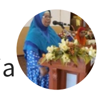
The scaling of hybrid method in solving unconstrained optimization method

Article · April 2016

DOI: 10.17654/MS099070983

4 authors, including:

Mohd Asrul Hery Ibrahim
Universiti Malaysia Kelantan Kota Bharu Malaysia



Puspa Liza Ghazali
Universiti Sultan Zainal Abidin

Zabidin Salleh
Universiti Malaysia Terengganu

Far East Journal of Mathematical Sciences (FJMS)

Published Online: March 2016
http://dx.doi.org/10.17654/MS099070983
Volume 99, Number 7, 2016, Pages 983-991 **ISSN: 0972-0871**

THE SCALING OF HYBRID METHOD IN SOLVING UNCONSTRAINED OPTIMIZATION METHOD

Mohd Asrul Hery Ibrahim, Mustafa Mamat, Puspa Liza Ghazali and Zabidin Salleh

Faculty of Entrepreneurship and Business
Universiti Malaysia Kelantan
Kampus Kota, Malaysia
e-mail: hery.i@umk.edu.my

Faculty of Informatics and Computing
Universiti Sultan Zainal Abidin
K. Terengganu, Malaysia

Faculty of Business Management and Accounting
Universiti Sultan Zainal Abidin
K. Terengganu, Malaysia

School of Informatics and Applied Mathematics
Universiti Malaysia Terengganu
Malaysia

Abstract

In this paper, the solution of unconstrained optimization problems has been suggested by using scaling of BFGS-SD method. This method is globally convergent with inexact line searches for general convex functions. The number of iterations and CPU-time of this method has been compared with those of the original BFGS method.

Received: July 2, 2015; Revised: September 14, 2015; Accepted: October 1, 2015
2010 Mathematics Subject Classification: 65K10.
Keywords and phrases: BFGS method, BFGS-SD method, global convergence, CPU-time.
Communicated by K. K. Azad

1. Introduction

Consider the unconstrained optimization problems:

$$\min_{x \in R^n} f(x) \tag{1.1}$$

and let $f : R^n \to R$ be continuously differentiable. The quasi-Newton methods are an iterative scheme with the ith iteration, an approximation point x_i and the $(i+1)$th iteration of x is given by

$$x_{i+1} = x_i + \alpha_i d_i, \tag{1.2}$$

where the search direction d_i is calculated by

$$d_i = -B_i^{-1} g_i \tag{1.3}$$

where g_i is a gradient of f_i. 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 (Nocedal and Wright [11]). Then the step size α_i in (1.2) was obtained using the Armijo line search rule (Armijo [2])

$$f(x_i) - f(x_i + \alpha_i d_i) \geq -\sigma \alpha_i g_i^T d_i, \tag{1.4}$$

where $s > 0$, $\beta \in (0, 1)$, $\sigma \in (0, 1)$ and $\alpha_i = \max\{s, s\beta, s\beta^2, \ldots\}$. Then the sequence of $\{x_i\}_{i=0}^{\infty}$ is converged to the optimal point x^* which minimizes f (Armijo [2]), while the updated Hessian approximation formula in (1.3) requires B_i as a positive definite and satisfying the quasi-Newton equation

$$B_{i+1} s_i = y_i, \tag{1.5}$$

where $s_i = \alpha_i d_i$ and $y_i = g_{i+1} - g_i$. Hence, the Broyden-Flether-Goldfarb-Shanno (BFGS) formula is used to calculate the approximation of Hessian which is satisfied for (1.5). The BFGS algorithm for unconstrained optimization problem uses the matrices B_i which are updated by the formula

$$B_{i+1} = B_i - \frac{B_i s_i s_i^T B_i}{s_i^T B_i s_i} + \frac{y_i y_i^T}{s_i^T y_i}, \tag{1.6}$$

which can be found in Dai [3], Mascarenhas [9], Yuan and Wei [12] and Ibrahim et al. [6, 7].

This paper is organized as follows. In Section 2, we elaborate the step size that is used in the modification of the Broyden's algorithm. Then the combination of search direction and the new procedure is also discussed in Section 2. The convergence result will be shown in Section 3 and an explanation about the numerical results is given in Section 4. The paper ends with a short conclusion in Section 5.

2. Scaling BFGS-SD

Mamat et al. [8] suggested to hybrid the search direction between the quasi-Newton methods and the steepest descent method while Han and Neumann [5] focused on combining the quasi-Newton and Cauchy direction. In both the researches, they proved that the new algorithms are globally convergent under line search conditions

$$f_{i+1} - f_i \leq \sigma_1 \alpha_i g_k^T d_k, \quad 0 < \sigma_1 < 1/2,$$

$$g_{i+1}^T d_i \geq \sigma_2 g_i^T d_i, \quad 0 < \sigma_1 < \sigma_2 < 1, \tag{2.1}$$

with assuming the level set $L = \{x \mid f(x) \leq f(x_0)\}$ is bounded. Mamat et al. [8] suggested the search direction for quasi-Newton method is

$$d_i = -\eta_i B_i^{-1} g_i - \delta_i g_i, \tag{2.2}$$

where $\eta_i > 0$ and $\delta_i > 0$.

Our primary motivation here is to develop a fixed parameter in (2.2) to reduce the number of iterations as well as the CPU-time compared to the original BFGS method. Besides, the objective of this research is also to form the properties of global convergence on the BFGS-SD method. Hence, we suggested a search direction with fixed parameter which is stated as follows:

$$d_i = -B_i^{-1} g_i + \rho g_i,$$

$$(B_i - \rho I) d_i = -g_i, \tag{2.3}$$

where I is the identity matrix and $\rho < 0.$ Hence, the new algorithm based on (2.3) is derived as follows:

Algorithm BFGS-SD

Step 0. Given a starting point x_0 and $B_0 = I_n.$ 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 (2.3).

Step 3. Calculate the step size α_i by (1.4).

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

Step 5. Update B_{i-1} by (1.6) to obtain $B_i.$

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

The difference between the BFGS-SD algorithm and the BFGS method is in Step 2.

3. Convergence Result

For the analysis of this section, we assume that the sequences $\{x_i\}$, $\{B_i\}$, $\{y_i\}$ and $\{s_i\}$ all generated by Algorithm BFGS-SD as given in Section 2 are positive definite. We also 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

$$g_i^T d_i \leq c_1 \| g_i \|^2 \tag{3.1}$$

for all $i \geq 0$, then the search directions satisfy the sufficient descent condition which can be proved in Theorem 3.1. Hence, we need to make a few assumptions based on the objective function.

Assumption 3.1.

(H1): The level set L is convex. Moreover, positive constants c_1 and c_2 exist satisfying

$$c_1 \| z \|^2 \leq z^T F(x) z \leq c_2 \| z \|^2$$

for all $z \in R^n$ and $x \in L$, where $F(x)$ is the Hessian matrix for f.

(H2): The Hessian matrix is Lipschitz continuous at the point x^*, that is, there exists the positive constant c_3 satisfying

$$\| g(x) - g(x^*) \| \leq c_3 \| x - x^* \|$$

for all x in a neighborhood of x^*.

Theorem 3.1. *Suppose that Assumption* 3.1 *holds and* $\{B_i\}$ *is bounded. Then the condition* (3.1) *holds for all* $i \geq 0$.

Proof. From (2.3), we see that

$$g_i^T d_i = -g_i^T B_i^{-1} g_i + \rho g_i^T g_i$$

$$g_i^T d_i \leq -\lambda_i \| g_i \|^2 + \rho \| g_i \|^2$$

$$\leq c_1 \| g_i \|^2,$$

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

Lemma 3.1 (See Nocedal and Wright [11]). *Suppose that Assumption* 3.1 (*H*2) *holds*. *Then the step size* α_i *determined by* (2.1) *satisfies that*

$$f_{i+1} - f_i \leq -c_2 \frac{(g_i^T d_i)^2}{\| d_i \|^2}, \tag{3.2}$$

where c_2 *is a positive constant*.

Based on Theorem 3.1 and Lemma 3.1, we can regard the new algorithm as having strong global convergence.

Theorem 3.2 (Global convergence). *Suppose that Assumption* 3.1 *and Theorem* 3.1 *hold*. *Then*

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

Proof. Combining Theorem 3.1 and Lemma 3.1 gives

$$\sum_{i=0}^{\infty} \frac{\| g_i \|^4}{\| d_i \|^2} < \infty. \tag{3.3}$$

Hence, from Theorem 3.1, we can define that $\| d_i \| \leq -c \| g_i \|$. Then (3.3) will be simplified as $\sum_{i=0}^{\infty} \| g_i \|^2 < \infty$. Therefore, the proof is completed. □

4. Numerical Results

In this section, we give some numerical result on scaled BFGS -SD algorithm. Our source code is written in Matlab 2010 using the Intel Pentium® Dual Core processor. Test functions are the standard unconstrained optimization problems adapted from More et al. [10] and Andrei [1].

As suggested by More et al. [10], for each of the test problems, the initial point x_0 will take further away from the minimum point. This represents a total of 77 test problems. In doing so, it leads us to test the global convergence properties and the robustness of our method. In (2.3), we use $\rho = 10^{-4}$ as BFGS-SD[1] method and $\rho = 10^{-5}$ as BFGS-SD[2]. While in the Armijo line search, $s = 1$, $\beta = 0.5$ and $\sigma = 0.1$. The stopping criteria used are $\| g_i \| \leq 10^{-6}$ and the number of iterations exceeds its limit, which is set to be 10000. The comparisons of the BFGS-SD[1] and BFGS-SD[2] results with BFGS are summarized in Table 1 for number of iterations (*ni*) and the CPU-time.

Table 1. Numerical results obtained by BFGS, BFGS-SD[1] and BFGS-SD[2] algorithm

Global characteristic	BFGS	BFGS-SD[1]	BFGS-SD[2]
Total iteration	213484	93254	122731
CPU-time	5580.3	1783.5	3264.1

Table 1 shows the suggested algorithm is better compared to the original BFGS in terms of the number of iterations and the CPU-time. The speed factor of the number of iterations is 2.29 for BFGS-SD[1] and 1.73 for BFGS-SD[2], while the speed factor of CPU-time is 3.12 for BFGS-SD[1] and 1.71 for BFGS-SD[2]. The performance profile results will be shown in Figures 1 and 2, respectively, using the performance profile introduced by Dolan and Moré [4]. 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.

Figures 1 and 2 show that the BFGS-SD[2] method has the best performance, since it can solve 98% of the test problems compared with the BFGS-SD[1] method (95%) and BFGS method (86%). Besides, we can say that the BFGS-SD[1] is the *fastest* solver for two characters which is the number of iterations and the CPU-time pertaining to Figures 1 and 2.

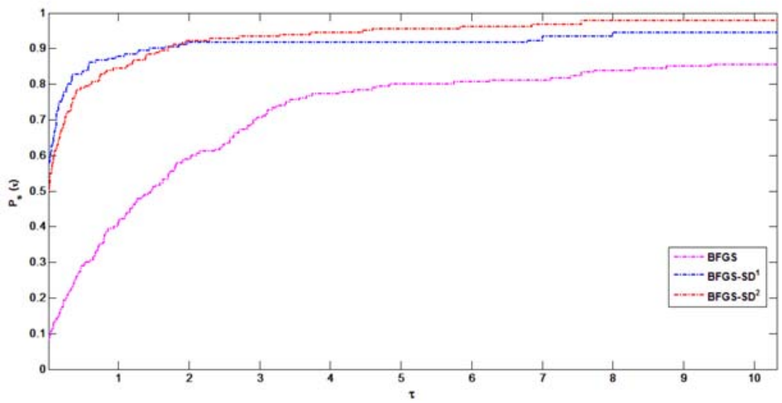


Figure 1. Performance profile based on the number of iterations.

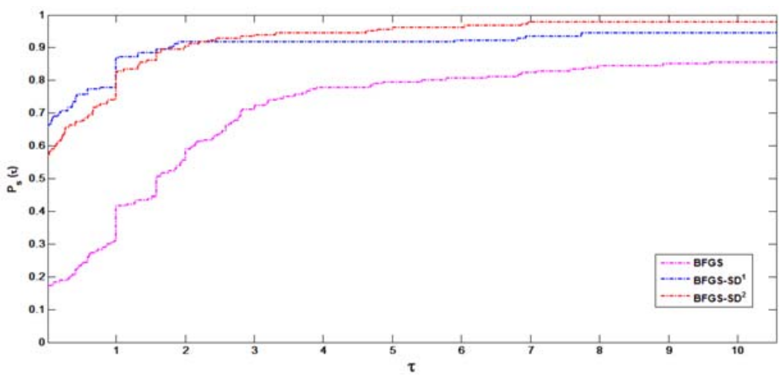


Figure 2. Performance profile based on CPU-time.

5. Conclusion

Unconstrained optimization problems have been solved by a scaling of BFGS-SD method. The efficiency of the proposed method is evident by numerical results of test problems. The solution of unconstrained optimization problems with exact line search is an interesting area for further investigation.

Acknowledgements

This research was supported by Fundamental Research Grant Scheme

(FRGS) Vote No. 59256 of Ministry of Higher Education Malaysia. The authors thank the anonymous referees for their valuable suggestions which led to the improvement of the manuscript.

References

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

[2] L. Armijo, Minimization of functions having Lipschitz continuous partial derivatives, Pacific J. Math. 16(1) (1966), 1-3.

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

[4] E. D. Dolan and J. J. Moré, Benchmarking optimization software with performance profiles, Math. Program. 91(2) (2002), 201-213. doi: 10.1007/s101070100263

[5] L. Han and M. Neumann, Combining quasi-Newton and Cauchy directions, Int. J. Appl. Math. 12(2) (2003), 167-191.

[6] M. A. H. Ibrahim, M. Mamat and L. W. June, BFGS method: a new search direction, Sains Malaysiana 43(10) (2014a), 1591-1597.

[7] M. A. H. Ibrahim, M. Mamat and L. W. June, The hybrid BFGS-CG method in solving unconstrained optimization problems, Abstract and Applied Analysis, Vol. 2014, Article ID 507102, 2014, 6 pages. doi: 10.1155/2014/507102

[8] M. Mamat, I. Mohd, L. W. June and Y. Dasril, Hybrid Broyden method for unconstrained optimization, Int. J. Numer. Meth. Appl. 1(2) (2009), 121-130.

[9] W. F. Mascarenhas, The BFGS method with exact line searches fails for non-convex objective functions, Math. Program. 99(1) (2004), 49-61. doi: 10.1007/s10107-003-0421-7

[10] J. J. More, B. S. Garbow and K. E. Hillstrom, Testing unconstrained optimization software, ACM Trans. Math. Softw. 7(1) (1981), 17-41. doi: 10.1145/355934.355936

[11] J. Nocedal and S. J. Wright, Numerical Optimization, Springer, 2006.

[12] G. Yuan and Z. Wei, Convergence analysis of a modified BFGS method on convex minimizations, Comput. Optim. Appl. 47(2) (2010), 237-255. doi: 10.1007/s10589-008-9219-0