**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 21, No. 3, March 2021, pp. 1830~1836
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v21.i3.pp1830-1836 

# A new class of self-scaling for quasi-Newton method based on the quadratic model

**Basim A. Hassan, Ranen M. Sulaiman**
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Iraq

**Article Info**

***Article history:***

Received Sep 23, 2020
Revised Nov 30, 2020
Accepted Dec 23, 2020

***Keywords:***

Global convergence
Self-scaling quasi-Newton
Unconstrained optimization

**ABSTRACT**

Quasi-Newton method is an efficient method for solving unconstrained optimization problems. Self-scaling is one of the common approaches in the modification of the quasi-Newton method. A large variety of self-scaling of quasi-Newton methods is very well known. In this paper, based on quadratic function we derive the new self-scaling of quasi-Newton method and study the convergence property. Numerical results on the collection of problems showed the self-scaling of quasi-Newton methods which improves overall numerical performance for BFGS method.



***Corresponding Author:***

Basim A. Hassan
Department of Mathematics
College of Computers Sciences and Mathematics
University of Mosul, Iraq
Email: basimah@uomosul.edu.iq, basimabas39@gmail.com

## 1. INTRODUCTION

All over the complete paper, we deem the unconstrained minimization problem as:

$$Min \quad f(x) \ , \ \mathrm{x} \in \mathrm{R}^{\mathrm{n}} \tag{1}$$

The most widely used general iterative suggestion for problem resolution (1) is given by:

$$x_0 \in R^n \ , \quad x_{k+1} = x_k + \alpha_k d_k \tag{2}$$

where $x_{k+1}$ is a new iteration point, $x_k$ is a current iterative point, $\alpha_k > 0$ is a step-size and $d_k$ is the search direction. (see, for example, [1, 2 and 3]). Quasi-Newton methods form an important class of numerical methods for solving optimization problems. For our purposes, we examine the general iterative scheme of quasi-Newton direction :

$$B_k d_k + g_k = 0 \tag{3}$$

where $B_k$ is an appropriate approximation of the inverse of the Hessian. If we apply the quasi-Newton equation to the chosen approximations $x_k$ and $x_{k+1}$, our task is to find a $B_{k+1}$ satisfying:

$$B_{k+1}\delta_k = y_k \tag{4}$$

where $\delta_k = x_{k+1} - x_k = \alpha_k d_k$ and $y_k = g_{k+1} - g_k$, see [4].

This key of such methods is a matrix updating procedure, of which the BFGS method is the most successful and widely used can be separated into :

$$B_{k+1}^{BFGS} = B_k - \frac{B_k \delta_k \delta_k^T {B_k}^T}{\delta_k^T B_k \delta_k} + \frac{y_k y_k^T}{\delta_k^T y_k} \tag{5}$$

Let $H_k$ be the inverse of $B_k$. Then the inverse update formula of (5) method is represented as :

$$H_{k+1}^{BFGS} = H_k - \frac{H_k y_k \delta_k^T + \delta_k y_k^T H_k}{\delta_k^T y_k} + \left[1 + \frac{y_k^T H_k y_k}{\delta_k^T y_k}\right] \frac{\delta_k \delta_k^T}{\delta_k^T y_k} \tag{6}$$

The Biggs modification of the BFGS formula to improve the performance of updates, can be written

$$H_{k+1} = H_k - \frac{H_k y_k \delta_k^T + \delta_k y_k^T H_k}{\delta_k^T y_k} + \left[\frac{1}{\rho_k} + \frac{y_k^T H_k y_k}{\delta_k^T y_k}\right] \frac{\delta_k \delta_k^T}{\delta_k^T y_k} \tag{7}$$

which will satisfy :

$$B_{k+1}\delta_k = \rho_k y_k, \quad \rho_k = \frac{4\delta_k^T g_k + 2\delta_k^T g_{k+1} + 6[f(x_k) - f(x_{k+1})]}{\delta_k^T y_k} \tag{8}$$

Is then called Self-scaling of quasi-Newton methods. More details can be found in [5]

The idea of variant self-scaling of quasi-Newton methods had been studied by many researchers for example, see (Oren, [6]); (Yuan, [7]) and (Basim and Hawraz [8]). A self-scaling of quasi-Newton algorithm was developed to decrease the number of iterations and preserves the global convergence on quasi-Newton algorithms. Interested researcher can refer to [9-12] for further studies and recent reference regarding quasi-Newton. Next, derivation of the a new self-scaling Quasi-Newton are described and tested.

## 2. A NEW SELF-SCALING QUASI-NEWTON METHODS

In order to we derive the new self-scaling of quasi-Newton method we consider the second-order Taylor approximation as :

$$f(x) = f(x_{k+1}) + g_{k+1}^T (x - x_{k+1}) + \frac{1}{2}(x - x_{k+1})^T Q (x - x_{k+1}) \tag{9}$$

Substituting $x_k$ in to $x$ in equation (9) and using exact line search $g_{k+1}^T d_k = 0$, then (9) we have:

$$f(x_{k+1}) - f(x_k) = -\frac{1}{2}\delta_k^T Q \delta_k \tag{10}$$

In fact, by using exact line search with this function, the optimal step size $\alpha_k$ is given by :

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k} \tag{11}$$

From (10) and (11) we get :

$$f(x_{k+1}) = f(x_k) - \frac{1}{2}\frac{(g_k^T d_k)^2}{d_k^T Q d_k} \tag{12}$$

Thus one of the possible choices in an approximation of $Qd_k$ can be given by :

$$Qd_k = -\frac{(g_k^T d_k)^2}{2(f(x_{k+1}) - f(x_k))(d_k^T y_k)} y_k \tag{13}$$

Since $\delta_k = x_{k+1} - x_k = \alpha_k d_k$, we obtain :

$$Q\delta_k = -\frac{\alpha_k (g_k^T d_k)^2}{2(f(x_{k+1}) - f(x_k))(d_k^T y_k)} y_k \tag{14}$$

A good approximation to the Hessian matrix $Q$ is a sequence of a positive definite matrices $B_{k+1}$ which will satisfy :

$$B_{k+1}\delta_k = -\frac{\alpha_k (g_k^T d_k)^2}{2(f(x_{k+1}) - f(x_k))(d_k^T y_k)} y_k \tag{15}$$

The above relation we obtain the scaling denoted as $\rho_k^{BR1}$ and can be written as :

$$B_{k+1}\delta_k = \rho_k^{BR1} y_k \tag{16}$$

where :

$$\rho_k^{BR1} = -\frac{\alpha_k (g_k^T d_k)^2}{2(f(x_{k+1}) - f(x_k))(d_k^T y_k)} \tag{17}$$

So – called $BR1$ method.

### 2.1. Outline of the new algorithm

The outline of the new algorithm is as follows :

Step 0 : Choose an initial point $x_0 \in R^n$, set $k = 1$.
Step 1 : If the stopping criterion is a satisfied stop.
Step 2 : Compute $d_k = -H_k g_k$.
Step 3 : Find a $\alpha_k$ which satisfies the Wolfe rule :

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k$$
$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k$$

Step 4 : Generate a new iteration point by $x_{k+1} = x_k + \alpha_k d_k$ and calculate the new updating formula (7) and using (17).
Step 5 : Set $k = k+1$ and go back to Step 1.

## 3. GLOBAL CONVERGENCE PROPERTY OF THE NEW ALGORITHM

Our main interest is convergence Property for the new Algorithm and defined by the relation :

$$\lim_{k\Rightarrow\infty} \inf \|g_k\| = 0 \tag{18}$$

The BFGS update generate identical conjugate gradient search direction provided that the function quadratic and exact line searches are used. To prove that the new updates generate identical conjugate gradient search directions. We first introduce a useful lemma it can be proved in a similar way to the proof of Lemma 5.1 in [13].

**Lemma (2.1):**

If the BFGS algorithm is applied to the quadratic Function:

$$f(x) = \frac{1}{2} x^T Q x + b^T x \tag{19}$$

Using the same starting point $x_0$ and initial symmetric positive definite matrix $H_0$ , then:

$$H_k g^* = H_0 g^* \tag{20}$$

The detailed proof was given by Nazareth [14].
Powell (Powell, [15]), showed that, the conjugate gradient method achieves the limit :

$$\lim_{k\Rightarrow\infty} \inf \|g_k\| = 0 \tag{21}$$

if the level set $\{\ x:\ f(x) \le f(x_*)\ \}$ is bounded and $\alpha_k$ is defined so that $\{\ g_{k+1}^T d_k = 0\ ,\ k \ge 1\}$ holds for all $k$ .

The following theorem are often used to explain the global convergence.

**Theorem (2.3)**

Assume that $f(x)$ be a quadratic function defined in (19) and that the line searches are exact : if $H_k$ is any symmetric positive definite matrix and for the new updating formula,

$$H_{k+1} = H_k - \frac{H_k y_k \delta_k^T + \delta_k y_k^T H_k}{\delta_k^T y_k} + \left[ \frac{1}{\rho_k^{BR1}} + \frac{y_k^T H_k y_k}{\delta_k^T y_k} \right] \frac{\delta_k \delta_k^T}{\delta_k^T y_k} \tag{22}$$

where $\rho_k^{BR1}$ defined in (17), then the search direction as :

$$d_{k+1}^{new} = -H_{k+1}^{new} g_{k+1} \tag{23}$$

is identical to the Conjugate Gradient direction [H/S direction ] $d_{CG}$ and defined by:

$$d_{k+1}^{CG} = -g_{k+1} + \frac{y_k^T g_{k+1}}{y_k^T d_k} d_k \quad for\ k \geq 1 \tag{24}$$

**Proof :**
The update (22) can be written as:

$$H_{k+1} = H_k - \frac{H_k y_k \delta_k^T}{\delta_k^T y_k} - \frac{\delta_k y_k^T H_k}{\delta_k^T y_k} + \left[\frac{1}{\rho_k^{BR1}} + \frac{y_k^T H_k y_k}{\delta_k^T y_k}\right] \frac{\delta_k \delta_k^T}{\delta_k^T y_k}$$

Now,

$$d_{k+1}^{new} = -H_k g_{k+1} + \frac{H_k y_k \delta_k^T g_{k+1}}{\delta_k^T y_k} + \frac{y_k^T H_k g_{k+1}}{\delta_k^T y_k} \delta_k - \frac{1}{\rho_k^{BR1}} \frac{\delta_k^T g_{k+1}}{\delta_k^T y_k} \delta_k - \frac{y_k^T H_k y_k \delta_k^T g_{k+1}}{(\delta_k^T y_k)^2} \delta_k \tag{25}$$

using the property $\delta_k^T g_{k+1} = 0$ quoted earlier which holds for line searches we get :

$$d_{k+1}^{new} = -H_k g_{k+1} + \frac{y_k^T H_k g_{k+1}}{\delta_k^T y_k} \delta_k \tag{26}$$

The vector $g_{k+1}$ can be substituted for $H_k g_{k+1}$ by using lemma (2.1). Therefore :

$$d_{k+1}^{new} = -g_{k+1} + \frac{y_k^T g_{k+1}}{\delta_k^T y_k} \delta_k \tag{27}$$

We also know that $d^{BFGS}$ and $d^{CG}$ are identical (Nazareth, [14,15]), and $d^{new}$ is identical to $d^{BFGS}$ with exact line searches. Hence as shown in (27) becomes :

$$d_{k+1}^{new} = -g_{k+1} + \frac{y_k^T g_{k+1}}{d_k^T y_k} d_k = d_{k+1}^{CG} \tag{28}$$

hence the proof.

## 4. Numerical results

In this section, several computational experiments are conducted on a series of unconstrained optimization test issues the specifics please see more [16] to explain the application and efficacy of the proposed process. Some other class from test problems was observed in [17-25]. Performance of a given methods was measured by two separate data: total number of iteration and total number of function evaluations, respectively. The stopping rule applied throughout was : "If $|f(x_k)| > 10^{-5}$, let $stop1 = |f(x_k) - f(x_{k+1})| \,/\, |f(x_k)|$; Otherwise, let $stop1 = |f(x_k) - f(x_{k+1})|$. For every problem, if $\|g_k\| < \varepsilon$ or $stop\ 1 < 10^{-5}$ is satisfied, the program will be stopped", see [26].

The graphs are plotted using data derived from numerical computations using the output model proposed by Dolan and More [27]. The suggested BR1 approach has the highest results in terms of both number of iterations as seen in Figure 1 and number of function evaluation Figure 2.

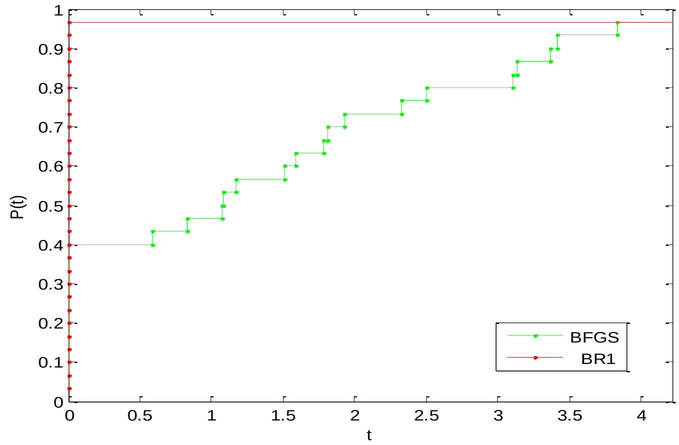


Figure 1. Number of iteration profile via dolan and more

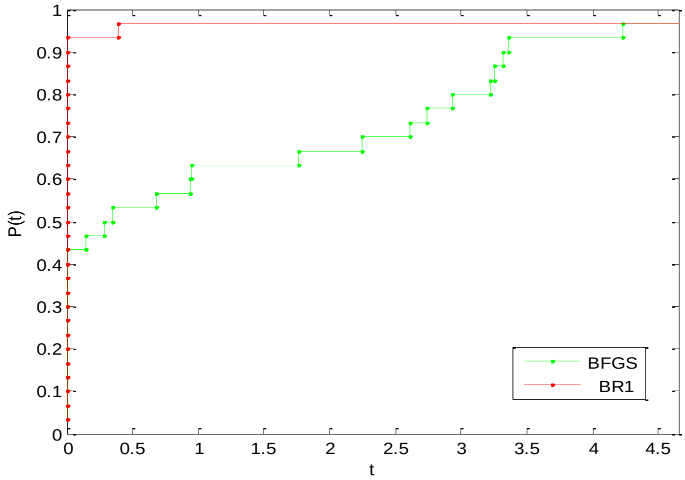


Figure 2. Number of function evaluation profile via dolan and more profile

From Figure 1 and Figure 2, it is concluded that the most efficient algorithm in terms of the number of iterations is our method, being the fastest for 75% of the problems, followed by total number of function evaluations, that was the most efficient for nearly 66% of the problems. This is a very interesting fact.

## 5. CONCLUSIONS

In this paper, a self-scaling quasi-Newton method has been derived to find minimum for unconstrained optimization problem. Self-scaling of quasi-Newton methods which improved overall numerical performance of these methods.

## ACKNOWLEDGMENT

The authors are very grateful to the University of Mosul / College of Computers Sciences and Mathematics for their provided facilities, which helped to improve the quality of this work".

## REFERENCES

[1] Cauchy, A., "Metode generale pour la resolution des systems equations simultanees," *C. R. Sci.*, vol. 25, pp. 46-89, 1847.
[2] Luenberg, D.G.,Ye, Y. *Linear and Nonlinear Programming*, Springer, New York, 2008.

[3] Sun,W., Yuan, Ya-Xiang, *Optimization Theory and Methods: Nonlinear Programming*, Springer, New York, 2006.
[4] Fletcher, R. *Practical Methods of Optimization*, John Wiley and Sons, ChiChester (New York), 1987.
[5] Biggs, M.C., "A note on Minimization algorithms making use of nonquadratic properties of the objective function," *Journal of the Institute of Mathematics and Its Application*, vol. 12, pp. 337-338, 1973.
[6] Oren, S.S. and Luenberger, D. "Self-Scaling Variable Metric Algorithm," *Part I, Management Science*, vol. 20, no. 5, pp. 845-862, 1974.
[7] Yuan, Y., "A modified BFGS algorithm for unconstrained optimization," *IMA Journal Numerical Analysis*, vol. 11, pp. 325-332, 1991.
[8] Basim A. Hassan and Hawraz N. Jabbar, " A New Transformed Biggs 's Self-Scaling Quasi-Newton Method for Optimization'," *ZANCO Journal of Pure and Applied Sciences*, vol. 31, pp. 1-5, 2018.
[9] Basim A. Hassan, "A modified quasi-Newton methods for unconstrained Optimization," Journal *of pure and applied mathematics*, vol. 2019, no. 42, pp. 504-511, 2019.
[10] Basim A. Hassan, "A new type of quasi-Newton updating formulas based on the new quasi-Newton equation SIAM J. N. A., control and optimization," *Numerical Algebra, Control & Optimization*,.vol.10 no. 2, pp.227-235, 2020.
[11] Basim A. Hassan and Mohammed W. T., "A New Variants of Quasi-Newton Equation Based on the Quadratic Function for Unconstrained Optimization," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 2, pp.701-708.
[12] Basim A. Hassan and Mohammed W. T, "A Modified Quasi-Newton Equation in the Quasi-Newton Methods for Optimization," *HIKARI Ltd, Applied Mathematical Sciences*, vol. 10, pp. 463-472, 2019.
[13] Zhiwei Q, "The relationships between CG, BFGS, and two limited memory algorthim," *Electronic Journal of Undergraduate Mathematics*, vol. 12, pp. 5-20, 2007.
[14] Nazareth L., "A relationship between BFGS and conjugate gradient algorithm and its implementation for new algorithms," *SIAM Journal Numerical Analysis*, vol. 61, pp. 794-800, 1979.
[15] Powell M. J., *Non-convex minimization calculations and the conjugate gradient method, in Lecture Notes in Mathematics 1066*, Springer-Berlag (Berlin), pp. 122-141, 1983.
[16] More J., Garbow B., and Hillstrome K., "Testing unconstrained optimization software," *ACM Trans. Math. Software*, vol. 7, pp. 17-41, 1981.
[17] Basim A. Hassan, "A Globally Convergence Spectral Conjugate Gradient Method for Solving Unconstrained Optimization Problems and Mohammed," *Raf. J. of Comp. & Math's.*, vol. 10, pp. 21-28, 2013.
[18] Basim A. Hassan, "Development a Special Conjugate Gradient Algorithm for Solving Unconstrained Minimization Problems," *Raf. J. of Comp. & Math's.*, vol. 9, pp. 73-84, 2012.
[19] Chen L.H., Deng N.Y., and. Zhang J.Z, "A modified quasi-Newton method for structured optimization with partial information on the Hessian," *Comput. Optim. Appl.*, vol. 35, pp. 5-18, 2006.
[20] Wei Z., Li G. and Qi L., "New quasi-Newton methods for unconstrained optimization problems," *Math. Program. Applied Mathematics and Computation*, vol. 175, pp. 1156-1188, 2006.
[21] Yuan G, Wei Z, "Convergence analysis of a modified BFGS method on convex minimizations," *Comput Optim Appl*, vol. 47, pp. 237-255, 2010.
[22] Yuan G, Wei Z, "BFGS trust-region method for symmetric nonlinear equations," *J Comput Appl Math*., vol. 230, pp. 44-58, 2009.
[23] Yuan G, Wei Z, Lu X, "Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search," *Appl Math Model*, vol. 47, pp. 811-825, 2017.
[24] Yuan G., Sheng Z., Wang B., Hu W. and, Li C., "The global convergence of a modified BFGS method for nonconvex functions," *Journal of Computational and Applied Mathematics*, vol. 327, pp. 274-294, 2018.
[25] Zhang JZ, Xu CX, "Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equation," *J Comput Appl Math*, vol. 137, pp. 269-278, 2001.
[26] Yuan Y. and Sun W., *Theory and Methods of Optimization*, Science Press of China, 1999.
[27] Dolan E. and. Moré J. J, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201-213, 2002.