



ISSN: 0067-2904

New Class of Rank 1 Update for Solving Unconstrained Optimization Problem

Saad Shakir Mahmood*, Jaafer Hmood Eidi

Mathematics Department, College of Education, Al-Mustansiriya University, Baghdad – Iraq

Received: 19/2/2021

Accepted: 18/4/2021

Abstract:

The focus of this article is to add a new class of rank one of modified Quasi-Newton techniques to solve the problem of unconstrained optimization by updating the inverse Hessian matrix with an update of rank 1, where a diagonal matrix is the first component of the next inverse Hessian approximation, The inverse Hessian matrix is generated by the method proposed which is symmetric and it satisfies the condition of modified quasi-Newton, so the global convergence is retained. In addition, it is positive definite that guarantees the existence of the minimizer at every iteration of the objective function. We use the program MATLAB to solve an algorithm function to introduce the feasibility of the proposed procedure. Various numerical examples are given.

Key words: Unconstrained optimization, Hessian matrix, modified Quasi-Newton condition, Positive definite.

الفئة الجديدة من تحديث الرتبة 1 لحل مسألة المثلية غير المقيدة

سعد شاكر محمود* , جعفر حمود عيدي

قسم الرياضيات, كلية التربية, الجامعة المستنصرية, بغداد- العراق

الخلاصة

تركز هذه المقالة على إضافة تحديث جديد من الرتبة 1 من تقنيات شبه نيوتن المحورة لحل مشكلة المثلية غير المقيدة عن طريق تحديث مصفوفة هيسين المعكوسة بتحديث الرتبة الأولى, حيث تكون المصفوفة القطرية هي المكون الأول للتقريب التالي لمعكوس مصفوفة هيسين, المصفوفة الناتجة عن الطريقة المقترحة متناظرة وتفي بشرط شبه نيوتن المحور, لذلك يتم الاحتفاظ بالتقارب العام لطرق شبه نيوتن المحورة. بالإضافة إلى ذلك إن المصفوفة العكسية الناتجة عن الطريقة المقترحة هي مصفوفة موجبة محددة لذلك تكون نقطة النهاية الصغرى موجودة في كل تكرار. قمنا باستخدام برنامج MATLAB لحل الأمثلة العددية لبعض الدوال التطبيقية لنبين فاعلية هذه الطريقة المقترحة وذلك بأعطاء أمثلة متعددة. الكلمات المفتاحية: المثلية غير المقيدة, مصفوفة هيسين, شرط شبه نيوتن, المحدد الموجب

1. Introduction:

There are many efforts to achieve a better approximation of the Hessian matrix. The modified Quasi-Newton (secant) condition is proposed by Zhang J, and Xu Ch. [1]. They use both gradient and function meaning knowledge to achieve a higher order accuracy in approximating the second curvature of the objective function.

*Email: saadshakirmahmood@uomustansiriyah.edu.iq

In [2], the authors find a revamped Broyden family composed of BFGS (Broyden, Fletcher, Goldfarb, and Shanno), which analogs to the changing that suggested by authors in [1,3]. The BFGS update was updated on the basis of the new Quasi-Newton condition $B_{k+1}s_k = y_k^*$ where $y_k^* = y_k + A_k s_k$, and A_k is a matrix.

The symmetric rank one update based on the extended quasi-Newton condition that was updated by authors in [4], and they provided that the method retains the symmetric and positive definite property and they also provided the proposed method's global and superlinear convergence.

In [5], the authors suggested the Broyden update to guarantee the positive definite property of Hessian matrix and to give the global convergence of the proposed process. In [6], the authors use the Taylor theorem for the modified BFGS approach. It was developed for solving system of nonlinear equations. In [7], the authors presented the modified BFGS approach for solving the nonlinear system. Updated quasi-Newton equation was suggested in [7] to get a more detailed approximation of the second curvature of the objective function. In [8], the authors introduced the modified DFP (Davidon–Fletcher–Powell) update based on modified quasi-Newton condition and provided the global and superlinear convergence of the proposed method. A new family of updated BFGS updates was proposed in [9] to resolve the unconstrained issue of optimization for non-convex functions based on a new modified weak Wolfe-Powell line search technique. In [10], the authors proposed the modified BFGS update (H-version) by updating the vector s (next solution-current solution) and they provided that the proposed method preserve the strong positive definite property and globally convergent.

The method of Newton $x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)$ is efficient because it uses the Hessian matrix that provides the useful curvature data, however the computational efforts of the Hessian matrix are very costly or it is difficult to test the Hessian matrix even though the Hessian matrix is not analytically usable this refers to a class of techniques that use only the values of the equation and the gradients of the objective function which are closely related to the method of Newton. Even the Hessian matrix is not analytically available, or the Hessian's estimation is challenging. These refer to a class of techniques that use only the values of the equation and the gradients of the objective function that are closely related to the method of Newton. The Quasi-Newton system is a class of techniques that does not need to calculate the Hessian method. Quasi-Newton is a class of methods that does not need to calculate the Hessian matrix, but it generates a series of Hessian approximations at the same time preserves a high rate of convergence. In this work we will construct a Hessian approximation instead of evaluating the Hessian matrix $\nabla^2 f(x)$, as well as we will also create a Hessian approximation. We predict that the Hessian matrix sequence of approximates $\{B_k\}$ has positive definiteness, it has a $p_k = -(B_k)^{-1} \nabla f(x^k)$ downward direction, and it works like the process of Newton. Moreover, it is also important for its measurement to be effortless to handle.

Consider the unconstrained optimization problem

$$\min_{x \in R^n} f: R^n \rightarrow R \tag{1}$$

The objective function f is assumed to be twice continuous and differentiable in a convex open set D , moreover f is uniformly convex. The most efficient quasi-Newton method is the (BFGS) method, which is introduced independently by Broyden, Fletcher, Goldfarb, and Shanno in [10]. Note that the Hessian approximation B_{k+1} can be modified in the BFGS approach

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \text{ and } H_{k+1} = \left(I - \frac{s_k y_k^T}{y_k^T s_k} \right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k} \right) + \frac{s_k s_k^T}{y_k^T s_k}$$

Where $B_{(k+1)}$ is the next Hessian matrix approximation, $H_{(k+1)}$ is the next inverse Hessian matrix approximation, $s_k = x^{k+1} - x^k$, x^k is the current solution, x^{k+1} is the next solution, $y_k = \nabla f(x^{k+1}) - \nabla f(x^k)$, and ∇f is the gradient of the objective function f ,

BFGS update is considered a common update to solve the unregulated problem of optimization where the BFGS update only preserves the positive definite property if $y_k^T s_k > 0$ at each iteration. The problem is to solve problem (1) by generating a symmetrical and positive definite inverse Hessian matrix sequence that satisfies the quasi-Newton condition at any iteration.

2. New class of Rank 1 Update

Consider problem (1) with next inverse Hessian matrix approximation H_{k+1} . The modified quasi-Newton equation is given as follows:

$$H_{k+1}y_k = s_k^* \tag{2}$$

The solution of Eq. (2) for H_{k+1} such that H_{k+1} is a diagonal matrix which is given in [10]

$$H_{k+1} = \frac{y_k^T s_k}{y_k^T y_k} I \tag{3}$$

Where I is the identity matrix and the best choice of the vector s_k^* which satisfy the equation

(2) is $s_k^* = \frac{y_k^T s_k}{y_k^T y_k} y_k$

Eq. (3) satisfies the modified quasi-Newton condition, symmetric, and positive definite if $y_k^T s_k > 0$, however, at every iteration it is only in the diagonal form, thus it is very minimal to solve the optimization problem with large dimensions. In this case, computation is very minimal, hence it is useful to solve the optimization problem with large dimensions such that the computation in this case is very minimum. Since the Hessian matrix is not always in the diagonal form, so Eq. (3) can be modified to get an update which have the same property of Eq. (3), but it is not diagonal so that it is called the new class of rank 1 update for unconstrained optimization.

Let us consider the rank 1 update formula

$$H_{k+1} = H_k + u_k v_k^T \tag{4}$$

Where $u_k, v_k \in R^n, n$ represents the dimension of the problem.

If H_k changed by $\frac{y_k^T s_k}{y_k^T y_k} I$ then, the new class of rank 1 update can be written as follows:

$$H_{k+1} = \frac{y_k^T s_k}{y_k^T y_k} I + u_k v_k^T \tag{5}$$

To determine u_k and v_k , let us consider the following equation:

$$0 = 1 - 1 = \frac{y_k^T s_k}{y_k^T s_k} - \frac{y_k^T H_k y_k}{y_k^T H_k y_k} = \left(\frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right)^T y_k = \left(\frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right) \left(\frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right)^T y_k,$$

and if $w_k = \frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k}$ then we have the following equation

$$w_k^T y_k = 0 \tag{6}$$

that leads to the solution of the Eq. (2) as follows:

$$H_{k+1} = \frac{y_k^T s_k}{y_k^T y_k} I + \left(\frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right) \left(\frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right)^T \tag{7}$$

From Eq. (7), it is clear that $u_k = v_k = \frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k}$

Then Eq.(7) is called the new rank 1 update formula.

3. Algorithm of the method

Step 1 Consider a starting point $x^0 \in R^n$, an initial symmetric positive definite matrix $H_0 \in R^{n \times n}$, an error tolerance $\epsilon > 0$; $k = 0$.

Step 2 Compute $\nabla f(x^k)$. If $\|\nabla f(x^k)\| < \epsilon$ then, the solution is x^k and stop, otherwise, compute $r_k = -H_k \nabla f(x^k)$

Step 3 Do exact or Inaccurate line search to identify a step size $\beta_k > 0$ such that $f(x^k + \beta_k r_k) < f(x^k)$.

Step 4 Set $x^{k+1} = x^k + \beta_k r_k$

Step 5 Set $s_k = x^{k+1} - x^k$ and $y_k = \nabla f(x^{k+1}) - \nabla f(x^k)$

Step 6 Compute H_{k+1} from equation (7)

Step 7 $k = k + 1$ and go to step 2

Theorem 1:

The new rank 1 update which is given in Eq. (7) is symmetric, Positive definite, and satisfies the modified quasi-Newton equation.

Proof:

Let $w_k = \frac{s_k}{y_k^T s_k} - \frac{H_k y_k}{y_k^T H_k y_k}$, hence Eq. (7) becomes

$H_{k+1} = \frac{y_k^T s_k}{y_k^T y_k} I + w_k w_k^T$, and since the first term is in the diagonal form then, Eq. (7) is

symmetric. By Eq. (3) and Eq. (6), it is clear that Eq.(2) is hold and hence the modified quasi-Newton condition is satisfied. Finally we need only to prove that $y_k^T s_k > 0$, that means

$$\begin{aligned} y_k^T s_k &= (\nabla f(x_{k+1}) - \nabla f(x_k))^T s_k \\ &= \nabla f(x_{k+1})^T s_k - \nabla f(x_k)^T s_k \\ &= \beta_k \nabla f(x^{k+1}) r_k + \beta_k r_k^T H_k^{-1} r_k \\ &= \beta_k \frac{\partial f(x^{k+1})}{\partial \beta_k} + \beta_k r_k^T H_k^{-1} r_k \end{aligned}$$

$= 0 + \beta_k r_k^T H_k^{-1} r_k > 0$, by exact line search and the positive definiteness of inverse of Hessian matrix. Hence H_{k+1} is generated by the new rank 1 update which is positive definite.

Powell developed the global convergence of the Quasi-Newton methods in [11]. These observations were applied to the small class of Broyden by Byrd, Nocedal and Yuan in [11].

Lemma 2: [10]

Let $f: R^n \rightarrow R$ that satisfies the assumption of problem (1), then

$$\frac{\|s_k\|}{\|y_k\|}, \frac{\|y_k\|}{\|s_k\|}, \frac{s_k^T y_k}{\|s_k\|^2}, \frac{s_k^T y_k}{\|y_k\|^2}, \text{ and } \frac{\|y_k\|^2}{s_k^T y_k} \text{ are bounded.}$$

Note that, from assumption of problem (1), and since $s_k^T B_k s_k$, and $y_k^T s_k$ are bounded from lemma (2) we have

$$\frac{\|y_k\|}{\|s_k\|}, \frac{\|s_k\|}{\|y_k\|}, \frac{y_k^T s_k}{\|y_k\|^2}, \frac{y_k^T s_k}{\|s_k\|^2}, \frac{\|s_k\|^2}{y_k^T s_k}, \frac{y_k^T H_k y_k}{\|y_k\|^2}, \frac{y_k^T B_k y_k}{\|y_k\|^2}, \frac{\alpha_k s_k y_k}{\|y_k\|^2}, \text{ and } \frac{\alpha_k^2 s_k^T B_k s_k}{\|y_k\|^2} \text{ are bounded,}$$

where α_k is any constant.

Lemma 3, [11]

Under exact line search, $\sum \|s_k\|^2$ and $\sum \|y_k\|^2$ are convergent.

Note that, from lemma 2 and lemma 3, clearly that $\sum \|s_k^T y_k s_k\|^2$ is convergent, which gives

that $\frac{\|s_k^T y_k s_k\|^2}{\|y_k\|^2}$ is convergent and bounded. Hence by the quasi-Newton methods (the proposed method) is global convergent.

4. Numerical Examples

In this section, numerical examples are solved by using the proposed method to introduce the effectiveness of the method, where the chosen problem represents an application functions in several science. The problems solved by using the proposed method with different dimensions and different starting points with initial inverse Hessian approximation equal to identity matrix. The problems are solved by using MATLAB program .

Example 1:[11]

$$1. \min f(x) = (1 - x_1)^2 + (x_2 - x_1)^2, H_0 = I$$

Table 1- Results of function 1 with different starting points

x^0	New rank one method		
	Iterations	f evaluation	x^*
$\begin{bmatrix} 5 \\ 9 \end{bmatrix}$	14	6.4285e-010	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$\begin{bmatrix} 50 \\ 19 \end{bmatrix}$	10	1.7209e-009	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$\begin{bmatrix} 21 \\ 128 \end{bmatrix}$	16	1.4847e-011	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$\begin{bmatrix} 1000 \\ 1234 \end{bmatrix}$	32	5.1670e-009	$\begin{bmatrix} 1 \\ 1.00001 \end{bmatrix}$

Example 2:[11]

2. $\min f = x_1^4 + x_1^3 + 10x_2^4 + x_2^3 + 10x_2^2 + 100x_3^4 + x_3^3 + 20x_3^2 + 1000x_4^4 + x_4^3 + 30x_4^2$,
 $H_0 = I$

Table 2- Results of function 2 with different starting points

x^0	New rank one method		
	Iterations	F evaluation	x^*
$\begin{bmatrix} 10 \\ 20 \\ 30 \\ 40 \end{bmatrix}$	32	-4.0752	$\begin{bmatrix} 0.0435 \\ -0.7456 \\ 0.3125 \\ 0.1222 \end{bmatrix}$
$\begin{bmatrix} 47 \\ 38 \\ 21 \\ 14 \end{bmatrix}$	47	-4.1420	$\begin{bmatrix} 0.0627 \\ -0.7457 \\ 0.3199 \\ 0.1230 \end{bmatrix}$
$\begin{bmatrix} 1050 \\ 2000 \\ 1800 \\ 1350 \end{bmatrix}$	46	-3.3643	$\begin{bmatrix} 0.0930 \\ 0.6700 \\ 0.3125 \\ 0.1223 \end{bmatrix}$
$\begin{bmatrix} 12345 \\ 23456 \\ 34567 \\ 45678 \end{bmatrix}$	44	-4.1384	$\begin{bmatrix} 0.0525 \\ -0.7455 \\ 0.3200 \\ 0.1222 \end{bmatrix}$

Example 3: [11]

3. $\min f(x) = \sum_{i=1}^n [(x_{2i-1} - 2)^4 + (x_{2i-1} - 2)^2 x_{2i}^2 + (x_{2i} + 1)^2]$, $H_0 = I$

Table 3- Results of function 3 with different starting points

x^0	New rank one method		
	Iterations	f evaluation	x^*
$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 9 \\ 10 \end{bmatrix}$	15	2.8807e-010	$\begin{bmatrix} 2 \\ -1 \\ \vdots \\ 2 \\ -1 \end{bmatrix}$
$\begin{bmatrix} 11 \\ 12 \\ \vdots \\ 19 \\ 20 \end{bmatrix}$	42	6.4966e-009	$\begin{bmatrix} 2 \\ -1 \\ \vdots \\ 2 \\ -1 \end{bmatrix}$

$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 19 \\ 20 \end{bmatrix}$	41	1.7220e-009	$\begin{bmatrix} 2 \\ -1 \\ \vdots \\ 2 \\ -1 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 39 \\ 40 \end{bmatrix}$	91	2.6314e-003	$\begin{bmatrix} 2 \\ -1 \\ \vdots \\ 2 \\ -1 \end{bmatrix}$

Example 4: [11]

4. $\min f(x) = \sum_{i=1}^n [(x_{2i-1} - 1)^2 + (x_{2i} - 11)^2 + (x_{2i-1} + x_{2i}^2 - 7)^2], H_0 = I$

Table 4- Results of function 4 with different starting points

x^0	New rank 1 method		
	Iterations	f evaluation	x^*
$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 9 \\ 10 \end{bmatrix}$	12	2.6864e-009	$\begin{bmatrix} 3 \\ 2 \\ \vdots \\ 3 \\ 2 \end{bmatrix}$
$\begin{bmatrix} 11 \\ 12 \\ \vdots \\ 19 \\ 20 \end{bmatrix}$	13	6.6255e-010	$\begin{bmatrix} 3 \\ 2 \\ \vdots \\ 3 \\ 2 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 19 \\ 20 \end{bmatrix}$	19	3.9368e-009	$\begin{bmatrix} 3 \\ 2 \\ \vdots \\ 3 \\ 2 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 39 \\ 40 \end{bmatrix}$	53	3.3691e-008	$\begin{bmatrix} 3 \\ 2 \\ \vdots \\ 3 \\ 2 \end{bmatrix}$

Example 5: [11]

Table 5- $\min f(x) = \sum_{i=1}^n [(n - \sum_{j=1}^n \cos x_j + i(1 - \cos x_i) - \sin x_i + e^{(x_i-1)^2}]$

x^0	New rank 1 method		
	Iterations	f evaluation	x^*
$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	4	7.5115e-008	$\begin{bmatrix} 0.0104 \\ 0.0083 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$	5	1.0552e-008	$\begin{bmatrix} 0.0039 \\ 0.0020 \\ 0.0010 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$	58	0.0784	$\begin{bmatrix} -0.0001 \\ 0 \\ -6.7413 \\ -44.4055 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$	19	0.03971	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ -11.7955 \\ -94.5720 \end{bmatrix}$

From previous examples, it is clear that the proposed method can solve all the problems with different dimensions and different starting points. The iterations and f evaluation explain that the proposed method is terminated at the minimizer of the objective function for all starting points.

5. Conclusions

The new rank one update was proposed to address the unconstrained question of optimization. This update satisfies the modified quasi-Newton condition and symmetric, moreover the positive definite property of inverse Hessian approximation holds. Four examples explain the affectivity of the proposed update.

Acknowledgements

The authors would like to thank the department of Mathematics, College of Education, University of Mustansiriyah, Baghdad-Iraq, for its support in the present work.

References

- [1] Zhang J. and Xu Ch, “ Properties and Numerical Performance of Modified quasi-Newton Methods Equations ,” *Elsevier, J. of Comp. and App. Math*, vol. 137, pp. 269-278, 2001.
- [2] Guo Q. and Liu J, “ Global convergence properties of two modified BFGS-type methods,” *Springer, J. of App. Math. and Comp*, vol. 23, pp. 311-319, 2007.
- [3] Yabe H, Ogasawara H and Yoshino M, “ Local and superlinear convergence of quasi-Newton methods based on modified secant conditions,” *Elsevier, J. of Comp. and App. Math*, vol. 205, no 1, pp. 617-32, 2007.
- [4] Mahmood S. S. and Farqad H, “ On Extended Symmetric Rank One Update for Unconstrained Optimization,” *J. of Education. Special issued*, pp. 206-220, 2017.
- [5] Mahmood S. S. Muhanah N. S., “ Symmetric and Positive Definite Broyden Update for Unconstrained Optimization,” *BSJ*, vol. 16, no.3, pp. 661-666, 2019.
- [6] Razieh D., Narges B. and Mohammad M. H., “ A new modified BFGS method for unconstrained optimization problems,” *Springer, J. of Comp. and App. Math*, vol. 37 pp. 5113-5125, 2018.
- [7] Razieh D., Narges B. and Mohammad M. H., “ A new modified BFGS method for solving system of nonlinear equations,” *Taylor and Francis LTD, J. of Interdisciplinary math*, vol. 22, pp. 75-89, 2019.
- [8] Mahmood S. S. and Shnywer Samira H., “ On Modified DFP Update for Unconstrained Optimization,” *American J. of App. Math*, vol. 5, no. 1, pp. 19-30, 2017.
- [9] Bojari S. and Eslachachi M.R., “ Global convergence of a family of modified BFGS methods under a modified weak-Wolfe-Powell line search for nonconvex functions,” *Springer. 4OR*, vol. 18, no. 2, pp. 219-244, 2020.
- [10] Mahmood S. S., “ Modified BFGS Update (H-Version) Based on the Determinant Property of Inverse of Hessian Matrix for Unconstrained Optimization,” *BSJ*, vol 17, no.3, pp. 994-1001, 2020.
- [11] Wenyu S. and Ya-Xiang Y., “ Optimization Theory and Method: Nonlinear Programming,” USA. 2006.