Ilorin Journal of Science
Volume 4, Number 1, 2017, pp. 92–113 (Printed in Nigeria)
ISSN: 2408 – 4840 © 2017 Faculty of Physical Sciences, University of Ilorin
https://doi.org/10.54908/iljs.2017.04.01.007

ILORIN
JOURNAL OF SCEINCE

ILJS-16-001

A Refinement of Quasi-Newton Iterative Method for Non-Linear Multivariable Optimization Problems

Ejieji*, C. N. and Ogunniran, M. O.
Department of Mathematics, University of Ilorin, Ilorin, Nigeria.

Abstract
The Newton's method and some of its modifications were reviewed. A new modification of the Newton's method, which is of the Quasi-Newton principle is constructed and presented. Some standard test problems were solved using the Newton' method, one of the existing modifications of the Newton's method (David Fletcher Powell) and the refined method. From the results obtained, the newly modified method compared favourably with the existing ones and proved superior to the David Fletcher Powell method.

Keywords: Newton's Method, Quasi-Newton, unconstrained optimization, multivariable optimization, non-linear optimization.

1. Introduction

Optimization can be defined as an act of obtaining the best result under any given circumstance or selecting the best solution from a set of available numerous solutions. In every practical situation the above statement translates to determining how best to allocate the available resources in order to either minimize the effort required or maximize the desired benefit. It is also true that the effort required or the benefit desired in every practical situation can be expressed as a function of certain decision variables and that most real life problems involve functions that are non-linear in nature. Hence researchers continue to seek for new methods or modifications of the existing ones for non-linear optimization problems.

In this work we reviewed the Newton's method and some of its modifications. We constructed a new modification for the Newton's method and compared its performance with the performance of some existing methods. The methods reviewed and the new modification derived are for the solution of the unconstrained optimization problems:

*Corresponding Author: Ejieji, C. N.
Email: ejieji.cn@unilorin.edu.ng

$$Minimize\; f = f(X);$$

where

$$X = (x_1, x_2, x_3, ..., x_n)^T \in \Omega \subseteq R^n, \quad f : \Omega \subseteq R^n \to R.$$

Newton's method is an iterative method that uses the evaluation of Hessians to solve problems of non linear programming. Newton in 1669 proposed his method for the solution of the unconstrained optimization problem:

Minimize $f = f(X)$,

where, $X = (x_1, x_2, ..., x_n)^T$ and $f: R^n \to R$.

Using Taylor series expansion, a quadratic approximation of the function $f(x)$ at $x = x_i$ is given by

$$f(x) = f(x_i) + (x - x_i) f'(x_i) + \frac{1}{2}(x - x_i)^2 f''(x_i) \; . \tag{1}$$

The necessary condition for $f(x)$ to have an optimum at x^* is that $f'(x^*) = 0$. Newton's method uses the direct root method which seeks to find the root (or solution) of the equation $f'(x^*) = 0$. Setting the derivative of the equation (1) to zero, we have:

$$f'(x) = f'(x_i) + f''(x_i)(x - x_i) = 0 \,. \tag{2}$$

If x_i denotes an approximation to the minimum of $f(x)$, (2) can be re-arranged to obtain an improved approximation as:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}. \tag{3}$$

The iterative process (3) above which is the Newton's method converges when the derivative $f'(x_{i+1})$ is close to zero or

$$|\, f'(x_{i+1}) \,| \;\le\; \varepsilon \;, \tag{4}$$

where ε is a small quantity called tolerance.

Extending the Newton's Method to a multi-variable function $f(X)$ using Taylor's series expansion we get

$$f(X) = f(X_i) + \nabla f_i^T (X - X_i) + \frac{1}{2}(X - X_i)^T [J_i](X - X_i) \ , \qquad (5)$$

where $[J_i] = [J] \ \ at \ \ X_i$ is the matrix of second partial derivatives (Hessian matrix) of f evaluated at the point X_i. By setting the partial derivative of (5) equal to zero we get the minimum of $f(X)$ by solving:

$$\frac{\partial f(X)}{\partial X_j} = 0; \qquad j = 1,2,...,n \qquad (6)$$

which gives from (5) and (6)

$$\nabla f = \nabla f_i + [J_i](X - X_i) = 0 \,. \qquad (7)$$

If $[J_i]$ is non singular i.e $|J_i| \neq 0$, (7) can be solved to obtain an improved approximation

at $(X = X_{i+1})$ as:

$$X_{i+1} = X_i - [J_i]^{-1} \nabla f_i \ . \qquad (8)$$

Since higher order terms have been neglected in (5), (8) is to be used iteratively to find the optimum solution X^*.

Algorithm of the Newton's Method

Step 0: Given X_0, set i =0, set ε as a stopping criterion,

Step 1: Set $d_i = -J(X_i)^{-1} \nabla f(X_i)$. If $d_i \leq \varepsilon$, stop else go to step 2.

Step 2: Choose step size $\lambda_i = 1$

Step 3: Compute $X_{i+1} = X_i + \lambda_i d_i$,

Step 4: Set i =i+1, go to step 1

This method was originally developed by Newton for solving non linear equations and later refined by Raphson in 1690 and hence the method is also known as Newton-Raphson method.

2. Early Modifications of the Newton's Method

Earlier modifications of the Newton's method have been reported in literature and prominent among them are the Quasi-Newton Methods which are also called the Variable metric methods.

Some of the widely known members of the Quasi-Newton methods for large scale unconstrained optimization include Broyden's (1967) method, the Symmetric Rank 1 Update (SR1)[(Davidon, 1959; Broyden's ,1967), the Davidon-Fletcher- Powell Method (DFP) (Davidon, 1959; Flecther and Powell, 1963), and the Broydon-Fletcher-Goldfarb-Shanno Method (BFGS) (Broyeden, 1970; Flecther, 1970; Goldfarb, 1970; Shanno, 1970), Yuan (1991), Zhang *et al* (1999), Hennig and Martin (2012). Xiaowei Jiang and Yueting Yang (2010) provided the self scaling Quasi-Newton method for large scale unconstrained optimization.

Baghmisheh *et al.* (2013) modified the Newton's method by using the Guass integration formula.

The basic iterative process used in the Newton's method is given by:

$$X_{i+1} = X_i - [J_i]^{-1} \nabla f(X_i) \ , \qquad (9)$$

where the Hessian matrix $[J_i]$ is composed of the second partial derivatives of f and varies with the decision vector X_i for a non quadratic (generally non linear) objective function, f. The main idea behind the Quasi Newton or Variable metric methods(including the ones mentioned earlier) is to approximate either $[J_i]$ by another matrix $[A_i]$ or $[J_i]^{-1}$ by another matrix $[B_i]$, using only the first partial derivatives of f. If $[J_i]^{-1}$ is approximated by $[B_i]$, (9) can be expressed as:

$$X_{i+1} = X_i - \lambda_i^* [B_i] \nabla f(x_i), \qquad (10)$$

where λ_i^* can be considered as the optimal step length along the direction:

$$S_i = -[B_i]\nabla f(X_i)\,. \qquad (11)$$

It can be seen that the steepest descent direction method can be obtained as a special case of (11) by setting $[B_i]=[I]$.

The Computation of $[B_i]$

To implement (9), an approximate inverse of the Hessian matrix, $[B_i]\equiv[A_i]^{-1}$, is to be computed. For this, we first expand the gradient of f about an arbitrary reference point, X_0, using Taylor's series as:

$$\nabla f(X) = \nabla f(X_0) + [J_0](X - X_0)\,. \qquad (12)$$

If we pick two points X_i and X_{i+1} and use $[A_i]$ to approximate $[J_0]$,(12) can be rewritten as:

$$\nabla f_{i+1} = \nabla f(X_0) + [A_i](X_{i+1} - X_0)\,, \qquad (13)$$

and

$$\nabla f_i = \nabla f(X_0) + [A_i](X_i - X_0)\,. \qquad (14)$$

Subtracting (14) from (13), we get

$$[A_i]d_i = g_i\,, \qquad (15)$$

where

$$d_i = X_{i+1} - X_i\,, \qquad (16)$$

$$g_i = \nabla f_{i+1} - \nabla f_i\,. \qquad (17)$$

The solution of (15) can be written as:

$$d_i = [B_i]g_i\,, \qquad (18)$$

where, $[B_i]=[A_i]^{-1}$ denotes an approximation to the inverse $[J_i]^{-1}$, of the Hessian matrix.

It can be seen that (18) represents a system of n-equations in n^2 unknown elements of the matrix $[B_i]$. Thus for $n>1$, the choice of $[B_i]$ is not unique and one would like to choose $[B_i]$ that is closest to $[J_0]$, in some sense.

Some techniques have been suggested in the literature for the computation of $[B_i]$ as the iterative process progresses (i.e for the computation of $[B_{i+1}]$ once $[B_i]$ is known) and each update is required to retain the symmetric and positive definiteness properties of $[B_i]$.

The general formula for updating the matrix $[B_i]$ can be written as:

$$[B_{i+1}]=[B_i]+[\Delta B_i], \tag{19}$$

where $[\Delta B_i]$ can be considered to be the update (or correction) matrix added to $[B_i]$. Some earlier suggestions for the update include:

(i) The Broyden (1957) Symmetric Rank1 Update

$$[B_{i+1}]=[B_i]+[\Delta B_i]\equiv[B_i]+\frac{(d_i-[B_i]g_i)(d_i-[B_i]g_i)^T}{(d_i-[B_i]g_i)^T g_i}, \tag{20}$$

(ii) The DFP: Davidon (1959)-Fletcher (1963)-Powell (1963) Update

$$[B_{i+1}]=[B_i]+[\Delta B_i]\equiv[B_i]+\frac{d_i d_i^T}{d_i^T g_i}-\frac{([B_i]g_i)([B_i]g_i)^T}{([B_i])^T g_i}, \tag{21}$$

which can also be expressed as

$$[B_{i+1}^{DFP}=[B_i]+\frac{\lambda_i^* S_i S_i^T}{S_i^T g_i}-\frac{[B_i]g_i g_i^T[B_i]}{g_i^T[B_i]g_i}, \tag{22}$$

since

$$X_{i+1}=X_i+\lambda_i^* S_i,$$

where S_i is the search direction, and $d_i=X_{i+1}-X_i$ can be rewritten as:

$d_i=\lambda_i^* S_i$ **;** and

(iii) The BFGS: Broyden (1969), Fletcher (1970), Goldfarb (1970) and Shanno (1970) Update

$$[B_{i+1}^{BFGS}] = [H_{i+1}]^{-1} = \left([B_i] + \frac{\Delta g_i \Delta g_i^T}{\Delta g_i^T \Delta x_i} - \frac{[B_i]\Delta x_i \Delta x_i^T [B_i]}{\Delta x_i^T [B_i] \Delta x_i} \right)^{-1} . \quad (23)$$

3. A New Modification of the Newton's Method

The New Computation of $[B_i]$

The basic iterative process in the Newton's method as given in equation (9) is

$$X_{i+1} = X_i - [J_i]^{-1} \nabla f(X_i)$$

and the first partial derivative of $\nabla f(X)$ when expanded in Taylor's series about an arbitrary reference point, X_0, is

$$\nabla f(X) = \nabla f(X_0) + [J_0](X - X_0) . \quad (24)$$

If $[J_i]^{-1}$ is approximated by $[B_i]$, (9) can be expressed as:

$$X_{i+1} = X_i - \lambda_i^* [B_i] \nabla f(x_i) , \quad (25)$$

where λ_i^* can be considered as the optimal step length along the direction:

$$S_i = -[B_i] \nabla f(X_i) . \quad (26)$$

If we pickthree points X_i, X_{i+1} *and* X_{i+2} and use $[A_i] \equiv [J_i]$ then the derivative of f in equation(24) can be written generally as:

$$\nabla f(X) \equiv \nabla f(X_0) + [A_i](X - X_0) . \quad (27)$$

Evaluating (27) at the three chosen points X_i, X_{i+1} *and* X_{i+2} we get

$$\nabla f_{i+2} = \nabla f(X_0) + [A_i](X_{i+2} - X_0) , \quad (28)$$

$$\nabla f_{i+1} = \nabla f(X_0) + [A_i](X_{i+1} - X_0) , \quad (29)$$

$$\nabla f_i = \nabla f(X_0) + [A_i](X_i - X_0). \tag{30}$$

Subtracting (29) and(30) from (28) gives

$$\nabla f_{i+2} - \nabla f_{i+1} - \nabla f_i = -\nabla f(X_0) + [A_i](X_{i+2} - X_{i+1} - X_i) + [A_i]X_0 \ . \tag{31}$$

Let

$$g_{i+2} = \nabla f_{i+2} - \nabla f_{i+1} - \nabla f_i \ ,$$

$$d_{i+2} = X_{i+2} - X_{i+1} - X_i \ ,$$

$$m_{i+2} = -\nabla f(X_0) + [A_i]X_0 \ .$$

Recall that: $[A_i] \equiv [J_i]$ and $[B_i] \equiv [J_i]^{-1}$.

Let

$$[A_i] \equiv [B_{i+2}]^{-1}$$

$\Rightarrow$

$$m_{i+2} = -\nabla f(X_0) + [B_{i+2}]^{-1} X_0 \qquad = [B_{i+2}]^{-1} X_0 - \nabla f(X_0)$$

(31) now becomes: $g_{i+2} = m_{i+2} + [A_i]d_{i+2}$ which implies that $d_{i+2} = [A_i]^{-1}(g_{i+2} - m_{i+2})$.

This can also be expressed as

$$d_{i+2} = [B_{i+2}](g_{i+2} - m_{i+2}) \ . \tag{32}$$

It can be seen that (32) represents a system of n-equations in n^3 -unknown elements of matrix $[B_{i+2}]$. As the iteration process progresses, we need to update $[B_{i+2}]$, which is of RANK 3 (because of the three points selected initially).

Therefore

$$[B_{i+3}] = [B_{i+2}] + [\Delta B_{i+2}],$$

where $[\Delta B_{i+2}]$ is the update (or correction) matrix added to $[B_{i+2}]$.

Let

$$[\Delta B_{i+2}] = C_1 Z_1 Z_1^T + C_2 Z_2 Z_2^T + C_3 Z_3 Z_3^T, \tag{33}$$

For (33) to satisfy the quasi-Newton condition

$$d_{i+2} = [B_{i+2}](g_{i+2} - m_{i+2}),$$

we must have

$$d_{i+2} = [B_{i+2}](g_{i+2} - m_{i+2}) + [\Delta[B_{i+2}](g_{i+2} - m_{i+2})]\}, \tag{34}$$

$$\left.\begin{aligned} = \ & [B_{i+2}](g_{i+2} - m_{i+2}) + C_1 Z_1 Z_1^T (g_{i+2} - m_{i+2}) + C_2 Z_2 Z_2^T (g_{i+2} - m_{i+2}) + \\ & C_3 Z_3 Z_3^T (g_{i+2} - m_{i+2}) \end{aligned}\right\}. \tag{35}$$

This implies that

$$d_{i+2} - [B_{i+2}](g_{i+2}) + [B_{i+2}](m_{i+2}) = C_1 Z_1 Z_1^T (g_{i+2} - m_{i+2}) + C_2 Z_2 Z_2^T (g_{i+2} - m_{i+2}) + C_3 Z_3 Z_3^T (g_{i+2} - m_{i+2}). \tag{36}$$

Since $Z_r^T (g_{i+2} - m_{i+2})$; $r = 1(1)3$ are scalar quantities, we rewrite (36) as:

$$\frac{d_{i+2} - [B_{i+2}](g_{i+2}) + [B_{i+2}](m_{i+2})}{Z_1^T (g_{i+2} - m_{i+2}) Z_2^T (g_{i+2} - m_{i+2}) Z_3^T (g_{i+2} - m_{i+2})}$$

$$= \frac{C_1 Z_1}{Z_2^T (g_{i+2} - m_{i+2}) Z_3^T (g_{i+2} - m_{i+2})} + \frac{C_2 Z_2}{Z_1^T (g_{i+2} - m_{i+2}) Z_3^T (g_{i+2} - m_{i+2})} +$$

$$\frac{C_3 Z_3}{Z_1^T(g_{i+2}-m_{i+2})Z_2^T(g_{i+2}-m_{i+2})}.$$

Equating terms in the above equationwe get

$$C_1 Z_1 = \frac{d_{i+2}}{Z_1^T(g_{i+2}-m_{i+2})},$$

$$Z_1 = d_{i+2} \quad and \quad C_1 = \frac{1}{d_{i+2}^T(g_{i+2}-m_{i+2})}. \tag{37}$$

Also

$$C_2 Z_2 = \frac{-[B_{i+2}]g_{i+2}}{Z_2^T(g_{i+2}-m_{i+2})},$$

$\Rightarrow$

$$Z_2 = [B_{i+2}]g_{i+2} \quad and \quad C_2 = \frac{-1}{[B_{i+2}]g_{i+2}^T(g_{i+2}-m_{i+2})} \tag{38}$$

and

$$C_3 Z_3 = \frac{[B_{i+2}]m_{i+2}}{Z_3^T(g_{i+2}-m_{i+2})}$$

$\Rightarrow$

$$Z_3 = [B_{i+2}]m_{i+2} \quad and \quad C_3 = \frac{1}{([B_{i+2}]m_{i+2})^T(g_{i+2}-m_{i+2})}. \tag{39}$$

Now the new update matrix $[B_{i+3}]$ is

$$[B_{i+3}] = [B_{i+2}] + [\Delta B_{i+2}]$$

$$= [B_{i+2}] + \frac{d_{i+2}d_{i+2}^T}{d_{i+2}^T(g_{i+2}-m_{i+2})} - \frac{[B_{i+2}]g_{i+2}([B_{i+2}]g_{i+2})^T}{([B_{i+2}]g_{i+2})^T(g_{i+2}-m_{i+2})} + \frac{[B_{i+2}]m_{i+2}([B_{i+2}]m_{i+2})^T}{([B_{i+2}]m_{i+2})^T(g_{i+2}-m_{i+2})}, \tag{40}$$

where

$$d_{i+2} = X_{i+2} - X_{i+1} - X_i \, ,$$

$$g_{i+2} = \nabla f_{i+2} - \nabla f_{i+1} - \nabla f_i \, ,$$

$m_{i+2} = [B_{i+2}]^{-1} X_0 - \nabla f_0$

Suppose:

$$[L_{i+2}] = \frac{d_{i+2} d_{i+2}^T}{d_{i+2}^T (g_{i+2} - m_{i+2})},$$

$$[M_{i+2}] = -\frac{[B_{i+2}] g_{i+2} ([B_{i+2}] g_{i+2})^T}{([B_{i+2}] g_{i+2})^T (g_{i+2} - m_{i+2})}$$

and

$$[N_{i+2}] = \frac{[B_{i+2}] m_{i+2} ([B_{i+2}] m_{i+2})^T}{([B_{i+2}] m_{i+2})^T (g_{i+2} - m_{i+2})} \, ,$$

then

$$[B_{i+3}] = [B_{i+2}] + [L_{i+2}] + [M_{i+2}] + [N_{i+2}] \, . \quad (41)$$

The Algorithm of the New Modification

The algorithm of the new method is as follows:

<u>Initialization</u>

Step 1:= Set $i = 0$, select X_i and compute ∇f_i, $[J_i]$

Step 2:= Compute X_{i+1}

and X_{i+2},

using the conventional Newton's method or any of its earlier modifications.

Compute

$$\nabla f_{i+1},\ [J_{i+1}]\ \ and\ \ \nabla f_{i+2}$$

Step 3:= Select $[B_{i+2}] = [I]$, the identity matrix.

Line Search

Step 4:= Compute

$$S_{i+2} = -[B_{i+2}]\nabla f_{i+2},$$

$$\lambda_{i+2} = \underset{\lambda \geq 0}{argmin} \quad f(X_{i+2} + \lambda S_{i+2}),\ \ \lambda_{i+2} \in R$$

$$X_{i+3} = X_{i+2} + \lambda_{i+2} S_{i+2},$$

$$\nabla f_{i+3} = \nabla f(X_{i+3}),$$

$$d_{i+3} = X_{i+3} - X_{i+2} - X_{i+1},$$

$$g_{i+3} = \nabla f_{i+3} - \nabla f_{i+2} - \nabla f_{i+1}.$$

Stopping Criterion

If $\| g_{i+3} \| = 0$, Stop, else compute

$$m_{i+3} = [B_{i+2}]^{-1} X_i - \nabla f_i,$$

Update $[B_{i+2}]$

$$[B_{i+3}] = [B_{i+2}] + \frac{d_{i+3}(d_{i+3})^T}{(d_{i+3})^T(g_{i+3} - m_{i+3})} - \frac{[B_{i+2}]g_{i+3}([B_{i+2}]g_{i+3})^T}{([B_{i+2}]g_{i+3})^T(g_{i+3} - m_{i+3})} + \frac{[B_{i+2}]m_{i+2}([B_{i+2}]m_{i+2})^T}{([B_{i+2}]m_{i+2})^T(g_{i+3} - m_{i+3})}$$

Step 5:= Set $i = i + 1$ and goto step 4.

4. Computational Experience with the new Modification

We now present some numerical results to demonstrate the performance of the new scheme.

4.1 Illustrative Examples

Example 4.1: The CUTE Quartic Function:

$$f(X) := \sum_{r=1}^{n} (x_r - 1)^4, \quad X_\circ = (2,2,2,\cdots,2)^T$$

For $n = 10$:

$$f(X) := \sum_{r=1}^{10} (x_r - 1)^4, \quad X_\circ = (2,2,2,2,2,2,2,2,2,2)^T$$

$$f(X) := (x_1 - 1)^4 + (x_2 - 1)^4 + (x_3 - 1)^4 + (x_4 - 1)^4 + (x_5 - 1)^4 + (x_6 - 1)^4 + (x_7 - 1)^4 +$$

$$(x_8 - 1)^4 + (x_9 - 1)^4 + (x_{10} - 1)^4 .$$

Table 4.1: Numerical Results for Example 4.1

Optimal Decision Variables/ Function value	Newton's Method	DFP Method	The New Method
x_1^*	1.087791495	1.000000000	1.000000000
x_2^*	1.087791495	1.000000000	1.000000000
x_3^*	1.087791495	1.000000000	1.000000000
x_4^*	1.087791495	1.000000000	1.000000000
x_5^*	1.087791495	1.000000000	1.000000000
x_6^*	1.087791495	1.000000000	1.000000000
x_7^*	1.087791495	1.000000000	1.000000000
x_8^*	1.087791495	1.000000000	1.000000000
x_9^*	1.087791495	1.000000000	1.000000000
x_{10}^*	1.087791495	1.000000000	1.000000000
f^*	.0005940319151	0.000000000	0.000000000

Example 4.2: The Extended Penalty Function:

$$f(X) := \sum_{r=1}^{n-1}(x_r - 1)^2 + (\sum_{r=1}^{n} x_r^2 - 0.25)^2, \quad X_\circ := (1,2,\ldots,n)^T$$

For $n = 10$:

$$f(X) := \sum_{r=1}^{9}(x_r - 1)^2 + (\sum_{r=1}^{n} x_r^2 - 0.25)^2, \quad X_\circ := (1,2,3,4,5,6,7,8,9,10)^T$$

$$f(X) := (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 + (x_4 - 1)^2 + (x_5 - 1)^2 + (x_6 - 1)^2 + (x_7 - 1)^2 +$$

$$(x_8 - 1)^2 + (x_9 - 1)^2 + (x_1{}^2 + x_2{}^2 + x_3{}^2 + x_4{}^2 + x_5{}^2 + x_6{}^2 + x_7{}^2 + x_8{}^2 + x_9{}^2 + x_{10}{}^2 - 0.25)^2$$

Table 4.2: Numerical Results for Example 4.2

Optimal Decision Variables /Function Value	Newton's Method	DFP Method	New Modification Method
x_1^*	0.357340138	0.310096800	0.324199644
x_2^*	0.357340149	0.325623036	0.337774361
x_3^*	0.357340162	0.341149270	0.351349079
x_4^*	0.357340174	0.356675499	0.364923796
x_5^*	0.357340186	0.372201728	0.378498513
x_6^*	0.357340198	0.387727957	0.392073231
x_7^*	0.357340210	0.403254185	0.405647948
x_8^*	0.357340222	0.418780413	0.419222665
x_9^*	0.357340234	0.434306648	0.432797382
x_{10}^*	$7.636906 \times 10^{(-7)}$	0.199011046	0.196083051
f^*	4.525718286	4.665980990	4.673038603

Example 4.3: The Diagonal 3 Function:

$$f(X) := \sum_{r=1}^{n}(exp(x_r) - rsin(x_r)), \quad X_{\circ} := (1,1,...,1)^T = (1,1,1,\cdots,1)^T$$

For $n = 10$:

$$f(X) := \sum_{r=1}^{10}(exp(x_r) - rsin(x_r)), \quad X_{\circ} := (1,1,1,1,1,1,1,1,1,1)^T$$

$$f(X) := \mathrm{e}^{x_1} - \sin(x_1) + \mathrm{e}^{x_2} - 2\sin(x_2) + \mathrm{e}^{x_3} - 3\sin(x_3) + \mathrm{e}^{x_4} - 4\sin(x_4) + \mathrm{e}^{x_5} - 5\sin(x_5)$$

$$+ \mathrm{e}^{x_6} - 6\sin(x_6) + \mathrm{e}^{x_7} - 7\sin(x_7) + \mathrm{e}^{x_8} - 8\sin(x_8) + \mathrm{e}^{x_9} - 9\sin(x_9) + \mathrm{e}^{x_{10}} - 10\sin(x_{10}).$$

Table 4.3: Numerical Results for Example 4.3

Optimal Decision Variables / Function Value	Newton's Method	DFP Method	New Modification Method
x_1^*	3.18159982×10^{-10}	0.590305204	0.00119340251
x_2^*	0.539785161	0.716056673	0.538428333
x_3^*	0.768578541	0.825675280	0.768628531
x_4^*	0.904788218	0.919436105	0.904818571
x_5^*	0.997576220	0.997805210	0.997615189
x_6^*	1.065758888	1.061643109	1.065856073
x_7^*	1.118389310	1.112073138	1.118592472
x_8^*	1.160454403	1.150372295	1.160749409
x_9^*	1.194962557	1.177883864	1.195214851
x_{10}^*	1.223851813	1.195951206	1.223779621
f^*	-21.20430570	-20.88807940	-21.20430136

Example 4.4: The Cosine Function

$$Max \quad f(X) := \sum_{r=0}^{n-1} \cos(-0.5x_{r+1} + x_r^2), \quad X_\circ := (1,1,...,1)^T$$

For $n = 10$, we have:

$$f(X) := \sum_{r=0}^{9} \cos(-0.5x_{r+1} + x_r^2), \quad X_\circ := (1,1,1,1,1,1,1,1,1,1)^T,$$

That is

$$f(X) := \cos(x_1^2 \;\; -0.5x_2) + \cos(x_2^2 \;\; -0.5x_3) + \cos(x_3^2 \;\; -0.5x_4) + \cos(x_4^2 \;\; -0.5x_5)$$

$$+\cos(x_5^2 \;\; -0.5x_6) + \cos(x_6^2 \;\; -0.5x_7) + \cos(x_7^2 \;\; -0.5x_8) + \cos(x_8^2 \;\; -0.5x_9) +$$

$$\cos(x_9^2 \;\; -0.5x_{10})$$

Table 4.4: Numerical Results for Example 4.4

Optimal Decision Variables / Function Value	Newton's Method	DFP Method	New Modification Method
x_1^*	0.501220390	0.510216963	0.523989718
x_2^*	0.502443752	0.521609578	0.530612019
x_3^*	0.504899430	0.533470733	0.536838779
x_4^*	0.509846844	0.550955128	0.540510825
x_5^*	0.519887586	0.560413001	0.544415543
x_6^*	0.540566189	0.534086962	0.554234382
x_7^*	0.584423604	0.559814383	0.579555174
x_8^*	0.683101895	0.613298949	0.660227161
x_9^*	0.933256390	0.729238766	0.841013042
x_{10}^*	1.741934913	1.076392099	1.421697946
f^*	8.999999998	8.998442574	8.999032696

4.2 Discussion of Results

The Newton' method, Davidon-Fletcher-Powell method and the new update for the Quasi-Newton method derived in this work were used to solve four standard test problems and the results are presented in Tables 4.1 to 4.4. From the results it can be observed that the new method compares favourably with the previous methods.

The problems considered are of the minimization type. Maximization problems are not beyond the scope of this work because a maximization problem can be converted to a minimization type by the relation:

$$max \;\; f(X) = min \; [-f(X)].$$

Very often performing an exact line search by a method such as as the bisection method or other classical methods is too expensive computationally in selecting a step size in an optimization algorithm. The Armijo's rule which is one of the inexact line search methodsand which guarantees a sufficient degree of accuracy was used to obtain the optimal step length at each iteration in the computation of the results presented.

At each step, the Armijo's rule requires that

$$f(x_k + \alpha S_k) \le f(x_k) + c\alpha \nabla f_k^T S_k, \quad c \in (0,1)$$

for some fixed $\alpha \in [0,1]$.

5. Summary and Conclusion

A new modification of the Newton's method which is based on the quasi-Newton principle has been presented in this work. The new scheme is a three-step method because three points are needed to implement it. The first point is the initial point which is usually given while the other two points are computed by using the classical Newton' method or any of its earlier modifications. The initial approximation to the Hessian matrix is sellected as the identity matrix $[B_i]=[I]$. For subsequent iterations the rank three update given in equation (40) is employed.

The classical Newton's method is a single-step iterative method while the new method is a multistep scheme. Apart from the fact that a multistep scheme gives a better result than a single step method the new modification of the Newton's method has the following advantages:

(a) The inversion of the Hessian matrix at every iteration is no longer needed;

(b) Computation and storage of second derivatives are no longer required.

It can therefore be observed that the new method has eliminated some of the drawbacks of the Newton's method.

Acknowledgement

The authors acknowledge the comments of the reviewers for their suggestions and comments in improving the quality of the manuscripts.

References

Baghmisheh, M., Mahmoudi, Y. and Jahangirirad, M. (2013): A new modification of the Newton's method by using the Gauss integration formula. ***Life Science Journal***. **10**(5s), 288-291.

Broyden, C. G. (1967): Quasi-Newton methods and their application to function minimization. ***Mathematics of Computation.*** **21**, 368-381.

Broyden, C. G. (1970): The convergence of a class of double-rank minimization algorithms 1. General Considerations. ***IMA Journal of of Applied Mathematics.*** **6**,76-90; The convergence of a class of double-rank minimization algorithms 2. The new algorithms. ***IMA Journal of of Applied Mathematics.*** **6**, 222-231.

Davidon W. C. (1959): Variable Metric Method of Minimization, Argonne National Laboratory Report No. ANL – 5999.

Fletcher, R. and Powell, M.J.D.(1963): A rapidly convergent descent method for minimization. ***Computer Journal***. **6**, 163-168.

Fletcher, R. (1970): A new approach for variable metric algorithms. ***Computer Journal***. **13**, 317 - 322.

Goldfarb, D. (1970): A family of variable metric methods derived by variational means, ***Mathematics of Computation***. **24**, 23-26.

Hennig Philipp and Martin Kiefel (2012): Quasi-Newton Methods: A New Direction. ***Journal of Machine Learning Research.*** **14**, 843-865.

Shanno, D. F. (1970): Conditioning of quasi-Newton methods for function minimization,

Mathematics of Computation. **24**, 647-656.

Xiaowei, J. and Yueting, Y. (2010): A self-scaling Quasi-Newton method for large scale unconstrained optimization. ***Journal of Information and Computational Science*** **7**: 8, 1739-1745.

Yuan Y. (1991): A modified BFGS algorithm for unconstrained optimization. ***IMA Journal of Numerical Analysis.*** **11**, 325-332.

Zhang J. Z., Deng, N. Y. and Chen L. H. (1999): A new Quasi-Newton equation and related methods for unconstrained optimization. ***Journal of Optimization and Applications.*** **102**, 147-167.