Available online at www.refaad.com
Gen. Lett. Math., 9(2) (2020), 101- 105
Research Article

General Letters in Mathematics (GLM)

Journal Homepage: http://www.refaad.com/views/GLM/home.aspx

ISSN: 2519-9277 (Online) 2519-9269 (Print)

A New conjugate gradient method for unconstrained optimization problems with descent property

Hussein Ageel Khatab [a*], Salah Gazi Shareef [b]

[a,b] Department of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region-Iraq
E- mails: [a] husseinkhatab632@gmail.com, [b] salah.gazi2014@gmail.com

Abstract

In this paper, we propose a new conjugate gradient method for solving nonlinear unconstrained optimization. The new method consists of three parts, the first part of them is the parameter of Hestenes-Stiefel (HS). The proposed method is satisfying the descent condition, sufficient descent condition and conjugacy condition. We give some numerical results to show the efficiency of the suggested method.

Keywords: *Unconstrained Optimization, Conjugate Gradient Method, Three Term Conjugate Gradient Algorithm, Descent Condition, Sufficient Descent Condition and Conjugacy Condition.*

1. Introduction

Consider the unconstrained optimization problem:

$$\text{Min } f(x), x \in R^n \tag{1.1}$$

where $f: R^n \to R$ is a real-valued, continuously differentiable function

A nonlinear conjugate gradient method for solving (1.1) are iterative methods of the form

$$x_{k+1} = x_k + \alpha_k d_k \tag{1.2}$$

starting from an initial guess $x_1 \in R^n$,

where $v_k = x_{k+1} - x_k$, the positive step size α_k is obtained by one dimensional line search, and d_k is a search direction. The search direction for the first iteration is the steepest descent direction, namely

$$d_1 = -g_1 \tag{1.3}$$

and the other search directions can be defined as:

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{1.4}$$

where $g_k = \nabla f(x_k)$ and β_k is a scalar. Some formulas for β_k are called Hestenes-Stiefel (HS) [8], Liu and Storey [9], Polak-Ribiere-Polyak (PRP) [11], Dai and Liao [2], Dai and Yuan (DY) [3], (CD)[4] and Fletcher-Reeves (FR) [5] Proposed by are given below:

$$\beta_k^{HS} = \frac{g_{k+1}^T(g_{k+1}-g_k)}{d_k^T(g_{k+1}-g_k)} \tag{1.5}$$

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k} \tag{1.6}$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T(g_{k+1}-g_k)}{\|g_k\|^2} \tag{1.7}$$

$$\beta_k^{DL} = \frac{g_{k+1}^T(y_k - t s_k)}{d_k^T y_k}, \quad t > 0 \tag{1.8}$$

---

*Corresponding author
Email address: husseinkhatab632@gmail.com (Hussein Ageel Khatab)
doi: 10.31559/glm2020.9.2.6
Received: 18 Nov 2019: Revised: 30 June 2020: Accepted 4 Aug 2020

$$\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T(g_{k+1}-g_k)} \tag{1.9}$$

$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k} \tag{1.10}$$

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \tag{1.11}$$

where $y_k = g_{k+1} - g_k$, symbol $\|.\|$ denotes the Euclidean norm of vectors. The global convergence results about Fletcher-Reeves (FR) method, Polak-Ribiere-Polyak(PRP) method, Hestenes-Stiefel(HS)method, Dai-Yuan(DY) method, Conjugate Descent(CD) method and Liu-Storey(LS) method can see [1,6,13,15,12,4,16].

Also, many parameters are suggested, for example, Hager and Zhang [7] suggested a new conjugate gradient method and called CG-DESCENT method. Zhang Li et al. [16-17] also suggested some modified conjugate gradient methods.

Conjugate directions which introduce in (1.4) have the property

$$d_{k+1}^T y_k = 0$$

For Quasi-Newton methods, the search direction d_{k+1} can be calculated in the form

$$d_{k+1} = -H_{k+1} g_{k+1} \tag{1.12}$$

By above equation and Quasi-Newton condition $H_{k+1} y_k = v_k$, we get

$$d_{k+1}^T y_k = -(H_{k+1} g_{k+1})^T y_k = -g_{k+1}^T (H_{k+1} y_k) = -g_{k+1}^T v_k \tag{1.13}$$

Perry replaced the conjugacy condition $d_{k+1}^T y_k = 0$ by the condition

$$d_{k+1}^T y_k = -g_{k+1}^T v_k \tag{1.14}$$

Dai and Liao introduce the following conjugacy condition:

$$\mathrm{d}_{k+1}^T \mathrm{y}_k = -\mathrm{t}\mathrm{g}_{k+1}^T \mathrm{v}_k \tag{1.15}$$

where $t \geq 0$ is a scalar.

This paper is organized as follow: in Section 2, we suggest a new conjugate gradient method. In Section 3, we prove the descent condition, sufficient descent condition and conjugacy condition of the new method. In Section 4, we present the numerical results and we give the conclusion in section 5.

2. Derivation of The New Method

There are many three-term conjugate gradient algorithms suggested for solving nonlinear unconstrained optimization, the first three-term nonlinear conjugate gradient algorithm was presented by Nazareth [5], in which the search direction is determined by

$$d_{k+1} = -y_k + \frac{y_k^T y_k}{y_k^T d_k} d_k - \frac{y_{k-1}^T y_k}{y_{k-1}^T d_{k-1}} d_{k-1} \tag{2.1}$$

With $d_{-1} = 0$, $d_0 = -g_0$

(ZZL) [6] proposed a computationally efficient three-term CG method with the following search direction:

$$d_{k+1} = -g_{k+1} + \beta_k^{HS} d_k - \frac{g_{k+1}^T d_k}{d_k^T y_k} y_k \tag{2.2}$$

To derive the new method, firstly, we suppose that

$$d_{k+1} = -g_{k+1} + \beta_k^{NEW} d_k - \mu \frac{g_{k+1}^T d_k}{\|g_k\|^2} g_k \quad , \quad \mu \in (0,1) \tag{2.3}$$

As a new three-term conjugate gradient method multiplying both sides of equation (2.3) by y_k, we have

$$d_{k+1}^T y_k = -g_{k+1}^T y_k + \beta_k^{NEW} d_k^T y_k - \mu \frac{g_{k+1}^T d_k}{\|g_k\|^2} g_k^T y_k \tag{2.4}$$

Now, from equation (1.14) and equation (2.4), we get

$$-g_{k+1}^T v_k = -g_{k+1}^T y_k + \beta_k^{NEW} d_k^T y_k - \mu \frac{g_{k+1}^T d_k}{\|g_k\|^2} g_k^T y_k \tag{2.5}$$

Implies that

$$\beta_k^{NEW} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k} + \mu \frac{g_{k+1}^T d_k}{d_k^T y_k \|g_k\|^2} g_k^T y_k \quad , \text{ where } \mu \in (0,1) \tag{2.6}$$

2.1 Algorithm of New Method

Step (1): Select x_1 and $\varepsilon = 10^{-5}$.
Step (2): Set $d_1 = -g_1$, $g_k = \nabla f(x_k)$, Set $k = 1$.
Step (3): Compute the step length $\alpha_k > 0$ satisfying the Wolfe line search

$$f(x_k + \alpha_k d_k) - f(x_k) \leq c_1 \alpha_k g_k^T d_k$$
$$|g_{k+1}^T d_k| \leq c_2 |g_k^T d_k|$$

where, $0 < c_1 < c_2 < 1$.

Step (4): Calculate

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

$$g_{k+1} = \nabla f(x_{k+1}), If \ \|g_{k+1}\| \le \varepsilon, \text{then stop.}$$

Step (5): Calculate β_k^{NEW} by (2.6)

Step (6): Compute $d_{k+1} = -g_{k+1} + \beta_k^{NEW} d_k$

Step (7): If $|g_{k+1}^T g_k| > 0.2\|g_{k+1}\|^2$then go to step 2.

Else

$k = k + 1$ and go to step 3.

Theorem 1: Suppose that the sequence $\{x_k\}$ is generated by (1.2), then the search direction in (1.4) with new conjugate gradient (2.6) satisfy the descent condition, i.e. $d_{k+1}^T g_{k+1} \le 0$ with exact and inexact line search.

Proof: From (1.4) and (2.6) we have,

$$d_{k+1} = -g_{k+1} + \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k} + \mu \frac{g_{k+1}^T d_k}{d_k^T y_k \|g_k\|^2} g_k^T y_k\right) d_k \tag{2.7}$$

after we multiplying both sides by g_{k+1}, we obtain

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} - \frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} + \mu \frac{(g_{k+1}^T d_k)^2}{d_k^T y_k \|g_k\|^2} g_k^T y_k \tag{2.8}$$

if the step size is chosen by an exact line search which is $d_k^T g_{k+1} = 0$, then, the proof is done.

Now, we prove that the equation (2.8) is satisfies the descent condition If $d_k^T g_{k+1} \neq 0$.

By mathematical induction, $d_0^T g_0 = -\|g_0\|^2 \le 0$, where $d_0 = -g_0$, to prove case $K + 1$, firstly, we assume that it is true for case k that is mean $d_k^T g_k \le 0$, and this is true if $d_k = -g_k$, now, since $d_k = -g_k$, then the equation (2,8) becomes,

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

since the parameter of (HS) is satisfies the descent condition, then, the first two terms of equation (2.8) are less than or equal to zero, and it is clear that the third and fourth terms are less than zero, since $\alpha_k, \ \mu, and \ d_k^T y_k$ are positive, so we get

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} - \frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} - \mu \frac{(g_{k+1}^T d_k)^2}{\|g_k\|^2} < 0. \blacksquare$$

Theorem 2: Assume that the sequence $\{x_k\}$ is generated by (1.2), then the search direction in (1.4) with new conjugate gradient (2.6) satisfies the sufficient descent condition.

$$d_{k+1}^T g_{k+1} \le -C\|g_{k+1}\|^2.$$

Proof: From (1.4) and (2.6) we get

$$d_{k+1} = -g_{k+1} + \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\alpha_k g_{k+1}^T d_k}{d_k^T y_k} + \mu \frac{g_{k+1}^T d_k}{d_k^T y_k \|g_k\|^2} g_k^T y_k \right) d_k \tag{2.9}$$

Multiply both sides of above equation by g_{k+1}, to get

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} - \frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} + \mu \frac{(g_{k+1}^T d_k)^2}{d_k^T y_k \|g_k\|^2} g_k^T y_k \tag{2.10}$$

since the parameter of (HS) is satisfies the descent condition a, then the equation (2,10) becomes

$$\begin{aligned} d_{k+1}^T g_{k+1} &\le -\frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} + \mu \frac{(g_{k+1}^T d_k)^2}{d_k^T y_k \|g_k\|^2} g_k^T y_k \\ &= -\frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} - \mu \frac{(g_{k+1}^T d_k)^2}{\|g_k\|^2} \\ &= -\left[\frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} + \mu \frac{(g_{k+1}^T d_k)^2}{\|g_k\|^2}\right] * \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \end{aligned}$$

Since,

$$= -\|g_{k+1}\|^2 \left[\frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} + \mu \frac{(g_{k+1}^T d_k)^2}{\|g_k\|^2}\right] * \frac{1}{\|g_{k+1}\|^2}$$

Let $C = \left[\frac{\alpha_k (g_{k+1}^T d_k)^2}{d_k^T y_k} + \mu \frac{(g_{k+1}^T d_k)^2}{\|g_k\|^2}\right] * \frac{1}{\|g_{k+1}\|^2}$ which is positive, then

$$d_{k+1}^T g_{k+1} \le -C\|g_{k+1}\|^2. \blacksquare$$

Theorem 3: Suppose that the sequence $\{x_k\}$ is generated by (1.2), then the search direction in (1.4) with new conjugate gradient (2.6) satisfy the conjugacy condition.

Proof: From (1.4), (2.6) and multiply both sides by y_k , we have

$$d_{k+1}^T y_k = -g_{k+1}^T y_k + \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\alpha_k g_{k+1}^T d_k}{d_k^T y_k} + \mu \frac{g_{k+1}^T d_k}{d_k^T y_k \|g_k\|^2} g_k^T y_k \right) d_k^T y_k \quad (2.11)$$

Implies that

$$d_{k+1}^T y_k = -g_{k+1}^T y_k + \left(g_{k+1}^T y_k - \alpha_k g_{k+1}^T d_k + \mu \frac{g_{k+1}^T d_k}{\|g_k\|^2} g_k^T y_k \right) \quad (2.12)$$

Then,

$d_{k+1}^T y_k = -\alpha_k g_{k+1}^T d_k + \mu \frac{g_{k+1}^T d_k}{\|g_k\|^2} g_k^T y_k$

Since $d_k = -g_k$, then,

$$d_{k+1}^T y_k = -g_{k+1}^T d_k \left(\alpha_k + \mu \frac{d_k^T y_k}{\|g_k\|^2}\right) \quad (2.13)$$

Hence,

$d_{k+1}^T y_k = -g_{k+1}^T v_k \frac{(\alpha_k + \mu \frac{d_k^T y_k}{\|g_k\|^2})}{\alpha_k}$

since $\frac{(\alpha_k + \mu \frac{d_k^T y_k}{\|g_k\|^2})}{\alpha_k} > 0$, let $t = \frac{(\alpha_k + \mu \frac{d_k^T y_k}{\|g_k\|^2})}{\alpha_k}$, so, we have,

$d_{k+1}^T y_k = -t g_{k+1}^T v_k = 0.$ ■

3. Numerical Results

In this section, we report the detailed numerical results of a number of problems by new method. We compare our method with Conjugate Gradient algorithms (HS) and (DY), the comparative tests involve nonlinear unconstrained problems (standard test function) with different dimensions$4 \leq n \leq 5000$, programs are written in FORTRAN90 language, the stopping condition for all cases is $\|g_{k+1}\| \leq 10^{-5}$. The results given in tables (1) and (2) specifically quote the number of function (NOF) and the number of iteration (NOI). More experimental results in tables (1) and (2) confirm that the new method is superior to standard Conjugate Gradient methods (HS) and (DY), with respect to the NOI and NOF.

Table (1): Comparative Performance of the three algorithms (HS, DY and New Conjugate Gradient Method)

Test function	Dim.	Algorithm of HS		Algorithm of DY		New algorithm	
		NOI	NOF	NOI	NOF	NOI	NOF
GCentral	4	22	159	18	127	19	128
	10	22	159	18	127	19	128
	50	22	159	19	138	19	128
	100	22	159	20	153	21	157
	500	23	171	23	192	21	157
	1000	23	171	23	192	22	170
	5000	28	248	24	205	26	216
Miele	4	28	85	36	115	25	72
	10	31	102	36	115	33	105
	50	31	102	45	156	33	105
	100	33	114	45	156	33	105
	500	40	146	53	188	35	119
	1000	46	176	60	222	35	119
	5000	54	211	66	257	40	140
Powell	4	38	108	50	128	28	72
	10	38	108	51	130	28	72
	50	38	108	51	130	31	91
	100	40	122	51	130	31	91
	500	41	124	51	130	31	91
	1000	41	124	51	130	31	91
	5000	41	124	52	132	31	91
Wood	4	30	68	28	65	26	61
	10	30	68	28	65	27	63
	50	30	68	28	65	27	63
	100	30	68	28	65	27	63
	500	30	68	29	68	28	65
	1000	30	68	29	68	28	65
	5000	30	68	29	68	28	65
Cubic	4	12	35	14	39	11	31
	10	13	37	15	43	11	31
	50	13	37	15	43	12	35
	100	13	37	15	43	12	35
	500	13	37	15	43	12	35
	1000	13	37	15	43	12	35
	5000	13	37	15	43	13	37
Extended Psc1	4	7	18	6	16	6	16
	10	6	16	6	16	6	16
	50	6	16	6	16	6	16
	100	7	18	6	16	6	16
	500	7	18	6	16	6	16
	1000	7	18	6	16	6	16
	5000	7	18	6	16	6	16
Sum	100	14	81	14	85	14	78

	500	21	124	21	118	22	121
	1000	23	128	24	125	19	83
Wolfe	100	49	99	45	91	49	99
	500	52	105	48	79	52	105
	1000	70	141	52	105	61	123
Total		1278	4513	1392	4729	1125	3853

Table (2): Percentage of Improving of the New Method

	Algorithm of HS	**New Algorithm**
NOI	100%	88.0281690141%
NOF	100%	85.375581653%

	Algorithm of DY	New Algorithm
NOI	100%	80.8189655172%
NOF	100%	81.4759991542%

4. Conclusion

We have suggested a new conjugate gradient method for unconstrained optimization problems. We proved the descent condition, sufficient descent condition and Conjugacy condition to the proposed method, the numerical tests were carried out on low and high dimensionality problems, and comparisons were made amongst different test functions. The new method has proven its efficiency through results in tables (1) and (2).

References

[1] E.G. Birgin and J.M. Martinez. *A spectral conjugate gradient method for unconstrained optimization*. Appl. Math. Optim., 43 (2001), 117-128.

[2] Y.H. Dai, and Liao, *New conjugacy conditions and related nonlinear conjugate gradient methods*. Appl. Math. Optim., 43(1) (2001), 87-101, https://doi.org/10.1007/s002450010019

[3] Y.H. Dai and Y. Yuan, *A nonlinear conjugate gradient with a strong global convergence properties*, SIAM J. Optim. 10 (1999), 177–182.

[4] R. Fletcher, *Practical Methods of Optimization* vol.1: Unconstrained Optimization. Jhon Wiley & Sons, New York, (1987).

[5] R. Fletcher and Reeves, C., *Function minimization by conjugate gradients*, Comput. J., 7(2) (1964), 149–154, https://doi.org/10.1093/comjnl/7.2.149

[6] J.C. Gilbert and J. Nocedal. *Global Convergence properties of conjugate gradient methods for optimization*. SIAM J. Optim., 2 (1992), 21-24.

[7] W. Hager, and Zhang H.C., *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM J. Optim., 16(1) (2005), 170-192, https://doi.org/10.1137/030601880

[8] M. R. Hestenes and E. Steifel*, Method of conjugate gradient for solving linear equations*, J.Res. Nat. Bur. Stand. 49 (1952), 409–436.

[9] Y. Liu, and C. Storey, *Efficient generalized conjugate gradient algorithms*, Part 1: Theory. JOTA, 69(1) (1991), 129-137, https://doi.org/10.1007/bf00940465

[10] L. Nazareth, *A conjugate direction algorithm without line searches*, Journal of Optimization Theory and Applications, 23 (3) (1977), 373–387, https://doi.org/10.1007/bf00933447

[11] E. Polak and G. Ribiere, *note sur la convergence de directions conjugees,* Rev. Francaise Inform. Recherche Operationelle 3 (1969), 35–43, https://doi.org/10.1051/m2an/196903r100351

[12] M. Raydan. *The Barzilain and Borwein gradient method for the large unconstrained minimization problem*. SIAM J. Optim. 7(1) (1997), 26-33, https://doi.org/10.1137/s1052623494266365

[13] J. Sun and J. Zhang. *Convergence of conjugate gradient methods without line search*. Annals Operations Research., 103 (2001), 161-173.

[14] L. Zhang, W. Zhou, and D. Li, *Some descent three-term conjugate gradient methods and their global convergence*, Optim. Methods Softw., 22(4) (2007), 697–711, https://doi.org/10.1080/10556780701223293

[15] L. Zhang, W. Zhou and D. Li. *Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search*. Numer. Math, 104 (2006), 561-572, https://doi.org/10.1007/s00211-006-0028-z

[16] L. Zhang, W.J. Zhou, D.H. Li, *A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence*, IMA Journal of Numerical Analysis, 26(4) (2006), 629-640, https://doi.org/10.1093/imanum/drl016

[17] L. Zhang, W.J. Zhou, D.H. Li, *Global convergence of a modified Fletcher-Reeves conjugate method with Armijo- type line search*, Numerische Mathematik, 104 (4) (2006), 561-572, https://doi.org/10.1007/s00211-006-0028-z