Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 26 (2023), No. 2, pp. 289–300
DOI : 10.47974/JIM-1493

Using a new class quasi-Newton equation for solving unconstrained optimization problems

Basim A. Hassan *
Department of Mathematics
College of Computers Sciences and Mathematics
University of Mosul
Mosul
Iraq

Hawraz N. Jabbar †
Department of Mathematics
College of Sciences
University of Kirkuk
Kirkuk
Iraq

Abstract

The quasi-Newton plays a basic role in constructing the Hessian approximation, and quasi-Newton algorithms are built based on it. In this paper, a new class quasi-Newton equation is presented. The proposed method always generates a new descent search direction. Under appropriate conditions, the new method is proved to possess global convergence. Comparative results show the computational efficiency of the proposed method.

Subject Classification: *90C30, 65K05, 90C53, 49M37.*

Keywords: *Unconstrined optimization, Quasi-Newton equations, Global convergence.*

1. Introduction

Our objective is the minimization of a smooth multi-variable nonlinear (convex or nonconvex) problem :

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

* *E-mail:* **basimah@uomosul.edu.iq (Corresponding Author)**
† *E-mail:* **hawrazmath@gmail.com**

Throughout this paper, f is function and its gradient g is available, see [16]. An interesting feature for quasi-Newton techniques are generally held to be the adapted efficient and robust minimization methods for solving problems (1), see [13]. The quasi-Newton techniques is a typical iterative method of the form :

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

where α_k is a step size and d_k is a descent direction. The precise line search formula, which is defined as reducing $f(x_k + \alpha_k d_k)$, is the best option α_k. In the quadratic case, we can find an explicit formula for α_k, which obtains :

$$\alpha_k = \frac{-g_k^T d_k}{d_k^T Q d_k}. \tag{3}$$

is often a good representation of a more general function, especially in the minimum. For more details see [1,8]. The step size α_k was obtained so as to satisfy the Wolfe conditions:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k \tag{4}$$

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k \tag{5}$$

where $0 < \delta < \sigma < 1$, see [12]. The search direction typical is obtained by solving:

$$B_k d_k + g_k = 0 \tag{6}$$

where B_k is an approximation of the Hessian matrix, should satisfy the secant equation :

$$B_{k+1} \delta_k = \gamma_k \tag{7}$$

we correct the correction formulae for the approximate Hessian matrix B_k and its inverse H_k can be expressed as:

$$B_{k+1}^{BFGS} = B_k - \frac{B_k \delta_k \delta_k^T B_k{}^T}{\delta_k^T B_k \delta_k} + \frac{\gamma_k \gamma_k^T}{\delta_k^T \gamma_k} \tag{8}$$

or :

$$H_{k+1}^{BFGS} = H_k - \frac{H_k \gamma_k \delta_k^T + \delta_k \gamma_k^T H_k}{\delta_k^T \gamma_k} + \left[1 + \frac{\gamma_k^T H_k \gamma_k}{\delta_k^T \gamma_k} \right] \frac{\delta_k \delta_k^T}{\delta_k^T \gamma_k} \tag{9}$$

is the BFGS method. For more details see [15].

However, the performance of the BFGS method is greatly improved in practice and globally convergent when f is convex; it has been studied by some authors (see Byrd et al. [9]; Toint [18]). Unlike that the global convergence is not guaranteed. To avoid a negative aspect may not globally be convergent for non-convex functions and to deduce adapted efficient and robust algorithms, which are numerically tractable and globally convergent, many researchers modified the secant equation to obtain a high-order precision in estimating the goal function's second-order curvature information (see Li and Fukushima [19]; Yuan and Wei [21]; Zhang et al. [23]). In the following, some of these methods are mentioned.

Wei-et al. [20]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k + \frac{2(f_k - f_{k+1}) + (g_{k+1} + g_k)^T \delta_k}{\delta_k^T \mu_k} \mu_k \tag{10}$$

Biglari-et al. [3]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k + \frac{4(f_k - f_{k+1}) + 2(g_{k+1} + g_k)^T \delta_k}{\delta_k^T \mu_k} \mu_k \tag{11}$$

Chen-et al. [10]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k + \frac{6(f_k - f_{k+1}) + 3(g_{k+1} + g_k)^T \delta_k}{\delta_k^T \mu_k} \mu_k \tag{12}$$

Basim [4]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \frac{1}{2}\gamma_k + \frac{(f_k - f_{k+1}) - 1/2(g_{k+1}^T \delta_k)}{\delta_k^T \mu_k} \mu_k \tag{13}$$

Basim and Mohamed [5]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k - \frac{\delta_{k+1}^T \delta_k}{\delta_k^T \mu_k} \mu_k. \tag{14}$$

Basim and Ghada [6]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k - \frac{2(f_k - f_{k+1})\delta_{k+1}^T \gamma_k}{\delta_k^T \mu_k} \mu_k. \tag{15}$$

Basim and Ranen [7]

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \frac{3}{2}\gamma_k + \frac{(f_{k+1} - f_k) - 3/2g_{k+1}^T\delta_k}{\delta_k^T\mu_k}\mu_k \tag{16}$$

where $\delta_k^T\mu_k \neq 0,$ but they often have good computational performance.

In addition, based on Yuan et al. [21], we expect has good studies of the other properties of BFGS method. The methods are developed by combining modified unified quasi-Newton equations of Ataee et al. [2] with the magnitude of the technique of Razieh et al. [17]. Based on this combination, we provide a novel class model of a modified quasi-Newton problem, will result to an efficient quasi-Newton equations.

2. A new class quasi-Newton equations and its algorithm

We propose an extension of several modified quasi-Newton equations in the following sections. The unified quasi-Newton equation was suggested by Ataee et al. [2] by extending (7):

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k + w\frac{2(f_k - f_{k+1}) + (g_{k+1} + g_k)^T\delta_k}{\delta_k^T\mu_k}\mu_k \tag{17}$$

where $w \in [0, 1, 2, 3]$ and μ_k is any vector satisfying $\delta_k^T\mu_k \neq 0.$ Razieh et al. [17], studied on this algorithms with additional important aspect, due to the magnitude of the $\|s_k\|$ if it's small, then we have:

$$f_{k+1} - f_k = \int_0^1 \nabla f(x_k + td_k)\,dt\ s_k = g_k^T s_k\,. \tag{18}$$

Thus, motivated by these facts, we present an updated version of it., as follow:

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k + w\frac{\gamma_k^T\delta_k}{\delta_k^T\mu_k}\mu_k \tag{19}$$

We now modify and derive the new modified quasi-Newton equation. Multiplying both side equation (19) by δ_k^T, we have:

$$\delta_k^T Q_{k+1}\delta_k = \delta_k^T\gamma_k + w\ \gamma_k^T\delta_k = (1+w)\gamma_k^T\delta_k \tag{20}$$

Thus, the equation (20) is equivalent to:

$$\frac{1}{(1+w)}\delta_k^T Q_{k+1}\delta_k = \delta_k^T\gamma_k \tag{21}$$

It yields

$$\delta_k^T Q_{k+1}\delta_k = \delta_k^T \gamma_k - \frac{w}{(1+w)}\delta_k^T g_k \tag{22}$$

As a result, one alternative for a fair approximation to Q_{k+1} is:

$$\delta_k^T B_{k+1}\delta_k = \delta_k^T \gamma_k - \frac{w}{(1+w)}\delta_k^T g_k \tag{23}$$

This results in the quasi-Newton equation:

$$\begin{aligned} B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k - \frac{w}{(1+w)}\frac{\delta_k^T g_k}{\delta_k^T \mu_k}\mu_k \\ = \gamma_k - \varpi\frac{\delta_k^T g_k}{\delta_k^T \mu_k}\mu_k \end{aligned} \tag{24}$$

where μ_k is any vector satisfying $\delta_k^T \mu_k \neq 0$ and $\varpi_k = w/(1+w)$ where we can chooses $w \in [0, 1, 2, 3]$. We provide the quasi-Newton equations that are deduced from (24) by various choice of the parameter $\mu_k = g_{k+1}$.

Now, we establish our modified BFGS algorithm as follows, and we call it new Algorithm.

Algorithm 2.1 (Modified BFGS algorithm)

Stage 1 : Give an $x_0 \in R^n$ and $H_0 = I.$ Let $k = 0.$

Stage 2 : Compute $g_k.$ If $\|g_k\| = 0,$ then stop.

Stage 3 : Compute a $d_k = -H_k g_k.$

Stage 4 : Find α_k by (4) and (5).

Stage 5 : Let $x_{k+1} = x_k + \alpha_k d_k.$ If $s_k^T \tilde{y}_k > 0,$ update H_{k+1} by the formula (9) and (24), otherwise let $H_{k+1} = H_k.$

Stage 6 : Set $k = k+1$ and go to stage 2.

In our analysis, the positive definite property plays a important roles.

Theorem 1: *If the curvature condition* $\delta_k^T \tilde{\gamma}_k > 0$ *is satisfied, then the updated matrix* B_{k+1} *is positive definite.*

Proof: By definition $\tilde{\gamma}_k$, one has that:

$$\delta_k^T \tilde{\gamma}_k = \delta_k^T \gamma_k - \frac{w}{(1+w)} \delta_k^T g_k \tag{25}$$

Since $\delta_k^T \gamma_k > 0$ and $-\delta_k^T g_k > 0$ are positive values, then $\delta_k^T \gamma_k - (w/(1+w))$ $\delta_k^T g_k > 0$, hence we have:

$$\delta_k^T \tilde{\gamma}_k > 0 \tag{26}$$

The last inequality implies B_{k+1} is positive definite.

3. Convergence analysis

To give the convergence result, the following assumptions are given.
Assumptions:

(1) Let Ω be the level set at the initial point x_0:

$$\Omega = \left\{ x \in R^n \,\middle|\, f(x) \le f(x_0) \right\} \tag{27}$$

is bounded.

(2) The f is smooth and the same order between its gradient and the variables holds, i.e.,

$$\left\| g(o) - g(c) \right\| \le L \left\| o - c \right\|, \quad \forall o, c \in \Omega. \tag{28}$$

The $\{f(x_k)\}$ is a non-increasing. Clearly, by Assumption 1, it holds that:

$$\lim_{k \to \infty} f_k = f^*. \tag{29}$$

(3) Continuity of $Q(x) = \nabla^2 f(x)$ indicate that a neighborhood exists Ω and two positive constants m and Msuch that:

$$m \left\| d \right\|^2 \le d^T Q(x) d \le M \left\| d \right\|^2, \; \forall d \in \Omega. \tag{30}$$

It holds for all $x \in S$ and $d \in R^n$. These assumptions are the same as those in [23].

Theorem 2: *If f satisfy Assumptions and* $\tilde{\gamma}_k$ *defined by (24). Then there's a constant M that says:*

$$\|\tilde{\gamma}_k\| \leq M\|\delta_k\|. \tag{31}$$

Proof: Let us start by definition $\tilde{\gamma}_k$ as:

$$B_{k+1}\delta_k = \tilde{\gamma}_k = \gamma_k - \frac{w}{(1+w)}\frac{\delta_k^T g_k}{\delta_k^T \mu_k}\mu_k \tag{32}$$

Chiefly, by using equation (28) and inequality (30), we obtain:

$$\begin{aligned} \|\tilde{\gamma}_k\| &\leq \left\|\gamma_k - \frac{w}{(1+w)}\frac{\delta_k^T g_k}{\delta_k^T . \mu_k}.\mu_k\right\| \\ &\leq \|\gamma_k\| + \frac{\left|(w/(1+w)\delta_k^T g_k\right|}{\|.\delta_k\|} \\ &\leq L\|\delta_k\| + \xi\|.\delta_k\| \end{aligned} \tag{33}$$

This inequalty implythe (31) with $\xi = \alpha_k(w/(1+w))$ and $M = L + \xi$.

Theorem 3: *Let* $\{x_k\}$ *be generated by new algorithm and there exist constants* $a_1 > 0$ *and* $a_2 > 0$ *we obtained:*

$$\|B_k\delta_k\| \leq a_1\|\delta_k\| \ \ and \ \ \delta_k^T B_k \delta_2 \geq a_2\|\delta_k\|^2, \tag{34}$$

For infinitely many k. Then:

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{35}$$

Proof: Theorem 3.2 of Basim [4].

However, if f is ageneral function, may happen $\delta_k^T\tilde{\gamma}_k < 0.$ For positive definiteness of the B_{k+1}, put :

$$K = \{k \ : \delta_k^T\tilde{\gamma}_k / \|\delta_k\|^2 \geq \beta\}, \ \ \beta > 0 \tag{36}$$

We will also need the important lemma, which is due to Byrd et al. see [10].

Lemma 1 : *Consider* $\{\delta_k\}$ *and* $\{\tilde{\gamma}_k\}$ *are sequences generated by new algorithm. Then, there exists m and Msuch that:*

$$\frac{\delta_k^T\tilde{y}_k}{\|\delta_k\|^2} \geq m \ \ and \ \ \|\tilde{y}_k\|^2 / \delta_k^T\tilde{y}_k \leq M. \tag{37}$$

Inequality, "(34) hold for at least $[t/2]$ *values of* $k \in \{1,2,.....,t\}$ *with any positive integer t"* .

Theorem 4: *Let* $\{x_k\}$ *be generated by new algorithm. Then the* $\liminf_{k\to\infty} \|g_k\| = 0$ *satisfies.*

Proof: By mathematical contradiction, putting that $\varepsilon > 0, \ \forall k \in K$.

$$\|g_k\| \geq \varepsilon. \tag{38}$$

Then it follows from (36) we get:

$$\delta_k^T \tilde{\gamma}_k \geq \beta \|\delta_k\|, \quad \forall k \in K. \tag{39}$$

Using equation (33) with equation (39) together, lead to :

$$\|\tilde{y}_k\|^2 / \delta_k^T \tilde{y}_k \leq M. \tag{40}$$

Applying lemma 1, to the matrix $\{B_k\}_{k\in K}$, obviously, there exist a_1 and a_2 then **(**34) hold satisfies for infinitely many *k*. Therefore, by theorem **3**, completes the proof.

4. Numerical results

The numerical comparisons between all variants of the BFGS approach based on our new quasi-Newton equation (24) and the classic quasi-Newton equation (7) are shown in Table 1. We used [14] as a source for our test problems. When one of the following three criteria is met, we stop the algorithm: "If $|f(x_k)| > 10^{-5}$, let $stop\,1 = |f(x_k) - f(x_{k+1})| \ / \ |f(x_k)|$; otherwise, let $stop\,1 = |f(x_k) - f(x_{k+1})|$. For every problem, if $\|g_k\| < \varepsilon$ or $stop\,1 < 10^{-5}$ are satisfied, the program will be stopped", as in [22].

Figures 1 and 2 indicate the performance of these four techniques for solving these 30 test problems in terms of the number of iterations (NI) and the number of function evaluations (NF), respectively, as measured by the profiles of Dolan and Moré [11]. So, the versions of the BFGS algorithms is encouraging and improved and we see that versions of the BFGS algorithms solves about (53-76)% and (50-76)% of the test problems with the fewest number of function evaluations and iterations.

Table 1

Comparison of the tested BFGS methods

		BFGS		$\varpi_k = 1/2$		$\varpi_k = 2/3$		$\varpi_k = 3/4$	
P.No.	n	NI	NF	NI	NF	NI	NF	NI	NF
Rose	2	35	140	5	37	5	59	48	335
Froth	2	9	26	11	30	6	11	5	16
Badscp	2	43	166	34	125	38	144	10	33
Badscb	2	3	30	3	30	3	30	3	30
Beale	2	15	50	14	42	11	35	9	30
Jensam	2	2	27	2	27	2	27	2	27
Helix	3	34	113	10	29	9	26	9	26
Bard	3	16	54	10	28	11	31	18	56
Gauss	3	2	4	2	4	2	4	2	4
Gulf	3	2	27	2	27	2	27	2	27
Box	3	2	27	2	27	2	27	2	27
Sing	4	20	60	14	44	20	60	15	47
Wood	4	19	61	6	21	6	21	6	21
Kowosb	4	21	65	30	115	10	31	11	33
Bd	4	17	54	5	17	5	17	5	17
Osb1	5	2	27	2	27	2	27	2	27
Biggs	6	25	72	33	98	22	61	12	64
Osb2	11	3	31	3	31	3	31	3	31
Watson	20	31	102	6	18	5	15	5	15
Singx	400	64	209	13	43	12	40	12	40
Pan1	.400	.2	.27	.2	.27	.2	.27	.2	.27
Pan2	.200	.2	.5	.2	.5	.2	.5	.2	5
Verdim	.100	.2	.27	.2	.27	.2	27	.2	.27
Trige	.500	9	.33	10	38	9	27	9	27
Bv	500	2	4	2	4	2	4	2	4
Ie	500	6	16	8	21	10	27	10	27
Band	500	57	281	8	28	6	43	10	60
Lino	500	2	4	2	4	2	4	2	4
Trig	500	3	7	3	7	3	7	3	7
Band	500	3	7	3	7	3	7	3	7
Total		453	1756	249	963	217	902	226	1101

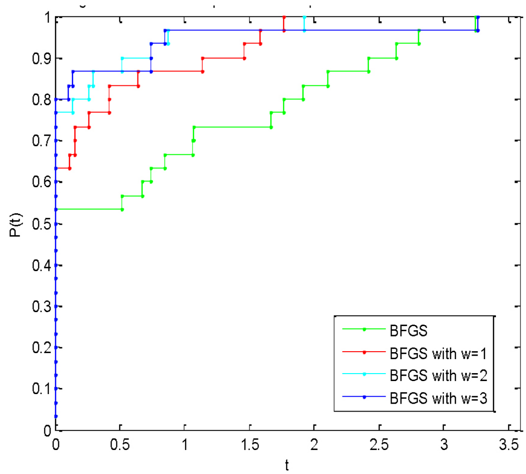

Figure 1

Performance profiles with respect to the number of iterations

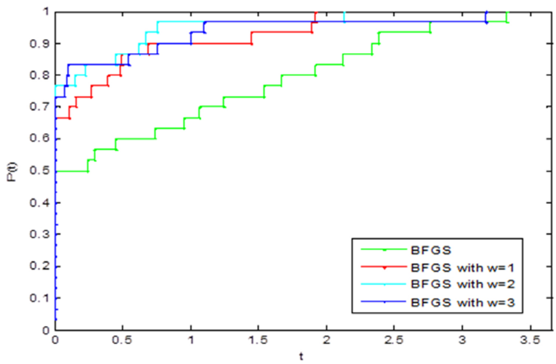

Figure 2

Performance profiles with respect to the number of function evaluations

5. Conclusions

We offer a new set of quasi-Newton equations in quasi-Newton techniques in this study. Under assumptions, we proved the global convergence property of our proposed quasi-Newton methods. Our numerical results have shown that our versions of the BFGS algorithms performs better than the BFGS method does, which requires less iterations and function evaluations. Moreover, preliminary experimental comparisons also indicate that our extension is very beneficial to the performance.

References

[1] Ashok D. B. and Tirupathi R. C., (2011), Optimization Concepts and Applications in Engineering, New York, Rowan University.

[2] Ataee D.T. and Reza M. P., (2015), A new regularized limited memory BFGS-type method based on modified secant conditions for unconstrained optimization problems, *J of Global Opt.*, 63, pp.709-728.

[3] Biglari F., Hassan M.A., and Leong W. J.,(2011),' New quasi-Newton methods via higher-order tensor models, *J. Comput. Appl. Math.* , (8) pp. 2412-2422.

[4] Basim A.H., (2020), A new type of quasi-Newton updating formulas based on the new quasi-Newton equation, *SIAM J. Numerical Algebra, Control, and Optimization*. 10, 2, pp.227-235.

[5] Basim A. Hassan and Mohammed W. T., (2020),'' A New Variants of Quasi-Newton Equation Based on the Quadratic Function for Unconstrained Optimization, I*ndonesian Journal of Electrical Engineering and Computer Science*, 19,(2), pp.701–708.

[6] Basim A.H. and Ghada M., (2020),'' A New Quasi-Newton Equation on the Gradient Methods for Optimization Minimization Problem, *Indonesian Journal of Electrical Engin. and Computer Science*, 19,(2), pp. 737-744.

[7] Basim A.H. and Ranen M. Sulaiman (2020),'' Using a New Type Quasi-Newton Equation for Unconstrained Optimization'', 7th International Engineering Conference Research & Innovation amid Global Pandemic, Erbil, Iraq, pp.188-122.

[8] Basim A. Hassan, Zeyad M. A., Hawraz N. J., (2019), A descent extension of the Dai - Yuan conjugate gradient technique, Indonesian *Journal of Electrical Engineering and Computer Science*, pp. 661-668.

[9] Byrd R, Nocedal J,YuanY.,(1987), Global convergence of a class of quasi-Newton methods on convex problems. *SIAM J. Numer Anal*, 1171-1189.

[10] Chen L.H., Deng N.Y., and. Zhang J.Z, (2006),'A modified quasi-Newton method for structured optimization with partial information on the Hessian, Comput. Optim. Appl. 35, pp. 5-18.

[11] Dolan E. and Mor´e J.J., (2002), Benchmarking optimization software with performance profiles, Math. Program.,91, pp.201-213.

[12] Hawraz N. Jabba and Basim A. Hassan, (2021), Two-versions of descent conjugate gradient methods for large-scale unconstrained optimization, *Indonesian Journal of Electrical Engineering and Computer Science* Vol. 22, No. 3, pp. 1643-1649.

[13] Jian H., Zong-Chuan W. and Qing C., (2018), An unconstrained optimization reformulation for the nash game, *Journal of Interdisciplinary Mathematics*, pp. 1303-1307.

[14] More J., Garbow B., and Hillstrome K., (1981),' Testing unconstrained optimization software, ACM Trans. Math. Software, 7, pp. 17-41.

[15] Mohd A. H., Mustafa M. and Leong W. J., (2014), ' BFGS Method: A New Search Direction ', Sains Malaysiana 43(10), pp. 1591–1597.

[16] Pater M. and P. Kaelo, (2019), A convergence modified HS-DY hybrid conjugate gradient method for unconstrained optimization problems. *Journal of Information and Optimization Sciences*, pp.97-113.

[17] Razieh D., Narges B. & Mohammad M., (2019), A new modified BFGS method for solving systems of nonlinear equations, *Journal of Interdisciplinary Mathematics,* No. 1, pp. 75-89.

[18] Toint PL (1986) Global convergence of the partioned BFGS algorithm for convex partially separable opertimization. Math Program 36:290–306.

[19] Li D. and Fukushima M.,(2001),' A modified BFGS method and its global convergence in nonconvex minimization, *J. Comp. Appl. Math.,* pp.15-35.

[20] Wei Z., Li G. and Qi L., (2006),' New quasi-Newton methods for unconstrained optimization problems. Math. Program. *Applied Mathematics and Computation,* 175, pp. 1156–1188.

[21] Yuan G, Wei Z (2010) Convergence analysis of a modified BFGS method on convex minimizations. Comput Optim Appl 47:237–255.

[22] Yuan, Y., Sun, W., (1999),' Theory and Methods of Optimization. Science Press of China, Beijing.

[23] Zhang JZ, Deng NY, Chen LH (1999) New quasi-Newton equation and related methods for unconstrained optimization. *J Optim Theory Appl* 102:147–167.

Received July, 2021

Revised February, 2022