NEW CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION WITH LOGISTIC MAPPING

Salah Gazi Shareef[1], Hussein Ageel Khatab[2] and Sawsan Sami Ismael[3]
[1,2] Department of Mathmatics, Faculty of Science, University of Zakho, Kurdistan Region – Iraq.
[3] Department of Mathmatics, College of Education, University of Mosul, Iraq.

(Accepted for publication: April 11, 2016)

ABSTRACT:

In this paper , we suggested a new conjugate gradient algorithm for unconstrained optimization based on logistic mapping, descent condition and sufficient descent condition for our method are provided. Numerical results show that our presented algorithm is more efficient for solving nonlinear unconstrained optimization problems comparing with (DY) .

KEY WORDS: *Unconstrained Optimization, Conjugate Gradient Method, Descent Condition , Logistic Mapping.*

1- INTRODUCTION

The nonlinear conjugate gradient method is designed to solve the following unconstrained optimization problem

$$\min_{x\in R^n} f(x) \quad (1.1)$$

Where $f: R^n \to R$ is a continuously differentiable nonlinear function whose gradient is denoted by g, i.e $g_k = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$. The iterativeformula for solving CG method is expressed as

$$x_{k+1} = x_k + \alpha_k d_k \ , \ k=0,1,2,\dots \ (1.2)$$

Where $\alpha_k d_k = x_{k+1} - x_k = v_k$, $\alpha_k > 0$ is the step length and d_k is the search direction defined by

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \qquad (1.3)$$

$d_0 = -g_0$, where β_k is a parameter known as conjugate gradient coefficient. Some well-known classical formulas for β_k are the Hestenes and Stiefel(HS) (1952), Polak (1969), Ribière and Polyak (PRP) (1969), Fletcher and Reeves (FR) (1964)., Dai and Yuan (DY) (1999), Liu and Storey (LS) ,(1992), and conjugate descent (CD) Fletcher (1987) are given below

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \quad (1.4)$$

$$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{\|g_{k-1}\|^2} (1.5)$$

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} (1.6)$$

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

$$\beta_k^{CD} = -\frac{\|g_{k+1}\|^2}{d_k^T g_k} (1.8)$$

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k} (1.9)$$

Where g_{k+1} and g_k are the gradients of *f(x)* at the point x_{k+1} and x_k respectively.Consider $\|.\|$ the Euclidean norm and define $y_k = g_{k+1} - g_k$. The global convergence of above conjugate gradient methods are studied by many researchers, see, for instance [Hager and Zhang (2006).] and references therein. To establish the convergence results of these methods , it is usually required that the step length α_k should satisfies the strong Wolfe conditions :

$$f(x_{k+1}) - f(x_k) \le \alpha_k \delta_1 d_k^T g_k (1.10)$$

$$|g_{k+1}^T d_k| \le -\delta_2 |d_k^T g_k| (1.11)$$

Where $0 < \delta_1 < \delta_2 < 1$ see [Jorge and Stephen (1999)] Some convergence analyses even require that α_k be computed by the exact line search, that is

$$f(x_k + \alpha_k d_k) = \min f(x_k + \alpha_k d_k), (1.12)$$

$$\alpha_k > 0,$$

on the other hand , many other numerical methods for unconstrained optimization are proved to be convergent under the Wolfe condition :

$$f(x_{k+1}) - f(x_k) \leq \alpha_k \delta_1 d_k^T g_k \quad (1.13)$$

$$g_{k+1}^T d_k \geq \delta_2 d_k^T g_k \quad (1.14)$$

This paper is organized as follows: In Sect.2, we propose the new conjugate gradient (CG) algorithm. In section 3 , we prove the descent condition and the sufficient descent condition of this new method. In Section 4, we give some the numerical results. In section 5, the conclusion is given.

2- New Conjugate Gradient Algorithm ($\boldsymbol{\beta_k^{NEW}}$)

In this section, we propose our new β_k which is known as β_k^{NEW}. The main idea is to use logistic mapping with β_k^{DY}. For more details about the logistic mapping see [LU et al. (2005)].

From the logistic mapping and (1.7), we have

$$\beta_k^{NEW} = \mu \beta_k^{DY}(1 - \beta_k^{DY}) \quad (2.1)$$

where $0 < \mu \leq 1$

Multiplying the second term from right hand side of (2.1) by scalar $\overline{K}$, we get

$$\beta_k^{NEW} = \mu \beta_k^{DY}(1 - \overline{K}\beta_k^{DY}) \text{ , where}$$

$$\overline{K} = \frac{g_{k+1}^T v_k}{d_k^T y_k}$$

Now , we suggest the following

$$\beta_k^{NEW} = \mu \frac{\|g_{k+1}\|^2}{d_k^T y_k}\left(1 - \frac{g_{k+1}^T v_k}{d_k^T y_k}\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right) \quad (2.2)$$

Algorithm of new formula ($\boldsymbol{\beta_k^{NEW}}$)

Step (1): Given initial point $x_0 \in R^n$.

Step (2): $k=0, g_0 = \nabla f(x_0), d_0 = -g_0$, if $g_0 = 0$, then stop.

Step(3) : compute α_k by using cubic line search to minimize $f(x_k + \alpha_k d_k)$, $i.e, \quad f_{k+1} \leq f_k$.

Step (4): $x_{k+1} = x_k + \alpha_k d_k$.

Compute $g_{k+1} = \nabla f(x_{k+1})$, if $\|g_{k+1}\| \leq 10^{-5}$, then stop. Else $v_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$.

Step(5): compute β_k^{NEW} by (2.2).

Step (6): compute d_k by (1.3) and (2.2).

Step (7): Use Powell restart

If $|g_{k+1}^T g_k| \geq 0.2\|g_{k+1}\|^2$

then go to step 2,

else, k=k+1 and go to step 3.

3- Descent Condition and Sufficient Descent Condition of New Formula (β_k^{NEW}):

in this section , we will study the descent condition and the sufficient descent condition of formula (β_k^{NEW}):

Theorem (3.1):-Assume that the sequence $\{x_k\}$ is generated by (1.2), then the search direction (1.3) with new formula (2.2) satisfies the descent condition .

Proof:- Multiplying equation (1.3) by g_{k+1} and by using (2.2), we get

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left[\mu \frac{\|g_{k+1}\|^2}{d_k^T y_k} - \mu \frac{g_{k+1}^T v_k}{d_k^T y_k}\left(\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right)^2\right] d_k^T g_{k+1} \quad (3.1)$$

this implies that

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \mu \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - \mu \frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k}\left(\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right)^2 \quad (3.2)$$

If the step-length α_k is chosen by an exact line search, that means

$d_k^T g_{k+1} = 0$, then the equation (3.2)

gives $d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 \leq 0$

Then the proof is completed.

If the step-length α_k is chosen by inexact line search, that means

$d_k^T g_{k+1} \neq 0$,since , $d_k^T g_{k+1} \leq d_k^T y_k$, then equation (3.2) gives

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \mu\|g_{k+1}\|^2 - \mu \frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k} \left(\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right)^2 \quad (3.3)$$

Since, $\mu \in (0,1]$ and $d_k^T y_k > 0$ so, (3.3) will be in the form

$$d_{k+1}^T g_{k+1} \leq \|g_{k+1}\|^2(\mu - 1) - \mu \frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k} \left(\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right)^2 \leq 0 \ .$$

Then the proof is complete.

Theorem (3.2):-Suppose that x_k and d_k are generated by the method of the form (1.2), (1.3) and (2.2), and the step size α_k is obtained by (1.10) and (1.11)then, the sufficient descent condition is satisfied, i.e

$$d_{k+1}^T g_{k+1} \leq -C\|g_{k+1}\|^2 \quad (3.4)$$

Proof :-We can write equation (3.3) as follows

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 \left[1 - \mu + \mu \frac{\alpha_k (d_k^T g_{k+1})^2}{\|g_{k+1}\|^2 d_k^T y_k} \left(\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right)^2\right] \quad (3.5)$$

$$\text{let } C = 1 - \mu + \mu \frac{\alpha_k (d_k^T g_{k+1})^2}{\|g_{k+1}\|^2 d_k^T y_k} \left(\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right)^2,$$

then (3.5) becomes

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

then the proof is complete.

4- NUMERICAL RESULTS

This section is devoted to test the implementation of the new method. The comparative tests involve well-known nonlinear problems (standard test functions) with different dimension 4≤n≤5000, all programs are written in FORTRAN95 language and for all cases the stopping condition is $\|g_{k+1}\| \leq 10^{-5}$. The results are given in Table (1) is specifically quote the number of functions NOF and the number of iteration NOI. Experimental results in Table (1) confirm that the new CG method is superior to standard CG method (DY) with respect to the NOI and NOF.

Table(1): Comparative performance of two algorithms (standard CG method (DY) and new CG method (β_k^{NEW})).

Test function	N	CG (DY)		NEW CG	
		NOI	NOF	NOI	NOF
Wood	4	28	65	27	62
	10	28	65	27	62
	50	28	65	28	64
	100	28	65	28	64
	500	29	68	28	64
	1000	29	68	28	64
	5000	29	68	30	68
G.Central	4	18	127	21	83
	10	18	127	21	83
	50	19	138	21	83
	100	20	153	24	114
	500	23	192	28	163
	1000	23	192	28	163
	5000	24	205	42	321
Powell3	4	14	33	14	32
	10	14	33	14	32
	50	14	33	14	32
	100	14	33	14	32
	500	14	33	14	32
	1000	14	33	14	32
	5000	15	35	15	34
Sum	4	3	11	3	11
	10	6	34	6	34
	50	11	60	11	60
	100	14	85	13	72
	500	21	118	20	102
	1000	24	125	16	77
Miele	4	36	115	32	95
	50	45	156	38	132
	500	53	188	45	170
	1000	60	222	52	210
	5000	66	257	57	236
Osp.	50	37	134	37	133
	500	138	439	129	401

Wolfe	500	48	97	47	95
	1000	52	105	51	103
Total		1057	3977	1037	3615

Table (2): Percentage of improving the New formula

Tools	CG(DY)	NEW CG
NOI	100%	98.10785
NOF	100%	90.89766

5- CONCLUSION

This paper gives a modified conjugate gradient method for solving nonlinear unconstrained optimization in formula (2.2) by logistic mapping, it is shown that the search direction with this formula satisfies the descent condition and sufficient descent condition. The numerical results show that the given modified method is competitive to the Dai-Yuan (DY) conjugate gradient method for some test problems.

6- REFERENCES

Dai Y. H., and Y. X. Yuan, (1999). A nonlinear conjugate gradient method with a strong global convergence property, SIAM Journal on optimization, 10, pp. 177-182.

Fletcher R. ,(1987). Practical Methods of Optimization, Vol. 1: Unconstrained Optimization(New York: Wiley and Sons).

Fletcher R. and C. Reeves ,(1964). Function minimization by conjugate gradients, Computer Journal, 7, pp. 149-154.

Hager W., H.C. Zhang,(2006). A survey of nonlinear conjugate gradient methods, Pacific Journal of Optimization, 2, pp. 35-58.

Hestenes M. R .and E. Stiefel,(1952). Method of conjugate gradient for solving linear equation, J. Res. Nat. Bur. Stand. 49, 409-436.

Jorge Nocedal and Stephen J. Wright,(1999). Numerical Optimization,Springer-Verlag New York, Inc.

Liu Y. and C. Storey,(1992). Efficient generalized conjugate gradient algorithms part 1: Theory, J. Comput. Appl. Math. 69 , pp. 129-137.

LU Hui-juan, ZHANG Huo-ming, MA Long-hua,(2005). A new optimization algorithm based on chaos, Zhejiang University, Hangzhou 310027, China.

Polayk B., (1969).The conjugate gradient method in extreme problems, USSR Computational Mathematics and Mathematical Physics, pp. 94-112.

Polak B. and G. Ribiere, (1969). Note sur la convergence des methods de directions conjuguees, Reue Francaise deRecherche Operationnele, 16,pp. 35-43.

پوخته

دڤێ ڤەکۆلینێدا ، مە پێشنیارکری پەیسکێن هاوشێوە بوو نموونەی یێن نەگرێدای. پشت بەستن ب(Logistic Mapping)مەرجێن لاری ولاریا کافی بو ڤی شێوازێ هاتیە بدەست خەستن. ئەنجامین ژمارییێ بوو ئەلگوریزمێی وەسا دیارکری کوبتریا چالاکە بوشیکارگرنێ نموونەیێن نەگرێدای ونەییێت هێلی ، بە راوردکرن دگەل (DY) .

الخلاصة:

في هذا البحث ، اقترحنا خوارزمية جديدة للتدرج المترافق للأمثلية غير المقيدة استناد على التطبيق اللوجستي،وتم تحقيق شرط الانحدار وشرط الانحدار الكافي لطريقتنا ، النتائج العددية تبين ان خوارزميتنا المقترحة لها تأثير اكثر في حل مسائل الأمثلية غير المقيدة وغير الخطية مقارنة مع (DY).