A New Symmetric Rank One Algorithm for Unconstrained Optimization

Abbas Y. Al-Bayati
profabbasalbayati@yahoo.com
College of Computer Sciences and Mathematics
University of Mosul/Iraq

Salah G. Shareef *College of Education*
University of Zahko
Iraq

Received on: 23/02/2005 **Accepted on: 25/06/2005**

ABSTRACT

In this paper, a new symmetric rank one for unconstrained optimization problems is presented. This new algorithm is used to solve symmetric and positive definite matrix. The new method is tested numerically by (7) nonlinear test functions and method is compared with the standard BFGS algorithm.

The new matrix used is symmetric and positive definite and it generates descent directions and satisfied QN-like condition.

Keywords: unconstrained optimization, descent direction, QN-like condition.

خوارزمية متناظرة المصفوفة من الرتبة الأولى في الأمثلية غير المقيدة.

عباس يونس البياتي
كلية علوم الحاسوب والرياضيات، جامعة الموصل

صلاح غازي شريف
كلية التربية، جامعة زاخو

تاريخ قبول البحث : 2005/6/25 **تاريخ استلام البحث : 2005/2/23**

الملخص

في هذا البحث تم استحداث خوارزمية مصفوفتها متناظرة ومن ذات الرتبة الأولى تستخدم في الأمثلية غير المقيدة. هذه الخوارزمية الجديدة للمصفوفات المستحدثة تستخدم في حل الأنظمة المتناظرة والموجبة التعريف. تم اختبار هذه الخوارزمية الجديدة عددياً باستخدام (7) دوال قياسية غير خطية وقورنت مع خوارزمية الـ BFGS القياسية وتم إثبات هذه الصيغة بأنها متناظرة وموجبة التعريف وتولد اتجاهات انحدارية وتحقق شرط QN الشبيهة.

الكلمات المفتاحية: الامثلية غير المقيدة، الأنظمة المتناظرة والموجبة التعريف، اتجاه انحداري، شرط QN الشبيهة.

1. Introduction:

In this section we take a numerical approach toward the development of a numerical algorithm lying some where intermediate to the steepest descent and Newton methods. The idea underlying quasi-Newton methods is to use an approximation to the inverse Hessian instead of the true inverse that is required in Newton's methods.

An efficient quasi-Newton method was proposed by Davidon (1959) and many others followed this pioneering work (see Dennis and More (1977), and Dennis and Schnable (1983)).

The search vector is calculated according to the following equation:

$$d_k = -H_k g_k$$

where H_k is in some way an approximation to G^{-1}.

The new point x_{k+1} is found by line searches, i.e. λ_k minimize $f(x_k + \lambda_k d_k)$ w.r.t. $\lambda_k g_{k+1}$ is then found and H_k is update to H_{k+1} as

$$H_{k+1} = H_k + E_k$$

where E_k is a matrix of rank one at most two, normally calculated from x_k, x_{k+1}, g_k, g_{k+1} and H_{k+1}. The initial approximation H_0 can be any positive definite matrix.

We required H_{k+1} to have some of the properties of the inverse Hessian matrix. Following the property $Gv_k = y_k$ (for a quadratic function), we choose E_k in such away that

$H_{k+1}y_k = v_k$

which is called the quasi-Newton condition.

The matrix H_{k+1} holds the same curvature information as G_k^{-1} in the direction d_k when the objective function is quadratic. In other words, the curvature information of f along d_k is given by $v_k^T G_k v_k$, which can be approximated with first-order information:

$v_k^T G_k v_k \approx v_k^T y_k$

where $v_k = x_{k+1} - x_k = \lambda_k d_k$ and this relationship is exact if f is quadratic.

Let

$$H_{k+1}y_k = v_k \qquad \text{…(1.1)}$$

where $y_k = g_{k+1} - g_k$ and $v_k = x_{k+1} - x_k$.

Let

$$H_{k+1} = H_k + E_k = H_k + auu^T \qquad \text{…(1.2a)}$$

where a is a constant and u a square positive definite matrix.

In which a symmetric rank one matrix $E = auu^T$ is added in to H_k. Note that $E_{ij} = au_i u_j$ so that E_k can be calculated by $n^2 + n$ multiplication's only. Now if the quasi-Newton condition (1.1) is to be satisfied, it follows that:

$$H_k y_k + auu^T y_k = v_k \qquad \text{…(1.2b)}$$

Hence u is proportional to $v_k - H_k y_k$. Since any change of length can be taken up in a, $u_k = v_k - H_k y_k$ is set, in which case $au^T y_k = 1$ must hold, which defines a thus the rank one formula is given by:

$$H_{k+1} = H_k + \frac{(v_k - H_k y_k)(v_k - H_k y_k)^T}{(v_k - H_k y_k)^T y_k} \qquad \text{…(1.3)}$$

The general formula, in a slight modification of Fletcher's (1987) Parameterization, is given by

$$H_{k+1} = H_k - \frac{H_k y_k (H_k y_k)^T}{y_k^T H_k y_k} + \frac{v_k v_k^T}{v_k y_k^T} + \varphi\, w_k w_k^T \qquad \text{…(1.4)}$$

where $\varphi \geq 0$ is the free parameter and

$$w = \frac{v_k}{v_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k} \qquad \text{…(1.5)}$$

Much effort, both analytic and computational, has been devoted to identifying the best quasi-Newton formula or even the best from the much wider class of variable metric methods ontroduced by Huang (1970). The choice with the widest support is the BFGS algorithm which was derived independently in 1970 in four different ways by Broyden, Fletcher, Goldfarb, and Shanno. In the parameterization of (1.4), this BFGS update formula corresponds to choosing

$$\varphi = y_k^T H_k y_k \qquad \text{…(1.6)}$$

2. New Symmetric Rank One Formula:

In this section we shall derive a new symmetric rank one update as follows if u and w are column vectors, then u^T is row vector and the product $u^T w$ is a scalar.

Suppose that G_{k+1} and G_k are square matrices, and u and w are vector with the property that:

$$G_{k+1} = G_k - uw^T \qquad \ldots(2.1)$$

Then the inverse of G_{k+1} becomes:

$$H_{k+1} = H_k + \alpha H_k u.w^T H_k \qquad \ldots(2.2)$$

where α is a scalar:

$$\alpha = \frac{1}{1 - w^T H_k u} \qquad \ldots(2.3)$$

Thus if G_k has an inverse and $w^T H_k u \neq 1$, then G_{k+1} has an inverse. See William (1988) for more details.

Now let $u = \frac{y_k}{v_k^T y_k}$ and $w = y_k$ are vectors then the formula (2.2) becomes:

$H_{k+1} = H_k - \alpha H_k \frac{y_k y_k^T}{v_k^T y_k} H_k$, then we have the new symmetric rank one formula as follows:

$$H_{k+1} = H_k - \alpha \frac{H_k y_k y_k^T H_k}{v_k^T y_k} \qquad \ldots(2.4)$$

where

$$\alpha = \frac{1}{1 - \frac{y_k^T H_k y_k}{v_k^T y_k}} \qquad \ldots(2.5)$$

Now if $G_{k+1} v_k = y_k$ then $v_k = H_{k+1} y_k$ it clear to prove that the new proposed algorithm will satisfy the QN condition: from (2.2) that:

$$\begin{aligned} v_k = H_{k+1} y_k &= \left(H_k + \alpha H_k u.w^T H_k\right) y_k \\ &= H_k y_k + \alpha H_k u.w^T H_k y_k \\ &= H_k y_k + \alpha\left(w^T H_k y_k\right) H_k u \end{aligned}$$

$$v_k = H_k y_k + \gamma H_k u \qquad \ldots(2.6)$$

where $\gamma = \alpha\left(w^T H_k y_k\right) = \frac{w^T H_k y_k}{1 - w^T H_k u}$ which implies that:

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

This indicates that the new proposed symmetric matrix is a positive definite and satisfies the QN condition.

3. The Outlines of New Algorithm:

Step (1): Set x_1, ε, $H_1 = I$.

Step (2): For $k = 1$ to n, set $d_1 = -H_1 g_1$.

Step (3): Compute $x_{k+1} = x_k + \lambda_k d_k$, where λ_k is optimal step size.

Step (4): Cheek if $\|g_{k+1}\| < \varepsilon$ then stop, else $y_k = g_{k+1} - g_k$, $v_k = x_{k+1} + x_k$.

Step (5): $\alpha = \dfrac{1}{1 - \dfrac{y_k^T H_k y_k}{v_k^T y_k}}$.

Step (6): $H_{k+1} = H_k - \alpha \dfrac{H_k y_k y_k^T H_k}{v_k^T y_k}$.

Step (7): $d_{k+1} = -H_{k+1} g_{k+1} + \beta_k d_k$ where $\beta_k = \dfrac{g_{k+1}^T H_{k+1} y_k}{d_k^T y_k}$.

Step (8): If $k = n+1$ or $d_{k+1}^T g_{k+1} > 0$, then go to step (1), else $k = k+1$ and go to step (3).

4. Numerical Results:

Seven test functions were tested with different dimensions $4 \le n \le 500$ all programs are written in FORTRAN 90 language and for all cases the stopping criterion is taken to be $\|g_{k+1}\| < 1 \times 10^{-5}$.

The line search routine used was cubic interpolation which uses function and gradient values and it is an adaptation of the routine published by Bundy (1984).

The results are given in the Table (1) is specifically quoting the number functions NOF and the number of iterations NOI.

Experimental results in Table (1) confirm that the new algorithm is superior to standard BFGS method.

In about 53% in NOI and 56% NOF.

Table (1). Comparative Performance of The Two Algorithms functions (classical BFGS and new update method).

Test function	*n*	*BFGS classical* NOI (NOF)	*New update* NOI(NOF)
Powell	4	21(86)	19(57)
	20	38(123)	19(57)
	100	71(197)	22(73)
	500	50(148)	22(73)
Wood	4	37(110)	40(96)
	20	84(244)	47(109)
	100	251(775)	51(119)
	500	283(791)	48(113)
Cubic	4	19(58)	16(45)
	20	35(99)	18(51)
	100	70(167)	18(51)
	500	53(124)	19(53)
Rosen	4	34(87)	32(83)
	20	34(87)	32(83)
	100	34(87)	32(83)
	500	34(87)	32(83)

Shallow	4 20 100 500	8(26) 8(26) 8(26) 8(26)	9(24) 9(24) 9(24) 9(24)
Non-diagonal	4 20 100 500	24(72) 48(115) 74(177) 78(188)	28(70) 39(97) 39(97) 40(99)
Gedger	4 20 100 500	6(17) 6(17) 6(17) 6(17)	28(70) 39(97) 39(97) 40(99)
Total		**1428(3994)**	**671(1748)**

Tools	BFGS	New Update
NOI	100	46.9
NOF	100	43.7

5. Appendix:

All the test function used in this paper are from general literature:

1. Generalized Powell function

$$f(x)=\sum_{i=1}^{n/4}\left(x_{4i-3}-10x_{4i-2}\right)^2+5\left(x_{4i-1}-x_{4i}\right)^2+\left(x_{4i-1}-2x_{4i}\right)^2$$
$$+10\left(x_{4i-9}-x_{4i}\right)^4+\left(x_{4i-2}-2x_{4i-1}-x_{4i}\right)^2,$$

Starting point $:(3,1,0,1,\dots)^T$

2. Generalized Wood function

$$f(x)=\sum_{i=1}^{n/4}100\left(x_{4i-2}-x_{4i-3}^2\right)^2+\left(1-x_{4i-3}\right)^2+90\left(x_{4i}-x_{4i-1}^2\right)^2+\left(1-x_{4i-1}^2\right)^2+1.0,$$

Starting point $:(-3,-1,-3,-1,\dots)^T$

3. Generalized Rosen Brock Banana function

$$f(x)=\sum_{i=1}^{n/2}100\left(x_{2i}-x_{2i-1}^2\right)^2+\left(1-x_{2i-1}\right)^2,$$

Starting point $:(-1.2,1,\dots,-1.2,1)^T$

4. Generalized non diagonal function

$$f(x)=\sum_{i=1}^{n/2}100\left(x_1-x_i^2\right)^2+\left(1-x_i\right)^2,$$

Starting point :$\left(-1,\ldots,-1\right)^T$

5. Generalized Beale function

$$f(x)=\sum_{i=1}^{n/2}\left[1.5-x_{2i}+\left(1-x_{2i}\right)\right]^2+\left[2.25-x_{2i-1}\left(1-x_{2i}^2\right)\right]^2+\left[2.625-x_{2i-1}\left(1-x_{2i}^2\right)\right]^2,$$

Starting point :$\left(-1,-1,\ldots,-1,-1\right)^T$

REFERENCES

[1] Bunday, B., "Basic Optimization Methods", Edward Arnold Square, London, (1984).

[2] Davidon, W.C., "Variable Metric Methods for Minimization" Argonn, National Laboratory, Rept-ANL-5990 (Rev)., (1959).

[3] Dennis, J.R. and Schnable, R.B., "Numerical Methods for Unconstrained Optimization and Nonlinear Equations", Prentice Hall, Englewood Cliffs, N.S., (1983).

[4] Dennis, J.R. and More, J., "Quasi-Newton Methods Motivation and Theory", SIAM Review, 19 (1977).

[5] Fletcher, R., "Practical Methods of Optimization", John Wiley and Son, Chichester, (1987).

[6] Huang, H.Y., "Unified Approach to Quadratically Convergent Algorithms for Function Minimization", JOTA, 5, PP. 405-423, (1970).

[7] William, W., "Applied Numerical Linear Algebra", Prentice Hall International Editions, (1988).