DOI: http://dx.doi.org/10.21123/bsj.2022.19.2.0297

Positive Definiteness of Symmetric Rank 1 (H-Version) Update for Unconstrained Optimization

Saad Shakir Mahmood ***Jaafer Hmood Eidi*** ***Jinan Adel Jasem***

Department of Mathematics, College of Education, University of Mustansiriyah, Baghdad, Iraq
Corresponding author: saadshakirmahmood@uomustansiriyah.edu.iq , drjaffarmath@uomustansiriyah.edu.iq , jinanadel78@uomustansiriyah.edu.iq
ORCID ID: https://orcid.org/0000-0001-5433-9710, https://orcid.org/0000-0002-9010-7527 , https://orcid.org/0000-0002-5128-1522

Received 24/4/2020, Accepted 3/12/2020, Published Online First 20/9/2021

Abstract:

Several attempts have been made to modify the quasi-Newton condition in order to obtain rapid convergence with complete properties (symmetric and positive definite) of the inverse of Hessian matrix (second derivative of the objective function). There are many unconstrained optimization methods that do not generate positive definiteness of the inverse of Hessian matrix. One of those methods is the symmetric rank 1(H-version) update (SR1 update), where this update satisfies the quasi-Newton condition and the symmetric property of inverse of Hessian matrix, but does not preserve the positive definite property of the inverse of Hessian matrix where the initial inverse of Hessian matrix is positive definiteness. The positive definite property for the inverse of Hessian matrix is very important to guarantee the existence of the minimum point of the objective function and determine the minimum value of the objective function.

Keywords: Hessian matrix, Positive definite, Quasi-Newton condition, Symmetric rank 1Update, Unconstrained optimization

Introduction:

Symmetric Rank 1 (SR1 H-version) update is important in theoretical research and practical computing. However, the drawback that SR1 (H-version) update does not retain the positive definiteness of updates hurts its performance in computing [1]. Fortunately, the drawback can be avoided if the modified quasi-Newton condition has been employed to modify the SR1 (H-version) update.

Zhang and Ch [2] introduced the modified quasi-Newton equation which uses both gradient and function value information in order to yield a higher order accuracy for approximating the second curvature of an objective function. Yabe, H and M [3] considered a modified Broyden family which includes the BFGS (Broyden–Fletcher–Goldfarb–Shanno) update. Guo and J [4] modified the BFGS update based on the new quasi-Newton equation, $B_{k+1}s_k = y_k + A_k s_k$,where A_k is a matrix.

Mahmood and H [5] Introduced the modified DFP (Davidon–Fletcher–Powell) update based on Zhang-Xu's condition and provided the global and superlinear convergence of the proposed method. Mahmood and S [6] proposed a modified Broyden update based on the positive definite property of Hessian matrix, via updating the vector y (the difference between the next gradient of the objective function and the current gradient of the objective function) and provided the global and superlinear convergence of the proposed method. Razieh, B and H [7] introduced the modified BFGS method for solving the system of non-linear equations by using Taylor theorem, this proposed method is derivative-free, so the gradient information is not needed at each iteration. Razieh, B and H [8] proposed a modified quasi-Newton equation to get a more accurate approximation of the second curvature of the objective function by using Chain rule. Then, based on this modified secant equation, they present a new BFGS method for solving unconstrained optimization problems. Bojari and R [9] proposed a new family of modified BFGS update to solve the unconstrained optimization problem for nonconvex functions based on a new modified weak Wolfe – Powell line search technique. Yuan [10] proposed a modified

BFGS algorithm which requires that the function value is matched, instead of the gradient value, at the previous iterate. This new algorithm preserves the global and local superlinear convergence properties of the BFGS algorithm.

In this research a modified update for the SR1 (H-version) update has been proposed to guarantees the positive definite property and preserves the symmetry property for the inverse of Hessian matrix via updating the vector s which represents the difference between the next solution and the current solution. The proof of convergence for the proposed method is given, and then tow numerical examples has been solved by the original SR1 (H-version) update and also solved by the proposed method.

Modified SRI (H-version) Update:

In this section the positive definite property for the inverse of Hessian matrix has been guarantee by updating the vector s_k. Then for this purpose let us consider the objective function $f: R^n \to R$ with the following assumptions:

i. f is twice continuously differentiable.

ii. f is uniformly convex, i.e. $\exists\ m_1, m_2 \in R^+$ $\ni m_1\|a\|^2 \le a^T\, \nabla^2 f\, a \le m_2\|a\|^2$,$\forall\ a \in R^n$

SR1 update (7) try to update the Hessian matrix by using the formula

$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}$, and by using Sherman-Morrison-Woodbury formula [11], the inverse of the Hessian matrix can be write as

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k} \qquad (1)$$

Which represent the solution of the quasi-Newton condition [12]

$$H_{k+1}\, y_k = s_k \qquad (2)$$

Where B_{k+1} is the next Hessian matrix, B_k is the current Hessian matrix, H_{k+1} is the next inverse of Hessian matrix, H_k is the current Hessian matrix, s_k is the difference between the current solution and the next solution($s_k = x_{k+1} - x_k$), and y_k is the difference between the current gradient and the next gradient of the objective function ($y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$).

Eq. 1 does not preserve the positive definite property because if $(s_k - H_k y_k)^T y_k < 0$ then, $Z^T H_{k+1} Z$ is not always positive for all $Z \in R^n$, that means there is no guarantee to minimize the objective function at each iteration, so if the current inverse of Hessian matrix is positive definite then, the next inverse of Hessian matrix may be not positive definite and hence this iteration must be deleted.

Now define:

$$s_k^* = \alpha_k s_k \qquad (3)$$

where $\alpha_k \in R$, and form Eq. 2

$$H_{k+1}\, y_k = s_k^* = \alpha_k s_k \qquad (4)$$

This is called the modified quasi-Newton condition.

The formula of inverse of Hessian matrix for the SR1 (H-version) update has been considered with replacing each s_k by $\alpha_k s_k$, in Eq. 1, and hence:

$$H_{k+1} = H_k + \frac{(\alpha_k s_k - H_k y_k)(\alpha_k s_k - H_k y_k)^T}{(\alpha_k s_k - H_k y_k)^T y_k} \qquad (5)$$

Set

$$w_k = (\alpha_k s_k - H_k y_k)$$

(6)

and by substitution Eq. 6 in Eq. 5, then

$$H_{k+1} = H_k + \frac{w_k w_k^T}{w_k^T y_k} \qquad (7)$$

Now, set $w_k^T y_k > 0$ and by Eq. 6, $(\alpha_k s_k - H_k y_k)^T y_k > 0$ and then, $\alpha_k > \frac{y_k^T H_k y_k}{y_k^T s_k}$

Now set

$$\alpha_k = 2\frac{y_k^T H_k y_k}{y_k^T s_k} \qquad (8)$$

Note that, one can choose another value of α_k but must be satisfies the inequality $\alpha_k > \frac{y_k^T H_k y_k}{y_k^T s_k}$.

Now to show that $\frac{y_k^T H_k y_k}{y_k^T s_k}$ is positive, where $s_k = x_{k+1} - x_k$, $x_{k+1} = x_k + \lambda_k p_k$, and the direction $p_k = -H_k \nabla f(x_k)$, and $\lambda_k > 0$, where λ_k is the step size [12], x_k is the current solution, x_{k+1} is the next solution, and $\nabla f(x_k) = \nabla f_k$ is the current gradient of the objective function *f*.

Now $\frac{s_k^T B_k s_k}{y_k^T s_k} = \frac{-\lambda_k s_k^T \nabla f_k}{s_k^T(\nabla f_{k+1} - \nabla f_k)} = \frac{-\lambda_k s_k^T \nabla f_k}{s_k^T \nabla f_{k+1} - s_k^T \nabla f_k}$, where $H_k = B_k^{-1}$ and $y_k^T s_k = s_k^T y_k$.

Since $s_k^T \nabla f_{k+1} = 0$ (conjugate direction property) [1] then, $\frac{s_k^T B_k s_k}{y_k^T s_k} = \frac{-\lambda_k s_k^T \nabla f_k}{-s_k^T \nabla f_k} = \lambda_k > 0$,

and since $B_k = H_k^{-1}$ is positive definite ($s_k^T B_k s_k > 0$ and $y_k^T H_k y_k > 0$), then $y_k^T s_k > 0$ and then $\frac{y_k^T H_k y_k}{y_k^T s_k} > 0$ which means that $\alpha_k = 2\frac{y_k^T H_k y_k}{y_k^T s_k} > 0$.

In addition, by more simplifying from Eq. 5 and Eq. 8, H_{k+1} can be write as follows:

$$H_{k+1} = H_k + \frac{\left((2\frac{y_k^T H_k y_k}{y_k^T s_k})s_k - H_k y_k\right)\left((2\frac{y_k^T H_k y_k}{y_k^T s_k})s_k - H_k y_k\right)^T}{y_k^T H_k y_k}$$

(9)

This is called the modified SR1(H-version) update. The sequence of inverse Hessian matrix produced by Eq. 9, never go to a near singular matrix which make the computation never break before get the minimizer of the objective function.

Theorem 1

The modified SR1(H-version) update generate a positive definite inverse of Hessian matrix if the current inverse of Hessian matrix is positive definite.

Proof:

Let $0 \neq z \in R^n$, then

$$z^T H_{k+1} z = z^T (H_k + \frac{\left((2\frac{y_k^T H_k y_k}{y_k^T s_k})s_k - H_k y_k\right)\left((2\frac{y_k^T H_k y_k}{y_k^T s_k})s_k - H_k y_k\right)^T}{y_k^T H_k y_k})z$$

$$z^T H_{k+1} z = z^T H_k z + z^T \frac{\left((2\frac{y_k^T H_k y_k}{y_k^T s_k})s_k - H_k y_k\right)\left((2\frac{y_k^T H_k y_k}{y_k^T s_k})s_k - H_k y_k\right)^T}{y_k^T H_k y_k} z \quad (10)$$

By substitution Eq. 6 in Eq. 10,

$$z^T H_{k+1} z = z^T H_k z + \frac{z^T w_k w_k^T z}{y_k^T H_k y_k} = z^T H_k z + \frac{\|w_k^T z\|^2}{y_k^T H_k y_k}$$

Since $y_k^T H_k y_k > 0$, and $z^T H_k z > 0$ by the positive definiteness of H_k, and $\|w_k^T z\|$ is always positive, therefore, $z^T H_{k+1} z > 0$ and H_{k+1} is positive definite.

SR1 (H-version) update algorithm:

1. Choose the starting point x_0 and the initial approximation $H_0 = I$, $error = \epsilon > 0$, set $k = 0$.
2. Compute $\nabla f(x_k)$
3. If $\|\nabla f(x_k)\| < \epsilon$ then, stop and x_k is the optimal solution, else continue to the next step.
3. Solve the system $p_k = -H_k \nabla f(x_k)$ for p_k.
4. Do line search to find $\lambda_k > 0, \ni f(x_k + \lambda_k p_k) < f(x_k)$.
5. Set $x_{k+1} = x_k + \lambda_k p_k$
6. Set $s_k = x_{k+1} - x_k$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$.
7. Compute H_{k+1} from Eq. 9.
8. Set $k = k + 1$ and go back to the step 2.

Convergence of the method:

In this section, the convergence of the modified SR1 (H-version) update is provided. The following assumptions are needed.

Assumption 1, [5]

(A): $f: R^n \to R$ *is twice continuously differentiable on convex set* $D \subset R^n$.

(B): $f(x)$ *is uniformly convex, i.e., there exist positive constants* c *and* C *such that for all* $x \in L(x) = \{x \mid f(x) \leq f(x_0)\}$, *where* x_0 *is starting point, we have* $c\|u\|^2 \leq u^T \nabla^2 f(x) u \leq C\|u\|^2$, $\forall\, u \in R^n$.

The assumption (B) implies that $\nabla^2 f(x)$ is positive definite on $L(x)$, and that f has a unique minimizer x^* in $L(x)$.

Lemma 2, [12]

Let $f: R^n \to R$ *satisfy Assumption 1, then* $\frac{\|s_k\|}{\|y_k\|}, \frac{\|y_k\|}{\|s_k\|}, \frac{s_k^T y_k}{\|s_k\|^2}, \frac{s_k^T y_k}{\|y_k\|^2}, \frac{\|y_k\|^2}{s_k^T y_k}$ *are bounded .*

Note that, from Assumption 1, and since $s_k^T B_k s_k, y_k^T s_k$ are bounded, and by lemma 2 then $\frac{\|y_k\|}{\|\alpha_k s_k\|}, \frac{\|\alpha_k s_k\|}{\|y_k\|}, \frac{y_k^T \alpha_k s_k}{\|y_k\|^2}, \frac{y_k^T \alpha_k s_k}{\|\alpha_k s_k\|^2}, \frac{\|\alpha_k s_k\|^2}{y_k^T \alpha_k s_k}, \frac{y_k^T H_k y_k}{\|y_k\|^2}, \frac{y_k^T B_k y_k}{\|y_k\|^2}, \frac{\alpha_k s_k y_k}{\|y_k\|^2}$, and $\frac{\alpha_k^2 s_k^T B_k s_k}{\|y_k\|^2}$ are bounded.

Lemma 3, [7]

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 $\sum\|\alpha_k s_k\|^2$ is convergent, which gives $\frac{\|\alpha_k s_k\|^2}{\|y_k\|^2}$ is convergent and bounded.

Theorem 4

For modified SR1 (H-version) update, the determinant of the next inverse of Hessian matrix is given by:

$$|H_{k+1}| = |H_k| \left[\frac{w_k^T r_k}{w_k^T s_k}\right], \quad where \quad w_k = (H_k y_k - \alpha_k s_k),\ r_k = y_k - B_k w_k$$

Proof:

From 5, set $w_k = (H_k y_k - \alpha_k s_k)$ then, $|H_{k+1}| = \left|H_k - \frac{w_k w_k^T}{w_k^T y_k}\right|$, since H_k is positive definite then, exist a triangular matrix $L_k \in R^{nxn} \ni H_k = L_k L_k^T$, and therefore

$$|H_{k+1}| = |H_k| \left|I - \frac{L_k^{-1} w_k w_k^T L_k^{T^{-1}}}{w_k^T y_k}\right| = |H_k| \left|I - \frac{L_k^{-1} w_k (L_k^{-1} w_k)^T}{w_k^T y_k}\right|$$

and apply Sherman-Morrison-Woudbury formula for the last Eq. 6, then

$$|H_{k+1}| = |H_k| \left[1 - \frac{(L_k^{-1} w_k)^T L_k^{-1} w_k}{w_k^T y_k}\right] = |H_k| \left[1 - \frac{w_k^T L_k^{-1^T} L_k^{-1} w_k}{w_k^T y_k}\right] = |H_k| \left[1 - \frac{w_k^T B_k w_k}{w_k^T y_k}\right],$$

Where $B_k = L_k^{-1^T} L_k^{-1}$, and

$$|H_{k+1}| = |H_k| \left[\frac{w_k^T y_k - w_k^T B_k w_k}{w_k^T y_k}\right] = |H_k| \left[\frac{w_k^T (y_k - B_k w_k)}{w_k^T y_k}\right]$$

Set $r_k = y_k - B_k w_k$, then

$$|H_{k+1}| = |H_k| \left[\frac{w_k^T r_k}{w_k^T y_k}\right] \quad (11)$$

One can use Eq. 11 to compute the determinant of the inverse of Hessian matrix at every iteration which must be always positive to ensure that the matrix is positive definite.

Theorem 5

Suppose that $f(x)$ satisfies Assumption 1, then the sequence $\{x_k\}$ generated by Eq. 5, is converges.

Proof:

Consider Eq. 5

$H_{k+1} = H_k + \frac{(\alpha_k s_k - H_k y_k)(\alpha_k s_k - H_k y_k)^T}{(\alpha_k s_k - H_k y_k)^T y_k} = H_k - \frac{(H_k y_k - \alpha_k s_k)(H_k y_k - \alpha_k s_k)^T}{(H_k y_k - \alpha_k s_k)^T y_k} = H_k - \frac{w_k w_k^T}{w_k^T y_k}$, where w_k as in theorem 4.

Define $\eta_k = \frac{w_k^T r_k}{y_k^T y_k}$, where w_k, and r_k are as in theorem 4, therefore

$$\eta_k = \frac{w_k^T r_k}{y_k^T y_k} = \frac{w_k^T (y_k - B_k w_k)}{y_k^T y_k} = \frac{y_k^T H_k y_k - \frac{y_k^T H_k y_k s_k^T B_k s_k}{\left(s_k^T y_k\right)^2}}{y_k^T y_k} \quad (12)$$

which is bounded by lemma 2 and lemma 3.

Define $\varphi(H_k) = tr(H_k) - ln\left(\left|H_k\right|\right) > 0$

By replacing H_k by H_{k+1} in $\varphi(H_k)$, then

$$0 < \varphi(H_{k+1}) = tr(H_{k+1}) - ln\left(\left|H_{k+1}\right|\right) = tr(H_k) - \frac{w_k^T w_k}{w_k^T y_k} - ln\left(\left|H_k\right| \frac{w_k^T r_k}{w_k^T y_k}\right) = tr(H_k) - \frac{w_k^T w_k}{w_k^T y_k} - ln\left(\left|H_k\right|\right) - ln(w_k^T r_k) + ln(w_k^T y_k)$$

$$0 < \varphi(H_{k+1}) = \varphi(H_k) - \frac{q_k}{cos^2\vartheta_k} - ln(w_k^T r_k) + ln(w_k^T y_k) \quad (13)$$

were $q_k = \frac{w_k^T y_k}{\|y_k\|^2}$, and $cos\vartheta_k = \frac{w_k^T y_k}{\|w_k\|\|y_k\|}$, by add and subtract $ln({y_k}^T y_k)$ to the right hand side of Eq. 13, therefor

$$0 < \varphi(H_{k+1}) = \varphi(H_k) - \frac{q_k}{cos^2\vartheta_k} - ln(w_k^T r_k) + ln(w_k^T y_k) + ln({y_k}^T y_k) - ln({y_k}^T y_k) = \varphi(H_k) - \frac{q_k}{cos^2\vartheta_k} - ln\frac{w_k^T r_k}{y_k^T y_k} + ln\frac{w_k^T y_k}{y_k^T y_k}$$

$$0 < \varphi(H_{k+1}) < \varphi(H_k) - \frac{q_k}{cos^2\vartheta_k} - ln(\eta_k) + ln(q_k) + 1 + ln(cos^2\vartheta_k) - ln(cos^2\vartheta_k)$$

$$0 < \varphi(H_{k+1}) < \varphi(H_k) + \left(1 - \frac{q_k}{cos^2\vartheta_k} + ln\ \frac{q_k}{cos^2\vartheta_k}\right) - ln(\eta_k) + ln(cos^2\vartheta_k) \quad (14)$$

Now consider the function $f(t) = 1 - t + \ln(t)$

$$f'(t) = -1 + \frac{1}{t} = 0$$

$\Rightarrow \frac{1}{t} = 1$

$\Rightarrow t = 1$ is extreme point.

$$f''(t) = -\frac{1}{t^2}$$

$at\ t = 1$

$$f''(t) = -\frac{1}{1^2} = -1 < 0$$

The function $f(t)$ has a maximum value at $t = 1$, thus, $\max\ f(t) = 1 - 1 + \ln(1) = 0,\ \forall t > 0$, Then, $f(t) = 1 - t + \ln(t) \le 0,\ \forall t > 0$, and hence $\left(1 - \frac{q_k}{cos^2\vartheta_k} + ln\ \frac{q_k}{cos^2\vartheta_k}\right) \le 0,\ \forall\ \frac{q_k}{cos^2\vartheta_k} > 0$ (15)

By substitution Eq. 15 in Eq. 14, then

$$0 < \varphi(H_{k+1}) < \varphi(H_k) - ln(\eta_k) + ln(cos^2\vartheta_k) \quad (16)$$

By summing Eq. 16 from $j = 0$ up to k

$$0 < \sum_{j=0}^k \varphi(H_{j+1}) < \sum_{j=0}^k \varphi(H_j) + \sum_{j=0}^k \left(-ln(\eta_j)\right) + \sum_{j=0}^k ln\left(cos^2\vartheta_j\right)$$

$0 < \sum_{j=0}^k \varphi(H_{j+1}) < \sum_{j=1}^k \varphi(H_j) + \sum_{j=0}^k C + \sum_{j=0}^k ln\left(cos^2\vartheta_j\right)$, where $C = -\ln(\eta_j)$

$$0 < \varphi(H_1) + \cdots + \varphi(H_{k+1}) < \varphi(H_0) + \cdots + \varphi(H_k) + C(k+1) + \sum_{j=0}^k ln\left(cos^2\vartheta_j\right)$$

$$0 < \varphi(H_{k+1}) < \varphi(H_0) + C(k+1) + \sum_{j=0}^k ln\left(cos^2\vartheta_j\right) \quad (17)$$

Where the constant C is assumed to be positive without loss of the generality. From Zoutendijk condition (1) (if f satisfy assumption 1, then $\sum cos^2\theta \|\nabla f\|^2 < \infty$) and hence $\lim_{k\to\infty} \|\nabla f\| cos\vartheta_k = 0$. If ϑ_k is bounded away from 90°, $\exists \mu \in R^+ \ni cos\vartheta_k > \mu > 0$, for k sufficient large and hence $\|\nabla f\| \to 0$ and by the first order necessary condition theorem (1), the prove is complete.

Now, assume by contradiction that $cos\vartheta_k \to 0$, then $\exists k_1 > 0 \ni \forall j > k_1$,

$$ln\left(cos^2\vartheta_j\right) < -2C \quad (18)$$

From Eq.17, and since $\sum_{j=0}^k ln\left(cos^2\vartheta_j\right) = \sum_{j=0}^{k_1} ln\left(cos^2\vartheta_j\right) + \sum_{j=k_{1+1}}^k ln\left(cos^2\vartheta_j\right)$, then

$$0 < \varphi(H_{k+1}) < \varphi(H_0) + C(k+1) + \sum_{j=0}^{k_1} ln\left(cos^2\vartheta_j\right) + \sum_{j=k_1+1}^k ln\left(cos^2\vartheta_j\right) \quad (19)$$

By substitution Eq.18 in Eq.19

$$0 < \varphi(H_{k+1}) < \varphi(H_0) + C(k+1) + \sum_{j=0}^{k_1} ln\left(cos^2\vartheta_j\right) + \sum_{j=k_1+1}^k (-2C)$$

$$0 < \varphi(H_{k+1}) < \varphi(H_0) + C(k+1) + \sum_{j=0}^{k_1} ln\left(cos^2\vartheta_j\right) - 2C(k - k_1)$$

$0 < \varphi(H_{k+1}) < \varphi(H_0) + \sum_{j=0}^{k_1} ln\left(cos^2\vartheta_j\right) + 2C.k_1 + C - C.k < 0$, for k sufficiently large, which is a contradiction, then $cos^2\vartheta_j \to 0$ is not true and $\lim_{k\to\infty} inf\|\nabla f\| \to 0$, and again by the first order necessary condition theorem (1) and the prove is complete.

Numerical Examples:

In this section, two numerical examples are studied by modifying SR1 (H-version) update. The results are compared with the results obtained by the original method.

Example 1:

In this example the objective function $f(x) = (1 - x_1)^2 + (x_2 - x_1)^2$ [1]*, has been solved by using the original method firstly, and then also solved by our method.*

Min. $f(x) = (1 - x_1)^2 + (x_2 - x_1)^2$, with $x^0 = \begin{bmatrix} 0 \\ -0.5 \end{bmatrix}$, $H_0 = I$, $error = 0.005$

Iteration 1

$\nabla f(x) = \begin{bmatrix} 4x_1 - 2x_2 - 2 \\ -2x_1 + 2x_2 \end{bmatrix}$, $\nabla f(x^0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$,

$P_0 = -H_0\nabla f(x^0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

$x^1 = x^0 + \lambda_0 P_0 = \begin{bmatrix} \lambda_0 \\ -0.5 + \lambda_0 \end{bmatrix}$, $f(x^1) = (1 - \lambda_0)^2 + (-0.5 + \lambda_0 - \lambda_0)^2 = (1 - \lambda_0)^2 + 0.25$,

$\frac{\partial f(x^1)}{\partial \lambda_0} = 1 - \lambda_0 = 0, \lambda_0 = 1$, and hence $x^1 = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$,

$s_0 = x^1 - x^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\nabla f(x^1) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$y_0 = \nabla f(x^1) - \nabla f(x^0) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $H_1 = H_0 + \frac{(s_0 - H_0y_0)(s_0 - H_0y_0)^T}{(s_0 - H_0y_0)^T y_0} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$, and since $\|\nabla f(x^1)\| > error$.

Iteration 2

$P_1 = -H_1\nabla f(x^1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $x^2 = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$, $s_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\nabla f(x^2) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $y_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and since $\|\nabla f(x^2)\| = 0 < error$, therefore the method is terminated at the solution $x^* = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$, and the minimum value of the objective function is $f(x^*) = 0.25$.

Now the same example has been solved by our method as follows:

Iteration 1

$\nabla f(x) = \begin{bmatrix} 4x_1 - 2x_2 - 2 \\ -2x_1 + 2x_2 \end{bmatrix}$, $\nabla f(x^0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$,

$P_0 = -H_0\nabla f(x^0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $x^1 = x^0 + \lambda_0 P_0 = \begin{bmatrix} \lambda_0 \\ -0.5 + \lambda_0 \end{bmatrix}$, $f(x^1) = (1 - \lambda_0)^2 + (-0.5 + \lambda_0 - \lambda_0)^2 = (1 - \lambda_0)^2 + 0.25$,

$\frac{\partial f(x^1)}{\partial \lambda_0} = 1 - \lambda_0 = 0, \lambda_0 = 1$, and hence $x^1 = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$,

$s_0 = x^1 - x^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\nabla f(x^1) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$y_0 = \nabla f(x^1) - \nabla f(x^0) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\alpha_0 = 2\frac{y_0^T H_0 y_0}{y_0^T s_0} = 4$,

$H_1 = H_0 + \frac{(\alpha_0 s_0 - H_0y_0)(\alpha_0 s_0 - H_0y_0)^T}{(\alpha_0 s_0 - H_0y_0)^T y_0} = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$, and since $\|\nabla f(x^1)\| > error$, and continue to the next iteration

Iteration 2

$P_1 = -H_1\nabla f(x^1) = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$, $x^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $s_1 = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}$, $\nabla f(x^2) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $y_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\alpha_1 = 12$, $H_2 = \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix}$, and since $\|\nabla f(x^2)\| = 0 < error$, therefore the method is terminated at the solution $x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and the minimum value of the objective function is $f(x^*) = (1 - 1)^2 + (1 - 1)^2 = 0$.

Example 2:

In this example the Freudenstein and Roth function (1),has been solved by using program MATLAB, and the final results firstly by the original method, and then by our method has been given

Min. $f(x) = \{-13 + x_1 + [(5 - x_2)x_2 - 2]x_2\}^2 + \{-29 + x_1 + [(x_2 + 1)x_2 - 14]x_2\}^2$,

with $x^0 = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$, $H_0 = I$, error=0.0005.

```
Maximum number of function evaluations exceeded;
  increase options.MaxFunEvals
x =
 -3514897
   8.2397
FVAL = 14596e+005
OUTPUT =
       iterations: 43
        funcCount: 202
         stepsize: -3.6779e-032
    algorithm: 'medium-scale: Quasi-Newton line search'
GRAD =
  1.0e+005 *
  -0.0105
   5352
INVHESSIAN =
  1.0e-005 *
   0.0908   -0.1254
  -0.1254    0.3665
```

Now the same example has been solved by our method:

x =
10.8384
3.8545
FVAL =
1.0383e-015
OUTPUT =
iterations: 24
funcCount: 202
stepsize: 2.9014e-016
algorithm: 'medium-scale: Quasi-Newton line search'
GRAD =
-0.0008
-0.0416
INVHESSIAN =
1.0e+005 *
12018 0.7665
0.7665 0.1424

From example 1, SR1 (H-version) update cannot terminate successfully at the minimum (min. f =0.25), because of the non-positive definite of inverse of Hessian matrix H_1 generated in first iteration ($|H_1| = 0$), and hence the method terminated at a saddle point which is not minimizer of the objective function, but clear that the modified SR1 (H-version) update can terminate successfully at the minimizer of the objective function. Moreover, the inverse of Hessian matrix generated by our method is positively definite at every iteration. From example 2, the function evaluation (FVAL) at the last iteration for the original method is very far from the exact value (min. f = 0), but in our method, it is clear that the function evaluation (FVAL) at the last iteration is very closely to the exact value (min. f =0). This means that the original method cannot successfully terminate at the minimum because of the not positive definite of inverse Hessian matrix generated by the method in iteration number 43 where

INVHESSIAN=
1.0e-005 *
-0.1227 0.2728
0.2728 -0.3763

is very closely to zero ($|H_{43}| = 0.0000000176$) or near singular matrix.

Conclusion:

In this paper, the SR1 (H-version) update has been modified to preserve the positive definite property for the next inverse of Hessian matrix at each iteration if the current inverse of Hessian matrix is positively definite which makes the computation continue until the objective function terminates at the minimum of the objective function. Moreover, theorem 1 proves the positive definiteness property and theorem 5 proves the convergence of our method and also two numerical examples are established to support our method.

Authors' declaration:

- Conflicts of Interest: None.
- Ethical Clearance: The project was approved by the local ethical committee in University of Mustansiriyah.

Author’s Contributions statement:

The first author makes a conception and design for this paper while the second author makes the analysis and the third author makes the revision and proofreading.

References:

1. Wenyu S, Ya-Xiang Y. Optimization Theory and Method: Nonlinear Programming. Springer, USA. 2006.
2. Zhang J, Xu Ch. Properties and Numerical Performance of Modified quasi–Newton Methods Equations. Elsevier, J. Comp. App. Math. 2001; 137: 269-278
3. Yabe H, Ogasawara H, Yoshino M. Local and superlinear convergence of quasi-Newton methods based on modified secant conditions. J. Comp. App. Math. 2007 Aug; 1;205(1):617-32.
4. Guo Q, Liu J. Global convergence properties of two modified BFGS-type methods. J. App. Math. Comp. 2007; 23: 311–319.
5. Mahmood S S, Shnywer S H. On Modified DFP Update for Unconstrained Optimization. American J. App. Math. 2017; 5(1): 19-30.
6. Mahmood S S, Muhanah N S. Symmetric and Positive Definite Broyden Update for Unconstrained Optimization. BSJ. 2019; 16(3): 661-666.
7. Razieh D, Narges B, Mohammad M H. A new modified BFGS method for solving system of nonlinear equations. J. Interdisciplinary Math. 2019; 22: 75-89.
8. Razieh D, Narges B, Mohammad M H. A new modified BFGS method for unconstrained optimization problems. J. Comp. App. Math. 2018; 37: 5113-5125.
9. Bojari S, Eslachachi M R. Global convergence of a family of modified BFGS methods under a modified weak-Wolfe–Powell line search for nonconvex functions. 4OR. 2020;18(2): 219-244.
10. Yuan Y X. Amodified BFGS algorithm for unconstrained optimization.IMA.1991; 11(3): 325-332.
11. Mahmood S S, Farqad H. On Extended Symmetric Rank One Update for Unconstrained Optimization. J. Edu. 2017; 1: 206-220.
12. Conn A R, Gould N I, Toint P L. Convergence of quasi-Newton matrices generated by the symmetric

rank one update. J. Math. Programming. 1991; 50(2): 177–195.

التحديد الايجابي لتحديث الرتبة 1 المتماثل (النسخة H) للامثلية غير المقيدة

سعد شاكر محمود **جعفر حمود عيدي** **جنان عادل جاسم**

قسم الرياضيات، كلية التربية، الجامعة المستنصرية، بغداد، العراق.

الخلاصة:

عدة محاولات بذلت لتحوير شرط كواسي نيوتن للمثلية غير المقيدة وذلك للحصول على تقارب اسرع مع خواص كاملة (التناظرية والموجبة) لمعكوس المصفوفة هسين (المشتقة الثانية لدالة الهدف)، هناك الكثير من طرق المثلية غير المقيدة التي لا تولد معكوس مصفوفة هيسين موجبة. احد هذه الطرق هو التحديث التناظري من الرتبة الاولى (النسخة H)، حيث ان هذا التحديث يحقق شرط كواسي نيوتن وايضا يحقق صفة التناظرية ولكنه لا يضمن خاصية الموجبة لمعكوس مصفوفة هيسين عندما تكون معكوس مصفوفة هيسين الابتدائية موجبة. ان الموجبة لمعكوس المصفوفة هيسين مهم لضمان وجود نقطة النهاية الصغرى لدالة الهدف وكذلك للحصول على اصغر قيمة لدالة الهدف.

الكلمات المفتاحية : مصفوفة هيسين، التحديد الموجب، شرط كواسي- نيوتن، التحديث من الرتبة 1 المتماثل (النسخة H)، الامثلية غير المقيدة.