Far East Journal of Mathematical Sciences (FJMS)

http://www.pphmj.com
http://dx.doi.org/10.17654/MS101030583
Volume 101, Number 3, 2017, Pages 583-593 **ISSN: 0972-0871**

A NEW SEARCH DIRECTION OF BROYDEN-CG METHOD

Mohd Asrul Hery Ibrahim[1], Zahratul Amani Zakaria[2], Mustafa Mamat[2], Ummie Khalthum Mohd Yusof[2] and Azfi Zaidi Mohammad Sofi[1]

[1]Faculty of Entrepreneurship and Business
Universiti Malaysia Kelantan
Kampus Kota
Malaysia

[2]Faculty of Informatics and Computing
Universiti Sultan Zainal Abidin
Kampus Tembila
Malaysia

Abstract

The Broyden family method is one of the well-known methods in quasi-Newton algorithm for solving unconstrained optimization problems. In this article, a new search direction for Broyden family method is proposed. The new search direction is developed by

Received: August 3, 2016; Accepted: September 24, 2016
2010 Mathematics Subject Classification: 13P25, 46N10, 45L05.
Keywords and phrases: Broyden family method, conjugate gradient method, search direction, global convergence.

hybridizing the two search directions in line search method known as quasi-Newton and conjugate gradient method under certain parameter. This method is popular as the Broyden-CG method. The suggested method has an attractive property whose search direction is sufficiently descent at every iteration. Under mild conditions, we prove that the proposed method has global convergence.

1. Introduction

Quasi-Newton methods are well-known methods in solving unconstrained optimization problems. These methods, which use the updating formulas for approximation of the Hessian, were introduced by Davidon in 1959, and later popularised by Fletcher and Powell in 1963 to give the Davidon-Fletcher-Powell (DFP) method. But the DFP method is rarely used nowadays. On the other hand, in 1970, Broyden, Fletcher, Goldfarb and Shanno developed the idea of a new updating formula, known as BFGS, which has become widely used. Recently, these are subject of several modifications.

During the last few decades, the convergence of the quasi-Newton methods received much attention. Powell [13] proved the global convergence of the BFGS method with the practical Wolfe line search in the case when the f is convex. His results have been extended to the restricted Broyden's family, except the DFP method by Byrd et al. [3]. The question that whether the BFGS method and other quasi-Newton methods are globally convergent for general objective functions remained open for several decades. Quite recently, Dai [4] and Mascarenhas [10] gave a negative answer by providing counter-examples independently.

Realizing the possible non-convergence for general objective functions, some authors have considered modifying quasi-Newton methods to enhance the convergence. For example, Li and Fukushima [6] modified the BFGS method by skipping the update when certain conditions are not satisfied and proved the global convergence of the resulted BFGS method with a "cautious update" (which is called the *CBFGS method*). However, their numerical tests show that the CBFGS method does not perform better than the ordinary

BFGS method. Then Mamat et al. [9] proposed a new search direction for quasi-Newton methods in solving unconstrained optimization problems. Generally, the search direction focused on the hybridization of quasi-Newton methods with the steepest descent method. The search direction proposed by Mamat et al. [9] is $d_i = -\eta B_i^{-1} g_i - \delta g_i$, where $\eta > 0$ and $\delta > 0$. They realized that the hybrid method is more effective compared with the ordinary BFGS in terms of computational cost. Hence, the delicate relationships between the conjugate gradient method and the BFGS method have been explored in the past. Two competing algorithms of this type are the L-BFGS method described by Nocedal [12] and the variable storage conjugate gradient (VSCG) method published by Buckley and Lenir [2].

2. Iteration Method

Consider the unconstrained optimization problem:

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

where $f : R^n \to R$ is continuously differentiable. The Broyden's method is an iterative method. On the *i*th iteration, for an approximation point x_i, the $(i+1)$th iteration of *x* is given by

$$x_{i+1} = x_i + \alpha_i d_i, \tag{2}$$

where the search direction d_i is calculated by

$$d_i = -B_i^{-1} g_i(x_i), \tag{3}$$

where g_i is a gradient of *f*. The search direction must satisfy the relation $g_i^T d_i < 0$, which guarantees that d_i is a descent direction of $f(x)$ at x_i. Then the step size, α_i in (2) is obtained by using the Wolfe line search (Wolfe [14, 15]) which is $\alpha_i > 0$ satisfying

$$f(x_i) - f(x_i + \alpha_i d_i) \geq -\delta \alpha_i g_i^T d_i$$

and

$$\sigma_1 g_i^T d_i \le (g(x_i + \alpha_i d_i))^T d_i \le -\sigma_2 g_i^T d_i, \tag{4}$$

where $0 < \delta < \sigma_i < 1$, $0 \le \sigma_2 < +\infty$ are constants. Then the sequence $\{x_i\}_{i=0}^{\infty}$ converges to the optimal point x^*, which minimizes. The updated Hessian approximation formula in (3) requires B_i to be positive definite satisfying the quasi-Newton equation

$$B_{i+1} s_i = y_i, \tag{5}$$

where

$$s_i = \alpha_i d_i,$$

$$y_i = g_{i+1} - g_i. \tag{6}$$

The Broyden's algorithm for unconstrained optimization problem uses the matrices B_i updated by the formula

$$B_{i+1} = B_i - \left(\frac{B_i s_i s_i^T B_i}{s_i^T B_i y_i} \right) + \frac{y_i y_i^T}{s_i^T y_i} + \phi_i (s_i^T B_i s_i) v_i v_i^T, \tag{7}$$

where ϕ is a scalar and

$$v_i = \left[\frac{y_i}{s_i^T} - \frac{B_i s_i}{s_i^T B_i s_i} \right].$$

This algorithm satisfies the quasi-Newton equation (7). The choice of the parameter ϕ is important, since it can greatly affect the perfomance of the method. When $\phi_i = 1$ in equation (7), we obtain the DFP algorithm and when $\phi_i = 0$, we get the BFGS algorithm. But, Byrd et al. [3] extended his result to $\phi \in (0, 1]$. Based on Nocedal [12], the Broyden's algorithm is one of the most efficient algorithms for solving the unconstrained optimization problem. Thus the algorithm for an iteration method of ordinary Broyden is described as follows:

Algorithm 1 (Broyden method)

Step 0. Given a starting point x_0 and $B_0 = I_n$. Choose values for *s*, β and σ.

Step 1. Terminate if $\| g(x_{i+1}) \| < 10^{-6}$.

Step 2. Calculate the search direction by (8).

Step 3. Calculate the step size α_i by the Armijo line search (7).

Step 4. Compute the difference $s_i = x_{i+1} - x_i$ and $y_i = g_{i+1} - g_i$.

Step 5. Update B_i by (3) to obtain B_{i+1}.

Step 6. Set $i = i + 1$ and go to Step 1.

This paper is organized as follows. In Section 3, we elaborate the new search direction of the Broyden family method. Then the numerical results are obtained in Section 4. The paper ends with a short conclusion in Section 5.

3. The Broyden-CG Algorithm

The modification of the quasi-Newton method based on a hybrid method has already been introduced by previous researchers. One of the studies is a hybridization of the quasi-Newton and Gauss-Siedel methods, aimed at solving the system of linear equations in Ludwig [7]. Luo et al. [8] suggested the new hybrid method, which can solve the system of nonlinear equations by combining the quasi-Newton method with chaos optimization.

Hence, the modification of the quasi-Newton method by previous researchers spawned the new idea of hybridizing the classical method to yield the new hybrid method. Hence, this study proposes a new hybrid search direction that combines the concept of search direction of the quasi-Newton and CG methods. It yields a new search direction of the hybrid method which is known as the Broyden-CG method. The search direction for the Broyden-CG method is

$$d_i = \begin{cases} -B_i^{-1} g_i, & i = 0, \\ -B_i^{-1} g_i + \eta(-g_i + \beta_i d_{i-1}), & i \geq 1, \end{cases} \tag{8}$$

where $\eta > 0$ and $\beta_i = \dfrac{g_i^T (g_i - g_{i-1})}{g_i^T d_{i-1}}$.

Hence, the complete algorithm for the Broyden-CG method is arranged as follows:

Algorithm 2 (Broyden-CG method)

Step 0. Given a starting point x_0 and $H_0 = I_n$. Choose values for s, β and σ and set $i = 1$.

Step 1. Terminate if $\| g(x_{i+1}) \| < 10^{-6}$ or $i \geq 10000$.

Step 2. Calculate the search direction by (8).

Step 3. Calculate the step size α_i by (4).

Step 4. Compute the difference between $s_i = x_i - x_{i-1}$ and $y_i = g_i - g_{i-1}$.

Step 5. Update H_{i-1} by (7) to obtain H_i.

Step 6. Set $i = i + 1$ and go to Step 1.

Based on Algorithms 1 and 2, we assume that every search direction d_i satisfied the descent condition

$$g_i^T d_i < 0$$

for all $i \geq 0$. If there exists a constant $c_1 > 0$ such that

$$g_i^T d_i \leq c_1 \| g_i \|^2 \tag{9}$$

for all $i \geq 0$, then the search directions satisfy the sufficient descent condition which is proved in Theorem 3.1.

Theorem 3.1. *Consider a hybrid method with search direction* (8), *then condition* (9) *holds for all* $i \geq 0$.

Proof. From (9), we see that

$$\begin{aligned} g_i^T d_i &= -g_i^T B_i^{-1} g_i + \eta g_i^T (-g_i + ((g_i - g_{i-1})^T g_i / g_i^T d_{i-1}) d_{i-1}) \\ &= -g_i^T B_i^{-1} g_i + \eta(-g_i^T g_i + ((g_i - g_{i-1})^T g_i / g_i^T d_{i-1}) g_i^T d_{i-1}) \\ &= -g_i^T B_i^{-1} g_i + \eta(-g_i^T + g_{i-1}). \end{aligned}$$

Based on Powell [13], $g_i^T g_{i-1} \geq \varepsilon \| g_i \|^2$ with ε lying in $(0, 1]$,

$$\begin{aligned} g_i^T d_i &= -g_i^T B_i^{-1} g_i + \eta(\varepsilon \| g_i \|^2) \\ &\leq -\lambda_i \| g_i \|^2 - \eta\varepsilon \| g_i \|^2 \\ &\leq c_1 \| g_i \|^2, \end{aligned}$$

where $c_1 = -(\lambda_i + \eta\varepsilon)$ which is bounded away from zero. Hence, $g_i^T d_i \leq c_1 \| g_i \|^2$ holds. The proof is thus complete. □

4. Numerical Results

In this section, we use a large number of test problems considered in Andrei [1], Michalewicz [16] and Moré et al. [11] in Table 1 to analyze the improvement of the Broyden-CG method with the Broyden method. The dimensions of the tests range between 2 and 1,000 only.

The comparison between Algorithm 1 (Broyden) and Algorithm 2 (Broyden-CG) uses the cost of computation based on the number of iterations and CPU-time. As suggested by Moré et al. [11], for each of the test problems, the initial point x_0 will take further away from the minimum point and we analyze three of initial points of each of the test problems. In doing so, it leads us to test the global convergence properties and the robustness of our method. For the Armijo line search, we use $s = 1$, $\beta = 0.5$ and $\sigma = 0.1$. In our implementation, the programs are all written in Matlab. The stopping criteria that we used in both algorithms are $\| g(x_{i+1}) \| \leq 10^{-6}$. The Euclidean norm is used in the convergence test to make these results comparable.

Table 1. A list of problem functions

Test problem	***N*-dimensional**	**Source**
Powell badly scaled	2	Moré et al. [11]
Beale	2	Moré et al. [11]
Biggs Exp6	6	Moré et al. [11]
Chebyquad	4,6	Moré et al. [11]
Colville polynomial	4	Michalewicz [16]
Variably dimensioned	4,8	Moré et al. [11]
Freudenstein and Roth	2	Moré et al. [11]
Goldstein pricepolynomial	2	Michalewicz [16]
Himmelblau	2	Andrei [1]
Penalty 1	2,4	Moré et al. [11]
Extended Powell singular	4,8	Moré et al. [11]
Extended Rosenbrock	2,10,100,200,500,1000	Andrei [1]
Trigonometric	6	Andrei [1]
Watson	4,8	Moré et al. [11]
Six-hump camel back polynomial	2	Michalewicz [16]
Extended shallow	2,4,10,100,200,500,1000	Andrei [1]
Extended strait	2,4,10,100,200,500,1000	Andrei [1]
Scale	2	Michalewicz [16]
Raydan 1	2,4	Andrei [1]
Raydan 2	2,4	Andrei [1]
Diagonal 3	2	Andrei [1]
Cube	2,10,100,200	Moré et al. [11]

Table 2. Performance of Algorithm 2 respect to Algorithm 1, 180 test problems

Global properties	**Algorithm 1**	**Algorithm 2**
Total iteration	99461	56512
Total CPU time (s)	9043.40	5759.70

Table 2 shows that the suggested algorithm is better compared to that of the original BFGS in terms of the number of iterations and the CPU-time.

The speed factor of the number of iterations is 176 for Algorithm 1, while the speed factor of CPU-time is 1.57 for Algorithm 2. The performance results are shown in Figures 1 and 2, respectively, using the performance profile introduced by Dolan and Moré [5]. The performance profile seeks to find how well the solvers perform relative to the other solvers on a set of problems. In general, $P(\tau)$ is the fraction of problems with performance ratio τ, thus, a solver with high values of $P(\tau)$ or one that is located at the top right of the figure is preferable.

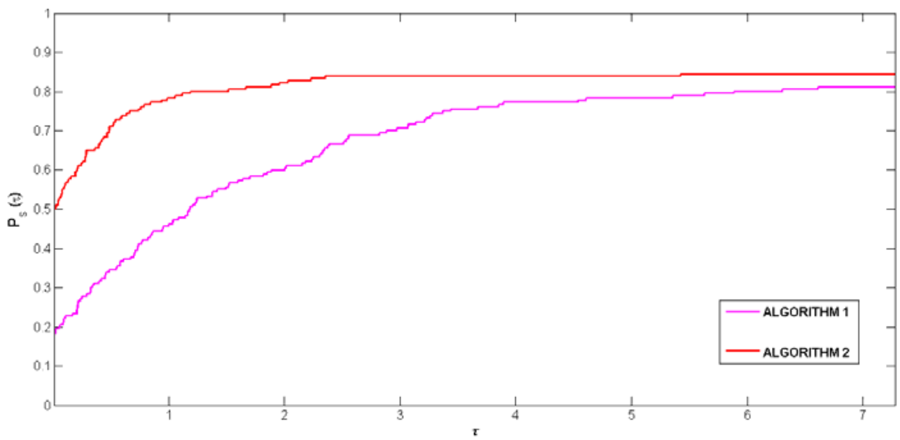

Figure 1. Performance profile based on the number of iterations.

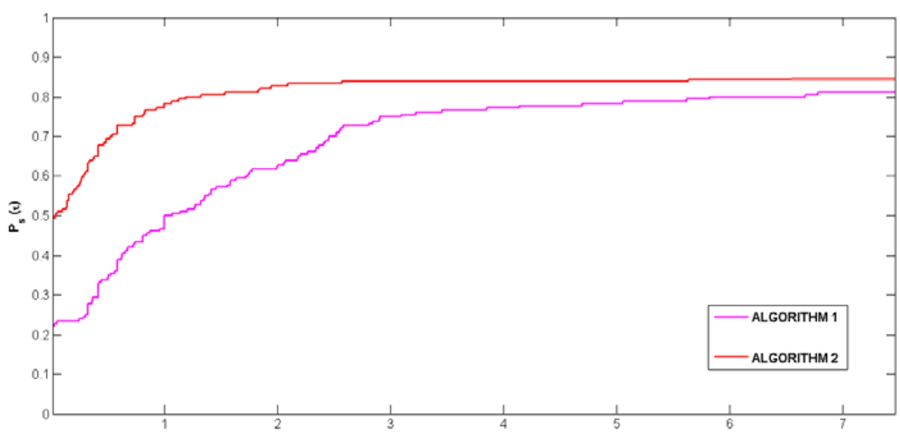

Figure 2. Performance profile based on CPU-time.

Figures 1 and 2 show that Algorithm 2 has better performance since it can solve 85% of the test problems compared with Algorithm 1 (81%). Besides, we can also say that Algorithm 2 is the fastest solver for two characters which is the number of iterations and the CPU-time pertaining to Figures 1 and 2.

5. Conclusion

A new search direction of Broyden family method based on hybridization of quasi-Newton and conjugate gradient's search direction was proposed. Based on numerical analysis, it is shown that our suggested method is more effective in solving unconstrained optimization problems.

Acknowledgement

The financial support of the Research Grant Scheme (Project No: UniSZA/14/GU), UniSZA is gratefully acknowledged.

References

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

[2] A. Buckley and A. Lenir, QN-like variable storage conjugate gradients, Math. Program. 27(2) (1983), 155-175.

[3] Richard H. Byrd, Jorge Nocedal and Ya-Xiang Yuan, Global convergence of a class of quasi-Newton methods on convex problems, SIAM J. Numer. Anal. 24(5) (1987), 1171-1191.

[4] Yu-Hong Dai, Convergence properties of the BFGS algorithm, SIAM J. Optim. 13(3) (2002), 693-702.

[5] Elizabeth D. Dolan and Jorge J. Moré, Benchmarking optimization software with performance profiles, Math. Program. 91(2) (2002), 201-213.

[6] Dong-Hui Li and Masao Fukushima, A modified BFGS method and its global convergence in nonconvex minimization, J. Comput. Appl. Math. 129 (2001), 15-24.

[7] A. Ludwig, The Gauss-Seidel-quasi-Newton method: a hybrid algorithm for solving dynamic economic models, J. Econom. Dynam. Control 31(5) (2007), 1610-1632.

[8] Y. Z. Luo, G. J. Tang and L. N. Zhou, Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-Newton method, Appl. Soft Comput. 8(2) (2008), 1068-1073.

[9] Mustafa Mamat, Ismail Mohd, Leong Wah June and Yosza Dasril, Hybrid Broyden method for unconstrained optimization, Inter. J. Numer. Meth. Appl. 1(2) (2009), 121-129.

[10] Walter F. Mascarenhas, The BFGS method with exact line searches fails for non-convex objective functions, Math. Program. 99(1) (2004), 49-61.

[11] Jorge J. Moré, Burton S. Garbow and Kenneth E. Hillstrom, Testing unconstrained optimization software, ACM Trans. Math. Software 7(1) (1981), 17-41.

[12] Jorge Nocedal, Updating quasi-Newton matrices with limited storage, Math. Comp. 35(151) (1980), 773-782.

[13] M. J. D. Powell, Some global convergence properties of a variable metric algorithm for minimization without exact line searches, Nonlinear Program. 9 (1976), 53-72.

[14] Philip Wolfe, Convergence conditions for ASCENT methods, SIAM Rev. 11(2) (1969), 226-235.

[15] Philip Wolfe, Convergence conditions for ASCENT methods. II: Some corrections, SIAM Rev. 13(2) (1971), 185-188.

[16] Zbigniew Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag, 1996.