International Journal of Computer Applications (0975 - 8887)
Volume 177 - No.14, October 2019

On the Numerical Performance of a New Conjugate Gradient Parameter for Solving Unconstrained Optimization Problems

Aliyu Usman Moyi
Department of Mathematical Sciences
Faculty of Science
Federal University Gusau, Zamfara State, Nigeria

Onwuka Blessing
Department of Mathematical Sciences
Faculty of Science
Federal University Gusau, Zamfara State, Nigeria

ABSTRACT

Nonlinear Conjugate gradient methods (CG) are widely used for solving unconstrained optimization problems. Their wide application in many Fields such as Engineering, Applied Sciences and Economics is due to their low memory requirements and global convergence properties. Numerous studies and modifications directed towards improving the efficiency of these methods have been conducted. In this paper, a new conjugate gradient parameter β_k that possess convergence properties is presented. We also present preliminary numerical results to show the efficiency of the proposed method.

Keywords

Unconstrained Optimization, Conjugate Gradient Method, Conjugate Gradient Coefficient, Global Convergence.

1. INTRODUCTION

In this paper, we deal with the conjugate gradient (CG) methods for the numerical solution of the unconstrained optimization problem

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

where $f : R^n \to R$ is assumed to be atleast twice continuously differentiable function and n is the dimension of x, which is assumed to be large. The iterates of the conjugate gradient methods are obtained by

$$x_{k+1} = x_k + \alpha_k d_k, \tag{2}$$

where x_k is the current iterate point, d_k is the search direction and $\alpha_k > 0$ is the step length. The success of any conjugate gradient method depends on the effective choices of both the search direction and the step length. Two strategies for calculating the step length is the exact line search and inexact line search method. The ideal choice would be the exact line search which is defined by

$$f(x_k + \alpha_k d_k) = \min_{\alpha \in R} f(x_k + \alpha_k d_k) \tag{3}$$

but in general it is computationally expensive to obtain since it requires too many evaluations of the objective function f and its gradient g. The other alternative choice is the inexact line search strategies such as the Armijo line search [7], Wolfe condition [] and Goldstein condition [2]. The conjugate gradient methods also define the search direction d_k by

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad d_0 = -g_0 \tag{4}$$

for $k \geq 1$, where the parameter $\beta_k \in R$ is a scalar known as conjugate gradient coefficient.

In the literature, several choices for β_k have been proposed which give rise to distinct conjugate gradient methods. The most well-known conjugate gradient methods are the Hestenes-Steifel (HS)[10], Fletcher-Reeves(FR)[11], Polak-Ribeire (PR) [3], Liu-Storey (LS),[13] Dai-Yuan (DY)[12] and Gilbert-Nocedal (PR+)[9] in which the update parameter of these methods is respectively specified as follows:

$\beta_k^{HS} = \frac{y_k^T g_{k+1}}{y_k^T d_k}$, $\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$, $\beta_k^{PR} = \frac{y_k^T g_{k+1}}{g_k^T g_k}$,
$\beta_k^{LS} = \frac{-y_k^T g_{k+1}}{g_k^T d_k}$, $\beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{y_k^T d_k}$, $\beta_k^{PR+} = \max\{\frac{y_k^T g_{k+1}}{g_k^T g_k}, 0\}$,

where $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$. Note that these formulae for β_k are equivalent to each other if the objective function is a strictly convex quadratic function and α_k is chosen through an exact line search. However, for general non-quadratic functions or under the inexact line search their behavior is quite different [6].

The rest of this paper is organized as follows. In the next section, we simply recall the conjugate gradient direction for solving unconstrained optimization problem and construct our new conjugate parameter subsequently. In section 3, we present the convergence of the proposed algorithm. Then we present some preliminary results of the new approach on some standard test problems and finally, we conclude the paper in section 5.

2. THE NEW CG PARAMETER

We proposed our modification based on the search direction from Ibrahim et.al [5] given as

$$d_{k+1} = -B_{k+1}^{-1} g_{k+1} + \lambda_{k+1} d_k, \tag{5}$$

where B_{K+1} is the BFGS updating matrix and $\lambda_{k+1} = \eta g_{k+1}^T g_{k+1}/g_{k+1}^T d_k$ with $\eta \in (0.1]$ is chosen to ensure conjugacy. from (4) and (5) we have,

$$-g_{k+1} + \beta_{k+1} d_k = -B_{k+1}^{-1} g_{k+1} + \lambda_{k+1} d_k, \tag{6}$$

multiply (6) by $s_{k+1}^T B_{k+1}$

$$-s_{k+1}^T B_{k+1} g_{k+1} + \beta_{k+1} d_k s_{k+1}^T B_{k+1} = -g_{k+1}^T s_{k+1} + \lambda_{k+1} d_k s_{k+1}^T B_{k+1} \tag{7}$$

But, $B_{k+1} s_{k+1} = y_{k+1}$, then,

$$-g_{k+1}^T y_{k+1} + \beta_{k+1} d_k^T y_{k+1} = -g_{k+1}^T s_{k+1} + \lambda_{k+1} d_k^T y_{k+1} \tag{8}$$

$$\beta_{k+1} d_k^T y_{k+1} = -g_{k+1}^T s_{k+1} + \lambda_{k+1} d_k^T y_{k+1} + g_{k+1}^T y_{k+1} \tag{9}$$

$$\beta_{k+1} = \frac{-g_{k+1}^T s_{k+1} + \lambda_{k+1} d_k^T y_{k+1} + g_{k+1}^T y_{k+1}}{d_k^T y_{k+1}} \tag{10}$$

Thus,

$$\beta_{k+1} = \frac{(\lambda_{k+1} d_k + g_{k+1}) y_{k+1}^T - g_{k+1}^T s_{k+1}}{d_k^T y_{k+1}} \tag{11}$$

Similarly, $d_k = s_k$

$$\beta_{k+1}^{NEW} = \frac{(\lambda_{k+1} s_k + g_{k+1}) y_{k+1}^T - g_{k+1}^T s_{k+1}}{s_k^T y_{k+1}}. \tag{12}$$

New Algorithm
We now present the basic steps of the algorithm for solving unconstrained optimization problems as;
Step 1 : Given an initial point $x_0 \in R^n$, set $d_0 = -g_0$ and $k = 0$
Step 2 : Test a criterion for stopping the iterations. If the test is satisfied, then stop; otherwise continue with step 3
Step 3 : Compute the search direction d_k by (4), with β_{k+1} defined in (12)
Step 4 : Find an acceptable steplength α_k, by using the following line search procedure. Given the constants $\eta \in (0, 1)$ and τ, τ' with $0 < \tau < \tau' < 1$

(i) Set $\alpha = 1$
(ii) Test the relation

$$f(x_k + \alpha d_k) \leq f(x_k) + \eta \alpha g_k^T d_k, \tag{13}$$

(iii) If (13) is not satisfied, choose a new α in $[\tau\alpha, \tau'\alpha]$ and go to (ii). If (13) is satisfied, set $\alpha_k = \alpha$ and $x_{k+1} = x_k + \alpha_k d_k$

Step 5 : Set $k := k + 1$, and go to step 2

3. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we present some numerical results from an implementation of our new conjugate gradient algorithm for solving unconstrained optimization problems, we evaluate the performance of oue new conjugate gradient parameter with that of

(1) Polak-Ribeire (PR)[3] ,
(2) Hestenes-Steifel (HS)[10] HS:

All the experiments are implemented on a PC using MATLAB version 7.13.0.564 (R 2011b), with double precision arithmetic. For each test function, we perform 40 numerical experiments with variable dimensions $50 \leq n \leq 1000$. As regards the stopping criteria used in our experiments, in all the algorithms, convergence is assumed if $\|g_k\| \leq \varepsilon$ where $\varepsilon = 10^{-4}$. We forced the algorithm to stop whenever the number of iterations exceeds 2000, and the symbol "-" is used to represent the failure. Test functions are the standard unconstrained optimization problems obtained from [1] as presentesd in the Table below.
Table 1 gives the performance of all the algorithms, where a total of 20 runs are performed. Based on the results, β_k^{New} solves 100% of the test problems while PR, HS, can only solve 80% and 90% of the test problems respectively. The performance of β_k^{New} over PR, is that New β_k^{New} needs 48% and 78% less in terms of the number of iterations than PR and HS. Overall, we believe that given the ratio of iteration counts for all the methods, we can conclude that our algorithm is promising even when less accurate line search strategy is employed.

4. CONCLUSION

Numerous studies on CG method have resulted in different conjugate gradient parameters. Although, these variety of methods have been shown to perform much better than the classical methods they have been reported to be complex and difficult to implement in practice. Therefore, identifying these shortcomings and rectifying them as modified methods for efficient performance is worthwhile. Thus, in this paper we have presented a new conjugate gradient parameter for solving unconstrained optimization problems. The numerical results for a small dimension of the test problems show that the new parameter is efficient and robust.

5. REFERENCES

[1] Andrei, N.: An unconstrained optimization test functions collection, Adv. Model. Optim. 10(1), 147–161 (2008)

[2] Goldstein,A. A.On Steepest Descent SIAM J. Control. 3(1965) 147-151.

[3] Polak, E.,Ribiere, Note sur la convergence de directions conjugees,Rev. Francaise Inform. Recherche Opetationelle 3 (1969) 36-43.

[4] Byrd, R. H., Nocedal, J.: A tool for the analysis of quasi-Newton methods with pplication to unconstrained minimization. SIAM J. Numer. Anal. 26(3), 727–739 (1989)

[5] Ibrahim, M. A , Mamat M. and Leong W, J.":BFGS Method: A New Search Direction. Sains Malaysiana 42(10) 2014, 1593-1599

[6] Dai, Y. H.: On the nonmonotone line search. J. Optim. Theory Appl. 112(2), 315–330 (2002)

[7] Armijo, L. Minimization of Functions having Lipschitz Continuous Partial Derivatives, Pacific J. Math 16(1966)1-3

[8] Dolan, E.D., More, J.J.: Benchmarking optimization software with performance profiles. Math. program. 91 (2), 201–213 (2002)

[9] Gilbert, J.C., Nocedal, J.: Global convergence properties of conjugate gradient methods for optimization. SIAM J. Optim. 2 (1), 21–42 (1992)

[10] Hestenes. M.R., Steifel, E. Method of Conjugate Gradient for Solving Linear Equations, J. Res. Nat. Bur. Stand.49 (1952) 409-436.

Table 1. Numerical Results of PR, HS, and New β_k

Test Fuction	Dim.	PR	HS	β_{k+1}^{New}
		NI	NI	NI
Extended	50	–	43	33
Freudeinstein	100	–	43	33
and Roth	500	–	47	34
	1000	–	47	36
Extended	50	115	47	39
Beal	100	117	47	39
	500	121	50	39
	1000	123	49	42
Extended	50	40	76	20
Block	100	35	69	21
Diagonal 1	500	38	70	19
	1000	40	70	19
Raydan 2	50	–	276	18
	100	–	352	18
	500	–	606	20
	1000	–	776	25
Generalized	50	31	104	20
Tridiagonal 1	100	33	105	21
	500	35	105	22
	1000	35	105	22
Diagonal 4	50	106	53	24
	100	129	55	24
	500	113	57	24
	1000	127	59	25
Extended	50	29	186	17
Himelblau	100	29	192	18
	500	31	205	19
	1000	31	211	19
Extended	50	33	24	10
Psc	100	33	24	10
	500	35	24	10
	1000	37	24	11
Extended	50	51	–	35
TCliff	100	51	–	35
	500	100	–	35
	1000	75	–	35
Extended	50	31	68	20
Three Exponential	100	33	70	21
Terms	500	35	76	22
	1000	35	80	22

[11] Fletcher, R. Reeves, C. Function Minimization by Conjugare Gradients, Comput. J. 7 (1964) 149-154.

[12] Dia, Y. H., Yuan, Y. Nonlinear Conjugate Gradient Method, Shanghai Scientific and Technical Publishers, Beijing , 1998

[13] Liu, Y., Storey, C. Efficient Generalized Conjugate Gradient Algorithms Part 1:Theory,J.Comput.Appl.Math. 69(1992) 129-137.

[14] Panier, E.R., Tits, A.L.: Avoiding the maratos effect by means of a nonmonotone line search I. General constrained problems. SIAM J. Numer. Anal. 28(4), 1183–1195 (1991)

[15] Perry, J.M.: A class of conjugate gradient algorithms with a two step variable metric memory. Center for Mathematical Studies in Economies and Management Science. Evanston Illiois: Northwestern University Press. 1977

[16] Sun, W., Han, J., Sun, J.: Global convergence of nonmonotone descent methods for unconstrained optimization problems. J. Comput. Appl. Math. 146(1), 89–98 (2002)

[17] Shi, Z.J., Wang, S., Xu, Z.: The convergence of conjugate gradient method with nonmonotone line search. Appl. Math. Comput. 217(5), 1921–1932 (2010)

[18] Yuan, G., Wei, Z.: Non monotone backtracking inexact BFGS method for regression analysis. Comm. Stat. Theory Meth. 42(2), 214–238 (2013)