International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Descent Condition for a Scalar Parameter of Spectral HS (SpHS) Conjugate Gradient Method

U A Yakubu, M Mamat, A V Mandara, A Iguda, S Murtala, M A Mohamed

***Abstract*: Spectral conjugate gradient method has been used in most cases as an alternative to the conjugate gradient (CG) method in order to solve nonlinear unconstrained problems. In this paper, we introduced a spectral parameter of HS conjugate gradient method resultant from the classical CG search direction and used some of the standard test functions with numerous variables to prove its sufficient descent and global convergence properties, the numerical outcome is verified by exact line search procedures.**

***Index Terms*: Unconstrained optimization, sufficient descent property, spectral conjugate gradient, global convergence.**

I. INTRODUCTION

The spectral CG method today is the standard method for solving large-scale unconstrained minimization problems, the method solves a large number of problems within a shortest possible time with few numbers of iteration and it requires a small storage location and less computational expensive as they do not use the Hessian matrix or its approximation. The spectral CG method has rapid global convergent properties and very effective in solving large-scale unconstrained problems [15]. Presently, [3] introduced a spectral CG method and computed their spectral parameter using the secant equation as originally given by [12]. Spectral CG method combines together the CG search direction and spectral parameter φ_k to construct a new search direction, see [1], [4], [8], [13], [14], [16], [17], [19], [20], [22] and [23], for the details.

In this paper, we formulate the spectral parameter via the CG search direction and a scalar parameter β_k of HS CG method. The general minimization problem is:

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

Manuscript published on November 30, 2019.
* Correspondence Author
Usman Abbas Yakubu*, Mathematics, Northwest University, Kano, Nigeria.
Mustafa Mamat Faculty of Informatics and Computing, University Sultan Zainal Abidin, Kuala Terengganu, Malaysia.
Abba Vulgue Mandara, Department of Mathematic, University of Maiduguri, Borno, Nigeria.
Abdul Iguda, Mathematical Sciences, Bayero University, Kano, Nigeria.
Salisu Murtala, Department of Mathematics, Federal University Dutse, Jigawa, Nigeria.
Mohamad Afendee Mohamed, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

where $f: R^n \to R$ is continuous and differentiable function, g_k is a gradient vector of function and $x_0 \in R^n$ is the initial point solved iteratively using recurrence expression below

$$x_{k+1} = x_k + \gamma_k d_k, \qquad k = 0,1,2,3,4,\dots \tag{2}$$

the vector x_{k+1} is a new iteration and x_k stand for a current iteration, the step size $\gamma_k > 0$ is obtained by line the search technique called exact line search given by

$$\gamma_k = arg \min_{\gamma>0} f(x_k + \gamma d_k) \tag{3}$$

d_k is stand for a search direction

$$d_k = \begin{cases} -g_k, & if\ k = 0 \\ -g_k + \beta_k d_{k-1}, & if\ k \geq 1 \end{cases} \tag{4}$$

$g_k = \nabla f(x)$ is the gradient vector, parameter β_k belong to the set of real number called CG coefficient. A classical β_k of HS is given below:

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})} \tag{5}$$

Thus, g_k and g_{k-1} in the above Hestenes-Stiefel (HS) parameter are gradient vectors of function at points x_k, x_{k-1} respectively, and $\|.\|$ represent a Euclidian norm. In this paper, a spectral HS (SpHS) CG method is presented without secant condition and proved its performance with recent spectral HS (RSHS) established by [13] and classical HS conjugate gradient methods and the new SpHS method can be applied on [24].

II. DETAILS OF SPHS CG METHOD

"Reference [1] introduced a Spectral Gradient Method (SGM)" by combining the Barzilai and Borwein spectral non-monotone techniques with classical projected gradient ideas for solving many variables problems. The recent articles were given by [13] and [14] inspire this piece of research. In order to determine spectral HS (SpHS) conjugate gradient method, the search direction is defined as

$$d_k = \begin{cases} -g_k, & if\ k = 0 \\ -\dfrac{1}{\varphi_k} g_k + \beta_k^{HS} d_{k-1}, & if\ k \geq 1 \end{cases} \tag{6}$$

From the search direction above, using the fact that $d_k = -g_k$ and substituting equation (5) we have,

$$\varphi_k = \left(1 - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}\right)^{-1} \tag{7}$$

Hence, a spectral scalar parameter φ_k is calculated by exact line search method.

Algorithm 1.0 (SpHS Method)

Step 1: Given a starting point $x_0 \in R^n$ set $k = 0$.

Step 2: Evaluate β_k as given in formula (5) above.
Step 3: Evaluate d_k given as in (6). If $\|g_k\| = 0$, then stop.
Step 4: Evaluate γ_k given in equation (3).
Step 5: Update the new point as given in the recurrence expression (2).
Step 6: If $f(x_{k+1}) < f(x_k)$ and $\|g_k\| < \varepsilon$ then stop, otherwise go to step 1 with $k = k + 1$

III. GLOBAL CONVERGENCE ANALYSIS

A. Sufficient descent condition

Sufficient descent condition guarantees that global convergence of iterative procedures is attained. The inequality below must hold true.

$$g_k^T d_k \le -C\|g_k\|^2 \quad for\ k \ge 0\ and\ C > 0 \tag{8}$$

Theorem 1.0. *Suppose a CG method with search direction* (6) *and* β_k^{HS} *given by* (5), *the condition* (8) *holds* $\forall\, k \ge 0$.

Proof. The proof of this theorem is by induction, since $g_0^T d_0 = -\|g_0\|^2$, the condition (8) satisfied as $k = 0$. Now we assume it is true for $k \ge 0$. Also, the inequality (8) hold, from the equation (6) multiply through by g_{k+1}^T and substitute equation (7) and it is known from the conjugacy conditions $g_{k+1}^T d_k = 0$

$g_{k+1}^T d_{k+1} = \|g_{k+1}\|^2$

Hence for constant $C = 1$ condition (8) is true for $k + 1$. ■

B. Global convergence properties

The assumptions 1.0 given below is used in the objective function to analyses the global convergence properties of the general CG method.

Assumptions 1.0. (i) A level set $\Omega = \{x \in R^n \mid f(x) \le f(x_0)\}$ is bounded, the function f is continuously differentiable in a neighborhood N of the level set Ω and x_0 is a starting point.

(ii) $g(x)$ is globally Lipschitz continuous in N that is $\exists$ a constant $L > 0$, such that $\|g(x) - g(y)\| \le L\|x - y\|$ for any $x, y \in N$.

Lemma 1.0. *Suppose the assumptions 1.0 holds and consider any recurrence expression* (2) *and search direction* (6), γ_k *is found in equation* (3). *Then Zoutendijk condition below holds.*

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{9}$$

The Proof of this lemma is given in [5].

Theorem 1.1. *Suppose the assumptions 1.0 holds, for any* $\{x_k\}$, $\{d_k\}$ *be given as SpHS CG method,* γ_k *are determined by equation* (3) *and* β_k *in equation* (5). *Then*

$$\lim_{k\to\infty} \|g_k\| = 0 \tag{10}$$

Proof. Square both sides of equation (6) above,

$$\|d_{k+1}\|^2 = (\beta_k^{HS})^2\|d_k\|^2 - \frac{2}{\varphi_k} g_{k+1}^T d_{k+1} - \frac{1}{\varphi_k^2}\|g_{k+1}\|^2 \tag{11}$$

Substituting equation (5) in equation (11), use the inequality (8) we can easily rewrite the equation (11) as follows

$$\|d_{k+1}\|^2 = \frac{\|g_{k+1}\|^4}{\|g_k\|^4}\|d_k\|^2 - \|g_{k+1}\|^2\left(\frac{1}{\varphi_k^2} - \frac{2C}{\varphi_k}\right) \tag{12}$$

Substitute (7) in (12) and multiply both sides of the equation by $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$ and note that for exact line search $g_{k+1}^T d_k = 0$. Therefore,

$$\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \tag{13}$$

From the Lemma 1.0. It implies that Theorem 1.1 does not hold true, then $\lim_{k\to\infty} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} = \infty$ and from equation (13) this is true $\infty \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$. So Theorem 1.1 is true for a sufficient large k. ■

IV. NUMERICAL RESULTS

The significant part of this step is to establish a comprehensive numerical experiment and relative procedure that will merriments the potentiality and efficiency of the developed spectral HS (SpHS) CG method when comparing its performances on the recent spectral HS (RSHS) introduced by Du, (2011) and classical HS conjugate gradient methods. The comparisons are decided on CPU time and the number of iterations only. The stopping criteria used for both methods is $\varepsilon = 10^{-6}$ as recommended by [21] and $\|g_k\| < \varepsilon$. A benchmark problem functions in [8], [9], [10], [15] and [20] have used to established test problems in table 1 with four different initial points by *MatlabR2015 subroutine* programming using Intel® Core™ *i*5-3317U (1.7GHz with 4 GB (RAM)). In this paper, failure of the test function in each method is characterized due to (i) memory requirement (ii) number of iterations exceed 1000 (iii) CPU time running in seconds reaches 1000 (iv) Line search method failed to find the step size or step length γ_k (v) If $\|g_k\|$ is not a number (NaN). The numerical outcomes revealed in the Fig.1 and Fig.2 are the eminent performance outline introduced by [2]. Performances are scaled based on a number of iterations and CPU time respectively. Thus, the maximum value of the percentage of probability $P_s(t)$ and the solver that swiftly reached the top as shown in the Fig.1 and Fig.2 resolve to be the best performing CG method for unconstrained optimization problems.

Table.1: Standard Test Problems functions

FUNCTIONS	DIMENSIONS	INITIAL POINTS
Trecanni	2	(2,2), (4,4), (5,5), (15,15)
Booth	2	(12,12), (-12,-12), (40,40), (-40,-40)
Zettl	2	(11,11), (-11,-11), (110,110), (-110,-110)
Leon	2	(3,3), (5,5), (7,7), (-7,-7)
Quartic	4	(22,22), (-22,-22), (50,50), (60,60)
Colville	4	(2,2), (13,13), (-13,-13), (19,19)
Wood	4	(5,5), (35,35),(50,50),(70,70)
Gen. Tridiagonal 1	10	(3, 3), (-3,-3), (5,5), (-5,-5)
Gen. Tridiagonal 2	10	(3,3), (-3,-3), (11,11), (21,21)
Fletcher	10	(2,2), (-2,-2), (21,21),), (-21,-21)
Freud. & Roth	2,4	(4,-4), (8,8), (9,9), (10,10)

Retrieval Number: B3003078219/2019©BEIESP
DOI:10.35940/ijrte.B3003.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Hager	2,4,10,100	(12,12), (-12,-12), (17,17), (-19,-19)
Quadratic Penalty QP1	2,4,10,100	(2,2), (-2,-2), (23,23), (-23,-23)
Raydan 1	2,4,10,100	(4,4), (-5,-5), (-7,-7), (-23,-23)
Ext. Tridiagonal 1	2,4,10,100,1000	(-5,-5), (21,21), (21,21), (25,25)
Extended Penalty	2,4,10,100,1000	(2,2), (8,8), (14,14), (25,-25)
Himmelblau	2,4,10,100,1000, 10000,100000	(8,8), (-8,-8), (11,11), (-11,-11)
White and Holst	2,4,10,100,1000, 10000,100000	(5,5), (7,7), (-7,-7), (10,10)
Shallow	2,4,10,100,1000, 10000,100000	(3,3), (-3,-3), (-11,-11), (20,20)
Rosenbrock	2,4,10,100,1000, 10000,100000	(15,15), (-15,-15), (17,17), (-17,-17)

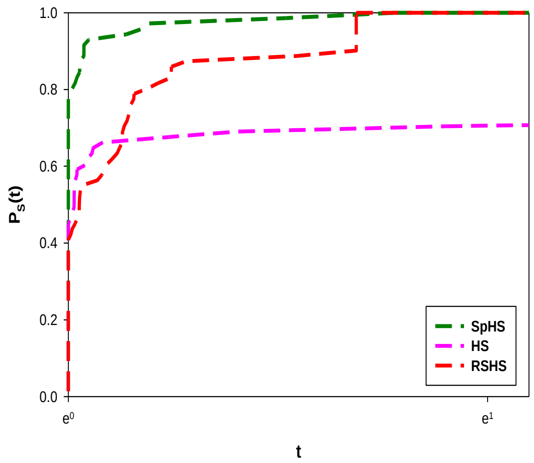

Fig.1: Performance outline based on the number of iterations

In this discussion the spectral HS (SpHS) CG method has done well since it solved the entire benchmark problems functions used in this paper. Fig.1 and Fig.2 shows that SpHS, HS and RSHS CG methods reached the solution point even though, the SpHS CG method has the maximum probability of being the fastest method on approximately = 80% and 65% respectively of the benchmark problems used in this research. Denoting the ‘‘successful’’ in the table 2 and table 3 meaning that SpHS CG method has less number of iterations and less CPU time compared to RSHS and HS CG methods while denoting the ‘‘Abortive’’ meaning that SpHS method produces more iterations and consumes more CPU time as compared to RSHS and HS CG methods. If the SpHS CG method provides the same number of iterations or CPU time with RSHS and HS CG methods then the method is equivalent to each other. For both table 2 and table 3, the results are obtainable in percentages as follows. The highest value of percentage showed that the SpHS CG method will be successful over the RSHS and HS CG methods.

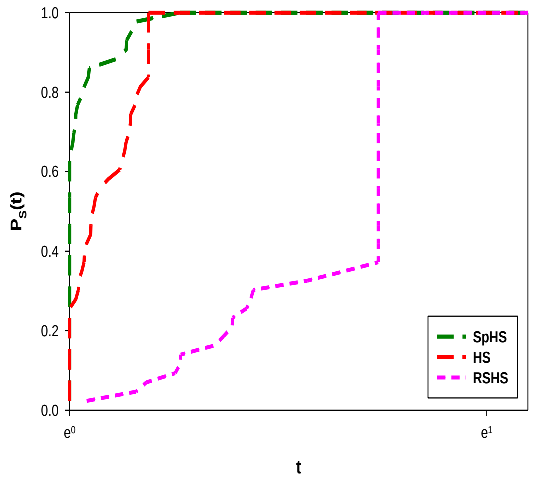

Fig.2: Performance outline based on CPU time.

Table 2: Percentage Performance of SpHS versus RSHS, HS CG Methods on Number of Iterations

Method	Comparison	RSHS	HS
SpHS	Successful	57.75%	50.70%
	Equivalent	32.29%	35.21%
	Abortive	9.96%	14.09%

In view of that, SpHS CG method in table 2 above solves the entire benchmark problems with the maximum percentage of success and its equivalent of 90.04% compared to RSHS CG method, 85.91% compared with HS CG method in terms of a number of iterations.

Table 3: Percentage Performance of SpHS versus RSHS, HS CG Methods on CPU time

Method	Comparison	RSHS	HS
SpHS	Successful	77.46%	76.06%
	Equivalent	0.0%	0.0%
	Abortive	22.54%	23.94%

Subsequently, the percentage of success and its equivalent of SpHS CG method based on CPU time is 77.46% as compared with the RSHS CG method and 76.06% compared to the HS CG method. Henceforth, the overall average percentage of success for both a number of iterations and CPU time of SpHS CG method is 87.96% and 76.76% for the average success and its corresponding equivalent. Ultimately, the new SpHS CG method has absolute capabilities of being the optimal solver as compared with RSHS and HS CG methods. Indeed, the effectiveness of the SpHS CG method is highly commendable.

V. CONCLUSION

Lastly, the new proposed SpHS CG method overcomes the weaknesses of recent spectral HS (RSHS) CG method, HS CG method and converges globally.

Retrieval Number: B3003078219/2019©BEIESP
DOI:10.35940/ijrte.B3003.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The numerical results show that the method performed excellently in terms of CPU time running in seconds and number of iterations reserved to achieve a sufficient minimum point in the objective function.

ACKNOWLEDGMENT

The authors thank the reviewers for their contributions towards achieving this important research and so also very grateful to the Malaysian Government and Universiti Sultan Zainal Abidin for funding this research under the Fundamental Research Grant Scheme (FRGS/1/2017/STG06/UniSZA/01/1).

REFERENCES

1. M. Raydan. The Barzilai and J.M. Borwein gradient methods for the large scale unconstrained minimization in extreme problems, SIAM. J. Optim. Vol. 7 (1), pp. 26-33, 1997.
2. E. Dolan, J.J More. Benchmarking optimization software with performance profile, Math. Prog. Vol. 91, pp. 201-213, 2002.
3. E.G. Birgin, J. M. Martinez. A spectral conjugate gradient method for unconstrained optimization, Appl. Math. Optim. Vol. 43 (2), pp. 117-128, 2011.
4. X. Wu. A new spectral Polak- Ribière -Polak conjugate gradient method, ScienceAsia. Vol. 41, pp. 345-349, 2015.
5. G. Zoutendijk. "Nonlinear programming, computational methods, in J Abadie (ED)", Integer and Nonlinear Programming, North-Holland, Amsterdam, 1970, pp. 37-86.
6. J. C. Gilbert, J. Nocedal. Global convergence properties of conjugate gradient methods for optimization, SIAM. J. Optim. 2, 21-42, 1992.
7. C. Hu, Z. Wan. An Extended Spectral Conjugate Gradient Method for unconstrained optimization problems, British Journal of Mathematics & Computer Science, Vol. 3, pp. 86-98, 2013.
8. U.A. Yakubu, M. Mamat, A.M. Mohamad, A.M. Sukono, and M. Rivaie. Secant free condition of a spectral PRP conjugate gradient method. International Journal of Engineering & Technology, Vol. 7 (3.28), pp. 325 – 328, 2018.
9. U.A. Yakubu, M. Mamat, A.M. Mohamad, A.M. Sukono, and M. Rivaie. Modification on spectral conjugate gradient method for unconstrained optimization. International Journal of Engineering & Technology, Vol. 7 (3.28), pp. 307 – 311, 2018.
10. U.A. Yakubu, M. Mamat, A.M. Mohamad, A.M. Puspa, L.G., and M. Rivaie. Secant free condition of a spectral Hestenes-Stiefel conjugate gradient method and its sufficient descent properties. International Journal of Engineering & Technology, Vol. 7 (3.28), pp. 312 – 315, 2018.
11. M.J.D. Powell. Restart procedures for the conjugate gradient method, Math. Program. Vol. 12, pp. 241-254, 1977.
12. J. Barzilai, J.M. Borwein. Two point step size gradient methods, IMA J Numer Anal. Vol. 8, pp. 141-148, 1988.
13. X. Du, J. Liu. Global convergence of a spectral HS conjugate gradient method, Procedia Engineering. Vol. 15, pp. 1487-1492, 2011.
14. N. Zull, M. Rivaie, M. Mamat, Z. Salleh, Z. Amani. Global convergence of a Spectral conjugate gradient by using strong Wolfe line search, Appl. Math. Sci. Vol. 63, pp. 3105-3117, 2015.
15. N. Andrei. An unconstrained optimization test functions collection, Adv. Modell. Optim. Vol. 10, pp. 147-161, 2008.
16. U.A. Yakubu, M. Mamat, M. Rivaie, A.M. Mohamad, and B.Y. Rabi'u. Secant free condition of a spectral WYL and its global convergence properties. Far East Journal of Mathematical science, Vol. 12, pp. 1889 – 1902, 2018.
17. U.A. Yakubu, M. Mamat, M. Rivaie, A.M. Mohamad, and J. Sabi'u. A recent modification on Dai-Liao conjugate gradient method for solving symmetric nonlinear equations. Far East Journal of Mathematical science, Vol. 12, pp. 1961 – 1974, 2018.
18. W. W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods, Pacific Journal of Optimization. Vol. 2 (1), pp. 35-58, 2006.
19. U.A. Yakubu and U. Yuksel. Necessary and sufficient conditions for first-order differential operators to be associated with disturbed Dirac operator in quaternionic analysis. Adv. appl. Clifford alg., Vol. 25 (1), pp. 1-12 2015.
20. A. Abashar, M. Mamat, M. Rivaie, and I. Mohd. Global convergence properties of a new class of conjugate gradient method for unconstrained optimization. Applied Mathematical Sciences Issue 65-68, pp. 3307-3319, 2014.
21. K.E. Hillstrom. A simulation test approach to the evaluation of the nonlinear optimization algorithm. *Journal ACM Trans.* Mathematics Software. Vol. 3 (4), pp. 305-315, 1977.
22. V. M. Abba, M. Mamat, M.Y. Waziri, A.M. Mohamad, and U.A. Yakubu. A new conjugate gradient coefficient with exact line search for unconstrained optimization Far East Journal of Mathematical science, Vol. 105, (2), pp. 193 – 206, 2018.
23. P. Faramarzi, K. Amini. A Modified Spectral Conjugate Gradient Method with Global Convergence. Journal of Optimization Theory and Applications. April 2019.
24. M. Mamat, Subiyanto, and A. Kartono. Mathematical model of cancer treatments using immunotherapy, chemotherapy and bio chemotherapy. Applied Mathematical Sciences Issue 5-8, Vol. 7, pp. 247-261, 2013.