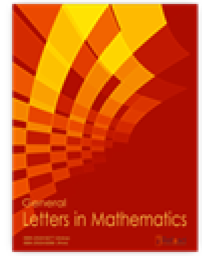

General Letters in Mathematics, Vol. 7, No.1 , Sep 2019, pp. 39 -44
e-ISSN 2519-9277, p-ISSN 2519-9269
Available online at http:// www.refaad.com
https://doi.org/10.31559/glm2019.7.1.5

A New Parameter Conjugate Gradient Method Based on Three Terms Unconstrained Optimization

Hussein Ageel Khatab [1] and Salah Gazi Shareef [2]

[1,2] Department of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region, Iraq

[1] husseinkhatab632@gmail.com , [2] salah.gazi2014@gmail.com

Abstract. In this paper, we suggest a new conjugate gradient method for solving nonlinear unconstrained optimization problems by using three term conjugate gradient method , We give a descent condition and the sufficient descent condition of the suggested method.

Keywords: *Unconstrained Optimization; Conjugate Gradient Method; Descent Condition; Three Term Conjugate Gradient Method.*

1. Introduction

We 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 generates a sequence $\{x_k\}$, k ≥ 1, starting from an initial guess $x_1 \in R^n$, using the recurrence

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

Or $v_k = \alpha_k d_k$
Where, $v_k = x_{k+1} - x_k$
The positive step size α_k is obtained by some line search, and d_k is a search direction. normally the search direction at the first iteration is the steepest descent direction , namely $d_1 = -g_1$ and the other search directions can be defined as:

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{1.3}$$

Where $g_k = \nabla f(x_k)$ and β_k is a scalar. There are many formulas for β_k , for example, Hestenes-Stiefel (HS)[8] , Conjugate descent (CD)[10], Polak-Ribiere-Polyak (PRP)[1] , Dai and Yuan (DY) [11] , (DPRP)[2] , RMIL [9] and β_k^1 [12], these formulas are as follows:

$$\beta_k^{HS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{d_k^T(g_{k+1} - g_k)} \tag{1.4}$$

$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k} \tag{1.5}$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{\|g_k\|^2} \tag{1.6}$$

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

$$\beta_k^{DPRP} = \beta_k^{RPR} - t\frac{g_{k+1}^T d_k \|y_k\|^2}{\|g_k\|^4} \tag{1.8}$$

where, $t > 1/4$

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

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

where, $t > 0$

$\mathbf{y_k} = \mathbf{g_{k+1}} - \mathbf{g_k}$, symbol $\|.\|$ denotes the Euclidean norm of vectors. The most well-studied properties of conjugate gradient methods are its global convergence properties. The convergence of conjugate gradient methods under different line searches has been studied by many authors, such as Gilbert and Nocedal [3]. Al-Baali has proven the descent property and global convergence of the Fletcher—Reeves Method with Inexact Line Search [7].

This paper is organized as follow: in Section 2, we will present a new conjugate gradient method. In Section 3, we prove the descent condition and sufficient descent condition of new method. In Section 4, some numerical experiments to this new conjugate gradient method are presented. In section 5, we will give the conclusion.

2. New conjugate gradient algorithm

In this section, we will derive the our suggestion based on a three term conjugate gradient method. J. Liu and X. Wu proposed a three-term conjugate gradient method, in which the search direction d_k has the form ,

$$d_{k+1} = -g_{k+1} - \beta_k s_k + \delta_k y_k \tag{2.1}$$

where $\beta_k = \frac{g_{k+1}^T s_k}{y_k^T s_k}$ and $\delta_k = \frac{g_{k+1}^T y_k}{y_k^T s_k}$, see[4]. Also, There are many three term conjugate gradient algorithm suggested, for instance:

$$d_{k+1} = -g_{k+1} + \beta_k^{HS} d_k - \frac{g_{k+1}^T d_k}{d_k^T y_k} y_k \quad \text{, see [6]} \tag{2.2}$$

$$d_{k+1} = -g_{k+1} + \beta_k^{PRP} d_k - \frac{g_{k+1}^T d_k}{\|g_k\|^2} y_k \quad \text{, see[5].} \tag{2.3}$$

We suggest a new three-term conjugate gradient as following:

$$d_{k+1} = -g_{k+1} + \beta_k^{NEW} d_k - \delta \mu y_k \tag{2.4}$$

where, $\delta \in (0,1)$ and $\mu = \frac{g_{k+1}^T d_k}{\|g_k\|^2}$

Now, from equation (2.3) and equation (2.4), we have

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

Multiplying both sides of above equation by d_k , we obtain

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

This implies that

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

So,

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

or

$$\beta_k^{NEW} = \beta_k^{RPR} + (\delta - 1) \frac{g_{k+1}^T d_k d_k^T y_k}{\|d_k\|^2 \|g_k\|^2} \tag{2.5}$$

where , $0 < \delta < 1$.

2.1 Algorithm of a new conjugate gradient method(β_k^{NEW})

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

$$f(x_k + \alpha_k d_k) - f(x_k) \le c_1 \alpha_k g_k^T d_k$$

$$|g_{K+1}^T d_k| \le c_2 |g_k^T d_k|$$

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

Step(4):- Compute

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

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

Step(5):- Compute β_k^{NEW} by (2.5)

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 2.1:- Assume that the sequence $\{x_k\}$ is generated by (1.2), then the search direction in (1.3) with new conjugate gradient method (2.5) satisfy the descent condition, i.e. $d_{k+1}^T g_{k+1} \le 0$ with exact and inexact line search.

Proof:- From (1.3) and (2.5) we have,

$$d_{k+1} = -g_{k+1} + \left(\frac{g_{k+1}^T y_k}{\|g_k\|^2} + (\delta - 1)\frac{g_{k+1}^T d_k d_k^T y_k}{d_k^T d_k \|g_k\|^2}\right) d_k \tag{2.6}$$

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

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

If the step length α_k is chosen by an exact line search which requires $d_k^T g_{k+1} = 0$. Then the proof is complete. If the step length α_k is chosen by inexact line search which requires $d_k^T g_{k+1} \neq 0$
the first two terms of equation (2.7) are less than or equal to zero because the parameter of (PRP) is satisfies the descent condition, and the third term clearly is less than or equal to zero, since $d_k^T y_k > 0$, so ,

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

Theorem 2.2:- :- Assume that the sequence $\{x_k\}$ is generated by (1.2), then the search direction in (1.3) with new conjugate gradient method (2.5) satisfy the sufficient descent condition, i.e.

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

Proof:- From (1.3) and (2.5) we have

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

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

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

Now, since the parameter of (PRP) is satisfies the descent condition, then the above equation becomes

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

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

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

Let $C = (1 - \delta)\frac{(g_{k+1}^T d_k)^2 d_k^T y_k}{\|g_{k+1}\|^2 \|d_k\|^2 \|g_k\|^2}$ which is positive, then

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

3. Numerical Results

This section is devoted to test the implementation of new method. We compare the our method with Conjugate Gradient methods ((PRP), (HS) , RMIL and β_k^1) the comparative tests involve Well-known nonlinear problems (standard test function) with different dimensions $4 \le n \le 5000$, all programs are written in FORTRAN90 language and for all cases the stopping condition is $\|g_{k+1}\| \le 10^{-5}$,the results given in table (1) and table (2) specifically quote the number of function NOF and the nuber of iteration NOI . More experimental results in table (1) and table(2) confirm that the new CG is superior to standard CG (PRP) , standard CG (HS) , RMIL and β_k^1 with respect to the NOI and NOF.

Table (1): Comparative Performance of the Three Algorithms (PRP , HS and New Conjugate Gradient Method)

Testfunction	N	Algorithm of PRP		Algorithm of HS		New algorithm	
		NOI	NOF	NOI	NOF	NOI	NOF
Cubic	4	15	45	12	35	12	36
	10	16	47	13	37	12	36
	50	16	47	13	37	12	36
	100	16	47	13	37	13	38
	500	16	47	13	37	13	38
	1000	16	47	13	37	13	38
	5000	16	47	13	37	13	38
Extended PSC1	4	7	18	7	18	6	16
	10	6	16	6	16	6	16
	50	6	16	6	16	6	16
	100	7	18	7	18	6	16
	500	7	18	7	18	6	16
	1000	7	18	7	18	6	16
	5000	7	18	7	18	6	16
GCentral	4	22	159	22	159	21	147
	10	22	159	22	159	21	147
	50	22	159	22	159	22	159
	100	22	159	22	159	22	159
	500	23	171	23	171	22	159
	1000	23	171	23	171	22	159
	5000	30	270	28	248	23	173
Miele	4	37	116	28	85	29	88
	10	37	116	31	102	29	88
	50	44	148	31	102	35	119
	100	44	148	33	114	35	119
	500	44	148	40	146	36	121
	1000	50	180	46	176	36	121
	5000	50	180	54	211	43	159
OSP	4	8	45	8	45	8	44
	10	13	57	13	58	13	58
	50	37	147	36	133	32	113
	100	49	176	49	185	47	166
	500	110	341	112	353	109	333
	1000	161	493	156	473	156	480
	5000	276	844	256	774	275	833
Powell	4	35	87	38	108	29	74
	10	35	87	38	108	29	74
	50	35	87	38	108	29	74
	100	35	87	40	122	29	74
	500	35	87	41	124	29	74
	1000	35	87	41	124	32	92
	5000	35	87	41	124	32	92
Wood	4	29	67	30	68	27	63
	10	29	67	30	68	27	63
	50	29	67	30	68	28	65
	100	30	69	30	68	28	65
	500	30	69	30	68	28	65
	1000	30	69	30	68	28	65
	5000	30	69	30	68	28	65
Total		1734	5922	1679	5856	1569	5322

Table (2): Comparative Performance of the two Algorithms (RMIL , β_k^1 and New Conjugate Gradient Method)

Testfunction	N	Algorithm of RMIL		β_k^1		New algorithm	
		NOI	NOF	NOI	NOF	NOI	NOF
Cubic	4	16	47	13	37	12	36
	10	16	47	13	37	12	36
	50	16	47	13	37	12	36
	100	16	47	13	37	13	38
	500	16	47	13	37	13	38
	1000	16	47	13	37	13	38
	5000	16	47	14	39	13	38
Extended PSC1	4	7	18	7	18	6	16
	10	6	16	6	16	6	16
	50	6	16	6	16	6	16
	100	7	18	7	18	6	16

	500	7	18	7	18	6	16
	1000	7	18	7	18	6	16
	5000	7	18	7	18	6	16
GCentral	4	33	197	22	159	21	147
	10	34	208	22	159	21	147
	50	39	265	22	159	22	159
	100	43	315	22	159	22	159
	500	48	380	23	171	22	159
	1000	51	420	23	171	22	159
	5000	56	489	28	248	23	173
Miele	4	52	164	33	104	29	88
	10	59	194	33	104	29	88
	50	67	229	33	104	35	119
	100	79	273	33	104	35	119
	500	90	317	33	104	36	121
	1000	90	317	33	104	36	121
	5000	106	395	33	104	43	159
OSP	4	8	45	8	44	8	44
	10	13	57	13	58	13	58
	50	39	152	34	134	32	113
	100	60	210	49	180	47	166
	500	236	745	107	328	109	333
	1000	471	F	146	448	156	480
	5000	F	F	257	763	275	833
Powell	4	F	F	38	108	29	74
	10	F	F	38	108	29	74
	50	F	F	38	108	29	74
	100	F	F	40	122	29	74
	500	F	F	40	122	29	74
	1000	F	F	40	122	32	92
	5000	F	F	40	122	32	92
Wood	4	96	199	30	68	27	63
	10	101	209	30	68	27	63
	50	103	213	30	68	28	65
	100	118	243	30	68	28	65
	500	128	263	30	68	28	65
	1000	128	263	30	68	28	65
	5000	148	303	30	68	28	65
Total		6655	12016	1630	5580	1569	5322

Notes : 1- $F > 1000$.
2- We took $F = 500$ for summation.

Table(3): Percentage of Improving of the New Method

	Algorithm of PRP	New Algorithm
NOI	100%	90.48442907657%
NOF	100%	89.8682877406%
	Algorithm of HS	New Algorithm
NOI	100%	93.4484812388%
NOF	100%	90.881147541%
	Algorithm of RMIL	New Algorithm
NOI	100%	23.5762584523%
NOF	100%	44.2909454061%
	β_k^1	New Algorithm
NOI	100%	96.2576687117%
NOF	100%	95.376344086%

4. Conclusion

In this paper ,we suggested a new conjugate gradient method for unconstrained optimization. Implemented and tested to some extent, while numerical tests were carried out, on low and high dimensionality problems, and comparisons were made amongst different test functions with inexact line search. Some of the numerical results have been reported.

References:

[1] E. Polak & G. Ribiere, Note sur la convergence de directions conjugees, *Rev. Francaise Inform Recherche Operationelle*, 3(16)(1969), 35–43, https://doi.org/10.1051/m2an/196903r100351

[2] G.H. Yu & L.T. Guan, Modified PRP Methods with sufficient desent property and their convergence properties. *Acta Scientiarum Naturalium Universitatis Sunyatseni(Chinese)*, 45(4)(2006), 11–14.

[3] J. C. Gilbert & J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM J. Optim*. 2(1)(1992), 21-42, https://doi.org/10.1137/0802003

[4] J. Liu & X. Wu, New three-term conjugate gradient method for solving unconstrained optimization problems, *Science asia*, 40(4)(2014), 295-300, https://doi.org/10.2306/scienceasia1513-1874.2014.40.295

[5] L. Zhang, W. Zhou & D.H . Li, A descent modified Polak-Ribi`ere-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

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

[7] M. , Al-Baali, Descent property and global convergence of the Fletcher Reeves method with inexact line search, *IMA J. Numer. Anal,* 5(1)(1985), 121-124, https://doi.org/10.1093/imanum/5.1.121

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

[9] M. Rivaie, M Mamat, W.J. Leong & M. Ismail, A new class of nonlinear conjugate gradient coefficients with global convergence properties, *Applied Mathematics and Computation*, 2189(2012), 11323-11332, http://dx.doi.org/10.1016/j.amc.2012.05.030

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

[11] Y.H. Dai & Y. Yuan, A nonlinear conjugate gradient with a strong global convergence properties, *SIAM J. Optim*. 10(1)(1999), 177–182.

[12] Y. Xiao, H. Song & Z. Wang, A modified conjugate gradient algorithm with cyclic Barzilai–Borwein steplength for unconstrained optimization, *Journal of Computational and Applied Mathematics*, 236(13)(2012), 3101-3110, https://doi.org/10.1016/j.cam.2012.01.032