Iraqi Journal of Statistical Science (18) 2010
p.p. [1-34]

A Modified Super-linear QN-Algorithm for Unconstrained Optimization

Abbas Y. AL-Bayati* **Ivan Subhi Latif****

ABSTRACT

In this paper, we have proposed a modified QN-algorithm for solving a self-scaling large scale unconstrained optimization problems based on a new QN-update. The performance of the proposed algorithm is better than that used by Wei, Li, Yuan algorithm. Our numerical tests show that the new proposed algorithm seems to converge faster as compared with a standard similar algorithm in many situations .

خوارزمية مطورة لأشبـاه نيوتن فوق الخطية في الامثلية غير المقيدة

الملخص

في هذا البحـث تم استحداث خوارزمـية مطورة إلى أشـباه نيوتن فوق الخطـية فـي مجـال الأمثـلية غـير المـقيدة. الفـكرة الأسـاسـ ـ ـية تعتـمد على أحـد نمـاذج تحـويل أشباه نيـوتن . الخوارزمـية الجـديدة ذات كفـاءة اعلى مـن الخـوارزمـية التي أستـخدمها Wei, Li, Yuan. والنتـائج الـعملية

---

*Prof. Dr/ Dept. of Mathematics/College of Computers Sciences and Mathematics/ University of Mosul. E-mail :profabbaslbayati@yahoo.com

**Dept. of Mathematics, College of Scientific Education, University of Salahaddin, Erbil-Iraq. E-mail :ivansubhi2001@yahoo.com

Received:28/8 /2006 ____________________Accepted: 17 /12 / 2006

توضـح بأن النمـوذج الجديد المـستخدم ذات كفـاءة عالـية تقـارب أفضـل مقـارنة بالخوارزميـات المماثلة في مجالات عـدة.

KEYWORDS: Unconstrained Optimization, Quasi-Newton updates, Super-liner convergence.

1. Introduction

This paper analyzes the convergence properties of self-scaling QN-methods for solving the unconstrained optimization problem

$$\text{Min } f(x) \qquad x \in \Re^n, \qquad (1)$$

where f is twice continuously differentiable function. The convergence of QN-methods for unconstrained Optimization has been the subject of much analysis. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is generally considered to be the most effective among other variable metric methods for unconstrained Optimization problem. One interesting property of BFGS method is its self correcting mechanism (A detailed explanation for example, in Nocedal (see [11]) with this self correcting property, Powell see[13] shows that the BFGS method with an inexact line search satisfies Wolfe conditions is globally super-linearly convergent for convex problem, and Byrd, Nocedal and Yuan (see [5]) extend Powell's analysis to the restricted Broyden class excluding the DFP method. AL-Bayati's (see [1]) presented a new self-scaling variable-metric algorithm which was based on a known two-parameter family of rank-two updating formulae. The best of these algorithm are also modified to employ inexact line searches with marginal effect thus Wei, Li and Qi (see [15]) have proposed some modified BFGS that the average performance of their algorithm was better than standard BFGS algorithm.

Wei, et al. (see [14]) proved the super-linear convergence of Wei, Li and Qi (see [15]) algorithm under some suitable conditions. In this paper, a new modified QN-algorithm is proposed. The Basic idea is based on the new QN-equation $V_k = H_k y_k^*$ where y_k^* is the sum of y_k and $A_k V_k$ and A_k is some matrix.

This paper is organized as follows in the next section; we represent some basic properties of the modified BFGS algorithm. In section3 we prove the super-linear convergence for the modified QN-algorithm under some reasonable conditions. The search direction in a VM-method is the solution of the system of equations

$$d_k = -H_k g_k \qquad (2)$$

where the matrix H_k is an approximate to G_k^{-1} the new approximation H_{k+1} is chosen to take account of this new curvature information which is done by satisfying the condition

$$H_{k+1} y_k = \zeta_k V_k \qquad \text{(called QN-like condition)} \qquad (3)$$

where ζ_k is a scalar, generally for the QN-methods $\zeta_k = 1$ and hence equation (3) reduces to

$$H_{k+1} y_k = V_k \qquad \text{(called the QN-condition)} \qquad (4)$$

since information has been gained about f only in one dimension (along d_k), H_{k+1} is allowed to differ from H_k by a correction matrix C_k of at most rank two, i.e.

$$H_{k+1} = H_k + C_k \qquad (5)$$

the matrix C_k is therefore the update to H_k there are an infinite number of possible rank-two updates which satisfy the QN-condition but our main interest is in updates which form the Broyden one-parametric class (see [3]). The matrix H_{k+1} is defined by:

$$H_{k+1} = H_k - \frac{H_k y_k y_k^{\mathrm{T}} H_k}{y_k^{\mathrm{T}} H_k y_k} + \theta_k W_k W_k^{\mathrm{T}} + \frac{V_k V_k^{\mathrm{T}}}{V_k^{\mathrm{T}} y_k} \qquad (6)$$

with

$$W_k = (y_k^{\mathrm{T}} H_k y_k)^{1/2} \left[\frac{V_k}{V_k^{\mathrm{T}} y_k} - \frac{H_k y_k}{y_k^{\mathrm{T}} H_k y_k} \right] \qquad (7)$$

where θ_k is a scalar chosen such that $\theta_k \in [0,1]$ different choice of θ_k then defined different updates. the Davidon-Fletcher-Powell(DFP)update (see [7]) is defined as equation(4) with $\theta_k = 0$ where the Broyden-Fletcher-Goldfarb-Shanno(BFGS) update corresponds to $\theta_k = 1$ (see[6] and [9]). Oren (see[13]) found that a proper scaling of the objective function improve the performance of algorithms that use Broyden family of updated. Hence Oren's family of self-scaling VM-updates can be expressed as:

$$H_{k+1} = \left[H_k - \frac{H_k y_k y_k^{\mathrm{T}} H_k}{y_k^{\mathrm{T}} H_k y_k} + \theta_k W_k W_k^{\mathrm{T}} \right] \eta_k + \frac{V_k V_k^{\mathrm{T}}}{V_k^{\mathrm{T}} y_k} \qquad (8)$$

where

$$\eta_k = \frac{y_k^{\mathrm{T}} V_k}{y_k^{\mathrm{T}} H_k y_k} \qquad (9)$$

This choice for the scalar parameter η_k was made primarily because in this case η_k requires the quotient of two quantities which are already computed in the updating formula. Al-Bayati (see [2]) found another interesting family of VM-updates by further scaling of Oren's family of updates with a scalar

$$\delta_k = \frac{1}{\eta_k} \qquad (10)$$

So that the updating formulas becomes

$$H_{k+1} = H_k - \frac{H_k y_k y_k^{\mathrm{T}} H_k}{y_k^{\mathrm{T}} H_k y_k} + W_k W_k^{\mathrm{T}} + \sigma_k \left(\frac{V_k V_k^{\mathrm{T}}}{V_k^{\mathrm{T}} y_k}\right) \qquad (11)$$

For more details see[8]

2. Modified BFGS Algorithm:

Wei, Li and Qi proposed a new QN-equation (see[15])

$$B_{k+1} V_k = y_k^* \qquad (12)$$

where $B_k = H_k^{-1}$ where $y_k^* = y_k + A_k V_k$ and A_k is some matrix. By using equation(12)they gave BFGS type updates

$$B_{k+1} = B_k - \frac{B_k V_k V_k^{\mathrm{T}} B_k}{V_k^{\mathrm{T}} B_k V_k} + \frac{y_k^* y_k^{*\mathrm{T}}}{V_k^{\mathrm{T}} y_k^*} \qquad (13)$$

where $y_k^* = y_k + A_k V_k$ and $A_k = \dfrac{2(f(x_k) - f(x_{k+1})) + (g(x_{k+1}) - g(x_k))^{\mathrm{T}}}{\|V_k\|^2} I$ using equation(13) and the following Wolf Powell step-size rule

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g(x_k)^{\mathrm{T}} d_k \qquad (14)$$

where $\delta \in (0, \tfrac{1}{2})$ and $\sigma \in (\delta, 1)$ and

$$g(x_k + \alpha_k d_k)^{\mathrm{T}} d_k \ge \sigma g(x_k)^{\mathrm{T}} d_k \qquad (15)$$

2.1. Outline of the Modified BFGS Algorithm (MBFGS):

Corresponding (MBFGS) the outliers of MBFGS algorithms may be listed

Step1: choose an initial point $x_0 \in \Re^n$ and an initial positive definite matrix B_1 set $K = 1$.

Step2: if $\|g_k\| = 0$ then stop! Go to step2.

Step3: solve $H_k d_k + g_k = 0$ to obtain a search direction d_k

Step4: find d_k by Wolf-Powell step-size rule

$f(x_k + \alpha_k d_k) \le f(x_k) + \delta\alpha_k g(x_k)^{\mathrm{T}} d_k$ and $g(x_k + \alpha_k d_k)^{\mathrm{T}} d_k \ge \sigma g(x_k)^{\mathrm{T}} d_k$ where $\delta \in (0, \frac{1}{2})$ and $\sigma \in (\delta, 1)$ and

Step5: set $x_{k+1} = x_k + \alpha_k d_k$ .calculate A_k and update B_{k+1} by formula

$$B_{k+1} = B_k - \frac{B_k V_k V_k^{\mathrm{T}} B_k}{V_k^{\mathrm{T}} B_k V_k} + \frac{y_k^* y_k^{*\mathrm{T}}}{V_k^{\mathrm{T}} y_k^*}$$

where $y_k^* = y_k + A_k V_k$ and $A_k = \frac{2(f(x_k) - f(x_{k+1})) + (g(x_{k+1}) - g(x_k))^{\mathrm{T}}}{\|V_k\|^2} I$

Step6:set $k = k + 1$, go to step1

2.2. Some Properties of the MBFGS Algorithm:

The global convergence of the MBFGS algorithm needs the following three assumptions

Assumption2.2.1: The level set $\Omega = \{x | f(x) \le f(x_0)\}$ is contained in a bounded convex set D

Assumption2.2.2: The function f is continuously differentiable on D and there exists constant $L \ge 0$ such that $\|g(x) - g(y)\| \le L\|x - y\|$,for all $x, y \in D$.

Assumption2.2.3: The function f is uniformly convex that is there are positive constants m_1 and m_2 such that

$$m_1\|z\|^2 \le z^{\mathrm{T}} G(x) z \le m_2\|z\|^2 \qquad (16)$$

for all $x, z \in \Re^n$,where G denotes the Hessian matrix of f .

Theorem2.1: Let $\{x_k\}$ be generated by MBFGS algorithm then we have (see[11])

$$\lim_{k\to 0}\inf\|g_k\| = 0. \qquad (17)$$

The super-linear convergence analysis of the MBFGS algorithm needs the following assumptions:

Assumption2.2.4: $x_k \to x^*$ at which $g(x^*) = 0$ and $G(x^*)$ is positive definite**.**

Assumption2.2.5: G is holder continuous at x^* that is, there exists constant $r \in (0,1)$ and $M > 0$ such that

$$\|G(x) - G(x^*)\| \le M\|x - x^*\|^r \qquad (18)$$

for all x in neighborhood of x^* since $\{f(x_k)\}$ is a decreasing sequence also the sequence $\{x_k\}$ generated by MBFGS is contained in Ω and there exists a constant f^* such that

$$\lim_{k\to\infty} f(x_k) = f^* \qquad (19)$$

3. A new Modified QN- Algorithm:

In this section we propose a new QN-method based on the following QN-condition

$$V_k = H_k y_k^* \qquad (20)$$

where $y_k^* = y_k + A_k V_k$ and A_k is some matrix defined by $A_k = \dfrac{y_k^T V_k}{y_k^T H_k y_k} I$ using equation(20)and taking H_{k+1} as Al-Bayati update (see[1]).

$$H_{k+1} = H_k - \frac{H_k y_k y_k^{\mathrm{T}} H_k}{y_k^{\mathrm{T}} H_k y_k} + W_k W_k^{\mathrm{T}} + \sigma_k \left(\frac{V_k V_k^{\mathrm{T}}}{V_k^{\mathrm{T}} y_k}\right)$$

where

$$\delta_k = \frac{y_k^{\mathrm{T}} H_k y_k}{y_k^{\mathrm{T}} V_k}, \; W_k = (y_k^{\mathrm{T}} H_k y_k)^{1/2} \left[\frac{V_k}{V_k^{\mathrm{T}} y_k} - \frac{H_k y_k}{y_k^{\mathrm{T}} H_k y_k}\right]$$

and using also the following Armijo condition

$$f_k - f_{k+1} \geq c_1 \left\| V_k^{\mathrm{T}} y_k \right\|, V_k^{\mathrm{T}} y_k \geq (1 - c_2)(V_k^{\mathrm{T}} y_k) \quad (21)$$

where $c_1 \in (0, 1/2)$, $c_2 \in (c_1, 1)$ in (MBFGS) yields a new QN-algorithm given as below:

3.1. Outline of the Modified QN- Algorithm (NEW):

The outliers of the new algorithm may be given as:

Step1: choose an initial point $x_0 \in \Re^n$ and use-update positive definite matrix H_1 set $K = 1$.

Step2: if $\|g_k\| = 0$ then stop! Go to step2.

Step3: solve $H_k d_k + g_k = 0$ to obtain a search direction d_k

Step4: find d_k by Armijo line search step-size rule

$$f_k - f_{k+1} \geq c_1 \left\| V_k^{\mathrm{T}} y_k \right\|, V_k^{\mathrm{T}} y_k \geq (1 - c_2)(V_k^{\mathrm{T}} y_k)$$

where $c_1 \in (0, 1/2)$, $c_2 \in (c_1, 1)$

Step5: set $x_{k+1} = x_k + \alpha_k d_k$ .calculate a new A_k and update H_{k+1} by the following formula

$$H_{k+1} = H_k - \frac{H_k y_k^* y_k^{*\mathrm{T}} H_k}{y_k^{*\mathrm{T}} H_k y_k^*} + W_k W_k^{\mathrm{T}} + \sigma_k \left(\frac{V_k V_k^{\mathrm{T}}}{V_k^{\mathrm{T}} y_k^*}\right) \qquad (22)$$

where

$$\delta_k = \frac{y_k^{*\mathrm{T}} H_k y_k^*}{y_k^{*\mathrm{T}} V_k}, W_k = (y_k^{*\mathrm{T}} H_k y_k^*)^{1/2} \left[\frac{V_k}{V_k^{\mathrm{T}} y_k^*} - \frac{H_k y_k^*}{y_k^{*\mathrm{T}} H_k y_k^*} \right]$$

and $y_k^* = y_k + A_k V_k$ and $A_k = \dfrac{y_k^{\mathrm{T}} V_k}{y_k^{\mathrm{T}} H_k y_k} I$ (newly defined) (23)

Step6:set $k = k+1$, go to step1.

3.2. Some Theoretical Properties of the New Algorithm:

To show the global and super-linear convergence of the new algorithm use the same assumption 2.2.1,2.2.2,2.2.3 and 2.2.4 where used for all x in neighborhood of x^* since $\{f(x_k)\}$is a decreasing sequence also the sequence $\{x_k\}$generated by new algorithm is contained in Ω and that there exists a constant f^* such that

$$\lim_{k\to\infty} f(x_k) = f^* \qquad (24)$$

lemma3.2.1: let $(\alpha_k, x_{k+1}, g_{k+1}, d_{k+1})$be generated by the new algorithm then H_{k+1}^{-1} is positive definite for all k provided that $V_k^{\mathrm{T}} y_k^* > 0$.

Proof:-

The new algorithm has the following QN-condition

$$H_{k+1} V_k = y_k \qquad (25)$$

and preserve positive definiteness of the matrices $\{H_k\}$ if α_k is chosen to satisfy the Armijo condition

$$f_k - f_{k+1} \geq c_1 \left\| V_k^{\mathrm{T}} y_k \right\|, \qquad V_k^{\mathrm{T}} y_k \geq (1-c_2)(V_k^{\mathrm{T}} y_k)$$

where f_k denoted $f(x_k)$, $c_1 \in (0,½), c_2 \in (c_1,1)$.Note that the second condition in (21) guarantees that $V_k^{\mathrm{T}} y_k^* > 0$ whenever $g_k \neq 0$.

lemma3.2.2: let $\{x_k\}$ be generated by the new algorithm then we have

$$m_1 \|V_k\|^2 \leq V_k^{\mathrm{T}} y_k^* \leq m_2 \|V_k\|^2 \ , \ k = 1,2,... \qquad (26)$$

and

$$\|y_k^*\| \leq (2L + m_2)\|V_k\| \ , \ k = 1,2,... \qquad (27)$$

Proof:-

Using assumption2.2.2 and equation (16)

$$m_1 \|z\|^2 \leq z^{\mathrm{T}} G(x) z \leq m_2 \|z\|^2$$

and tailors formula we have

$$\begin{aligned} y_k^{*\mathrm{T}} V_k &= (g_{k+1} - g_k)^{\mathrm{T}} V_k \\ &= \mathrm{V}_{\mathrm{k}}^{\mathrm{T}} G(\zeta_1) V_k \end{aligned}$$

where $\zeta_1 \in (x_k, x_{k+1})$ thus (26) holds

$$m_1 \|V_k\|^2 \leq V_k^{\mathrm{T}} y_k^* \leq m_2 \|V_k\|^2 .$$

To prove (27) using the equation (26)

$$V_k^{\mathrm{T}} y_k^* \leq m_2 \|V_k\|^2$$

therefore

$\|y_k^*\| \le m_2 \|V_k\|^2$ (28)

also use assumption 2.2.2

$\|g(x) - g(y)\| \le L\|x - y\|$,for all $x, y \in D$.

Thus

$\|y_k^*\| \le L\|V_k\|$ (29)

from equation (28),(29) we get

$\|y_k^*\| \le (2L + m_2)\|V_k\|$.

Theorem3.1 let $\{x_k\}$ be generated by the new algorithm then x_k tends to x super-linearly.(see[2])

lemma3.2.3: suppose that $(\alpha_k, x_{k+1}, g_{k+1}, d_{k+1})$ be generated by the new algorithm and that G is continuous at x^* then we have

$\lim_{k\to\infty}\|A\| = 0$ (30)

Proof:-

By using Taylor's formula, we have

$$\begin{aligned} y_k^{*T}V_k &= (g_{k+1} - g_k)^T V_k \\ &= V_k^T G(\zeta_1) V_k \end{aligned}$$

and

$$f_k - f_{k+1} = (g_{k+1} - g_k)^{\mathrm{T}} V_k$$
$$= \mathrm{V}_k^{\mathrm{T}} G(\zeta_2) V_k$$

and

$$f_k - f_{k+1} = -g_{k+1}^{\mathrm{T}} V_k + \tfrac{1}{2} V_k^{\mathrm{T}} G(\zeta_2) V_k$$

where

$$\zeta_1 = x_k + \theta_{1k}(x_{k+1} - x_k)$$
$$\zeta_2 = x_k + \theta_{2k}(x_{k+1} - x_k)$$

and $\theta_{1k}, \theta_{2k} \in (0,1)$ .from the definition of A and lemma3.2.1 we get

$$A = \frac{V_k^{\mathrm{T}} H_{k+1}^{-1} V_k - V_k^{\mathrm{T}} G(\zeta_1) V_k}{y_k^{\mathrm{T}} H_k y_k}$$

and

$$V_k^{\mathrm{T}} H_{k+1}^{-1} V_k = V_k^{\mathrm{T}} G(\zeta_1) V_k$$

hence

$\|A_k\| \le \|G(\zeta_1) - G(\zeta_2)\|$ Therefore (30) holds.

lemma3.2.4: let $\{x_k\}$ be generated by the new algorithm denoted $Q = G(x^*)^{-1/2}$ then there are positive constants b_i , $\mathrm{i} = 1,2,3,4$ and $\eta \in (0,1)$ such that for all large k

$$\left\|H_{k+1}^{-1} G(x^*)^{-1}\right\| \le \left(\sqrt{1 - \rho W_k^2} + b_1 \tau_k + b_2 \|A\|\right) \left\|H_k - G(x^*)^{-1}\right\| + b_3 \tau_k + b_4 \|A_k\| \quad (31)$$

where

$\|A\| = \left\|Q^{\mathrm{T}} A Q\right\|_F$, $\|.\|_F$ is the forbenius norm of a matrix and W_k is defined as

$$W_k = \frac{\left\| Q^{-1}(H_k - G(x^*)^{-1} y_k^* \right\|}{\left\| H_k - G(x^*)^{-1} \right\| \left\| Q y_k^* \right\|} \quad (32)$$

In particular $\left\| H_k \right\|$, $\left\| H_k^{-1} \right\|$ are bounded

Proof:-

To prove (32)

$$H_{k+1} = H_k + \frac{(V_k - H_k y_k^*)V_k^{\mathrm{T}} + V_k (V_k - H_k y_k^*)^{\mathrm{T}}}{y_k^{*\mathrm{T}} V_k} + \frac{(y_k^{*\mathrm{T}})(V_k - H_k y_k^*) V_k V_k^{\mathrm{T}}}{(y_k^{*\mathrm{T}} V_k)^2})$$

$$= \left(I - \frac{V_k y_k^{*\mathrm{T}}}{y_k^{*\mathrm{T}} V_k} \right) H_k \left(I - \frac{y_k^* V_k^{\mathrm{T}}}{y_k^{*\mathrm{T}} V_k} \right) + \frac{V_k V_k^{\mathrm{T}}}{y_k^{*\mathrm{T}} V_k}$$

It is the dual form of the DFP type algorithm in the sense that $H_{k+1} \to H_{k+1}^{-1}$ and $V_k \to y_k$ we also have

$$\left\| Q y_k^* - Q^{-1} V_k \right\| \le \left\| Q \right\| . \left\| y_k^* - Q^{-2} V_k \right\|$$

$$= \left\| Q \right\| . \left[\left\| y_k^* - G(x^*) V_k \right\| \right]$$

$$\le \left\| Q \right\| \left(\left\| \int_0^1 G(x_k + \eta V_k) V_k \, d\eta - G(x^*) V_k \right\| + \left\| A V_k \right\| \right)$$

$$\le \left\| Q \right\| . \left\| V_k \right\| \left(\left\| \int_0^1 G(x_k + \eta V_k) V_k - G(x^*) \right\| d\eta + \left\| A \right\| \right)$$

$$\le \left\| Q \right\| . \left\| V_k \right\| \left(M_2 \int_0^1 G(x_k + \eta V_k) V_k - G(x^*) d\eta + \left\| A \right\| \right)$$

$$\le \left\| Q \right\| . \left\| V_k \right\| \left(M_2 \int_0^1 (\left\| x_k - x^* + \eta V_k \right\|^r) d\eta + \left\| A \right\| \right)$$

$$\leq \|Q\|.\|V_k\|\left(M_2\int_0^1\left(\eta\|x_k - x^*\| + (1-+\eta)\|x_k - x^*\|\right)^r d\eta + \|A\|\right)$$

$$\leq \|Q\|.\|V_k\|\left(M_2\tau_k + \|A\|\right)$$

since $\tau_k \to 0$ and $\|A\| \to 0$ it is clear that when k is large enough

$$\|Qy_k^* - Q^{-1}V_k \leq \beta\|QV_k\|\|,$$

for some constant $\beta \in (0, ½)$,therefore from lemma3.2.1 (with identification $V \to y_k$, $y \to V_k$, $H^{-1} \to H_k$, $A \to G(x^*)^{-1}$ and $M \to Q^{-1}$ there are constants $p \in (0,1)$ and $b_5, b_6 > 0$ such that

$$\|H_{k+1} - G(x^*)^{-1}\| \leq \left(\sqrt{1-\rho W_k^2} + b_5\frac{\|Q^{-1}V_k - Qy_k^*\|}{\|Qy_k^*\|}\right)\|H_k - G(x^*)^{-1}\| + b_6\frac{\|V_k - G(x^*)^{-1}y_k^*\|}{\|Qy_k^*\|}$$

where W_k is defined as more over ,there exists a constant b_7 such that for all k large enough

$$\|y_k^*\| = \|Qg_{k+1} - g_k + AV_k\|$$

$$\geq \|Qg_{k+1} - g_k\| - \|A\|\|QV_k\|$$

$$\geq b_7\|x_{k+1} - x_k\| - \|A\|Q\|V_k\|$$

$$= (b_7 - \|A\|\|Q\|)\|V_k\|$$

using $\|A\| \to 0$ the above inequality implies that there is a constant C such that when k is sufficiently large

$$\|Qy_k^*\| \geq C\|V_k\|$$

so we may obtain that

$$\frac{\|Qy_k^* - Q^{-1}V_k\|}{\|Qy_k^*\|} \le C^{-1}\|Q\|(M_2\tau_k + \|A\|) \qquad (33)$$

and

$$\frac{\|V_k - G(x^*)^{-1}y_k^*\|}{\|Qy_k^*\|} = \frac{\|V_k - Q^2 y_k^*\|}{\|Qy_k^*\|}$$

$$= \frac{\|Q(Qy_k^* - Q^{-1}V_k\|}{\|Qy_k^*\|}$$

$$\le \frac{\|Q\|\|Qy_k^* - Q^* V_k\|}{C\|V_k\|}$$

$$\le C^{-1}\|Q\|^2 (M_2\tau_k + \|A\|)$$

from which and (33),we get (31)

lemma3.2.5: Let $\{x_k\}$be generated by the new algorithm then the following Dennis More condition holds for the new technique

$$\lim_{k\to\infty} \frac{\|H_k^{-1} - G(x^*)V_k\|}{\|V_k\|} = 0 \qquad (34)$$

Proof:-

Using $\tau_k \to 0$, $\|A\| \to 0$ and $\|H_k\|$ is bounded and following inequality

$$\sqrt{1-\tau} \le 1 - \tfrac{1}{2}\tau \ , \ \forall \tau \in (0,1)$$

we can deduce that there are positive constants M_1 and M_2 such that for all large k]

$$\|H_k - G(x^*)^{-1}\| \le (1 - \tfrac{1}{2}\rho W_k^2)\|H_k - G(x^*)^{-1}\| + M_1\tau_k + M_2\|A_k\|$$

that is

$$\tfrac{1}{2}\rho W_k^2\left\|H_k - G(x^*)^{-1}\right\| \le \left\|H_k - G(x^*)^{-1}\right\| - \left\|H_{k+1} - G(x^*)^{-1}\right\| + M_1\tau_k + M_2\left\|A_k\right\|$$

summing the above inequality over k ,we get

$$\tfrac{1}{2}\rho\sum_{k=k_0}^{\infty} W_k^2\left\|H_k - G(x^*)^{-1}\right\| < +\infty$$

where k_0 is sufficiently large index such that (31) holds for all $k \ge k_0$ .In particular, we have

$$\lim_{k\to\infty} W_k^2\left\|H_k - G(x^*)^{-1}\right\| = 0$$

that is

$$\lim_{k\to\infty}\frac{\left\|Q^{-1}(H_k - G(x^*)^{-1}y_k^*\right\|}{\left\|Qy_k^*\right\|} = 0 \qquad (35)$$

Moreover we have

$$\left\|Q^{-1}(H_k - G(x^*)^{-1}y_k^*\right\| = \left\|Q^{-1}(H_k(G(x^*) - H_k^{-1})\right\|G(x^*)^{-1}y_k^*$$

$$\ge \left\|Q^{-1}(H_k(G(x^*) - H_k^{-1})V_k\right\| - \left\|Q^{-1}(H_k(G(x^*) - H_k^{-1})(V_k - G(x^*))^{-1}y_k^*\right\|$$

$$\ge \left\|Q^{-1}(H_k(G(x^*) - H_k^{-1})V_k\right\| - \left\|Q^{-1}(H_k(G(x^*) - H_k^{-1})(V_k - G(x^*))^{-1}y_k^*\right\| -$$

$$\left\|A\right\|\left\|Q^{-1}(H_k(G(x^*) - H_k^{-1})G(x^*)^{-1}V_k\right\|$$

using the fact that $\left\|H_k\right\|$ and $\left\|H_{k+1}\right\|$ are bounded and that $G(x)$ is continuous , we have

$$\left\|Q^{-1}(H_k(G(x^*)-H_k^{-1})V_k - G(x^*)^{-1}y_k^*\right\|$$
$$=\left\|Q^{-1}(H_k(G(x^*)-H_k^{-1})G(x^*)^{-1}(G(x^*)V_k - y_k^*\right\|$$
$$=\left\|Q^{-1}(H_k(G(x^*)-H_k^{-1})G(x^*)^{-1}[(G(x^*)-G(x_k))V_k + (G(x_k)V_k - y_k)]\right\|$$
$$\le \left\|Q^{-1}(H_k(G(x^*)-H_k^{-1})G(x^*)^{-1}\right\|\left\|(G(x^*)-G(x_k)\right\|\left\|V_k\right\|\left\|G(x_k)V_k - y_k\right\|$$
$$= O\left\|V_k\right\|$$

and

$$\left\|A\right\|\left\|Q^{-1}(H_k(G(x^*)-H_k^{-1})V_k\right\|$$
$$\le \left\|A\right\|\left\|Q^{-1}(H_k(G(x^*)-H_k^{-1})G(x^*)^{-1}\right\|\left\|V_k\right\|$$
$$= O(\left\|V_k\right\|) \qquad (36)$$

therefore ,there exists a positive constant $k>0$ such that

$$\left\|Q^{-1}(H_k - G(x^*)^{-1}y_k^*\right\| \ge K\left\|G(x^*)-H_k^{-1}\right\|V_k - O(\left\|V_k\right\|)$$

on the other hand, from equation (36)and lemma 3.2 we have

$$\left\|Qy_k^*\right\| \le \left\|Q\right\|\left\|y_k^*\right\| \le (2L+m_2)\left\|Q\right\|\left\|V_k\right\|$$

from the above inequality (35) and (36) we conclude that the Dennis-More condition holds. □

Theorem3.2

Let $\{x_k\}$ be a sequence generated by the new algorithm then the x_k tends to x super-linearly.

Proof:-

We will verify that $\alpha_k = 1$ for large k .since the sequence H_k is bounded we have

$\|d_k\| = \|H_k g_k\| \le \|H_k\| \|g_k\| \to 0$

by Taylor's expansion, we get

$$f_{k+1} - f_k - \delta g_k^{\mathrm{T}} d_k = (1-\delta) g_k^{\mathrm{T}} d_k + \tfrac{1}{2} d_k^{\mathrm{T}} G(x_k + \theta d_k) d_k$$

$$= -(1-\delta) d_k^{\mathrm{T}} H^{-1} d_k + \tfrac{1}{2} d_k^{\mathrm{T}} G(x_k + \theta d_k) d_k$$

$$= -(\tfrac{1}{2} - \delta) d_k^{\mathrm{T}} H^{-1} d_k - \tfrac{1}{2} d_k^{\mathrm{T}} (H_k^{-1} - G(x_k + \theta d_k)) d_k$$

$$= -(\tfrac{1}{2} - \delta) d_k^{\mathrm{T}} G(x^*) d_k + O(\|d_k\|^2)$$

where $\theta \in (0,1)$ and the last equality follows the Dennis More condition(34)thus

$$f_{k+1} - f_k - \delta g_k^{\mathrm{T}} d_k \le 0$$

for all large k .In other words $\alpha_k = 1$ the firs inequality of the Armijo equation(21) for all k sufficiently large on the other hand

$$g_k^{\mathrm{T}} d_k - \delta g_k^{\mathrm{T}} d_k = (g_{k+1} - g_k)^{\mathrm{T}} d_k + (1-\sigma) g_k^{\mathrm{T}} d_k$$

$$= d_k^{\mathrm{T}} G(x_k + \theta d_k) d_k - (1-\delta) g_k^{\mathrm{T}} H_k^{-1} d_k$$

$$= d_k^{\mathrm{T}} G(x_k + \theta d_k) d_k - (1-\delta) d_k^{\mathrm{T}} G(x_k) d_k + O(\|d_k\|^2)$$

$$= \delta d_k^{\mathrm{T}} G(x_k) d_k + O(\|d_k\|^2)$$

where $\theta \in (0,1)$ so we have

$$g_{k+1}^{\mathrm{T}} d_k \ge \delta g_k^{\mathrm{T}} d_k$$

which means that $\alpha_k = 1$ satisfies the Armijo equation(21) for all sufficiently large .Therefore we assert that $\alpha_k = 1$ for large k. Consequently, we can deduce that x_k converges super-linearly.

4. Numerical Results:

In this section, we compare the numerical behavior of the new algorithm with the MBFGS algorithm for different dimensions of test functions. Comparative tests were performed with (41)(specified in the Appendices 1 and 2) well-known test functions (see [10]). All the results are show in Table (1), (2) while Table (3) give the percentage of NOI and NOF. All the results are obtained with newly-programmed FORTRAN routines which employ double precision. The comparative performances of the algorithms taken in the usual way by considering both the total number of function evolutions (NOF) and the total number of iterations required to solve the problem (NOI) .In each case the convergence criterion is that the value of $\|g_k\| < 1\times10^{-5}$ the Armijo fitting by Frandsen (see [2]) and Powell line search (see [3]) used as the common linear search subprogram.

Each of the function was solved using the following algorithms

(1) MBFGS Algorithm :

(2)The new algorithm

The important thing is that the new algorithm needs less iteration, fewer evaluations of $f(x)$ and $g(x)$ than MBFGS. We can see that other algorithm may fail in some cases while the new algorithm always converges. Moreover numerical experiments also show that the new algorithm always convergence stabiles. Namely there are about (60-87) % improvements of NOI for all dimensions Also there are (30 -78) % improvements of NOF for all test functions.

Table (1):Comparison between the New algorithm and MBFGS algorithms using different value of 12< N <4320 for 1st test function.

N. OF Test	TEST FUNCTION	MBFGS NOI(NOF)					NEW NOI(NOF)				
		12	36	360	1080	4320	12	36	360	1080	4320
1	GEN-Shallow	33 11	79 26	108 26	188 52	354 101	15 11	15 11	15 11	15 11	15 11
2	GEN-Beal	26 14	33 14	53 21	463 226	103 28	15 12	15 12	15 12	16 13	16 13
3	Arwhad	17 7	24 8	42 13	31 9	75 16	27 23	14 9	17 11	17 10	17 10
4	GEN-Edger	17 8	23 10	34 8	34 8	62 12	11 8	11 8	11 8	12 9	12 9
5	Digonal4	13 2	14 2	22 5	30 4	36 5	12 8	12 8	12 8	12 8	12 8
6	EX-Denschnb	45 21	47 22	53 25	55 26	57 27	27 24	28 25	30 27	31 28	32 29
7	EX-BD1	384 181	372 185	1954 938	508 253	508 253	68 64	70 66	76 72	79 75	82 78
8	Digonal5	5 3	6 4	6 4	6 4	6 4	5 3	6 4	6 4	6 4	6 4
9	GEN-Strail	43 13	68 16	103 21	114 21	321 44	15 10	15 10	15 10	15 10	16 11
10	Digonal6	7	8	8	8	8	7	7	8	8	8

		5	6	6	6	6	5	5	6	6	6
11	Digonal7	12 7	12 7	13 8	13 8	13 8	12 7	12 7	13 8	13 8	13 8
12	EX-Denschnf	16 7	25 8	61 15	72 16	116 24	12 8	12 8	12 8	13 9	13 9
13	GEN-PSC1	32 14	34 15	36 16	36 16	38 18	29 25	30 26	30 26	33 28	33 28
14	GEN-Quadratic	39 22	40 19	62 24	123 49	117 42	12 9	15 12	17 14	17 14	21 18
15	Digonal8	7 4	7 4	8 5	8 5	8 5	7 4	7 4	8 5	8 5	8 5
16	GEN-penal1	19 8	25 11	139 52	527 180	527 180	7 4	7 4	32 26	32 26	32 26
17	GEN-TRI	37 17	101 45	78 76	222 87	252 89	19 15	19 16	20 16	19 15	20 16
General TOTAL of first 17 functions		752 344	918 402	2780 1263	2438 970	2601 862	300 240	295 235	337 272	346 279	356 289

Table (2):Comparison between the New algorithm and MBFGS algorithms using different value of 12< N <4320 for 2nd test function .

N. OF Test	TEST FUNCTION	MBFGS NOI(NOF)					NEW NOI(NOF)				
		12	36	360	1080	4320	12	36	360	1080	4320
1	EX-freudenst	F	F	F	F	F	26 (20)	26 (20)	27 (21)	27 (21)	27 (21)
2	Shanno	F	F	F	F	F	44 (34)	19 (13)	24 (17)	25 (17)	29 (21)
3	Liarwhd	F	F	F	F	F	402 (395)	1078 (1071)	22 (16)	70 (64)	70 (64)
4	EX-BD2	F	F	F	F	F	23 (16)	24 (17)	25 (18)	25 (18)	25 (18)
5	WX-Powell	F	F	F	F	F	56 (50)	57 (51)	58 (52)	59 (53)	59 (53)
6	Engval	F	F	F	F	F	117 (113)	114 (110)	97 (93)	81 (77)	80 (76)
7	Cosin	F	F	F	F	F	26 (20)	17 (13)	27 (22)	87 (75)	16 (12)
8	Biggs	F	F	F	F	F	17 (14)	33 (30)	233 (230)	669 (660)	669 (660)
9	GEN-Cubic	F	F	F	F	F	39 (34)	39 (34)	40 (35)	40 (35)	40 (35)
10	EX-Himmebil	F	F	F	F	F	17 (8)	18 (9)	18 (9)	18 (9)	18 (9)
11	EX-Host	F	F	F	F	F	39	39	40	40	40

							34	34	35	35	35
12	EX-three expontial	F	F	F	F	F	7 4	8 5	8 5	8 5	8 5
13	dquadratic	F	F	F	F	F	25 20	21 16	17 12	17 12	18 13
14	Perturbed	F	F	F	F	F	23 18	19 14	16 11	16 11	16 11
15	Raydan1	F	F	F	F	F	15 13	28 25	98 93	164 163	169 165
16	GEN-Helical	F	F	F	F	F	46 39	46 39	47 40	47 40	47 40
17	EX-Fred	F	F	F	F	F	24 16	24 16	24 16	24 16	24 16
18	GEN-Non digonal	F	F	F	F	F	71 55	74 57	70 54	65 50	83 61
19	Maratos	F	F	F	F	F	158 122	166 125	161 121	158 124	158 124
20	Full Hessian	F	F	F	F	F	8 4	7 3	8 3	9 3	9 3
21	Sincos	F	F	F	F	F	224 220	234 230	256 252	266 262	266 262
22	GQ2	F	F	F	F	F	86 79	62 54	55 47	73 68	73 68
23	Raydon2	F	F	F	F	F	7 5	7 5	8 6	8 6	8 6

Table(3): Percentage performance of the new algorithm against MBFGS algorithm for 100% in both NOI and NOF.

N	Costs	NEW
12	NOI NOF	60.11 30.23
36	NOI NOF	67.87 41.54
360	NOI NOF	87.88 78.46
1080	NOI NOF	85.81 71.24
4320	NOI NOF	86.31 66.47

5. Conclusions:

In this Paper, a new modified QN-algorithm for solving a self-scaling algorithm for solving large-scale unconstrained optimization problems is proposed .The new algorithm is a self-scaling QN- algorithm. The basic idea is based on a new QN-update proved to have super-linear convergence property. Our numerical results supports our claim and also indicate that the new algorithm may be competitive with the MBFGS algorithm in most cases of test function.

Appendix1:

All the test functions used in Table (1) for this paper are from general literature:

1. Generalized Shallow Function:

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

$x_0=[-2.,-2.,...,-2.,-2.]$.

2. Generalized Beale Function:

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

, $x_0=[-1.,-1.,...,-1.,-1.]$.

3. Arwhead Function (CUTE):

$$f(x)=\sum_{i=1}^{n-1}\left(-4x_i+3\right)+\sum_{i=1}^{n-1}\left(x_i^2+x_n^2\right)^2 \ ,$$

$x_0=[1.,1.,...,1.,1.]$.

4. Generalized Edger Function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}-2)^4+(x_{2i-1}-2)^2x_{2i}^2+(x_{2i}+1)^2 \ ,$$

$x_0=[1.,0.,...,1.,0.]$.

5. Diagonal4Function:

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

$x_0=[1.,1.,...,1.,1.]$, $\mathrm{c}=100$.

6. Extended Denschnb Function (CUTE):

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

$x_0=[0.1,0.1,...,0.1,0.1]$.

7. Extended Diagonal BDI Function:

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

$x_0=[0.1,0.1,...,0.1,0.1]$.

8. Diagonal5 Function:

$$f(x)=\sum_{i=1}^{n}\log\left(\exp(x_i)+\exp(-x_i)\right),$$

$x_0=[1.1,1.1,...,1.1,1.1]$.

9. Generalized Strait Function:

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

$x_0=[-2.,...,-2.]$.

10. Diagonal 6 Function:

$$f(x)=\sum_{i=1}^{n}(\exp(x_i)-(1+x_i))\ ,$$

$x_0=[1.,1.,...,1.,1.]$.

11. Diagonal 7 Function:

$$f(x)=\sum_{i=1}^{n}(\exp(x_i)-2x_i-x_i^2), \quad x_0=[1.,1.,...,1.,1.].$$

12. Extended Denschnf Function (CUTE):

$$f(x)=\sum_{i=1}^{n/2}\left(2(x_{2i-1}+x_{2i})^2+(x_{2i-1}-x_{2i})^2-8\right)^2+\left(5x_{2i-1}^2+(x_{2i}-3)^2-9\right)^2,$$

$x_0=[2.,0.,2.,0.,...,2.,0.]$.

13. Generalized pscl Function:

$$f(x)=\sum_{i=2}^{n-1}(x_i^2+x_{i+1}^2+x_i x_{i+1})^2+\sin^2(x_i)+\cos^2(x_i),$$

$x_0=[3.,0.1,...,3.,0.1]$.

14. Generalized quartic Function GQ1

$$f(x)=\sum_{i=1}^{n-1}x_i^2+(x_{i+1}+x_i^2)^2,$$

$x_0=[1.,1.,...,1.,1.]$.

15. Diagonal 8 Function:

$$f(x)=\sum_{i=1}^{n}x_i\exp(x_i)-2x_i-x_i^2 ,$$

$x_0=[1.,1.,...,1.,1.]$.

16. Generalized Penal1 Function:

$$f(x)=\sum_{i=1}^{n}(x_i-1)^2+eps(x_i^2-0.25)^2 \text{ ,}$$

$x_0=[1.,2.,...,n]$, $\text{eps}=1.\text{E-}5$.

17. Generalized Tridiagonal-1 Function:

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

$x_0=[2.,2.,...,2.,2.]$.

Appendix2:

All the test functions used in Table (2) for this paper are from general literature:

1. Extended Freudenstein & Roth Function:

$$f(x)=\sum_{i=1}^{n/2}\left(-13+x_{2i-1}+((5-x_{2i})x_{2i}-2)x_{2i}\right)^2+\left(-29+x_{2i-1}+((x_{2i}+1)x_{2i}-14)x_{2i}\right)^2$$

,

$x_0=[0.5.,-2.,0.5,-2.,...,0.5,-2.]$.

2. Nondia (Shanno-78) Function (Cute):

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

$x_0=[-1.,-1.,...,-1.,-1.]$.

3. Liarwhd Function (cute):

$$f(x)=\sum_{i=1}^{n}4(-x_1+x_i^2)^2+\sum_{i=1}^{n}(x_i-1)^2,$$

$$x_0=[4.,4.,...,4.].$$

4. Extended Block-Diagonal BD2 Function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}^2+x_{2i}^2-2.)^2+(\exp(x_{2i-1}-1.)+x_{2i}^3-2.)^2,$$

$$x_0=[1.5,2.,...,1.5,2.].$$

5. Extended Powell Function:

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

$$x_0=[3,-1,0,1...,3,-1,0,1].$$

6. Engval1 Function (CUTE):

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

$$x_0=[2.,2.,...,2.].$$

7. Cosine Function (CUTE):

$$f(x)=\sum_{i=1}^{n-1}\cos(-0.5x_{i+1}+x_i^2),$$

$$x_0=[1.,1.,...,1.,1.].$$

8. Biggsb1 Function (CUTE):

$$f(x) = (x_i - 1)^2 + \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 + (1 - x_n)^2 ,$$

$x_0 = [1.,1.,...,1.,1.]$.

9. Generalized Cubic function:

$$f(x) = \sum_{i=1}^{n/2} [100 \; (x_{2i} - x_{2i-1}^3)^2 + (1 - x_{2i-1})^2] ,$$

$x_0 = [-1.2,1.,...,-1.2,1.]$.

10. Extended Himmelblau Function:

$$f(x) = \sum_{i=1}^{n/2} \left(x_{2i-1}^2 + x_{2i} - 11\right)^2 + \left(x_{2i-1} + x_{2i}^2 - 7\right)^2 ,$$

$x_0 = [1.1,1.1,...,1.1,1.1]$.

11. Extended White & Holst Function:

$$f(x) = \sum_{i=1}^{n/2} c(x_{2i} - x_{2i-1}^3)^2 + (1 - x_{2i-1})^2 ,$$

$x_0 = [-1.2,1.,...,-1.2,1]$, $\text{c} = 100$.

12. Extended Three Exponential Terms Function:

$$f(x) = \sum_{i=1}^{n/2} \left(\exp(x_{2i-1} + 3x_{2i} - 0.1) + \exp(x_{2i-1} - 3x_{2i} - 0.1) + \exp(-x_{2i-1} - 0.1)\right),$$

$x_0 = [0.1,0.1,...,0.1,0.1]$.

13. Dqudrtic Function (CUTE):

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

$x_0=[3.,3.,...,3.,3.]$, $c=100, d=100$.

14. Perturbed Penalty Function:

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

$x_0=[0.5,0.5,...,0.5,0.5]$.

15. Raydan 1 Function:

$$f(x)=\sum_{i=1}^{n}\tfrac{i}{10}\left(\exp(x_i)-x_i\right),$$

$x_0=[1.,1.,...,1.,1.]$.

16. General Helical Function:

$$f(x)=\sum_{i=1}^{n/3}\left(100x_{3i}-10*H_i\right)^2+100(R_i-1)^2+x_{3i}^2,$$

where $R_i=sqrt(x_{3i-2}^2+x_{3i-1}^2), H_i=\begin{cases}(2\pi)^{-1}\tan^{-1}\dfrac{x_{3i-1}}{x_{3i-2}} & if\ x_{3i-2}>0\\ 0.5+(2\pi)^{-1}\tan^{-1}\dfrac{x_{3i-1}}{x_{3i-2}} & if\ x_{3i-2}<0\end{cases}$

$x_0=[-1.,0.,0....,-1.,0.],0..$

17. Extended Fred Function:

$$f(x)=\sum_{i=1}^{n/2}(-13+x_{2i-1}+(5-x_{2i})+(x_{2i}-2)(x_{2i}))^2+\sum_{j=1}^{n/2}(-29+x_{2i-1}+(1-x_{2i})+(x_{2i}-14)(x_{2i}))^2,$$

$x_0=[1.,2.,...,n]$

18. Generalized Non diagonal function:

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

$x_0=[-1.,\ldots,-1.]$.

19. Extended Martos Function:

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

$x_0=[1.1,0.1,\ldots,1.1,0.1]$.

20. Full Hessian Function:

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

$x_0=[1.,1.,\ldots,1.,1.]$.

21. SINCOS Function:

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

$x_0=[3.,0.1,\ldots,3.,0.1]$.

22. Generalized Quartic Function GQ2:

$$f(x)=(x_1^2-1)^2+\sum_{i=2}^{n}(x_i^2-x_{i-1}-2)^2 ,\quad x_0=[1.,1.,\ldots,1.,1.].$$

23. Raydan 2 Function:

$$f(x)=\sum_{i=1}^{n}\left(\exp(x_i)-x_i\right),\quad x_0=[1.,1.,\ldots,1.,1.].$$

References:

[1] A. Y. AL-Bayati. "*Anew Non-Quadratic Model for Unconstrained Non-liner optimization method* " Natural and applied series Mu'tah J. Research & Studies, Vol.(8), pp. 133-155, (1993).

[2] A. Y. AL-Bayati. "*A new family of Self-Scaling VM-algorithms*"**,** Journal of Education and Science, Mosul University, Mosul, Iraq (1991).

[3] B. Bunday. "*Basic Optimization Methods* " Edward Arnold. Bedford square, London, UK.(1984).

[4] C. G. Broyden. "*The convergence of a class of double rank minimization algorithms* ", JIMA, Vol.(6), pp. 76-90, (1970).

[5] R. H. Byrd; J. Nocedal. and Y. Yuan. "*Global convergence of a class of quasi Newton methods on convex problems*"**, S**IAM J. Numer. Anal. Vol.(24), pp.1171-1190, (1987).

[6] R. Fletcher. "*A new approach to variable metric algorithm*", J. computer, Vol.(13), pp. 317-322, (1970).

[7] R. Fletcher and M. J. D. Powell." *A rapidly Convergent Descent Method for Minimization*", J. Computer, Vol.(6),PP.163-168, (1963).

[8] P. E. Fran; K. Jon, H. B. Niel and O .Ting. "*Un constrained Optimization test function*", Informatics and mathematical Modeling Technical university of Denmark.,3.edition,March , (2004).

[9] P. E. Gill and W. Murray. "*conjugate Gradient Method for Large-Scale Nonlinear Optimization*", Report so, Department of Operation Research, Stand ford University, California, pp.79-159, (1979).

[10] J. Nocedal. "*Unconstrained Optimization Test function Research Institute for Informatics*", Center for advanced Modeling and Optimization, Bucharest1, Romania. April 28, (2005).

[11] J. Nocedal. "*Theory of Algorithms for Unconstrained Optimization* ", Acta. Numer., Vol.(1), pp.199-242, (1992).

[12] S .S. Oren ."*Self-Scaling Variable Metric Algorithm*", Part II, Management Sci., Vol.(20), pp.863-874, (1974).

[13] M. J. D. Powell. "*Some Global Convergence Properties of A variable Metric Algorithm for Minimization Without Exact Line Searches*" . In R. W. Cottle and C. E. Lenka, editors, Nonlinear programming ,J. Amer. Math. Soc., pp.53-72, (1976).

[14] Z. Wei; G. Yu; G. Yuan and Z., Lian. "*The Super-linear Convergent of A modified BFGS Type Method for Unconstrained Optimization* ". J. Comp. Optim. & Appl. Vol.(29), pp.315-332, (2004).

[15] Z. Wei, G., Li and L. Qi, "*New Quasi-Newton Methods for Unconstrained Optimization Problems*" To appear in Mathematical Programming, Republic of china, October(2000).