NEW QUASI-NEWTON (DFP) WITH LOGISTIC MAPPING

Salah Gazi Shareef and Bayda Ghanim Fathi
Dept. of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region-Iraq.
(Accepted for publication: February 25, 2016)

Abstract:
In this paper, we propose a modification of the self-scaling quasi-Newton (DFP) method for unconstrained optimization using logistic mapping. We shoe that it produces a positive definite matrix. Numerical results demonstrate that the new algorithm is superior to standard DFP method with respect to the NOI and NOF.

Keywords: *Unconstrained optimization, Quasi-Newton methods, DFP method, Logistic mapping.*

1- Introduction

The quasi-Newton algorithms for minimizing a function $f(x), x \in R^n$, are iterative accelerated gradient methods which use past positions and functional values rather than an analytically or numerically calculated one to approximate the inverse of the Hessian matrix H of the function. This is accomplished by selecting an initial approximation H_0 to the inverse Hessian, as well as an initial approximation x_0 to the minimum of $f(x)$, and then finding at each step α_k, the scalar parameter which minimizes $f(x_k - \alpha_k \mathrm{H_k} g_k)$ where $g_k = g(x_k) = \nabla f(x)$.

It is known that the search direction of the quasi-Newton algorithms is

$$d_k = -\mathrm{H_k} g_k, \quad (1.1)$$

and the approximate matrix H_k is updated by

$$H_{k+1} = H_k + D_k, \quad (1.2)$$

where D_k is the correction matrix.

The Davidon Fletcher Powell (DFP) algorithm was the first quasi-Newton algorithm created (Shanno and Kettler, 1970). In this technique, substituting $\frac{v_k v_k^T}{v_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}$ where $v_k = x_{k+1}$ - x_k and $y_k = g_{k+1} - g_k$ for D_k and giving

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{v_k v_k^T}{v_k^T y_k}. \quad (1.3)$$

The following theorem will be used later.

Theorem (1.1). (Edwin and Stanislaw, 2001**).** Let a function $f \in C, x_k \in R^n, g_k = 0,$ and H_k is an $n \times n$ real symmetric positive definite matrix. If we set $x_{k+1} = x_k - \alpha_k \mathrm{H_k} g_k$, where $\alpha_k = \arg\min_{\alpha} f(x_k - \alpha H_k g_k)$, then $\alpha_k > 0$, and $f(x_{k+1}) < f(x_k)$.

2- A new self- scaling quasi-Newton (DFP) formula

For a control parameter, μ, the logistic mapping (Lu et al., 2006) is defined by

$$z_{k+1} = \mu z_k (1 - z_k) \quad (2.1)$$

Let us consider the quasi-Newton condition

$$H_{k+1} y_k = v_k, \quad (2.2)$$

where $v_k = \alpha_k \mathrm{d_k} = x_{k+1}$ - x_k, $\mu, \gamma \in (0,1)$ and $y_k = \Delta g_k = g_{k+1} - g_k$.

A new self-scaling quasi-Newton (DFP) formula can be defined as

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{\mu\gamma(1-\gamma) v_k v_k^T}{v_k^T y_k} \quad (2.3)$$

Algorithm: A New DFP Algorithm

Step (1):- Set $k = 0$; select x_0, and a real symmetric positive definite $\mathrm{H}_0(\mathrm{H}_0 = \mathrm{I})$.
Step (2):- If $\mathrm{g}_\mathrm{k} = 0$, stop ; else $\mathrm{d}_\mathrm{k} = -\mathrm{H}_\mathrm{k}\mathrm{g}_\mathrm{k}$, where $g(x) = \nabla f(x)$
Step (3):- Compute $\alpha_k = arg\ min\ f(x_k + \alpha d_k)$
$x_{k+1} = x_k + \alpha_k \mathrm{d}_\mathrm{k}$.
Step (4):- Compute $v_k = \Delta x_k = \alpha_k d_k$
$y_k = \Delta g_k = g_{k+1} - g_k$
$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{\mu\gamma(1-\gamma) v_k v_k^T}{v_k^T y_k}$, where $\mu, \gamma \in (0,1)$
Step (5):- Set $k = k + 1$; go to step 2.

Theorem (2.1). If the new self-scaling quasi-Newton (DFP) formula (2.3) applied to the quadratic function with Hessian $G = G^T$, then $H_{k+1}\Delta g_i = \mu\gamma(1-\gamma)\Delta x_i$ for $0 \le i \le k$ where $v_k = \Delta x_k = x_{k+1} - x_k$ and $y_k = \Delta g_k = g_{k+1} - g_k = G v_k$.

Note: $d_k^T G d_i$=0.

Proof. We prove this theorem by using induction criteria. For $k = 0$, we have

$$\mathrm{H}_1 y_0 = \mathrm{H}_0 y_0 - \frac{\mathrm{H}_0 \mathrm{y}_0 \mathrm{y}_0^\mathrm{T} \mathrm{H}_0}{\mathrm{y}_0^\mathrm{T} \mathrm{H}_0 \mathrm{y}_0} y_0 + \frac{\mu\gamma(1-\gamma)\mathrm{v}_0 \mathrm{v}_0^\mathrm{T}}{\mathrm{v}_0^\mathrm{T} \mathrm{y}_0} y_0$$
$$= \mu\gamma(1-\gamma)\mathrm{v}_0 .$$

Assume the result is true for $k-1$; that is $H_k \Delta g_i = \mu\gamma(1-\gamma)\Delta x_i$, $0 \le i \le k-1$.
We now show that $H_{k+1}\Delta g_i = \mu\gamma(1-\gamma)\Delta x_i$, $0 \le i \le k$. First consider $i = k$, we have

$$H_{k+1} y_k = H_k y_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} y_k + \frac{\mu\gamma(1-\gamma) v_k v_k^T}{v_k^T y_k} y_k,$$

implies that
$H_{k+1}\ y_k = \mu\gamma(1-\gamma)\ v_k$.
It remains to consider the case $i < k$. Using the hypothesis, we have

$$H_{k+1} y_i = H_k y_i - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} y_i + \frac{\mu\gamma(1-\gamma) v_k v_k^T}{v_k^T y_k} y_i$$

$$= \mu\gamma(1-\gamma) v_i - \frac{H_k y_k}{y_k^T H_k y_k}\left(y_k^T v_i\right) + \frac{\mu\gamma(1-\gamma) v_k}{v_k^T y_k}\left(v_k^T y_i\right).$$

Since

$$v_k^T y_i = v_k^T G v_i = \alpha_k \alpha_i d_k^T G d_i = 0$$

and

$$y_k^T v_i = v_k^T G v_i = \alpha_k \alpha_i d_k^T G d_i = 0.$$

Hence,

$$H_{k+1} y_i = \mu\gamma(1-\gamma) v_i.$$

The proof is completed

Theorem (2.2). Suppose that $g_k \neq 0$. In the new self-scaling quasi-Newton (DFP) formula (2.3), if H_k is positive definite, then so is H_{k+1}.

Proof. Multiply both sides of (2.3) by x^T from left and by x from right, we get

$$x^T H_{k+1} x = x^T H_k x - \frac{x^T H_k y_k y_k^T H_k x}{y_k^T H_k y_k} + \frac{\mu\gamma(1-\gamma) x^T v_k v_k^T x}{v_k^T y_k}$$
$$= x^T H_k x - \frac{(x^T H_k y_k)^2}{y_k^T H_k y_k} + \frac{\mu\gamma(1-\gamma)(x^T v_k)^2}{v_k^T y_k}.$$

We can define

$a = H_k^{1/2} x$ and $b = H_k^{1/2} y_k$,

where $H_k = H_k^{1/2}\, H_k^{1/2}$.

Now, using the definition of a and b, we obtain

$x^T \mathrm{H_k} x = x^T H_k^{1/2}\, H_k^{1/2} x = a^T a,$

$x^T \mathrm{H_k y_k} = x^T H_k^{1/2}\, H_k^{1/2} \mathrm{y_k} = a^T b,$

and

$\mathrm{y_k^T H_k y_k} = \mathrm{y_k^T} H_k^{1/2}\, H_k^{1/2} \mathrm{y_k} = b^T b.$

Hence

$$x^T \mathrm{H_{k+1}} x = a^T a - \frac{(a^T b)^2}{b^T b} + \frac{\mu\gamma(1-\gamma)(x^T v_k)^2}{v_k^T y_k}$$
$$= \frac{\|a\|^2 \|b\|^2 - (a^T b)^2}{\|b\|^2} + \frac{\mu\gamma(1-\gamma)(x^T v_k)^2}{v_k^T y_k}.$$

We know that $\mu\gamma(1-\gamma)$ is positive and we have $v_k^T y_k = v_k^T(g_{k+1} - g_k) = -v_k^T g_k$ because $v_k^T g_{k+1} = \alpha_k d_k^T g_{k+1} = 0$ by (In the conjugate direction algorithm, $g_{k+1}^T d_i = 0$ for all k, $0 \leq k \leq n-1$, and $0 \leq i \leq k$ (Edwin and Stanislaw, 2001)).

Since $v_k = \alpha_k d_k = -\alpha_k H_k g_k$, we get

$v_k^T y_k = -v_k^T g_k = \alpha_k g_k^T H_k g_k.$

The above yields

$$x^T H_{k+1} x = \frac{\|a\|^2 \|b\|^2 - (a^T b)^2}{\|b\|^2} + \frac{\mu\gamma(1-\gamma)(x^T v_k)^2}{\alpha_k g_k^T H_k g_k} \qquad (2.4)$$

The fractional terms on the right-hand side of (2.4) are nonnegative, the first term is nonnegative because of the Cauchy-Schwarz inequality, and the second term is nonnegative because $\mathrm{H_k}$, $\alpha_k > 0$ by Theorem (1.1) and $\mu\gamma(1-\gamma) > 0$. Therefore, to show that $x^T \mathrm{H_{k+1}} x > 0$ for $x \neq 0$, we only need to demonstrate that these terms do not vanish simultaneously. The first term vanishes only if a and b are proportional, that is if $a = \beta b$ for a scalar β).

To complete the proof it is enough to show that if $a = \beta b$, then $\frac{\mu\gamma(1-\gamma)(x^T v_k)^2}{\alpha_k \mathrm{g}_\mathrm{k}^\mathrm{T}\,\mathrm{H_k} g_k} > 0$.
First observe that

$H_k^{1/2}x = a = \beta b = \beta H_k^{1/2} y_k = H_k^{1/2}(\beta y_k)$.
Hence,

$$x = \beta y_k$$

Using the above expression for x and $v_k^T y_k = -\alpha_k g_k^T H_k g_k$, we obtain

$$\frac{\mu\gamma(1-\gamma)(x^T v_k)^2}{\alpha_k \mathrm{g}_\mathrm{k}^\mathrm{T}\,\mathrm{H_k} g_k} = \frac{\mu\gamma(1-\gamma)\beta^2(y_k^T v_k)^2}{\alpha_k \mathrm{g}_\mathrm{k}^\mathrm{T}\,\mathrm{H_k} g_k} = \frac{\mu\gamma(1-\gamma)\beta^2(\alpha_k \mathrm{g}_\mathrm{k}^\mathrm{T}\,\mathrm{H_k} g_k)^2}{\alpha_k \mathrm{g}_\mathrm{k}^\mathrm{T}\,\mathrm{H_k} g_k}$$
$$= \mu\gamma(1-\gamma)\beta^2\alpha_k \mathrm{g}_\mathrm{k}^\mathrm{T}\,\mathrm{H_k} g_k > 0.$$

Thus, for all $x \neq 0$

$x^T \mathrm{H_{k+1}} x > 0$.

Then the proof is completed.

3- Numerical Results

This section is devoted to test the implementation of the new method. We compare standard formula of DFP and new formula of self-scaling Q-N (DFF), the comparative tests involve well-known nonlinear problems (standard test function) with different dimensions $4 \leq n \leq 100$, all programs are written in FORTRAN95 language and for all cases the stopping condition is $\|g_{k+1}\|_\infty \leq 10^{-5}$. Efficiency of the new DFP algorithm has been tested by means of 10 standard problems. Experimental results in Table (1) represent the number of function evaluations NOF and the number of iterations NOI. Table (2) shows the percentage of improving the new algorithm and confirms that the new method is superior to standard method with respect to the NOI and NOF.

Table (1): Comparison between the performance of the standard DFP update and new DFP update.

Test fun.	n	Standard formula		New formula	
		NOI	NOF	NOI	NOF
Powell	4	23	126	17	89
	100	80	467	35	142
	500	60	328	35	137
	1000	44	230	38	151
Wood	4	39	250	37	192
	100	243	1380	251	1178
	500	751	3439	700	2575
	1000	1192	4758	1106	3548
Wolfe	4	7	18	7	16
	100	72	145	55	111
	500	82	165	61	123
	1000	95	191	68	137
Cubic	4	18	76	17	62
	100	34	114	46	132
	500	54	166	41	123
	1000	60	183	47	135
Rosen	4	36	145	34	116
	100	247	1017	219	767
	500	605	2240	348	1038
	1000	984	3570	459	1347
Mile	4	26	119	24	95
	100	38	174	30	123
	500	34	152	31	125
	1000	44	193	41	164
Beale	4	8	22	8	21
	100	10	27	10	26
	500	10	27	10	26
	1000	10	27	10	26
Gedger	4	6	18	5	14
	100	6	18	6	16
	500	6	18	6	16
	1000	6	18	6	16
shallow	4	13	40	13	36
	100	15	45	15	40
	500	16	46	15	40
	1000	16	46	15	40
G. central	4	21	146	11	52
	100	21	146	16	97
	500	22	154	16	97
	100	22	154	16	97
Total		5076	20598	3925	13286

Table (2): Percentage of improving the new algorithm

Tools	Standard formula	New formula
NOI	100 %	77.3 %
NOF	100 %	64.5 %

4- Conclusion

A new formula for updating quasi-Newton matrices based on DFP and which uses logistic mapping is presented. It is shown that the new algorithm produces positive definite matrices. Numerical experiments indicate that our algorithm is better than the original DFP with respect to the NOI and NOF.

5- References

Edwin, K. P. Chong and Stanislaw H. Zak: *An Introduction To Optimization*. Second Edition. John Wiley & Sons, Inc. United States of America, 2001.

Lu Hui-juan, ZHANG Huo-ming and MA Long-hua: A new optimization algorithm based on chaos. *Journal of Zhejiang University SCIENCE A*, 7(4), 2006.

Shanno, D. F. and Kettler, P. C.: Optimal Conditioning of Quasi- Newton methods. *Mathematics of Computation*. Volume 24, number 111, 1970.

پوخته :

هاته پێش بینی کرن دڤێ ڤەکولینێ دا پێش ئیخستنا رێکا (Quasi- Newton) یا بەرنیاس ب (**DFP**) بو چارەسەرکرنا کێشا نموونەیێن ئاسان بکار ئینانا هاوکێشا لوجستی ، مە دا دیار کرن کو ئەو رێزکراوێن نێزیک بو خوارزمیا نی ئەوێن موجەب و دەست نیشان کری بجێ بەجێکرنا گەلەک ئەنجاما بو مە هاتە دیارکرن کو خوارزمیا نی باشترە ژ خوارزمییا پیڤەری سەبارەت هژمارتن ، ژمارا چەندایەتی (**NOI**) و هژمارتنا هاوکێشە (**NOF**).

الملخص:

اقترحنا في هذا البحث تطوير جديد لطريقة شبه نيوتن المعروفة ب (DFP) لحل مسائل الامثلية الغير مقيدة باستخدام الدالة اللوجستية . وقد منا برهانا على ان مصفوفة التقريب للخوازمية الجديدة هي مصفوفة موجبة محددة. اثبتت النتائج العددية بان الخوارزمية الجديدة المقترحة افضل من الخوارزمية القياسية من حيث حساب عدد التكرارت (NOI) وعدد مرات حساب الدالة (NOF).