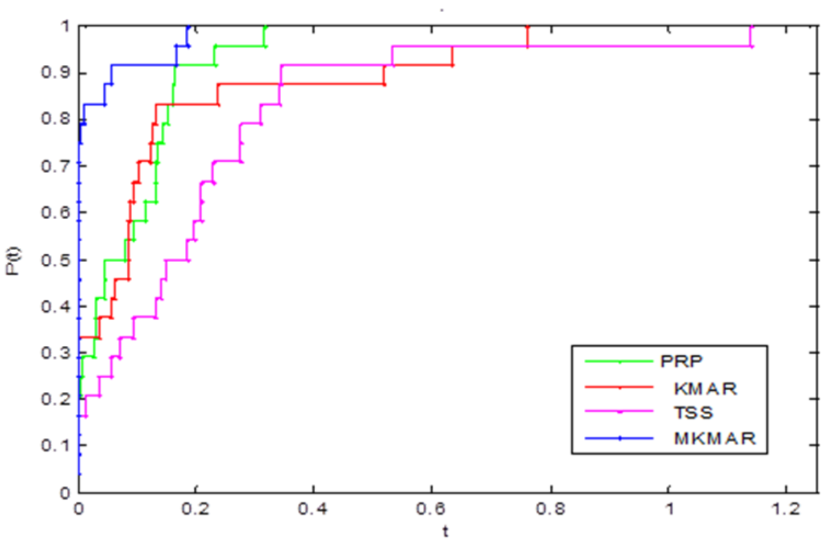

Figure 1: The performance profile of iteration.

For our comparison, we keep track of the number of iterations denotes (No.I), function evaluation denotes (No.F), finally the test function denotes (T. Fn.).

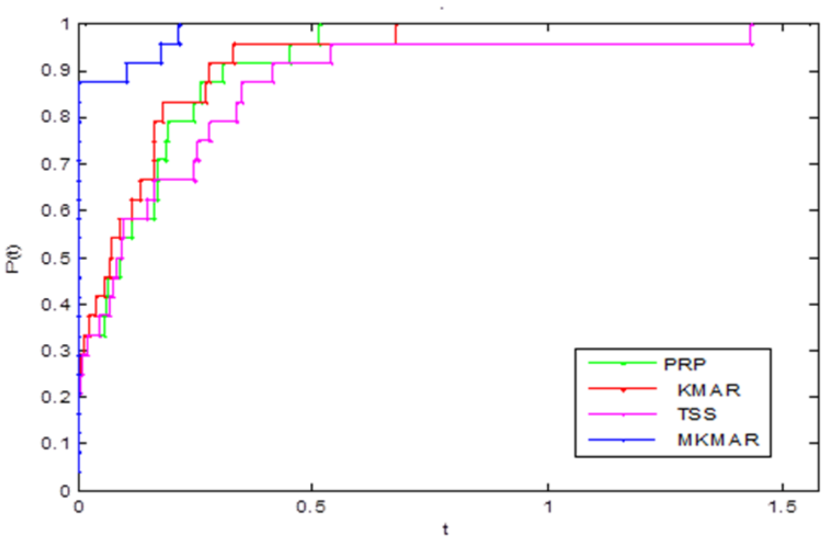

Figure 2: The performance profile of function evaluation.

Table 1: Compares Two Terms ($\beta_k^{KMAR}, \beta_k^{PRP}$), Three Terms Shanno and the Proposed Method TTKMAR, Based on No. I and No. F. ($N = 1000, 10000, 100000$)

T. Fn.	Two-terms β_k^{PRP}		Two-terms β_k^{KMAR}		Three-terms Shanno		Three-terms β_k^{MKMAR}	
	No. I	No. F	No. I	No. F	No. I	No. F	No. I	No. F
Ex. Block diagonal-1	161	337	155	332	156	335	151	319
Wolfe	321	652	319	648	328	675	307	632
Diagonal-6 (Cute)	9	28	8	25	8	26	8	26
Ex. Beal	58	151	57	148	62	156	51	135
DIXMAANAB (Cute)	19	56	19	49	19	52	17	45
DIXMAANAF (Cute)	19	45	16	41	19	47	16	41
NOND	155	422	180	592	178	598	149	413
Powell	119	339	141	392	124	421	112	332
OSP	1030	1868	1029	1880	1035	2120	1021	1750
PQ	1010	1901	1007	1890	1010	1938	1008	1894
Cubic	54	131	72	197	48	149	45	127
Shallow	27	69	26	67	32	80	24	63
Wood	94	220	103	235	99	273	101	227
Ex. Wood	290	612	245	590	298	622	234	565
Strait	18	42	18	42	18	42	18	42
Ex. Three Exponential	41	81	34	117	81	152	30	69
Fred	65	179	63	172	76	196	60	162
Quadratic Fn. QF1	1191	2335	1190	2332	1193	2347	1185	2328
Cosine	34	86	34	86	34	86	34	86
Gen. Tridiagonal-2	123	256	120	251	125	275	118	253
Diagonal-5	11	40	11	40	11	40	11	40
Helical	105	242	103	238	92	218	107	248
DENSCHNB	62	130	66	138	65	144	70	146
DENSCHNF	60	183	51	177	61	190	42	167
Total **No. I**	5076		4967		5172		4919	
No. F	11176		10691		11182		10110	

Table 2: The Work of the Proposed Method is Shown in Percentage.

Measures	Two-terms β_k^{PRP}	Two-terms β_k^{KMAR}	Three-terms Shanno	Three-terms β_k^{MKMAR}
No.I	98.14%	96.04%	100%	95.11%
No.F	99.95%	95.61%	100%	90.41%

In terms of percentage performance, Table 2 shows that the three-term β_k^{MKMAR} method outperforms the classic two-term β_k^{PRP}, two-term β_k^{KMAR}, and three-term Shanno methods. We discovered that the three-term β_k^{MKMAR} method saves (No. I, 4.89%), (No.F, 9.59%), the two-term β_k^{PRP} method saves (No.I, 1.86%), (No.F No.F, 0.05%), and the two-term β_k^{KMAR} method saves (No.I, 3.96%). (No.F, 4.39%). This behavior can be explained by making a minor change and adding a third term to the direction of the two-term β_k^{KMAR} method, so that the generated direction always meets the sufficient descent and globally convergent conditions.

6. Conclusions

We suggested a modification of spectral β_k^{KMAR} using three-term conjugate gradient method which defined by (16) and prove that our proposed method satisfied global convergent and desent condition. The numerical results show that the new method (TTK-MAR) is better than the classic two-term β_k^{PRP}, two-term β_k^{KMAR}, and three-term Shanno methods and more effective in practically.

References

[1] Ghada M. Al-Naemi. A new modified hs algorithm with strong powell-wolfe line search for unconstrained optimization. *Eastern-European Journal of Enterprise Technologies*, 116:14–21, 2022.

[2] Ghada M Al-Naemi and Ahmed H Sheekoo. New scaled algorithm for non-linear conjugate gradients in unconstrained optimization. *Indonesian Journal of Electrical Engineering and Computer Science*, 24(3):1589–1595, 2021.

[3] Neculai Andrei. Test functions for unconstrained optimization. *Research Institute for informatics. Center for Advanced Modeling and Optimization*, pages 8–10, 2004.

[4] Neculai Andrei. An unconstrained optimization test functions collection. *Adv. Model. Optim*, 10(1):147–161, 2008.

[5] Neculai Andrei. *Nonlinear Conjugate Gradient Methods for Unconstrained Optimization*. Springer International Publishing, 2020.

[6] Saman Babaie-Kafaki, Reza Ghanbari, and Nezam Mahdavi-Amiri. Two new conjugate gradient methods based on modified secant equations. *Journal of computational and applied Mathematics*, 234(5):1374–1386, 2010.

[7] I. Bongartz, A. R. Conn, Nick Gould, and Ph. L. Toint. CUTE. *ACM Transactions on Mathematical Software*, 21(1):123–160, mar 1995.

[8] Yu-Hong Dai and Yaxiang Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10(1):177–182, 1999.

[9] Elizabeth D. Dolan and Jorge J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, jan 2002.

[10] R. Fletcher. Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154, feb 1964.

[11] Roger Fletcher. *Practical methods of optimization, 1987*. 1987.

[12] M.R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409, dec 1952.

[13] Kamilu Kamfa, Sulaiman Ibrahim, MY Waziri, and Abdelrahman Abashar. Another improved three term prp-cg method with global convergent properties for solving unconstrained optimization problems. *Malaysian Journal of Computing and Applied Mathematics*, 1(1):1–10, 2018.

[14] Y. Liu and C. Storey. Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of Optimization Theory and Applications*, 69(1):129–137, apr 1991.

[15] E. Polak and G. Ribiere. Note sur la convergence de méthodes de directions conjuguées. *Revue française d'informatique et de recherche opérationnelle. Série rouge*, 3(16):35–43, 1969.

[16] Boris Teodorovich Polyak. The conjugate gradient method in extremal problems. *USSR Computational Mathematics and Mathematical Physics*, 9(4):94–112, 1969.

[17] Philip Wolfe. Convergence conditions for ascent methods. *SIAM Review*, 11(2):226–235, apr 1969.

[18] Yanlin Wu. A modified three-term prp conjugate gradient algorithm for optimization models. *Journal of inequalities and applications*, 2017(1):1–14, 2017.

[19] Hiroshi Yabe and Masahiro Takano. Global convergence properties of nonlinear conjugate gradient methods with modified secant condition. *Computational optimization and Applications*, 28(2):203–225, 2004.

[20] Jianzhong Zhang and Chengxian Xu. Properties and numerical performance of quasi-newton methods with modified quasi-newton equations. *Journal of Computational and Applied Mathematics*, 137(2):269–278, 2001.

[21] JZ Zhang, NY Deng, and LH0991 Chen. New quasi-newton equation and related methods for unconstrained optimization. *Journal of Optimization Theory and Applications*, 102(1):147–167, 1999.

[22] Li Zhang, Weijun Zhou, and Dong-Hui Li. A descent modified polak–ribière–polyak conjugate gradient method and its global convergence. *IMA Journal of Numerical Analysis*, 26(4):629–640, oct 2006.

[23] Li Zhang, Weijun Zhou, and Donghui Li. Some descent three-term conjugate gradient methods and their global convergence. *Optimization Methods and Software*, 22(4):697–711, aug 2007.

[24] G Zoutendijk. Nonlinear programming, computational methods. *Integer and nonlinear programming*, pages 37–86, 1970.