ZANCO Journal of Pure and Applied Sciences
The official scientific journal of Salahaddin University-Erbil
https://zancojournals.su.edu.krd/index.php/JPAS

ISSN (print):2218-0230, **ISSN (online):** 2412-3986, **DOI:** **http://dx.doi.org/10.21271/zjpas**

RESEARCH PAPER

A New Transformed Biggs 's Self-Scaling Quasi-Newton Method for Optimization

Basim A. Hassan 1, Hawraz N. Jabbar 2

1- Department of Mathematics, College of Computers Sciences and Mathematics, Mosul University, Mosul, Iraq.

2- Department of Mathematics, College of Sciences, Kirkuk University, Kirkuk, Iraq.

A B S T R A C T:

In our work, we have proposed a new transformation Biggs's self-scaling Quasi-Newton update which is based on the simple idea of approximation for the inverse Hessian matrix. This transformation has implemented both theoretically and numerically and tested on some well-known test cases. Numerical experiments indicate that this transformation is more effective than the standard BFGS-method.

KEY WORDS: Quasi Newton Method,Self-Scaling Variable Metric,Global convergence
.DOI: http://dx.doi.org/10.21271/ZJPAS.31.s2.1.
ZJPAS (2019) , 31(s2);1-5 .

INTRODUCTION :

Quasi-Newton (variable metric) techniques are awfully helpful implements for solving unrestrained problems:

$$\min \{ f(x) \mid x \in R^n \}. \qquad (1)$$

where f is twice continuously differentiable function of a large number of variables n. At iteration k of a quasi-Newton method a search direction d_k is found by :

$$d_k = -B_k^{-1} g_k, \qquad (2)$$

where B_k approximates, in some sense, the Hessian matrix $\nabla^2 f(x_k)$. A line search is then occurrence $x_{k+1} = x_k + \alpha_k d_k$, for some α_k that fulfills the line search standard. The step size is calculated by implementing some line pursuit for instance the accurate line search where the form is:

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T G d_k}. \qquad (3)$$

Information at this new point is used to engender a new estimated Hessian matrix B_{k+1}. Quasi-Newton (QN) methods require B_k non negative definite and fulfilling the QN-equation: $B_{k+1} s_k = y_k$ (4)

where

$$s_k = x_{k+1} - x_k = \alpha_k d_k \text{ and } y_k = g_{k+1} - g_k. \qquad (5)$$

One documented update formula is that the BFGS formula that updates B_{k+1} from B_k, y_k and s_k in the following way:

$$B_{k+1}^{BFGS} = B_k - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} \qquad (6)$$

* **Corresponding Author:**
Basim A. Hassan

E-mail: basimabas39@gmail.com

Article History:
Received: 06/05/2018
Accepted: 10/07/2018
Published: 12/05 /2019

If the inverse of B_k is denoted by H_k then the inverse update formula of (6) method is represented as:

$$H_{k+1}^{BFGS} = H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{s_k^T y_k} + \left[1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right] \frac{s_k s_k^T}{s_k^T y_k} \quad (7)$$

More details can be found in (Byatt D., Coope. D. and Price C. J., 2004, Hassan, M.A., June, L.W and Monsi, M., 2005).

Underneath, I will now change the derivation of Biggs's method and analyze its convergence.

2. NEW TRANSFORMED BIGGS'S SELF-SCALING QUASI-NEWTON

For general function, corresponds to Hessian matrix, i.e. we are using 2nd derivatives. Then we have the following relation:

$$s_k^T \nabla^2 f_{k+1} s_k = 4 s_k^T g_k + 2 s_k^T g_{k+1} + 6\left[f(x_k) - f(x_{k+1})\right] \quad (8)$$

The over situation is called interpolation stipulation in **(**Fletcher, R., 1987**,** Hassan, M.A., June, L.W and Monsi, M., 2005**).**

That's why it is balanced to need that the estimated Hessian gratify condition:

$$s_k^T B_{k+1} s_k = 4 s_k^T g_k + 2 s_k^T g_{k+1} + 6\left[f(x_k) - f(x_{k+1})\right] . \quad (9)$$

Biggs (Biggs, M.C. , 1971, Biggs, M.C. , 1973) provides the update of H_{k+1} having the estimate ρ_k selected hence that (9) holds. The esteemed value of ρ_k is given by:

$$\rho_k = \frac{4 s_k^T g_k + 2 s_k^T g_{k+1} + 6\left[f(x_k) - f(x_{k+1})\right]}{s_k^T y_k} \quad (10)$$

In tidy to satisfy (4), the BFGS technique is obtain by:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k} + \rho_k \frac{y_k y_k^T}{s_k^T y_k} \quad (11)$$

See **(**Hassan, M.A., June, L.W and Monsi, M., 2005**,** Hassan, M.A., June, L.W and Monsi, M., 2006**)** for any other enquiries.

Thus we can obtain another modified parameter from (10) by considering the following relation:

$$s_k^T B_{k+1} s_k = y_k^T s_k + 6(f_k - f_{k+1}) + 3(g_{k+1} + g_k)^T s_k \quad (12)$$

From some algebra, we get:

$$s_k^T B_{k+1} s_k = \frac{1}{2} y_k^T s_k + 3(f_k - f_{k+1}) + \frac{3}{2} g_{k+1}^T s_k + g_k^T s_k . \quad (13)$$

Now, equation (10) can be rewritten as:

$$\rho_k = \left[\frac{1}{2} y_k^T s_k + 3(f_k - f_{k+1}) + \frac{3}{2} g_{k+1}^T s_k + g_k^T s_k\right] / y_k^T s_k . \quad (14)$$

He showed that a modified QN-algorithm could be writes as follows:

$$H_{k+1} = H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{s_k^T y_k} + \left[\frac{1}{\rho_k} + \frac{y_k^T H_k y_k}{s_k^T y_k}\right] \frac{s_k s_k^T}{s_k^T y_k} \quad (15)$$

Now we are ready to state the steps of the new BFGS algorithm:

2.1. New Algorithm:

Step 1: Data $x_0 \in R^n$ and $H_0 = I$. Let $k = 0$.

Step 2: Stop if $\|g_k\| = 0$.

Step 3: Determine d_{k+1} by :

$$d_k = - H_k g_k .$$

Step 4: Find α_k satisfying the :

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

Step 5: Let the next iterative be

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

Step 6: Update H_0 for times to get H_{k+1} by (15) and (14).

Step 7: Let $k = k + 1$. Go to step 2.

3. GLOBAL CONVERGEGNCE RESULT

Throughout this section, the subsequent assumptions are going to be thought of concerning the target perform f .

Assumption 3.1. Let G be the matrix of second derivatives of f .

i. The objective function f is smooth in a neighborhood N of the level set" $D = \{x \in R^n : f(x) \le f(x_1)\}$ "and bounded below in R^n.

ii. The gradient is Lipschitz continuous, i.e., there exists a constant $L > 0$ such that:

$$\|\nabla f(x_1) - \nabla f(x_2)\| \le L\,\|x_1 - x_2\| \;; \quad \forall\; x_1, x_2 \in N \qquad (16)$$

For more details can be found in (Farzin M., Abu Hassan M. and Wah J., 2011).

Let θ_k mean the slant between s_k and $B_k s_k$, i.e.:

$$\cos\ \theta_k = \frac{s_k^T B_k s_k}{\|s_k\|\|B_k s_k\|} = -\frac{s_k^T g_k}{\|s_k\|\|g_k\|}\,. \qquad (17)$$

Then there are constants $a_1 > 0$ and $a_2 > 0$ such that :

$$a_2\|g_k\|\cos\ \theta_k \le \|s_k\| \le a_1\|g_k\|\cos\ \theta_k \qquad (18)$$

For more details can be found in (Byrd R.H., Nocedal J. and Yuan Y., 1987).

Theorem 3.1.

Let x_0 be a initial point for which f satisfies

Assumption 3.1. Think $\{x_k\}$ the series of points generated by the updating design $x_{k+1} = x_k + \alpha_k d_k$ where the sequence $\{B_k\}$ is generated by new algorithm and α_k satisfies the Wolfe conditions $(4-5)$. Then:

$$\sum_{k\ge 1}^{\infty}\|g_k\|^2\ \cos^2\theta_k < \infty\,. \qquad (19)$$

Proof:

From (1) we have:

$$\begin{aligned}&\sum_{k=1}^{\infty}\left(f(x_k) - f(x_{k+1})\right) = \\ &\lim_{N\to\infty}\sum_{k=1}^{N}\left(f(x_k) - f(x_{k+1})\right) = \\ &\lim_{N\to\infty}\left(f(x_1) - f(x_{k+1})\right) = f(x_1) - f^*\end{aligned} \qquad (20)$$

Thus:

$$\sum_{k=1}^{\infty}\left(f(x_k) - f(x_{k+1})\right) \le +\infty \qquad (21)$$

Which combing with

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta_1 \alpha_k d_k^T g_k\ , \qquad (22)$$

Yields:

$$\sum_{k=1}^{\infty}(-\alpha_k d_k^T g_k) \le +\infty\,. \qquad (23)$$

Then from (23) we have:

$$\begin{aligned}\infty &> \sum_{k=1}^{\infty}(-g_k^T d_k) \ge \sum_{\forall k}(-g_k^T d_k) \\ &= \sum_{\forall k}\frac{1}{\alpha_k}s_k^T B_k s_k = \sum_{\forall k}\alpha_k\frac{\|g_k\|^2}{\|B_k s_k\|^2}s_k^T B_k s_k\end{aligned} \qquad (24)$$

From (17) and (18) we get:

$$\frac{\|B_k s_k\|^2}{s_k^T B_k s_k} = \frac{\|B_k s_k\|}{\|s_k\|\cos\ \theta_k} = \alpha_k\frac{\|g_k\|}{\|s_k\|\cos\ \theta_k} \qquad (25)$$

This implies that:

$$\begin{aligned}\infty &> \sum_{k=1}^{\infty}(-g_k^T d_k) \ge \sum_{\forall k}(-g_k^T d_k) \\ &= \sum_{\forall k}\alpha_k\|g_k\|^2\frac{s_k^T B_k s_k}{\|B_k s_k\|^2} = \sum_{\forall k}\alpha_k\|g_k\|^2\frac{\|s_k\|\cos\ \theta_k}{\alpha_k\|g_k\|} \\ &= \sum_{k=1}^{\infty}\|g_k\|\|s_k\|\cos\ \theta_k\end{aligned} \qquad (26)$$

Which concludes the proof

4. NUMERICAL EXPERIMENTS

In this section, the statement some numerical experiments. We test Algorithm 2.1 with different value of ρ_k and the standard BFGS method on some problems in

(More J., Garbow B., and Hillstrome K., 1981). All codes be written in Matlab code. The stop rule " 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 each problem, if $\|g_k\| < \varepsilon$ or $stop\ 1 < 10^{-5}$ was satisfied, for more details can be found in (Yuan, Y., 1991). We set parameters in Algorithm 2.1 as follows: $\delta = 0.001$ and $\sigma = 0.9$".

Table 1 lists numerical results : "In Table 1, "problem" and "n" stand for the test function name and the dimension of the test function, respectively. " NI / NF" are the total number of the iterations and the function evaluations, respectively".

Table 1 : Comparison of different BFGS-algorithms with different test functions and different dimensions

		BFGS algorithm		**Biggs's algorithm**		**M Biggs's algorithm**	
P. No.	**n**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**
1	2	35	140	6	64	4	35
2	2	9	26	3	8	3	31
3	2	43	166	16	299	4	34
4	2	3	30	3	7	3	30
5	2	15	50	6	88	15	48
6	2	2	27	2	27	2	27
7	3	34	113	10	146	16	88
8	3	16	54	8	117	16	50
9	3	2	4	2	4	2	4
10	3	2	27	2	27	2	27
11	3	2	27	2	27	2	27
12	4	20	60	3	8	39	122
13	4	19	61	3	8	3	31
14	4	21	65	10	74	19	58
15	4	17	54	3	8	3	31
16	5	2	27	2	27	2	27
17	6	25	72	12	177	9	71
18	11	3	31	3	31	3	31
19	20	31	102	9	76	43	132
20	400	64	209	3	8	3	31
21	400	2	27	2	27	2	27
22	200	2	5	2	5	2	5
23	100	2	27	2	27	2	27
24	500	9	33	29	99	29	99
25	500	2	4	2	4	2	4
26	500	6	61	7	65	7	19
27	500	57	281	3	8	3	31
28	500	2	4	2	4	2	4
29	500	3	7	3	30	3	7
30	500	3	7	3	30	3	7
Total		453	1756	163	1530	248	1165

Problems numbers indicant for : "1. is the Rose, 2. is the Froth, 3. is the Badscp, 4. is the Badscb, 5. is the Beale, 6. is the Jensam, 7. is the Helix, 8. is Bard, 9. is the Gauss, 10. is the Gulf, 11. is the Box, 12. is the Sing, 13. is the Wood, 14. is the Kowosb, 15. is the Bd. 16. is the Osb1, 17. is the Biggs, 18. is the Osb2, 19. is the Watson, 20. is the Singx, 21. is the Pen1, 22. is the Pen2, 23. is Vardim, 24. is the Trig, 25. is the Bv, 26. is the Ie, 27. is the Band, 28. is the Lin, 29. is the Lin1, 30. is the Lino".

4. CONCLUSIONS

Supported a Bigg's Self-scaling technique; we tend to derive a replacement Self-scaling technique. World convergence of those techniques are established. Finally, a number of the numerical results are rumored, that explain the usefulness of the new self-scaling technique.

REFERENCES

BIGGS, M.C. 1971, Minimization algorithms making use of non- quadratic properties of the objective function. *Journal of the Institute of Mathematics and Its Application,* 8, 315-327.

BIGGS, M.C. 1973. A note on Minimization algorithms *making use of nonquadratic properties of the* objective function. *Journal of the Institute of Mathematics and Its Application*, 12, 337-338.

BYRD R.H., NOCEDAL J. and YUAN Y. 1987. Global convergence of a class of variable metric algorithms, *SIAM Journal on Numerical Analysis*, 24, 1171-1190.

BYATT D., COOPE. D. and PRICE C. J. 2004. Performance of BFGS methods with limited precision second order information, *ANZIAM J.* ,45, 511–522.

Farzin M., Abu Hassan M. and Wah J. 2011. Improved Hessian approximation with modified secant equations for symmetric rank-one method, Journal of Computational and Applied Mathematics, 235,pp. 2423–2431.

FLETCHER, R. 1987. Practical Methods of Optimization .*John Wiley and Sons*, Chi Chester (New York).

HASSAN, M.A., JUNE, L.W and MONSI, M. 2005. Modified of the limited method BFGS algorithm for large-scale nonlinear optimization. Mathematics *Journal Okayama Univ*., 47, 175-188.

HASSAN, M.A., JUNE, L.W and MONSI, M. 2006 Convergence of the Modified BFGS Method, *MATEMATIKA*, 22, 17-24.

MORE J., GARBOW B., and HILLSTROME K. 1981. Testing unconstrained optimization software, *ACM Trans. Math. Software*, 7, 17-41.

YUAN, Y. 1991. A modified BFGS algorithm for unconstrained optimization. *IMA Journal Numerical Analysis*, 11, 325-332..