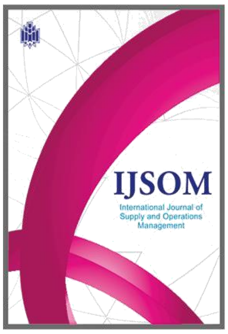
International Journal of Supply and Operations Management

IJSOM

November 2020, Volume 7, Issue 4, pp. 344-349
ISSN-Print: 2383-1359
ISSN-Online: 2383-2525
www.ijsom.com



A New Hestenes-Stiefel and Fletcher-Reeves Conjugate Gradient Method with Descent Properties for Optimization Models

Saleh Nazzal Alsuliman[a], Sulaiman Ibrahim Mohammed [a,*], Mustafa Mamat[a], Deiby T Salaki[b] and Nelson Nainggolan[b]

[a]*Faculty of Informatics and Computing, Sultan Zainal Abidin University, Kuala Terengganu, Malaysia*
[b]*Department of Mathematics, Faculty of Mathematics and Natural Science, Sam Ratulangi University, Indonesia*

Abstract

The conjugate gradient (CG) scheme is regarded as among the efficient methods for large-scale optimization problems. Several versions of CG methods have been presented recently owing to their rapid convergence, simplicity, and their less memory requirements. In this article, we construct a new CG algorithm via the combination of the classical methods of Fletcher-Reeves (FR), and Hestenes-Stiefel (HS). The new CG method possesses the descent properties and converge globally provided the exact minimization condition is satisfied. The tests of the new CG method using MATLAB are analysed in terms of iteration number and CPU time. Numerical results have been reported which shows that the proposed CG method performs better compare to other CG methods.

Keywords: Conjugate gradient parameter; Line search procedure; Unconstrained optimization.

1. Introduction

The CG techniques is known as one of the most efficient optimization procedures for solving applications problems in the field of medicine, science, engineering, and many more. The method also plays a significant role for unconstrained optimization problems (UOP) (Umar et al., 2020). The general UOP is stated as follows,

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

where R^n refers to the n-dimensional Euclidean space, $f: R^n \to R$ is smooth, $x \in R^n$ is a vector and $f(x)$ is an objective function (Sulaiman et al., 2020a). The efficiency of the any CG method is the less memory storage and the ability to obtain the solution of the problem defined in (1) (Yuan & Sun 1999; Hamoda et al., 2015). The CG methods are computed using iterative procedures

$$x_{k+1} = x_k + \alpha_k d_k, \quad k \geq 0, \tag{2}$$

with the step length $\alpha_k > 0$ and d_k is the direction of search (Sulaiman et al., 2020). The step-length can be obtained using either the exact or inexact line search procedure (Sulaiman et al., 2015a; 2015b). Recently, many researchers tend to employ the inexact procedure due to it rapid convergence (Mamat et al., 2020). However, this process only produces an approximate solution rather the real solution. Thus, in this paper, the exact minimization procedure is selected for computing the step-length. Basically d_k is obtained using:

* Corresponding author email address: sulaimanib@unisza.edu.my

$$d_k = \begin{cases} -g_k, & \text{if} \quad k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if} \quad k \geq 1, \end{cases} \tag{3}$$

β_k and g_k is the CG parameter of $f(x)$ and the gradient at x_k respectively. We have $\beta_k \in R$ is a scalar that differentiate various CG methods while $g_k = \nabla f(x_k)$ at the point x_k. Some well-known CG formulas are:

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \tag{4}$$

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \tag{5}$$

$$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}} \tag{6}$$

$$\beta_k^{RMIL} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T d_{k-1}} \tag{7}$$

$$\beta_k^{CD} = \frac{g_k^T g_k}{d_{k-1}^T g_{k-1}} \tag{8}$$

$$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}} \tag{9}$$

$$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}} \tag{10}$$

The conjugate gradient coefficients $\beta_k \in R$ are scalars, which determine different CG methods. Some known formulas for CG coefficients HS (Hestenes & Stiefel 1952), FR (Fletcher & Reeves 1964), LS (Liu-Storey 1992), RMIL (Rivaie et al., 2012), CD (Fletcher 1987), PR (Polak-Ribiere 1969) and the lastly DY (Dai-Yuan 1999). In this paper, we run the results on the convergence analysis using exact minimization procedures. An important property of convergence is choosing a suitable step-length α_k (Sulaiman et al., 2019). The most commonly used search is done under exact line search:

$$f(x_k + \alpha_\kappa d_k) = \min_{\alpha \geq 0} f(x_k + \alpha\, d_k) \tag{11}$$

In this study, we developed a simple CG parameter β_k. In Section 2, there is the algorithm with our new CG parameter. The descent and convergence of the proposed coefficient under exact line search technique is established in Section 3. Section 4, contains the numerical results, the selected benchmark functions and the discussion. Finally, our conclusion in Section 5.

2. New CG Coefficient

This section presents the proposed β_k^{SM} based on combination of FR and HS methods where SM denotes Saleh, Sulaiman, and Mamat as below:

$$\beta_k^{SSM} = \frac{\beta_k^{HS} + \beta_k^{FR}}{2} \tag{12}$$

which can be rewritten

$$\beta_k^{SSM} = \frac{\dfrac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} + \dfrac{g_k^T g_k}{g_{k-1}^T g_{k-1}}}{2}$$

The Algorithm of the proposed SSM coefficient is as follows.

Algorithm 1.
Stage 1: Given x_0, fixe $k=0$. If $\| g_k \|=0$, stop.
Stage 2: Determine β_k by (13).
Stage 3: Compute d_k by (3)
Stage 4: Determine α_k by (11).
Stage 5: Update x_{k+1} by (2)
Stage 6: Check If $\|g_k\| \leq \varepsilon$, stop. Else, go to Stage 2 and set $k=k+1$.

Convergence Analysis
This section discussed the convergence of β_k^{SM}. The convergence of FR parameter has been established under various line search (Dai et al., Dai & Yuan, 1999). To prove the convergence of the proposed method, we assumed that d_k should possess the following condition

$$g_k^T d_k < 0 \quad (13)$$

for all $k \geq 0$. If $\exists\, c > 0$, then, d_k would satisfy the following condition known as the sufficient descent condition

$$g_k^T d_k \leq -C\|g_k\|^2 \quad (14)$$

Theorem 1:

For any CG method (2) and (3) with CG coefficient β_k^{SSM} given as (3) and (12) respectively, then (13) holds for all $k \geq 0$.

Proof:

The proof of this Theorem would be by induction. That is, if $k=0$, it is obvious $g_0^T d_0 = -c \| g_0 \|^2$. Thus, (13) is true. Now, we want to prove (13) holds for $k \geq 1$.

Multiply (3) by g_{k+1}^T,

$$g_{k+1}^T d_{k+1} = g_{k+1}^T(-g_{k+1} + \beta_{k+1}^{SM} d_k)$$

$$= -\| g_{k+1} \|^2 + \beta_{k+1}^{SSM} g_{k+1}^T d_k$$

For exact line search, $g_{k+1}^T d_k = 0$. Then,

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

Hence, (14) holds true for $k+1$. Alternatively, we would show that HS method can reduce to FR under (11) as the following proof:

$$\beta_k^{SSM} = \frac{\beta_k^{HS} + \beta_k^{FR}}{2}$$

From the HS method,

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

It is known that by using (11), $g_k^T d_{k-1} = 0$, hence

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

From (13), $g_k^T d_k \leq -c\|g_k\|^2$ where $c > 0$ is a constant, therefore,

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{-d_{k-1}^T g_{k-1}} \leq \frac{g_k^T(g_k - g_{k-1})}{-(-c\|g_k\|^2)} \leq \frac{g_k^T(g_k - g_{k-1})}{c\|g_k\|^2}$$

That is mean,

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T g_k - d_{k-1}^T g_{k-1})} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{d_{k-1}^T g_k - d_{k-1}^T g_{k-1}}$$

Note that $g_k^T d_{k-1} = 0$ and $g_{k-1} = -d_{k-1}$ thus,

$$\beta_k^{HS} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \leq \beta_k^{FR}$$

Then,

$$\beta_k^{SSM} \approx \frac{\beta_k^{FR} + \beta_k^{FR}}{2} \approx \beta_k^{FR}$$

Therefore, the convergence properties of β_k^{SSM} will follow β_k^{FR}. This complete the proof.

Numerical Experiment

This section compares the efficiency of the new algorithm SSM with the methods of HS, CD, FR and RMIL based on iteration number and CPU time. All the test functions considered are taken from Andrei (2008) and Molga (2005). The termination condition was set as $\|g_k\| \leq 10^{-6}$. MATLAB R2018a was used in the computation which was run on an Intel Core i3 with RAM 3GB operation system. The list of test problems, the starting points, and their dimension are presented in Table 1 below. The researcher adopted four initial points with four dimensions for the computations of each test functions used ranging from points close to the solution points to points far away (Hilstrom 1977).

Table 1. List of test functions

N	Functions	Dimensions	Initial Points
1	Extended White & Holst	2, 4, 10, 100	(-3,...,-3), (3,…,3), (-12,…,-12), (12,…,12)
2	Dixon and Price	4, 8, 20, 60	(3,…,3), (6,…,6), (10,…,10), (13,…,13)
3	FLETCHCR	4, 10, 50, 100	(5,…,5), (10,…,10), (20,…,20), (30,…,30)
4	Generalized Quartic	4, 10, 50, 100	(5,…,5), (10,…,10), (15,…,15), (20,…,20)
5	Generalized Tridiagonal 1	4, 8, 10, 50	(5,…,5), (-5,…,-5), (13,…,13), (-10,…,-10)
6	Generalized Tridiagonal 2	2, 8, 10, 20	(5,…,5), (10,…,10), (15,…,15), (20,…,20)
7	Extended Block Diagonal BD1	2, 10, 100, 1000	(-2,…,-2), (2,…,2), (5,…,5), (7,…,7)
8	Raydan 1	4, 10, 50, 100	(-9,…,-9), (-6,…,-6), (6,…,6), (9,…,9)
9	Power	4, 10, 20, 80	(5 …, 5), (10…,10), (15,…,15), (20,…,20)
10	Sum Squares	2, 4, 10, 100	(3,…,3), (6,…,6), (9,…,9), (12,…,12)
11	Quadratic QF2	10, 20, 50, 100	(5,…,5), (10,…,10), (15,…,15), (20,…,20)
12	Extended Trigonometric	4, 10, 50, 100	(-16,…,-16), (-4,…,-4), (4,…,4), (16,…,16)
13	Extended Beale	2, 4, 10, 100	(3,…,3), (5,…,5), (7,…,7), (9,…,9)

Under exact line search, Fig. 1 gives the iteration number graph and Fig. 2 is the graph to show the performance profile in terms of CPU time for SM, HS, CD, FR and RMIL methods. This is based on performance profile introduced by [13]. Obviously, the method of SM has the fastest followed by methods of HS, RMIL, CD, and the least of them is the FR method. This proposed method can be applied to an application problem. See Sulaiman et al., (2020b); Kazeem & Mohammed (2019); Hamid & Fahad (2019); Ali et al., (2018).

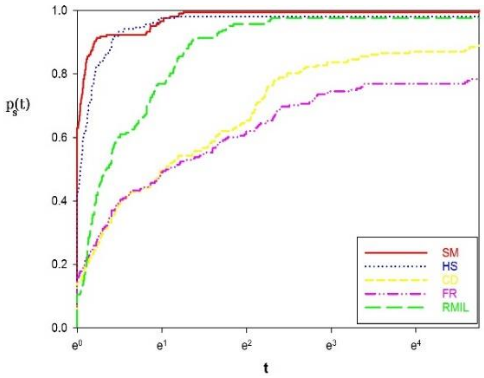


Figure 1. Performance based on number of iterations

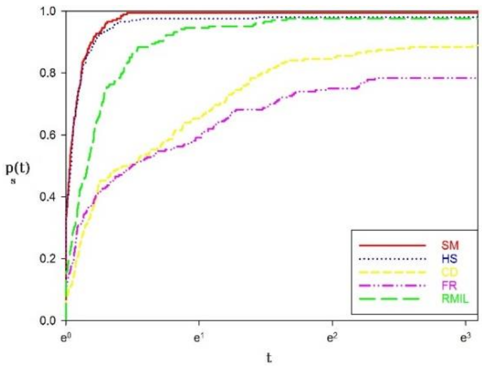


Figure 2. Performance based on the CPU times

Conclusion

In this article, the authors presented a new CG coefficient based on the combination of the methods of FR and HS for obtaining the solution of optimization problems with emphasis on unconstrained functions using exact minimization procedures. Based on the result, our parameter satisfied the sufficiently descent and global convergence properties. The results show that SM method gives the fastest performance in terms of CPU time and solves problems with minimum number of iterations. SM method gives the best performance compared to HS, CD, FR and RMIL. Thus, SM is a good alternative for to other existing methods.

Acknowledgments

The authors are grateful to Malaysian government for supporting this research and Universiti Sultan Zainal Abidin.

References

Ali, S., Mohammad, M., and Hadi, M. (2018). A Multi-objective Competitive Location Problem within Queuing Framework. *Int J Supply Oper Manage (IJSOM*), Vol.5 (1), pp. 42-65.

Andrei N. (2008). An unconstrained optimization test functions collection, *Adv. Model. Optim*, Vol. 10(1), pp. 147–161.

Dai Y. H and Yuan Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property, *SIAM J. Optim.*, Vol. 10(1), pp. 177–182.

Dolan E. D and Moré J. J. (2002). Benchmarking optimization software with performance profiles, *Math. Program.*, Vol. 91(2), pp. 201–213.

Fletcher R. (1980). *Practical Methods Of Optimization:* Vol. 1 Unconstrained Optimization. John Wiley & Sons.

Fletcher R and Reeves C M. (1964). Function minimization by conjugate gradients, *Comput. J.*, Vol. 7(2), pp. 149–154.

Hamid, S. and Farhad, H. (2019). Optimizing Total Delay and Average Queue Length Based on on Fuzzy Logic Controller in Urban Intersections. *International Journal of Supply and Operations Management,* Vol. 6(2), pp. 142-158.

Hamoda M, Rivaie M, Mamat M and Salleh Z. (2015). A new nonlinear conjugate gradient coefficient for unconstrained optimization, *Appl. Math. Sci.*, Vol. 9(37), pp. 1813–1822.

Hestenes M. R and Stiefel E. (1952). *Methods of conjugate gradients for solving linear systems*, NBS Washington, DC. Vol.49, (1), pp. 409-436.

Hillstrom K. E. (1977). Simulation test approach to the evaluation of nonlinear optimization algorithms, *ACM Trans. Math. Software ;(United States)*, Vol. 3(4), pp. 305-315.

Kazem, F. and Mohammad, R. G. (2019). Optimizing the Safety Stock with Guaranteed Service Model in Reverse Logistics Considering Internal and External Returns. *International Journal of Supply and Operations Management*, Vol. 6 (3), pp. 188-199.

Liu Y and Storey C. (1991). Efficient generalized conjugate gradient algorithms, part 1: theory, *J. Optim. Theory Appl.*, Vol. 69(1), pp. 129–137.

Mamat, M., Sulaiman, I.M., Maulana, M., Sukono, Zakaria, Z.A. (2020). An efficient spectral conjugate gradient parameter with descent condition for unconstrained optimization. *Journal of Advanced Research in Dynamical and Control Systems.* Vol. 12(2), pp. 12(02), pp. 2487-2493.

Molga M. and Smutnicki C. (2005). Test functions for optimization needs, *Test Funct. Optim. needs*, Vol. 101, pp. 1-43.

Polak E and Ribiere G. (1969). Note sur la convergence de méthodes de directions conjuguées, *ESAIM Math. Model. Numer. Anal. Mathématique Anal. Numérique*, Vol. 3(R1), pp. 35–43.

Rivaie M, Mamat M, June L. W and Mohd I. (2012). A new class of nonlinear conjugate gradient coefficients with global convergence properties, *Appl. Math. Comput.*, Vol. 218(22), pp. 11323–11332.

Sulaiman, I. M., Sukono, Sudradjat, S., and Mamat, M. (2019). New class of hybrid conjugate gradient coefficients with guaranteed descent and efficient line search. *IOP Conf. Ser.: Mater. Sci. Eng*. 621(012021).

Sulaiman I. M, Mamat M, Abashar A, Zabidin S. (2015a). A Modified Nonlinear Conjugate Gradient Method for Unconstrained Optimization, *Applied Mathematical Science* Vol. 9(54) pp. 2671–2682.

Sulaiman I. M, Mamat M, Abashar A, and Rivaie M. (2015b). The global convergence properties of an improved conjugate gradient method *Applied Mathematical Science*, Vol 9(38), pp. 1857-1868.

Sulaiman, I.M., Mamat, M., Owoyemi, A.E., Olowo, S.E., Zamri, N. (2020a). A class of spectral conjugate gradient method with descent condition for unconstrained optimization. *Journal of Advanced Research in Dynamical and Control Systems*, Vol. 12(02), pp. 2480-2486.

Sulaiman, I.M., Usman, A. Y., and Mamat, M. (2020b). Application of Spectral Conjugate Gradient Methods for Solving Unconstrained Optimization Problems. *An International Journal of Optimization and Control: Theories & Applications*, Vol.10 (2), pp. 198-205.

Umar, A.O., Mamat, M., Sulaiman, I.M., Waziri, M.Y., and Mandara, A.V. (2020). A new modification of conjugate gradient parameter with efficient line search for nonconvex function. *International Journal of Scientific and Technology Research.* Vol. 9(03), pp. 6572- 6575.

Yuan Y and Sun W. (1999). *Theory and methods of optimization.* Science Press of China, Beijing.

Zainal Abidin N, Mamat M, Dangerfield B, Zulkepli JH, Baten MA, Wibowo A (2014). Combating Obesity through Healthy Eating Behavior: A Call for System Dynamics Optimization. *PLoS ONE* Vol. 9(12): e114135. https://doi.org/10.1371/journal.pone.0114135.