A New Investigation for Conjugate Gradient Methods to Solve large problems in Unconstrained Optimization

ASEEL M. QASIM
Department of Mathematics, College of education of pure sciences
University of Mosul
IRAQ

Abstract:- I modified a new formula for the standard conjugate gradient (CG) technique in this paper. This new formula will help us to improve and develop the numerical results that we need to be more effective and realized for our functions, which will be used in the future in applicants, the new factor that we hybridize with HZ coefficient is used to create the new method of conjugate gradient technique and the global convergent property will be derived for the proposed algorithm. The numerical results imply that the new suggested algorithm is more suitable and efficient when it compared to other well-known standard CG methods, such as the FR method.

Key-Words: - Search direction, low memory, Quasi newton, nonlinear optimization, Matlab.

Received: July 23, 2023. Revised: February 27, 2024. Accepted: March 13, 2024. Published: May 22, 2024.

1 Introduction

Conjugate Gradient (CG) techniques are a set of nonlinear optimization algorithms that require little memory storage, fast speed,and strongly local and global properties features are satisfied to achieve the desired outcome [1]. Specifically, the scaled nonlinear optimization problems of the following kind are intended to be solved using the CG method Minimize

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

It is common knowledge that the function f(x) is a smooth nonlinear function that is continuously differentiable in the form f:R→R. The array of the first partial derivative, g(x) is defined as ∇f(x). The CG algorithms are between the most efficient methods for minimizing the function in the form (1), particularly for complex problems[1], where x is an n-dimensional real vector. The iterative iteration of the CG algorithm would be used to create a equence of points $\{x_i\}_{i=0}$ to tackle the unconstrained optimization problem(1), starting with an estimate of $x_0 \in R^n$ and repeated iteration sequences will be computed from the next iteration as an equation below:

$$x_{i+1} = x_i + \omega_i d_i \tag{2}$$

Where $d_{i+1}$is the search direction and ω_i is the step length calculated using a decide line search method, and d_{i+1} is defined as follows:

$$d_{i+1} = \begin{cases} -g_{i+1} + B_i d_i & i \geq 1 \\ -g_{i+1} \ ; \ i = 0 \end{cases} \tag{3}$$

Where g_{i+1} and B_i are the gradient and the coefficient of the conjugate gradient, respectively [2]. Furthermore, there are several kinds of this coefficient correspond to varied scalar parameters selections[3-8]

$$B_i^{FR} = \frac{\|g_{i+1}\|^2}{\|g_i\|^2} \tag{4}$$

$$B_i^{PRP} = \frac{g_{i+1}^T y_i}{\|g_i\|^2} \tag{5}$$

$$B_i^{HS} = \frac{g_{i+1}^T y_i}{t_i^T y_i} \tag{6}$$

$$B_i^{DY} = \frac{\|g_{i+1}\|^2}{t_i^T y_i} \tag{7}$$

Which is determining and effect for the conjugate gradient method, several references are existed for these formulas, see for example [9-14]. The selected line search condition is described as follows for the step calculation:

$$f(x_i + \omega_i ɖ_i) - f(x_i) \le p_1 \omega_i \, g_i^T ɖ_i \tag{8}$$

$$g(x_i + \omega_i ɖ_i)^T k_i \ge p_2 \, g_i^T ɖ_i \tag{9}$$

Additionally, the strong Wolfe condition (SWC) which satisfies conditions (8) and

$$|g(x_i + \omega_i ɖ_i)^T ɖ_i| \le -p_2 g_i^T ɖ_i \tag{10}$$

Where $0 < p_1 < p_2 < 1$, more formulas have been presented by many researchers such as the line search of Armijo, Goldstein condition, for more details see sources [15-19].

The following is the order of this paper's sections: The revised formula for coefficient B is generated in the following section, followed by sections 3 and 4, which discuss sufficient descent property and global convergence, while section 5 and 6 contains numerical experience and conclusion respectively.

2 A Novel Scalar Equation for the version Parameter

Our new principle of this research based on the idea of unconstrained optimization methods with the condition of Quasi-Newton and the limited memory BFGS method [20], when the search direction is a Quasi-Newton direction $ɖ_i$ can be written as equation (11) below:

$$ɖ_{i+1} = -G_{i+1}^{-1} g_{i+1} \tag{11}$$

Where G_k is satisfying the positive definite symmetric matrix properties that meet the requirements of the quasi-Newton equation.

We provide a newer version of the parameter in this section of the article ʋ which is used to scale the direction of the search, the suggestion problem solves the iteration method using the method of quasi Newton for the current iteration where the objective function's gradient is denoted by g_i, and the matrix H_i is approximating $G(x_i)$ where G_i is the real Hessian function. Traditionally, the Hessian matrix approximation iteration must satisfy the alleged secant equation $H_{i+1}s_i = y_i$ for

$$y_i = g_{i+1} - g_i \tag{12}$$

$$s_i = x_{i+1} - x_i \tag{13}$$

Now if we use the coefficient of HZ [21] to derive a new scale ʋ for the direction search where B_i^{HZ} is:

$$B_i^{HZ} = (y_i - 2ɖ_i \frac{\|y_i\|^2}{ɖ_i^t y_i})^t \frac{g_{i+1}}{ɖ_i^t y_i} \tag{14}$$

and for any scalar Ȼ then we will have :

$$ɖ_{i+1} = -Ȼ\, g_{i+1} + B_i^{HZ} s_i \tag{15}$$

If the direction is Newton direction then by use equation (11) we have:

$$-G_{i+1}^{-1} g_{i+1} = -Ȼ g_{i+1} + B_i^{HZ} s_i \tag{16}$$

Now if we substitute the value of B_i^{HZ} (14) in equation (15):

$$-G_{i+1}^{-1} g_{i+1} = -Ȼ g_{i+1} + (y_i - 2ɖ_i \frac{\|y_i\|^2}{ɖ_i^t y_i})^t \frac{g_{i+1}}{ɖ_i^t y_i} s_i \tag{17}$$

And multiplying both sides of the equation (17) by G_{i+1} :

$$-g_{i+1} = -Ȼ G_{i+1} g_{i+1} + (y_i - 2ɖ_i \frac{\|y_i\|^2}{ɖ_i^t y_i})^t \frac{g_{i+1}}{ɖ_i^t y_i} G_{i+1} s_i$$

Then if we multiply the above equation by s_i^t we get:

$$-s_i^t g_{i+1} = Ȼ s_i^t G_{i+1} g_{i+1} + (y_i - 2ɖ_i \frac{\|y_i\|^2}{ɖ_i^t y_i})^t \frac{g_{i+1}}{ɖ_i^t y_i} . s_i^t G_{i+1} s_i \tag{18}$$

To simplify Ȼ multiply both sides of the equation (18) by $ɖ_i^t y_i$

$$Ȼ s_i^t G_{i+1} g_{i+1} ɖ_i^t y_i = s_i^t g_{i+1} ɖ_i^t y_i + (y_i - 2ɖ_i \frac{\|y_i\|^2}{ɖ_i^t y_i})^t g_{i+1} s_i^t G_{i+1} s_i \tag{19}$$

then,

$$Ȼ = \frac{s_i^t g_{i+1}.ȡ_i^t y_i + (y_i - 2ȡ_i \frac{\|y_i\|^2}{ȡ_i^t y_i})^t g_{i+1}.s_i^t G_{i+1} s_i}{s_i^t G_{i+1} g_{i+1}.\, ȡ_i^t y_i} \quad (20)$$

Now using the secant relation $G_{i+1} s_i = y_i$

$$Ȼ = \frac{s_i^t g_{i+1}.ȡ_i^t y_i + (y_i - 2ȡ_i \frac{\|y_i\|^2}{ȡ_i^t y_i})^t g_{i+1}.s_i^t y_i}{y_i^t g_{i+1}.ȡ_i^t y_i} \quad (21)$$

The last equation (21) represent the value of the new parameter Ȼ which is updating by every calculation of y and this will give a good effect in the numerical results as we have seen in the last section of this research. Therefore Ȼ is most recent to determine the worth of y as we will notice in the section of numerical results.

3 The global convergence theorem:

Before we begin of the proof of our new algorithm, we most state the following basic assumptions of the objective function [22]:

(3.1)Assumptions:

In this section we state the principles concept of the statement that we need to complete the proof of the global convergence theorems:

(3.1.1) The level set $\psi = \{x_0 \epsilon\, R^n: f(x) \leq f(x_0)\}$ is bounded where x_0 is a stationary point namely, there is a constant Q that makes $\|x\| \leq Q, \forall\, \hat{x} \in R^2$.

(3.1.2) Assume ψ is some neighborhood of Ḛ then the function is continually diffable and its gradient Lipschitz continuous, that is there exist $k > 0$ such that for all x.

$\|g(x) - g(\hat{x})\| \leq k\|x - \hat{x}\|$

$\forall x, \hat{x} \in$ Ḛ, from assumption (3.1.1) and (3.1.2) we can directly get that$\|g_i\|$is bounded[23].

(3.2) Theorem(descent property)

Suppose that assumption (3.1.1) holds and ω satisfies the line search equations (8) and (9) and B is named in (14) then the search direction (3) satisfies the descent property.

Proof:

Substitute equation (14) and (21) in (15) we have :

$$ȡ_{i+1} = -Ȼ g_{i+1} + B_i s_i$$

$$ȡ_{i+1} g_{i+1}^t = -Ȼ\|g_{i+1}\|^2 + B_i g_{i+1}^t s_i \quad (22)$$

$$ȡ_{i+1} g_{i+1}^t = -\frac{s_i^t g_{i+1} ȡ_i^t y_i + (y_i - 2ȡ_i \frac{\|y_i\|^2}{ȡ_i^t y_i})^t g_{i+1}.\, s_i^t y_i}{y_i^t g_{i+1}.\, ȡ_i^t y_i}\|g_{i+1}\|^2 + (y_i - 2ȡ_i \frac{\|y_i\|^2}{ȡ_i^t y_i})^t \frac{g_{i+1}}{ȡ_i^t y_i} g_{i+1}^t s_i$$

$$ȡ_{i+1} g_{i+1}^t = -\frac{s_i^t g_{i+1}.\, ȡ_i^t y_i + (y_i - 2ȡ_i \frac{\|y_i\|^2}{ȡ_i^t y_i})^t g_{i+1}.\, s_i^t y_i}{y_i^t g_{i+1}.\, ȡ_i^t y_i}\|g_{i+1}\|^2 + (y_i^t g_{i+1} - 2ȡ_i^t g_{i+1} \frac{\|y_i\|^2}{ȡ_i^t y_i}) \frac{g_{i+1}^t s_i}{ȡ_i^t y_i} \quad (23)$$

Now for some s= mȡ:

$$\|g_{i+1}\|^2 \leq -\frac{mȡ_i^t g_{i+1} ȡ_i^t y_i + (y_i^t - 2ȡ_i^t \frac{\|y_i\|^2}{ȡ_i^t y_i}) g_{i+1} mȡ_i^t y_i}{y_i^t g_{i+1}.\, ȡ_i^t y_i}\|g_{i+1}\|^2 + (y_i^t g_{i+1} - 2ȡ_i^t g_{i+1} \frac{\|y_i\|^2}{ȡ_i^t y_i}) \frac{m g_{i+1}^t ȡ_i}{ȡ_i^t y_i}$$

$$\leq -\frac{mȡ_i^t g_{i+1} + m(y_i^t g_{i+1} - 2ȡ_i^t g_{i+1} \frac{\|y_i\|^2}{ȡ_i^t y_i})}{y_i^t g_{i+1}}\|g_{i+1}\|^2 + y_i^t g_{i+1} \frac{m g_{i+1}^t ȡ_i}{ȡ_i^t y_i}$$

$$\leq -^{m}\|g_{i+1}\|^2 + \frac{m(2ȡ_i^t g_{i+1} \frac{\|y_i\|^2}{ȡ_i^t y_i})}{y_i^t g_{i+1}}\|g_{i+1}\|^2$$

$$+ y_i^t g_{i+1} \frac{m g_{i+1}^t ȡ_i}{ȡ_i^t y_i}$$

from Wolfe and Powell restart[24]

$|g_k^T g_{k+1}| \geq 0.2 g_{k+1}^2$ and $ð\|g_i\|^2 = -ð g_i^t ȡ$

$$\leq -m\|g_{i+1}\|^2 + \frac{(-2ðg_i^t d_i\|y_i\|^2)}{0.2\|g_{i+1}\|^2.(ð-1)g_i^t d_i} m\|g_{i+1}\|^2 + my_i^t g_{i+1}\frac{-ðg_i^t d_i}{(ð-1)g_i^t d_i}$$

$$\leq -m\|g_{i+1}\|^2 + \frac{-10ð.lm}{(ð-1)\|g_{i+1}\|^2}\|g_{i+1}\|^2 + m\frac{-ð}{(ð-1)}\|g_{i+1}\|^2$$

$$\leq -m\|g_{i+1}\|^2 + \frac{-mð}{(ð-1)}\|g_{i+1}\|^2$$

$$\leq -c\|g_{i+1}\|^2 \quad (24)$$

(3.1.3)Theorem:

If we suppose that Assumptions (3.1.1) and (3.1.2) hold, Consider any conjugate gradient method in the form (3) is decent direction for all k ≥ 0, and to meet the Wolfe criteria, the step length w and the search direction conditions fulfil (8), (9) and (10) then the objective function is general function, then

$$\lim_{k\to\infty} \inf\| g_{i+1}\| = 0$$

Proof:

If we consider the direction of the equation (15) and the coefficient of (21) satisfy the following condition: $|ᑕ| =$

$$\left|\frac{s_i^t g_{i+1} d_i^t y_i + (y_i g_{i+1} - 2d_i\, g_{i+1}\frac{\|y_i\|^2}{d_i^t y_i})^t m d_i^t y_i}{y_i^t g_{i+1}.d_i^t y_i}\right|$$

$$= \left|\frac{md_i^t g_{i+1} + (y_i g_{i+1} - 2d_i\, g_{i+1}\frac{\|y_i\|^2}{d_i^t y_i})^t m}{y_i^t g_{i+1}}\right| \quad (25)$$

$$= \left|\frac{md_i^t g_{i+1}.d_i^t y_i + y_i g_{i+1} d_i^t y_i - 2d_i\;\; g_{i+1}\|y_i\|^2)^t m}{y_i^t g_{i+1}.d_i^t y_i}\right|$$

$$|ᑕ| = \left|1 + \frac{md_i^t g_{i+1}}{y_i^t g_{i+1}} - \frac{2d_i^t g_{i+1}\|y_i\|^2 m}{y_i^t g_{i+1}.d_i^t y_i}\right| \quad (26)$$

From Wolfe (8), (9) and Powell [24] and $d_i^t y_i \geq (ʊ-1).d_i^t g_i$

$$= \left|1 - \frac{mʊd_i^t g_i}{0.2\|g_{i+1}\|^2} + \frac{2mʊd_i^t g_i\|y_i\|^2 m}{0.2(1-ʊ)\|g_{i+1}\|^2.d_i^t g_i}\right|$$

$$= \left|1 - \frac{mʊd_i^t g_i}{0.2\|g_{i+1}\|^2} + \frac{2mʊ\|y_i\|^2 m}{0.2(1-ʊ)\|g_{i+1}\|^2}\right|$$

$$|ᑕ| = \left|1 - \frac{mʊ\ell^2}{0.2\mathcal{E}^2} + \frac{2mʊð^2}{0.2(1-ʊ)\mathcal{E}^2}\right| \quad (27)$$

$|ᑕ| = |ꞃ|$=F

Furthermore, by combining (27) and (14), we get:

$$d_{i+1} \leq |ᑕ|\|g_{i+1}\| + |B|\|s_i\|$$
$$\leq F + G = J$$

$$\sum_{k\geq 1}\frac{1}{\|d_{i+1}\|^2} \geq \frac{1}{J}\sum 1 = \infty$$

4 Numerical Experience:

We now analyze the efficiency of the suggested method by representing the numerical experiments that our method got it from some test problem selected by Mor'e, et al. [25] and Andrei [26]. Table 1 displays the test difficulties as follows: the columns n and iterate represent the test issue dimension and the number of iterations, while the time operator and the function's evaluation indicate respectively in the column of Time and NOF. The issues' dimensions span from 2 to 1000, And we used E=1×10^-5 as the stopping condition, meaning that the algorithm methods were terminated once the line search condition ||g||< 1×10^-5 was met.

5 Discussion

When we contrast our new recommended conjugate gradient method, with FR method we note that the first one method exhibits superior improvement than FR method. We use the MATLAB program to code each approach, and the number of iterations, function evaluations, and CPU time were used to compare the numerical results. Figures 1 and 2 illustrate the performance of the new approach in comparison with the new method, and Table 1 indicates that the findings we obtain were more accurate and effective than the FR method.

6 Conclusions

In this study, we use this unique combination of conjugate gradient methods to establish a modern form of condition based on a new Quasi-Newton equation. It has been established that these formulas are convergent globally. Lastly, several numerical findings have been published to demonstrate the new formula's efficacy.

Figure2:The iteration performance

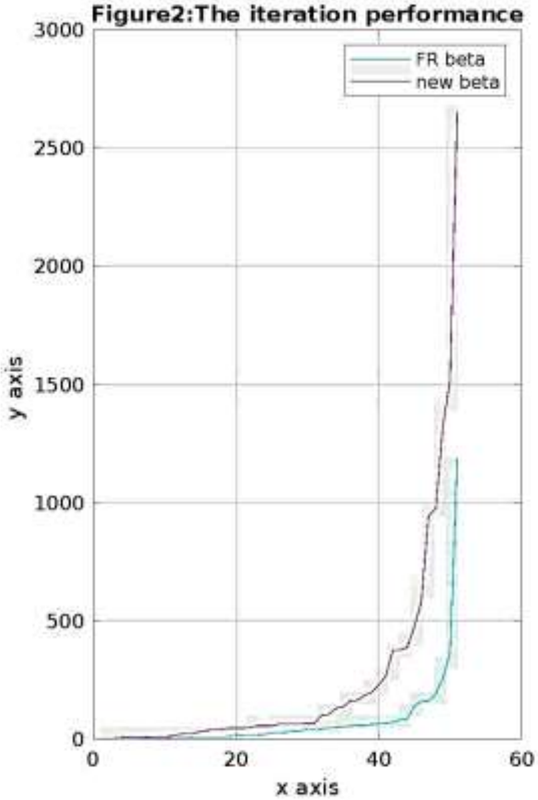
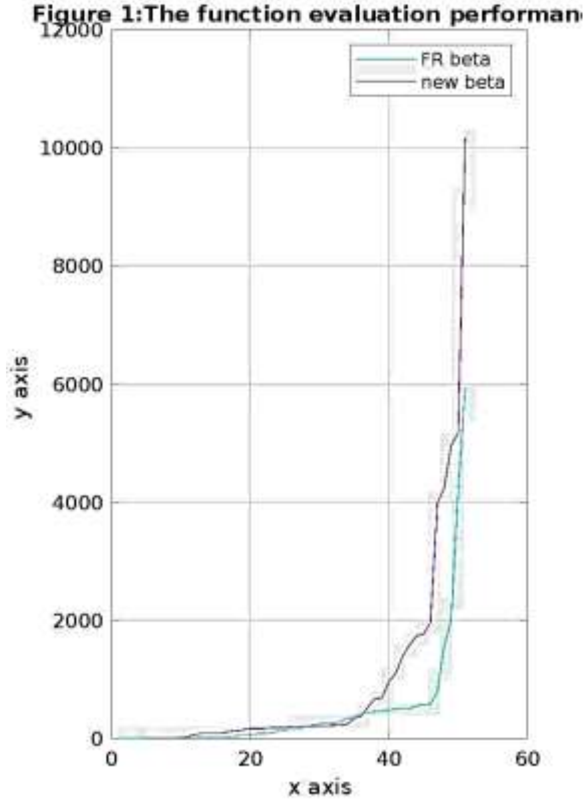


Figure 1:The function evaluation performan



functions	n	New			Origin FR		
		N	Time	NOF	NOI	Time	NOF
Extended Tridiagonal_1function	100	59	7.7224	344	46	7.6057	95
	500	71	8.673	460	46	5.5736	95
	1000	81	10.7371	446	46	8.2086	95
Extended three exponential	2	14	7.9331	50	135	9.0494	205
	4	11	8.9141	37	13	4.9572	23
	10	10	6.3319	31	23	10.5806	66
EXTROSNB (CUTE)	4	85	4.6533	487	585	6.9324	1736
	10	259	8.1376	2013	1517	4.958	5166
Full HESSIEAN	100	4	6.0968	16	65	9.8477	205
	500	4	5.9767	20	185	7.6375	686
	1000	4	6.4932	24	20	7.8385	120
Diagonal 7	100	7	6.6468	20	7	17.243	12
	500	7	6.645	20	8	9.2658	14
	1000	7	9.3462	20	8	5.935	14
Extended Freudenstein & Roth	4	47	14.3741	170	198	10.1078	378
Extended Beale	100	28	6.1582	125	161	10.7745	542
	500	66	4.9034	286	388	6.9016	1587
	1000	68	7.146	355	474	8.9469	1960
SINQUAD (CUTE)	100	365	8.6162	4178	975	9.271	4251
	10	42	5.8437	251	102	7.38	230
	4	55	9.154	267	129	11.788	212
SINCOS	100	18	7.494	86	56	6.3021	175
	500	19	5.9818	500	56	5.998	175
	1000	17	7.3724	131	56	9.1596	175
HIMMELBH (CUTE)	100	3	7.4456	27	10	6.3332	10
	500	3	6.5568	27	10	5.0704	10
	1000	4	7.3943	54	10	7.4184	10
Extended Block-Diagonal BD2	100	26	8.29	96	64	8.87	209
	500	54	9.58	227	65	9.3	223
	1000	34	12.39	170	49	7.67	149
HIMMELBG	100	2	16.3985	2	2	8.4636	2
	500	2	7.2897	2	2	10.6655	2
	1000	2	8.9534	2	2	8.7331	2
BIGGSB1 (CUTE)	100	162	7.3317	581	376	6.1117	1394
	500	1186	7.8274	5935	1323	7.7813	4965
	1000	3	9.2692	5	2652	21.187	10165

Generalized quartic GQ1 function	100	11	27.3525	27	63	8.151	184
	500	11	8.1733	27	63	6.184	184
	1000	11	8.8738	27	63	10.3052	184
Dixon	100	35	15.2	499	39	7.283	219
	500	43	9.47	556	39	8.8	219
	1000	41	9.4	466	37	13.57	244
Generalized quartic GQ2 function	100	62	5.8047	426	266	6.9297	1108
	500	59	6.4372	580	164	5.9456	678
	1000	189	8.9385	1574	228	6.9949	965
Penalty	4	16	12.4	54	24	7.08	40
	10	26	8.65	84	31	7.1	90
	100	49	9.422	205	946	8	3975
sum	500	140	7.626	504	377	9.0881	1769
	1000	160	6.7658	800	104	5.7982	347
Total		3694	437.377	23566	12314	430.7244	45696

Table 1: Numerical results of the new method vs. FR method

References

[1] Al-Baali, Mehiddin. "Descent property and global convergence of the Fletcher—Reeves method with inexact line search." *IMA Journal of Numerical Analysis* 5.1 (1985): 121-124.

[2] Chen, Yuanyuan. "Global convergence of a new conjugate gradient method with Wolfe type line search." *Journal of Information and Computing Science* 7.1 (2012): 67-71.

[3] Hestenes, Magnus R., and Eduard Stiefel. "Methods of conjugate gradients for solving linear systems." *Journal of research of the National Bureau of Standards* 49.6 (1952): 409-436.

[4] Fletcher, Reeves, and Colin M. Reeves. "Function minimization by conjugate gradients." *The computer journal* 7.2 (1964): 149-154

[5] Polak, Elijah, and Gerard Ribiere. "Note sur la convergence de méthodes de directions conjuguées." Revue française d'informatique et de recherche opérationnelle. Série rouge 3.16 (1969): 35-43.

[6] Polyak, Boris Teodorovich. "The conjugate gradient method in extremal problems." USSR Computational Mathematics and Mathematical Physics 9.4 (1969): 94-112.

[7] Dai, Yu-Hong, and Yaxiang Yuan. "A nonlinear conjugate gradient method with a strong global convergence property." *SIAM Journal on optimization* 10.1 (1999): 177-182.

[8] Al-Baali, Mehiddin, Yasushi Narushima, and Hiroshi Yabe. "A family of three-term conjugate gradient methods with sufficient descent property for unconstrained optimization." *Computational Optimization and Applications* 60 (2015): 89-110.

[9] Andrei, Neculai. "A new three-term conjugate gradient algorithm for

unconstrained optimization." *Numerical Algorithms* 68 (2015): 305-321.

[10] Babaie-Kafaki, Saman. "Two modified scaled nonlinear conjugate gradient methods." *Journal of computational and applied mathematics* 261 (2014): 172-182.

[11] Dong, XiaoLiang, Hongwei Liu, and Yubo He. "A self-adjusting conjugate gradient method with sufficient descent condition and conjugacy condition." *Journal of Optimization Theory and Applications* 165 (2015): 225-241.

[12] Dong, Yunda. "A practical PR+ conjugate gradient method only using gradient." *Applied Mathematics and Computation* 219.4 (2012): 2041-2052.

[13] Sulaiman, I. M., et al. "Some Three-Term Conjugate Gradient Algorithms with Descent Condition for Unconstrained Optimization Models." *J. Adv. Res. Dyn. Control Syst* 12 (2020): 2494-2501.

[14] Luo, Wenjun, and Zezhong Wu. "Two modified adaptive parameter conjugate gradientmethods for unconstrained optimization." (2023).

[15] Yamakawa, Yuya, Hiroyuki Sato, and Kensuke Aihara. "Modified Armijo line-search in Riemannian optimization with reduced computational cost." *arXiv preprint arXiv:2304.02197* (2023).

[16] Dai, Zhifeng, and Fenghua Wen. "Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property." *Applied Mathematics and Computation* 218.14 (2012): 7421-7430.

[17] Liu, Hailin. "A mixture conjugate gradient method for unconstrained optimization." *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*. IEEE, 2010.

[18] Zhang, Li, Weijun Zhou, and Donghui Li. "Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search." *Numerische mathematik* 104.4 (2006): 561-572.

[19] Yuan, Gonglin, and Xiwen Lu. "A modified PRP conjugate gradient method." *Annals of Operations Research* 166.1 (2009): 73-90.

[20] Byrd, Richard H., Jorge Nocedal, and Ya-Xiang Yuan. "Global convergence of a cass of quasi-Newton methods on convex problems." *SIAM Journal on Numerical Analysis* 24.5 (1987): 1171-1190.

[21] Hager, William W., and Hongchao Zhang. "A survey of nonlinear conjugate gradient methods." *Pacific journal of Optimization* 2.1 (2006): 35-58.

[22] Zoutendijk, G. "Nonlinear programming, computational methods." *Integer and nonlinear programming* (1970): 37-86.

[23] -H. Dai, Y., and L. -Z. Liao. "New conjugacy conditions and related nonlinear conjugate gradient methods." *Applied Mathematics and optimization* 43 (2001): 87-101.

[24] Powell, Micheal JD. "On the convergence of the variable metric algorithm." *IMA Journal of Applied Mathematics* 7.1 (1971): 21-36.

[25] Dolan, Elizabeth D., and Jorge J. Moré. "Benchmarking optimization software with performance profiles." Mathematical programming 91 (2002): 201-213.

[26] Andrei, Neculai. "A new three-term conjugate gradient algorithm for unconstrained optimization." *Numerical Algorithms* 68 (2015): 305-321.