International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 7 Issue 6, June-2018, Impact Factor: 4.059

Global Convergence properties of scalar of conjugate gradient methods for minimization problem

Eilaf Sulaiman Khaleel

Department of Mathematics. College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq

ABSTRACT

In this paper anew scaled is proposed for solving unconstrained optimization problems. Under mild condition the sufficiently descent property and global convergence of the proposed new method is proved with using strong Wolfe line search. Numerical comparison show that new proposed method is effective by comparing with Perry method.

Keywords: unconstrained optimization, conjugate gradient method, global convergence.

HOW TO CITE THIS ARTICLE

Eilaf Sulaiman Khaleel, "Global Convergence properties of scalar of conjugate gradient methods for minimization problem", International Journal of Enhanced Research in Science, Technology & Engineering, ISSN: 2319-7463, Vol. 7 Issue 6, June -2018.

1. INTRODUCTION

The function$f: R^n \to R$ is a continuously differentiable function whose gradient is denoted by g . the problem consider is [1]

$$\min f(x) \qquad x \in R^n \quad \ldots\ldots\ldots\ldots\ldots..(1)$$

The iterative formula of the conjugate gradient method is given by:

$$x_{i+1} = x_i + \alpha_i d_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

where $\alpha$is step length which is compute by using some line search $\alpha_i > 0$ which satisfies the Wolf conditions (W.C.):

$$f(x_i + \alpha_i d_i) - f(x_i) \le \delta \alpha_i g_i^T d_i \quad \ldots\ldots\ldots\ldots..(3)$$
$$|g_{i+1}^T d_i| \le -\sigma g_i^T d_i \quad \ldots\ldots\ldots\ldots.(4)$$

with $0 < \delta < \sigma < 1$, and $f_i = f(x_i)$, $g_i = g(x_i)$, g_i is the gradient of f evaluated at the new iterate x_i.[6] and $d_i$isthe search direction defined by :

$$d_i = \begin{cases} -g_i & for\ i = 1 \\ -g_i + \beta_i d_{i-1} & for\ i \ge 2 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots (5)$$

where β_i is ascalar , $g_i$denotes g(x) . [4]

2. NEW ALGORITHM FOR SCALED CONJUGATE GRADIENT METHOD

We proposed new scalar $\theta$of conjugate gradient method (CG), which d_i is the search direction defined by

$$d_i = -\theta_i g_i + \beta_i d_{i-1} \quad \ldots\ldots\ldots.(6)$$

and β_i is defined by [2]

$$\beta_i = \frac{g_i^T (I + \rho I - H_i^{-1}) y_{i-1}}{d_{i-1}^T y_{i-1}} \quad \ldots..(7)$$

Where I is the identity matrix and p>0 .the matrix H is updated by the BFGS formula as[7]

$$H_{i+1} = H_i - \frac{H_i s_i s_i^T H_i}{s_i^T H_i s_i} + \frac{y_i y_i^T}{s_i^T y_i} \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

which satisfied the QN condition :

$$H_{i+1} y_i = s_i$$

from eq.(6)

$\theta_i g_i = \beta_i d_{i-1} - d_i \quad *g_i^T \quad \ldots\ldots(9)$

$\theta_i \|g_i\|^2 = \beta_i g_i^T d_{i-1} - g_i^T d_i$

By $g_i^T d_i = -\|g_i\|^2$

$= \beta_i g_i^T d_{i-1} + \|g_i\|^2 \quad \ldots\ldots.. (10)$

$\theta_i = \frac{\|g_i\|^2}{\|g_i\|^2} + \beta_i \frac{g_i^T d_{i-1}}{\|g_i\|^2}$

$= 1 + \beta_i \frac{g_i^T d_{i-1}}{\|g_i\|^2} \quad \ldots\ldots\ldots.. (11)$

since

$\beta_i = \frac{g_i^T (I + \rho I - H_i^{-1}) y_{i-1}}{d_{i-1}^T y_{i-1}}$

$\theta_i = 1 + \frac{g_i^T (I + \rho I - H_i^{-1}) y_{i-1}}{d_{i-1}^T y_{i-1}} \frac{g_i^T d_{i-1}}{\|g_i\|^2} \quad \ldots\ldots\ldots(12)$

If we use ELS then θ is equal to 1

3. ALGORITHEM

Step (1): put a starting point x_1 and $H_1 = I_n$ and compute$f(x_1), g(x_1), d_1 = -g_1$ and i=1
Step (2): test convergent
If $\|g_i\| \le \epsilon \quad , \epsilon = 10^{-7}$ stop x_i is the optimal solution else go to step (3)
Step (3): computation α_i by line search and update variable $x_{i+1} = x_i + \alpha_i d_i$
Step (4): update (H_i) by eq.(8)
Step (5): computation direction d_i by eq.(6) where β is defend by eq.(7)and θ_i is defend by eq.(11).
Step (6): calculate $f_{i+1}, g_{i+1}, y_i, s_i$
Step (7): set i=i+1 and go to step (2)

4. CONVERGENCE

We need to make a few assumption based on the objective function to prove the convergence .

Assumption 4.1 :[3]

A- The level set L is convex moreover , positive constants C_1, C_2
$C_1 \|Z\|^2 \le Z^T H(x) Z \le C_2 \|Z\|^2$for all $Z \in R^n$ and $x \in L$, where H(x) is the Hessian matrix for f.
$B -$ The Hessian matrix is lipschitz continuous at the point X^*, that is there exists the positive constant C_3 satisfying
$\|g(x) - g(x^*)\| \le c_3 \|x - x^*\| \;\; for\ all\ x\ in\ aneighborhood\ of\ x^*$

Theorem 4.1:
Assume that the sequence $\{H_i\}$, $\{y_i\}$,$\{X_i\}$ all created by algorithm BFGS-SD are positive definite . then the search direction $d_i$defined by eq.(6) which β_i from eq.(7) and θ_i from eq.(12) , satisfied the descentcondition , for all i≥ 0 ,There exists a constant $c_1 > 0$ such that

$$\frac{g_i^T d_i}{\|g_i\|^2} \le -c_1 \; for\ all\ i \ge 0 \ldots\ldots.. \quad (13)$$

Proof :
From $d_i = -\theta_i g_i + \beta_i d_{i-1}$
$g_i^T d_i = -\theta_i \|g_i\|^2 + \beta_i g_i^T d_{i-1}$

$$= -\left(1 + \frac{g_i^T(I+\rho I - H_i^{-1})y_{i-1}}{d_{i-1}^T y_{i-1}} \frac{g_i^T d_{i-1}}{\|g_i\|^2}\right)\|g_i\|^2 + \left(\frac{g_i(I+\rho I - H_i^{-1})y_{i-1}}{d_{i-1}^T y_{i-1}}\right) g_i^T d_{i-1}$$
$$= -\|g_i\|^2 - \left(\frac{g_i^T(I+\rho I - H_i^{-1})y_{i-1}}{d_{i-1}^T y_{i-1}} g_i^T d_{i-1}\right) + \left(\frac{g_i{}^T(I+\rho I - H_i^{-1})y_{i-1}}{d_{i-1}^T y_{i-1}}\right) g_i^T d_{i-1}$$

since$g_i^T d_i = -\|g_i\|^2$
let c=1

$$\frac{g_i^T d_i}{\|g_i\|^2} \le -cc > 0$$

Theorem 4:2 (Global Convergence)

Suppose the search direction d_i defined by eq.(6) which β_i from eq.(7) and θ_i from eq.(12) , satisfied the descent condition and assumption 4.1 hold then

$$\liminf_{i\to\infty} \|g_i\|^2 = 0$$

Proof :

$$d_i = -\theta_i g_i + B_i d_{i-1}$$
$$\|d_i\| \le \|\theta_i g_i\| + \|B_i d_{i-1}\|$$
$$\le \left\|1 + \frac{g_i^T(I+\rho I - H_i^{-1})y_{i-1}}{d_{i-1}^T y_{i-1}} \frac{g_i^T d_{i-1}}{\|g_i\|^2}(g_i)\right\| + \left\|\frac{g_i{}^T(I+\rho I - H_i^{-1})y_{i-1}}{d_{i-1}^T y_{i-1}} d_{i-1}\right\|$$
$$\le \|g_i\| + \frac{|g_i^T(I+\rho I - H_i^{-1})y_{i-1}|}{|d_{i-1}^T y_{i-1}|} \frac{|g_i^T d_{i-1}|}{\|g_i\|^2}\|g_i\| + \frac{|g_i{}^T(I+\rho I - H_i^{-1})y_{i-1}|}{|d_{i-1}^T y_{i-1}|}\|d_{i-1}\|$$

By strong Wolfe Condition [8]

$|g_i^T d_{i-1}| \le -\sigma g_i^T d_i$
$g_i^T d_i = -\|g_i\|^2$

$$\le \|g_i\| - \frac{|g_i^T(I+\rho I - H_i^{-1})y_{i-1}|}{|d_{i-1}^T y_{i-1}|} \frac{\sigma\|g_i\|^2}{\|g_i\|^2}\|g_i\| + \frac{|g_i{}^T(I+\rho I - H_i^{-1})y_{i-1}|}{|d_{i-1}^T y_{i-1}|}\|d_{i-1}\|$$
$$\le \|g_i\| + \|(I+\rho I - H_i^{-1})\|\left[-\sigma \frac{|g_i^T y_{i-1}|}{|d_{i-1}^T y_{i-1}|}\|g_i\| + \frac{|g^T{}_i y_{i-1}|}{|d_{i-1}^T y_{i-1}|}\|d_{i-1}\|\right]$$
$$\le \|g_i\| + \|(I+\rho I - H_i^{-1})\|\left[-\sigma \frac{\|g_i\|\|y_{i-1}\|}{\|d_{i-1}\|\|y_{i-1}\|}\|g_i\| + \frac{\|g_i\|\|y_{i-1}\|}{\|d_{i-1}\|\|y_{i-1}\|}\|d_{i-1}\|\right]$$
$$\le \|g_i\|\left[1 + \|(I+\rho I - H_i^{-1})\|\left[-\sigma\frac{\|g_i\|}{\|d_{i-1}\|} + 1\right]\right]$$

let $c_1 = -\sigma \frac{\|g_i\|}{\|d_{i-1}\|} + 1 > 0$is constant
sinceH_i is bounded
implies that
$c_2 = \|(I+\rho I - H_i^{-1})\|$
$c = 1 + c_2 c_1$
$\|d_i\| \le c\|g_i\|$=k
$$\sum_{i\to 1}^{\infty} \frac{1}{\|d_i\|^2} \ge \frac{1}{k^2}\sum_{i\to 1}^{\infty} 1 = \infty$$
$\therefore \liminf_{i\to\infty} \|g_i\|^2 = 0$

5. NUMERICAL RESULTS

Several standard nonlinear unconstrained test function were minimize to compare the new algorithm with Perry algorithm show that the numerical result the comparative performance for all these algorithm are evaluated by considering both the (Number of Function "NOF") which is number of function evaluations and the (Number of Iteration "NOI")which is number of iterations and all these methods terminate when the following stooping criterion i,$\|g_i\| \le \epsilon$ $\epsilon = 10^{-7}$. In Table(1) we have compared our new algorithm with Perry algorithm by using different dimension, while Table(2) taking the average dimension of n=1000, 2000,…, 10000.

Table(1): Numerical Comparisons between the method and New method.

N	Test fu.	Dim	Perry	New
1	Trigonometric	1000 10000	29(18) 16(15)	29(19) 16(14)
2	Rosenbrock	1000 10000	39(22) 37(21)	34(18) 36(20)
3	White & holst	1000 10000	32(17) 33(19)	35(20) 35(20)
4	Penalty	1000 10000	2(2) 2(2)	2(2) 2(2)
5	perturbed quadratic	1000 10000	230(84) 1089(307)	311(84) 1223(337)
6	Raydon 1	1000 10000	414(169) 2001(1470)	327(157) 2001(1450)
7	Raydon 2	1000 10000	4(4) 4(4)	4(4) 4(4)
8	Generalized tridiag	1000 10000	24(5) 24(8)	27(14) 24(9)
9	Extended three exponential terms	1000 10000	10(6) 12(7)	8(4) 8(4)
10	Diagonal 4	1000 10000	4(2) 6(3)	4(2) 4(2)
11	Extended Himmeblau	1000 10000	19(10) 19(10)	10(6) 19(10)
12	Extended maratos	1000 10000	69(34) 66(33)	63(29) 65(30)
13	Extended hiebert	1000 10000	79(50) 79(50)	79(50) 77(49)
14	Extended quadratic penalty QP2	1000 10000	37(20) 42(21)	35(20) 40(20)
15	Quadratic function qf2	1000 10000	356(99) 1383(403)	357(106) 1261(373)
16	Extended ep1	1000 10000	2(2) 20(20)	2(2) 17(17)
17	Tridia	1000 10000	1160(319) 2001(541)	964(268) 2001(556)
18	Nondia	1000 10000	14(7) 4(3)	10(5) 4(3)
19	Broyden tridiagonal	1000 10000	40(19) 75(29)	35(15) 79(30)
20	Tridiagonal perturbed quadratic	1000 10000	301(85) 1297(342)	382(107) 1051(283)

Table (2): The total Numerical Comparisons between the method and New method.

N	Test fu.	perry	New
1	Trigonometric	204(188)	203(169)
2	Rosenbrock	386(216)	350(190)
3	White & holst	326(177)	350(200)
4	Penalty	40(40)	40(40)
5	perturbed quadratic	8042(2204)	7932(2187)
6	Raydon 1	12005(8140)	11035(2187)
7	Raydon 2	42(42)	42(42)
8	Generalized tridiag	306(149)	298(143)
9	Extended three exponential terms	109(65)	81(41)

10	Diagonal 4	44(22)	40(20)
11	Extended Himmeblau	190(100)	172(92)
12	Extended maratos	668(333)	666(304)
13	Extended hiebert	788(499)	776(493)
14	Extended quadratic penalty QP2	415(209)	407(207)
15	Quadratic function qf2	9200(2687)	8860(2603)
16	Extended ep1	45(45)	42(42)
17	Tridia	19070(5189)	18538(5103)
18	Nondia	62(38)	58(36)
19	Broyden tridiagonal	705(260)	698(251)
20	Tridiagonal perturbed quadratic	8175(2224)	8007(2218)

CONCLUSION

This paper used the scalar to modified the conjugate gradient methods for finding the minimum of unconstrained problems. With strong Wolfe condition this proposed method is proved to satisfy the sufficient descent condition and global convergence.

APPENDIX

The specifics the test functions that used in search can be find in [3]

1- Trigonometric function :
2- Rosenbrock function :
3- White & holst function :
4- Penalty function :
5- Perturbed quadratic function
6- Raydan1 Function
7- Raydan2 Function
8- Generalization tridiag function :
9- Extended three exponential terms
10- Diagonal-4 Function:
11- Extended Himmelblau Function:
12- Extended maratos
13- Extendedhiebert
14- Extended quadratic penalty QP1 function :
15- Quadrati function QF
16- Extended EP1
17- Tridia function
18- Nondia function
19- Broyden tridiagonal function
20- Tridiagnol perturbeded quadratic

REFERENCES

[1]. Al Bayati, A. Y. , And Hassan, S. K.,"New Extended Polak- Ribiere CG-Methods For Nonlinear Unconstraint Optimization", Canadian Journal On Science And Engineering Mathematics ,Vol.2, No.1 , 2011.
[2]. Alkailany, M.,A.,"the modified method for solving non-linear programming problems", International Journal of Enhanced Research in science technology & Engineering, vol.6,Issue8, 2017.
[3]. Andrei,A. " An Unconstrained Optimization Test Functions Collection" ,Advanced Modeling and Optimization (A.M.O.) ,ISSN: 1841-4311, Vol. 10, No. 1, 2008.
[4]. Fan,H., Zhu,Z., and Zhou,A., "Anew Descent nonlinear conjugate gradient method for unconstrained optimization" , Applied Mathematics, no.2, pp.1119-1123, 2011.
[5]. Ibrahim,M.A. , Mamat, M. , and Leong, W.J. , "The Hybrid BFGS -CG Method In Solving Unconstrained Optimization Problems" , Vol.2014, Article ID 507102, 2016.
[6]. Wolfe , p., " convergence conditions for a scent methods " , SIAM Rev., Vol.11 , pp.226-235, 1969.
[7]. Yuan, G. And Wei,Z. "Convergence Analysis Of A Modified BFGS Method On Convex Minimizations", Springer , Com. Optim. App47:237-255,2010.
[8]. Zeng-xin, w., yan-rong,T., "Global convergence of a modified LS conjugate gradient method with-strong Wolfe conditions", journal of guangxi university(Nat Sei Ed). Vol.33, no.2, 2008.