International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 6 Issue 11, November-2017, Impact Factor: 4.059

A New Spectral Conjugate Gradient Methods Based on the Descent Condition

Dr. Basim A. Hassan[1], Alaa A. Mahmood[2]

[1]Assist Professor / College of Computers Sciences and Math.\ University of Mosul / Iraq
[2]Graduate student / College of Computers Sciences and Math.\ University of Mosul / Iraq

ABSTRACT

A new spectral methods based on the descent condition. This methods has the property that the generated search direction is sufficiently descent without utilizing the line search. The proof of the convergence of the proposed method is given. Numerical experiments show that the method is efficient and feasible.

Keyword : Conjugate gradient method, Spectral conjugate gradient, Global convergent property.

INTRODUCTION

Conjugate gradient methods are a class of important methods for solving unconstrained optimization problem:

$$\min f(x) \ , \ x \in R^n \qquad \text{..........} \quad (1)$$

especially if the dimension n is large. They are of the form :

$$x_{k+1} = x_k + \alpha_k d_k \ , \quad k = 0,1,.\ .. \qquad \text{..........} \quad (2)$$

where α_k is a step size obtained by a line search, and d_k is the search direction defined by :

$$d_{k+1} = \begin{cases} -g_{k+1} & \text{if } k = 0 \\ -g_{k+1} + \beta_k d_k & \text{if } k > 0 \end{cases} \qquad \text{..........} \quad (3)$$

where β_k is a parameter, and g_{k+1} denotes $\nabla f(x_{k+1})$. Well-known conjugate gradient methods are the Fletcher-Reeves method, Hestenes-Stiefe method, Polak-Ribiere-Polyak method, Conjugate-Descent method and Dai-Yuan method etc.. More details can be found in [14]. In the Fletcher-Reeves (FR) method, the parameter β_k is specified by :

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \qquad \text{..........} \quad (4)$$

More details can be found in [16]. The convergence behavior of the different conjugate gradient methods with some line search conditions has been widely studied, including Armijo [5], Al-baali [4], Dai [7] and Zhang [10-11] etc.. There are many convergence results of the conjugate gradient methods [2,3,6]. The sufficient descent condition :

$$d_{k+1}^T g_{k+1} < -c\|g_{k+1}\|^2 \qquad \text{..........} \quad (5)$$

where $c > 0$ is a constant, is crucial to insure the global convergence of the conjugate gradient methods. Line search in the conjugate gradient algorithms often is based on the standard Wolfe conditions:

$$f(x_k) - f(x_k + \alpha_k d_k) \geq -\delta\alpha_k g_k^T d_k \qquad \text{..........} \quad (6)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \qquad \text{..........} \quad (7)$$

where $0 < \delta < \sigma < 1$. More details can be found in [12, 13].

Zhang et al. [9] proposed a modified FR method (called MFR), in which the direction d_{k+1} is defined by:

$$d_{k+1} = -\theta_k^{MFR} g_{k+1} + \beta_k^{FR} d_k, \quad \text{.......... (8)}$$

where

$$\theta_k^{MFR} = \frac{y_k^T d_k}{g_k^T g_k}. \quad \text{.......... (9)}$$

which is a descent direction independent of the line search.
The rest of this paper is organized as follows. In Sect. 2, we will state the idea to propose a new spectral conjugate gradient method in detail. Then, a new algorithm is developed and descent property in Section 3. Global convergence is established in Section 4. Section 5 is devoted to numerical experiments. Finally we present the conclusion in the last part.

A new spectral methods based on the descent condition

Due to the descent condition (5) is a very important property in the literature to analyze the global convergence of the conjugate gradient methods.

Enlightened by Hideaki and Yasushi (HY) [9], proposed the variant of the CG methods which we call the HY methods. The β_k in the HY method is defined as:

$$\beta_k^{HY} = \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})}. \quad \text{.......... (10)}$$

Suppose the current search direction d_{k+1} is a descent direction, namely, $d_{k+1}^T g_{k+1} < 0$. Now we need to find θ_k that defines a descent direction d_{k+1}. This requires that :

$$-\theta_k g_{k+1}^T g_{k+1} + \beta_k g_{k+1}^T d_k < 0. \quad \text{.......... (11)}$$

From (11) we get :

$$\beta_k g_{k+1}^T d_k < \theta_k \|g_{k+1}\|^2$$
$$g_{k+1}^T d_k < \theta_k \frac{\|g_{k+1}\|^2}{\beta_k} \quad \text{.......... (12)}$$

Immediately, we have :

$$-g_k^T d_k + g_{k+1}^T d_k = \theta_k \frac{\|g_{k+1}\|^2}{\beta_k}. \quad \text{.......... (13)}$$

Since

$$-g_k^T d_k + g_{k+1}^T d_k = \theta_k \frac{\|g_{k+1}\|^2}{\beta_k}. \quad \text{.......... (14)}$$

Additionally, (13) and (14) imply that :

$$y_k^T d_k = \theta_k (2/\alpha_k)(f_k - f_{k+1})$$
$$\theta_k = \frac{y_k^T d_k}{(2/\alpha_k)(f_k - f_{k+1})} \quad \text{.......... (15)}$$

which yields :

$$\theta_k^{SBA} = \frac{y_k^T d_k}{(2/\alpha_k)(f_k - f_{k+1})}. \quad \text{.......... (16)}$$

Thus, we obtain the following iterative direction :

$$d_{k+1} = -\left[\frac{y_k^T d_k}{(2/\alpha_k)(f_k - f_{k+1})}\right] g_{k+1} + \beta_k^{HY} d_k. \quad \text{.......... (17)}$$

International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 6 Issue 11, November-2017, Impact Factor: 4.059

NEW ALGORITHM AND DESCENT PROPERTY

In this section, we give the specific form of the proposed spectral conjugate gradient algorithm as follows and prove its descent property.

Now, we can outline our new algorithm as follows:

Outline of the SHY Algorithms:

Step 0 : Given $x_0 \in R^n$, $\varepsilon = 0.0001$, $\delta_1 \in (0,1)$, $\delta_2 \in (0, 1/2)$.

Step 1 : Computing g_k ; if $\|g_k\| \le \varepsilon$, then stop ; else continue.

Step 2 : Set $\beta_k = \beta_k^{HY}$ with θ_k^{SBA} respectively.

Step 3 : Set $x_{k+1} = x_k + \alpha_k d_k$, (Use SW-conditions to compute α_k).

Step 4 : Compute $d_{k+1} = -\theta_k g_{k+1} + \beta_k d_k$,

Step 5 : Go to Step 1 with new values of x_{k+1} and g_{k+1}.

We end this section by showing some properties of the new search direction, which will be showed in the following theorem.

Theorem 1.

Let the search direction d_{k+1} be generated by (17). Then the sufficient descent condition holds, i.e.,

$$g_{k+1}^T d_{k+1} \le -c\|g_{k+1}\|^2 . \qquad \text{..........} \quad (18)$$

Proof :

We prove this theorem by induction. Since $d_0 = -g_0$ we have $g_0^T d_0 = -\|g_0\|^2 < 0$. Suppose that $g_k^T d_k < -c_1\|g_k\|^2$ for all $k \in n$. Multiplying (17) by g_{k+1}, we have :

$$g_{k+1}^T d_{k+1} = -\theta^{SBA} g_{k+1}^T g_{k+1} + \frac{g_{k+1}^T g_{k+1}}{(2/\alpha_k)(f_k - f_{k+1})} g_{k+1}^T d_k \qquad \text{..........} \quad (19)$$

$$g_{k+1}^T d_{k+1} = \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})}\left[-y_k^T d_k + g_{k+1}^T d_k\right]$$

$$= \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} g_k^T d_k \qquad \text{..........} \quad (20)$$

$$= \frac{g_k^T d_k}{(2/\alpha_k)(f_k - f_{k+1})}\|g_{k+1}\|^2$$

Since $g_k^T d_k < -c_1\|g_k\|^2$, then we have :

$$g_{k+1}^T d_{k+1} < -c_1 \frac{\|g_k\|^2}{(2/\alpha_k)(f_k - f_{k+1})}\|g_{k+1}\|^2 . \qquad \text{..........} \quad (21)$$

where $c = c_1(\|g_k\|^2 /(2/\alpha_k)(f_k - f_{k+1}))$. Since c_1, $(2/\alpha_k)(f_k - f_{k+1})$ and $\|g_k\|^2$ are positive values, then c is also positive value.

$$g_{k+1}^T d_{k+1} \le -c\|g_{k+1}\|^2 . \qquad \text{..........} \quad (22)$$

This completes the proof.

The global convergence of the algorithm

In order to achieve the convergence of the algorithm ,we give some Assumptions and Lemmas as follow :

Assumption:

i- The level set $L = \{x \in R^n \mid f(x) \le f(x_0)\}$ is bounded.

ii- In some neighborhood U and L, $f(x)$ is continuously differentiable and its gradient is Lipschitz continuous, namely, there exists a constant $\mu_1 > 0$ such that :

$$\|g(x_{k+1}) - g(x_k)\| \le \mu_1 \|x_{k+1} - x_k\|, \ \forall \ x_{k+1}, x_k \in U \ . \qquad \text{..........} \ (23)$$

More details can be found in [8].

Under this assumption, we have the following lemma, which was proved by Zoutendijk [15].

Lemma 1.

Suppose that Assumption 1 holds. Consider any method $(1)-(3)$, where d_{k+1} satisfies $g_{k+1}^T d_{k+1} \le 0$ and α_k satisfies the Wolfe line search. Then :

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \ . \qquad \text{..........} \ (24)$$

The following theorem is based on Lemma 4.1.

Theorem 2.

Suppose that Assumption 1 holds. Let the sequences $\{g_{k+1}\}$ and $\{d_{k+1}\}$ be generated by Algorithm 3.1, and let the α_k be determined by the Wolfe line search (6) and (7). Then :

$$\lim_{k \Rightarrow \infty} \inf \|g_k\| = 0 \ . \qquad \text{..........} \ (25)$$

Proof :

To prove theorem 2, we use contradiction. That is, if theorem 2 is not true, then a constant $\varepsilon_1 > 0$ exists, such that :

$$\|g_{k+1}\| > \varepsilon_1 , \qquad \text{..........} \ (26)$$

Rewriting (17) as :

$$d_{k+1} + \theta_k^{SBA} g_{k+1} = \beta_k^{HY} d_k , \qquad \text{..........} \ (27)$$

And squaring both sides of the equation, we get :

$$\|d_{k+1}\|^2 + (\theta_k^{SBA})^2 \|g_{k+1}\|^2 + 2\theta_k^{SBA} d_{k+1}^T g_{k+1} = (\beta_k^{HY})^2 \|d_k\|^2 \qquad \text{..........} \ (28)$$

from (32), we get :

$$\|d_{k+1}\|^2 = (\beta_k^{HY})^2 \|d_k\|^2 - 2\theta_k^{SBA} d_{k+1}^T g_{k+1} - (\theta_k^{SBA})^2 \|g_{k+1}\|^2 . \qquad \text{..........} \ (29)$$

In fact, the direction d_{k+1} generated by the HY method [9] satisfies :

$$g_{k+1}^T d_{k+1} \le \beta_k^{HY} g_k^T d_k \ , \ \beta_k^{HY} \le \frac{g_{k+1}^T d_{k+1}}{g_k^T d_k} . \qquad \text{..........} \ (30)$$

From the above equation and (29), we have :

$$\|d_{k+1}\|^2 \le \left(\frac{g_{k+1}^T d_{k+1}}{g_k^T d_k} \right)^2 \|d_k\|^2 - 2\theta_k^{SBA} d_{k+1}^T g_{k+1} - (\theta_k^{SBA})^2 \|g_{k+1}\|^2 . \qquad \text{..........} \ (31)$$

Dividing the both of the equation by $(g_{k+1}^T d_{k+1})^2$, we have :

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

$$\le \frac{\|d_k\|^2}{(g_k^T d_k)^2} - (\theta_k^{SBA})^2 \frac{\|g_{k+1}\|^2}{c^2\|g_{k+1}\|^4} - 2\theta_k^{SBA} \frac{1}{c\|g_{k+1}\|^2} - \frac{1}{\|g_{k+1}\|^2} + \frac{1}{\|g_{k+1}\|^2} \quad \text{..........} \quad (32)$$

$$\le \frac{\|d_k\|^2}{(g_k^T d_k)^2} - \left(\theta_k^{SBA} \frac{\|g_{k+1}\|}{c\|g_{k+1}\|^2} + \frac{1}{\|g_{k+1}\|}\right)^2 + \frac{1}{\|g_{k+1}\|^2}$$

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

Using (32) recursively and noting that $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$, we get :

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \sum_{i=1}^{k} \frac{1}{\|g_i\|^2} \quad . \quad \text{..........} \quad (33)$$

Then we get from (33) and (26) that :

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \ge \frac{\varepsilon_1^2}{k}, \quad \text{..........} \quad (34)$$

which indicates :

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \ge \sum_{k=1}^{\infty} \frac{\varepsilon_1^2}{k} = \infty \quad . \quad \text{..........} \quad (34)$$

This contradicts the Zoutendijk condition in Lemma 4.1. The proof is completed.

NUMERICAL EXPERIMENTS

In this section, we report some results of the numerical experiments. We chose 15 test problems with the dimension n = 100 and n=1000 and initial points from the literature [1]. We test new Algorithm and compare its performance with those of FR method whose results be given by [16].

All these algorithms are implemented with the standard Wolfe line search conditions with $\delta_1 = 0.001$ and $\delta_2 = 0.9$. The termination condition is $\|g_{k+1}\| \le 10^{-6}$ or the iteration number exceeds 1000 or the function evaluation number exceeds 1000. In the following table, the numerical results are written in the form NI/NF, where NI, NF denote the number of iteration and function evaluations respectively. Dim denotes the dimension of the test problems. So, from the limited numerical experiments indicate that the proposed method is potentially efficient.

CONCLUSION

We have presented a new spectral conjugate gradient algorithm for solving unconstrained optimization problems. The algorithm satisfies the sufficient descent condition, independent of the line search. Under standard Wolfe line search conditions we proved the global convergence of the algorithm. SBA method gives the best results.

Table 1: Comparison of different CG-algorithms with different test functions and different dimensions

		FR algorithm		SBA algorithm		HY algorithm	
P. No.	n	NI	NF	NI	NF	NI	NF
1	100	19	35	18	34	19	36
	1000	38	65	38	65	34	62
2	100	43	88	34	68	F	F
	1000	46	92	33	71	36	90
3	100	32	64	10	21	13	25
	1000	77	129	15	29	19	34
4	100	180	313	84	160	74	145
	1000	F	F	91	171	80	151

5	**100**	**124**	**231**	**53**	**94**	**51**	**87**
	1000	**445**	**711**	**191**	**333**	**158**	**275**
6	**100**	**71**	**110**	**31**	**59**	**26**	**50**
	1000	**47**	**84**	**26**	**51**	**25**	**50**
7	**100**	**101**	**217**	**80**	**176**	**85**	**206**
	1000	**101**	**214**	**80**	**173**	**85**	**203**
8	**100**	**32**	**65**	**24**	**51**	**26**	**56**
	1000	**53**	**116**	**37**	**90**	**37**	**87**
9	**100**	**40**	**65**	**38**	**57**	**35**	**57**
	1000	**43**	**68**	**F**	**F**	**39**	**60**
10	**100**	**398**	**605**	**346**	**535**	**369**	**585**
	1000	**F**	**F**	**F**	**F**	**F**	**F**
11	**100**	**121**	**218**	**86**	**134**	**F**	**F**
	1000	**345**	**634**	**235**	**367**	**234**	**362**
12	**100**	**32**	**52**	**14**	**27**	**14**	**29**
	1000	**22**	**42**	**13**	**26**	**16**	**32**
13	**100**	**9**	**18**	**8**	**16**	**8**	**16**
	1000	**12**	**82**	**7**	**35**	**F**	**F**
14	**100**	**23**	**45**	**17**	**33**	**18**	**33**
	1000	**27**	**55**	**20**	**44**	**19**	**44**
15	**100**	**25**	**43**	**22**	**44**	**22**	**44**
	1000	**46**	**741**	**31**	**252**	**50**	**877**
Total		**2333**	**4746**	**1464**	**2808**	**1473**	**3485**

Fail : The algorithm fail to converge.

Problems numbers indicant for : 1. is the Trigonometric, 2. is the Extended White & Holst, 3. is the Extended Tridiagonal 1, 4. is the Extended Powell, 5. is the Quadratic Diagonal Perturbed, 6. is the Extended Wood, 7. is the Extended Hiebert, 8. is the Extended Quadratic Penalty, 9. is the Extended Tridiagonal 2, 10. is the TRIDIA (CUTE), 11. is the DIXMAANE (CUTE), 12. is the Extended Beale, 13. is the ARWHEAD (CUTE), 14. is the LIARWHD (CUTE), 15. is the Generalized Tridiagonal 1.

REFERENCES

[1]. Andrie N. (2008) ' An Unconstrained Optimization Test functions collection ' Advanced Modeling and optimization. 10, pp.147-161.

[2]. Andrei N.,(2008),' Another hybrid conjugate gradient algorithm for uncontrained optimization. Numerical Algorithm. , 47(2), pp.143-156.

[3]. Andrei N.,(2009),' Hybrid conjugate gradient algorithm for uncontrained optimization. Journal of Optimization Theory and Applications., 141(2), pp.249-264.

[4]. Al-Baali, (1985), 'Descent property and global convergence of the Fletcher–Reeves method with inexact line search', IMA J. Numer. Anal. 5 pp. 121–124.

[5]. Armijo L. Minimization of functions having Lipschitz continuous first partial derivatives. Pacific Journal of Mathematics. 1966; 16(1), pp.1-3.

[6]. Birgin E, Martinez J M. A spectral conjugate gradient method for unconstrained optimization. Appl.Math.Optim.. 2001; 43(2), pp.117-128.

[7]. Dai Y H, Yuan Y. A nonlinear conjugate gradient method with a strong global convergence property. SIAM J. Optim..1999; 10(1), pp.177-182.

[8]. Hager W. W. and Zhang. H, (2006), "A surrevy of nonlinear conjugate gradient methods "Paaific Journal of optimization, 2, pp.35-85.

[9]. Hideaki I., and Yasushi N., (2011)," Conjugate gradient methods using value of objective function for unconstrained optimization Optimization Letters, V.6, Issue 5, pp. 941-955.

[10]. Zhang L, Zhou W J, Li D H. Global convergence of a modified Fletcheer-Reeves conjugate gradient method with Arnijo-type line search. Numerical Mathematics. 2006; 104(4):561-572.

[11]. Zhang L, Zhou W J, Li D H. A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence. IMA Journal of Numerical Analysis. 2006; 26(4):629-640.

[12]. Wolfe P., (1969) ,'Convergence conditions for ascent methods, SIAM Rev. 11, pp. 226-235.

[13]. Wolfe P., (1971),'Convergence conditions for ascent methods II: some corrections, SIAM Rev. 13, pp. 185-188.

[14]. Zeng W.Q. and Liu H. L., (2015),'The Global Convergence of a New Spectral Conjugate Gradient Method, International Conference on Artificial Intelligence and Industrial Engineering, pp.484-487.

[15]. Zoutendijk, G., (1970) ' Nonlinear programming, computational methods. In: Abadie, J. (eds.) Integerand Nonlinear Programming' . North-Holland, Amsterdam .pp. 37–86.

[16]. Fletcher R. and Reeves C. M.,(1964). Funtion minimization by conjagate gradients, Computer Journal 7, pp. 149-154.