ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

A new Modified Nonlinear Conjugate Gradient Coefficients with Global Convergent Properties

Abdelrhaman Abashar

Department of Basic Sciences, Faculty of Engineering, Red Sea University, Sudan.

Abstract

Nonlinear conjugate gradient (CG) methods are currently considered to be one of most important techniques for solving large- scale unconstrained optimization problems. The main objective of present work is to propose a new method to improve the CG coefficient that possesses sufficient descent and global convergent properties under exact line searches. The numerical outcome established on number of iteration and CPU time shows that the proposed coefficient is more effective when compared with other CG formulas.

Keywords: Conjugate gradient, exact line search, sufficient descent condition.

1. Introduction

Consider the following unconstrained optimization problem

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

where $f: R^n \to R$ is continuously differentiable function. Conjugate gradient method is best methods for solving (1), particularly when the dimension n is large. The iterates of conjugate gradient methods are given by

$$x_{k+1} = x_k + \alpha_k d_k; \qquad (2)$$

where x_k is present iterate point and α_k is step size, which is computed by carrying out some line search, especially in exact line search

$$\alpha_k = \arg min_{\alpha \geq 0} f(x_k + \alpha d_k)\,. \quad (3)$$

ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

where d_k is the search direction defined by

$$d_k = \begin{cases} -g_k & if \quad k=0 \\ -g_k + \beta_k d_{k-1} & if \quad k \geq 1 \end{cases} \qquad (4)$$

where g_k is the gradient of $f(x)$ at the point x_k , β_k a scalar . We know classical formula's for β_k are the Hestenes-Stiefe (HS) method [5]. The Fletcher – Reeves (FR),[6].The Polak-Ribiere (PR) method [7,8].The conjugate descent method (CD),[9]. The Liu – Storey (LS), [10] and the Dai – Yuan (DY) method [11], the parameters β_k of these methods are as follows

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \quad (5) \qquad \beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \quad (6)$$

$$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}} \quad (7) \qquad \beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}} \quad (8)$$

$$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}} \quad (9) \qquad \beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}} \quad (10)$$

Zoutendijk [2] showed that the FR method with exact line search is globally convergent. Al-Baali [1] extended this result to the strong Wolfe-Powell line search. Powell [12] proved that the sequence of gradient norms could be bounded away from zero only when

$$\sum_{k \geq 0} \frac{1}{\|d_k\|} < +\infty \qquad (11)$$

In recent past, many researchers have been presented good comparative studies of some new CG methods such as, Andrei [3] Rivaie et al [13], Sun and Zhang [4], Wei et al. [16], A. Abashar et al. [15] and Hager and Zhang [18].

In this paper, we propose a new formula for the coefficient β_k and algorithm in section 2. We show that this parameter possesses sufficient descent property and satisfy global convergence analysis under exact line search in section 3. In Section 4, we do some numerical comparisons with other well-known conjugate gradient methods and discuss the results. Finally, in Section 5 we present our conclusion

2. The new modification

Recently, Rivaie et al. [13], proposed a new formula

$$\beta_k^{RMIL} = \frac{g_k^T\,(g_k - g_{k-1})}{\|d_{k-1}\|^2} \qquad (12)$$

Based on this formula, a new modification defined by

$$\beta_k^{*} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_{k-1}\|^2, \mu\|g_{k-1}\|^2)} \qquad (13)$$

where $\mu > 1$ is real number, the following algorithm describes the steps of using the CG method with (13) and exact line search to obtain the solution for the optimization functions.

Algorithm 2.1

Step 0. Initialization, given x_0.

Step 1. If $\|g_k\|^2 \leq \varepsilon$, then out put x_k ,stop.

Step 2. Compute β_k based on (6,12,13).

Step 3.Compute d_k based on (4).

Step 4. Compute $\alpha_k$based on (3).

Step 5. Update a new point based as (3).

Step 6. Set $k = k + 1$, go to step 2.

3. Global Convergence Analysis

In this section, the convergent analysis of above algorithm is presented. We show that the

ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

algorithm satisfy the sufficient descent condition and the global convergence properties.

The sufficient descent condition is defined as follows

$$g_k^T d_k \le -c \|g_k\|^2 \quad , \qquad c \in (0,1) \,. \qquad (14)$$

The following theorem shows that our new formula β_k^* , with exact line search possess the sufficient descent condition.

Theorem 1.

Suppose that the sequences $\{g_k\}$ and $\{d_k\}$ be generated by the CG method presented in equation (2), and the step size α_k is determined by the exact line search (3) then the condition (14) holds true for all $k \ge 0$.

Proof .

If $k = 0$ then $g_0^T d_0 \le -c \|g_0\|^2$

Hence condition holds true. We need to show that for $k \ge 1$, condition is true

From (4) multiply by g_{k+1}^T then

$$g_{k+1}^T d_{k+1} = g_{k+1}^T(-g_{k+1} + \beta_{k+1} d_k) = -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k \quad (15)$$

For exact line search, we know that $g_{k+1}^T d_k = 0$. Thus

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

Hence this condition is true for $k+1$

Lemma 1

The coefficient β_k^* satisfies the following

$$0 \le \beta_k^* \le \frac{\|g_k\|^2}{\mu \|g_{k-1}\|^2} \le \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

Proof

Firstly,we show that β_k^* is always not less than zero. By using Cauchy- Schwartz inequalities, we get

$$\beta_k^* = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_k\|^2, \mu\|g_{k-1}\|^2)} \geq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\|g_k\|\|g_{k-1}\|}{max(\|d_k\|^2, \mu\|g_{k-1}\|^2)} \geq 0$$

which shows $\beta_k^* \geq 0$

Secondly from definition of β_k^* ,

$$\beta_k^* = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_k\|^2, \mu\|g_{k-1}\|^2)} \leq \frac{\|g_k\|^2}{max(\|d_k\|^2, \mu\|g_{k-1}\|^2)} \leq \frac{\|g_k\|^2}{\mu\|g_{k-1}\|^2} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

This implies that $\beta_k^* \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$

(16)

In order to show the global convergence properties of CG methods. We need the following assumption.

Assumption 1

(i) The level set $\Omega = \{x \in R^n \; f(x) \leq f(x_0)\}$ is bounded, where x_0 is the starting point.

(ii) In some neighborhood N of Ω , the objective function is continuously differentiable, and its gradient is Lipschitz continuous, (there exists a constant $l > 0$ such that

ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

$$\|g(x) - g(y)\| \le l\,\|x - y\|, \forall\, x, y \in N).$$

The following lemma is based on Assumption 1. This lemma is also known as the Zoutendijk condition which plays an important role in global convergence of CG method with inexact line search and exact line search.

Lemma 2

Suppose that Assumption 1 holds. Consider any CG methods in the form (2) and (4) where d_k is a decent search direction and α_k satisfies the exact line search then the following holds,

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

The proof of Lemma 2 can be found in Zoutendijk [2].

Theorem 2:

Suppose that Assumption 1 is holds. Consider CG method of the form (4), α_k is obtained by the exact line search (3, the sufficient descent condition hold and suppose that β_k^* is obtained by (13)

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

Proof. To prove Theorem 2, we use contradiction. That is, if Theorem 2 is not true, then a constant $c > 0$ exists, such that

$$\|g_k\| \ge c \qquad (17)$$

Rewriting (4) to get

ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

$$d_k + g_k = \beta_k d_{k-1} \tag{18}$$

Squaring both sides of (18) to get

$$\|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + \beta_k^2 \|d_{k-1}\|^2 \tag{19}$$

From (16)

$$\|d_k\|^2 \leq -\|g_k\|^2 - 2g_k^T d_k + \left(\frac{\|g_k\|^2}{\|g_{k-1}\|^2}\right)^2 \|d_{k-1}\|^2$$

For exact line search $g_k^T d_k = -\|g_k\|^2$

$$\|d_k\|^2 \leq -\|g_k\|^2 + 2\|g_k\|^2 + \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_{k-1}\|^2$$

$$\|d_k\|^2 \leq \|g_k\|^2 + \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_{k-1}\|^2$$

Dividing both sides by $\|g_k\|^4$ to obtain

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4}$$

Noting that $\frac{\|d_0\|^2}{\|g_0\|^4} = \frac{1}{\|g_0\|^2}$

$$\frac{\|d_1\|^2}{\|g_1\|^4} \leq \frac{1}{\|g_1\|^2} + \frac{1}{\|g_0\|^2}$$

Then for all k we have,

 ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \sum_{i=0}^{k} \frac{1}{\|g_i\|^2} \qquad (20)$$

There for it follows from (17) and (20)

$$\frac{\|g_k\|^4}{\|d_k\|^2} \geq \frac{c^2}{k}$$

This implies that

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = \infty$$

This contradicts the Zoutendijk condition in Lemma 2. Therefore, the proof is completed.

4. Numerical results and discussion

In this section, we carried out some numerical experiments to test Algorithm 2.1; we use test problem considered in Andrei [14]. To analyze the efficiency of our new formula as compared with other CG methods mention in (6) and (7). The comparisons are based on the number of iterations and CPU time. We considered $\varepsilon = 10^{-6}$ and $\|g_k\| \leq \varepsilon$ as stopping criteria as presented in Hillstrom [17]. All problems are ruined under MATLAB code 7.10.0 (R 2010a). The CPU processor used was Core™ i3- 2328M(2.2GHZ,3MB L3 Cache), with 6GB DDR3 RAM. The performance results shown in Figs.1 and 2, respectively, are based on the performance profile introduced by Dolan and More [19]. The performance results are shown in Figs.1 and 2 for exact line search. This performance profile provided

ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

the means to evaluate and compare the performance of the set of solvers S on a test problem. Assuming n_s solvers and n_p problem exist, for each problem p and solver , they defined

$t_{p,s}$= computing time required to solve problems p by solver s (the number of iteration or CPU time)

We require baseline for comparisons, we compare the performance on problem p by solver s with the best performance by any solver on this problem, that we use the performance ratio

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in S\}}$$

We assume that parameter $r_M \geq r_{p,s}$ for all p, s, and $r_{p,s} = r_M$ if and only if solver s does not solve problem p. The performance of solver s on any given problem might be of interest, but we would like to obtain overall assessment of the performance of the solver. If we define

$$p_s(t) = \frac{1}{n_p} \, size \, \{p \in P : r_{p,s} \leq t\}$$

Then $p_s(t)$ is the probability for solver $s \in S$ that a performance ratio $r_{p,s}$ is within a factor $t \in R$ of the best possible ration. The function p_s is the cumulative distribution for the performance ratio. The performance profile $p_s : R \mapsto [0\,,1]$ for a solver is non-decreasing , piecewise constant function , continuous from the right at each break point , the value of $p_s(1)$ is the probability that the solver will win over the rest of the solvers.

ISSN: 1858-7658 (Online)
ISSN: 1858-7690 (Print)

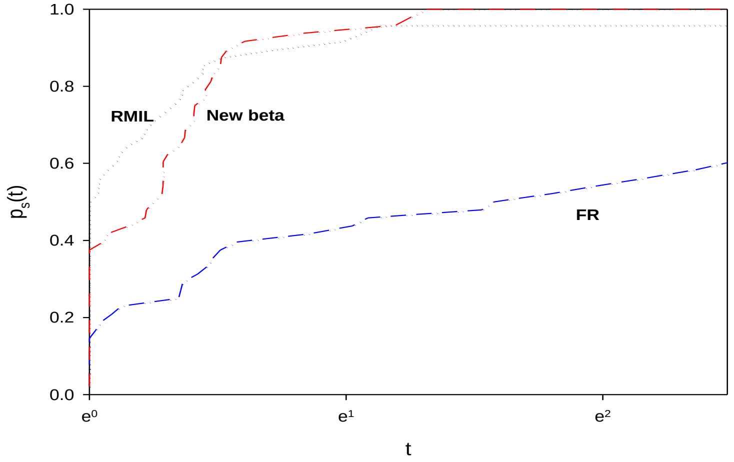

Fig.1.Performance profile based on the number of iteration

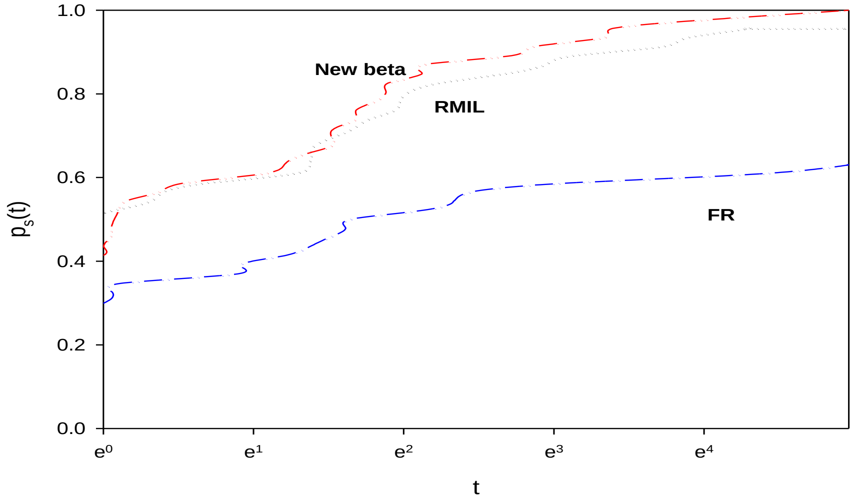

Fig.2.Performance profile based on the CPU time

In general , a solver with high value of p (t) or at the top right of the figure are represent the best solver from Figs. 1 and 2 shows that our new proposed has the best performance since it can solve all of the test problem, FR which solve 63% of test problem, RMIL which solve 95% of test problem.

5. Conclusion

In this paper we proposed a new and simple β_k that has a global convergence property. Numerical results have shown that the new β_k performs better than FR and RMIL. In the future, we intend to test our new formula using inexact line search, scaling and three terms Conjugate gradients.

Acknowledgements

The authors would like to thank the Red Sea University, Sudan.

Reference

1. M. Al-Baali, *Descent property and global convergence of Fletcher-Reeves method with inexact line search*, IMA. J. Numer. Anal. 5 (1985) 121–124.
2. G. Zoutendijk, *Nonlinear programming computational methods*, in:J.Abadie (Ed.), Integer and Nonlinear Programming, North-Holland, Amsterdam, 1970, pp. 37–86.
3. N. N. Andrei, Open problems in nonlinear conjugate gradient algorithms for unconstrained optimization, ICI Technical, Report, 13/08(2008).
4. M. J. Sun, J. Zhang, *Global convergence of conjugate gradient methods without line search*, Ann. Oper. Res. 103 (2001) 161–173.
5. M.. M. R. Hestenes, E.L. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Natl. Bur. Stand. Sec. B 49 1952, pp. 409–432

6. R.F R. Fletcher ,C.Reeves , *Function minimization by conjugate gradient* ,comput .J .**7,** 1964, pp.149-154.
7. B.E B.T. Polyak, The conjugate gradient method in extreme problems, USSR Comp .Math. Phys. 9 (1969) 94–112
8. E. E. Polak, G. Ribie`re, *Note Sur la convergence de directions conjuge`es*, Rev. Francaise Informat Recherche Operationelle, 3e Anne`e 16 (1969) 35–43.
9. R.F R.Fletcher. *Practical Method of Optimization*, second ed., Unconstrained Optimization, vol. I, Wiley,New York, 1987.
10. M. Y. Liu, C. Story, J. Comput. Appl. Math. 69, 129-137 (1992).
11. Y.H. Dai, Y. Yuan, A nonlinear conjugate gradient with a strong global convergence properties ,SIAM J. Optim. 10 (1999) 177–182.
12. M.J.D. Powell, Restart procedures of the conjugate gradient method, *Math. Program*. 2, 1997, pp.241–254.
13. M. Rivaie, M. Mamat, M. Ismail and M. Fauzi. Modified Hestenes-Steifel conjugate gradient coefficient for unconstrained optimization. *Journal of Interdisciplinary Mathematics*, 13(3), 2010, pp. 241-251
14. N. Andrei, *An unconstrained optimization test functions collection*, Adv. Model. Optim. 10 (2008) 147–161.
15. A. Abashar, M. Mamat, M., Rivaie, and M. Ismail. Global Convergence Properties of A new Class of Conjugate Gradient Method for Unconstrained Optimization. *J. Appl .Math.Sciences*. 8(67), 2014, pp. 3307-3319.
16. Z. Wei, S. Yao, L. Liu, The convergence properties of some new conjugate gradient methods, *Appl. Math. Comput*. 183 2006, pp. 1341–1350.
17. K .E. Hilstrom,*A simulation test approach to the evaluation of nonlinear optimization algorithms*, ACM. Trans. Math. Softw. 3 (1977) 305–315.
18. W.W. Hager, H.C. Zhang, *A new conjugate gradient method with guaranteed descent and efficient line search*, SIAM J. Optim. 16 (2005) 170–192.
19. E. Dolan, J.J.More, Benchmarking optimization software with performance profile, Math. Prog. 91 (2002) 201–213.