A CLASS OF DERIVATIVE-FREE ALGORITHMS

FOR UNCONSTRAINED MINIMIZATION

Ann-Sheng A. Chien

TR 74-220

August 1974

Department of Computer Science
Cornell University
Ithaca, New York 14853

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would first of all like to express my gratitude to Prof. John E. Dennis, Jr., my thesis advisor, for his excellent guidance and encouragement throughout the author's graduate studies. I am also indebted to Prof. John H. Williams and Alfred N. Schatz for serving on my special committee.

I would like to thank Prof. J. J. Moré who gave me many useful suggestions when this work was carried out. I would also like to thank the Dept. of Computer Science for supporting my graduate education for the academic year 1971-1972 and the National Science Foundation for support throughout the period 1972-1974.

Finally I would like to thank my fiancé, Way, without his patience, forbearance, and encouragement, I could not have done this work.

I. INTRODUCTION

This thesis will be concerned with the problem of calculating the least value of a given function $f(x)$, $x = (x_1, x_2, \ldots, x_n)$, in the case that only values of f are available. We expect that f is twice continuously differentiable but suppose that it is perferable not to calculate even one derivative.

In fact, our purpose is to find an algorithm that can find the minimum of f by using only values of f and that has some desirable superlinearly convergent properties. It seems as though if one wants to obtain a superlinear rate of convergence, one can in theory follow the method of a superlinearly convergent algorithm that requires the evaluation of derivatives, replacing derivatives by differences, using a step-size that tends to zero at the right speed.

Several people have given algorithms that calculate the minimum value of f by using $f(x)$ and its first derivatives. Among these algorithms some, for example the Davidon-Fletcher-Powell algorithm (1963) and Powell's dogleg algorithm (1970), use the calculated first derivative to estimate the second derivative matrix. The approximation to the second derivative matrix is to be revised for each iteration by some update formula. These two algorithms have been proved to be superlinearly convergent under some reasonable conditions and seem to be two good methods to follow. Stewart (1967) described a method which modified the Davidon-Fletcher-Powell algorithm by using finite differences to approximate the first derivative. Gill and Murray (1970) gave another modification to the same algorithm. Numerical results show that both methods are quite good. But neither of them gave any theoretical convergence results of their algorithms.

Recently, Broyden-Dennis-Moré proved that several well-known direct prediction quasi-Newton methods have local super-linearly convergent properties. Powell (1974) has proved some uncommon properties of a class of minimization algorithms which maintain an approximation to the Hessian and further require the change in x to be subject to a bound that is also revised automatically. Our present concern is extending all the algorithms included in Broyden-Dennis-Moré's paper and Powell's paper to accept the finite difference approximation to the gradient. Since both papers have proved the convergence property under the assumption that the second derivative approximations do not have to converge to the true Hessian at the solution, it is desirable to prove our extended algorithms will converge in the same case. Hence, it is proved in this thesis that all the algorithms after the modification preserve all the convergence properties. And most of the results established in Powell's paper including the one concerning the global convergence behavior of the algorithms have been extended to the case that can be applied to our algorithms.

The organization is as follows. Chapter II contains some definitions, short descriptions of all the quasi-Newton methods under consideration here and the description of how the modification is to be carried out. In Chapter III, we prove some general convergence results. In Chapter IV, we discuss the convergence of the single-rank methods. In Chapter V, we discuss the double-rank methods. In Chapter VI, we first describe the modified dogleg algorithm then we outline a class of minimization algorithms to which the modified Powell algorithm belongs. We will then prove a global convergence property of this class of algorithm under very mild conditions of f(x). And superlinearly convergence of this class of algorithms is proved even though the second derivative approximation may not converge to the true Hessian at the solution. Finally we apply our results to the modified Powell algorithm and conclude its convergence. There are two appendices following Chapter VI. Appendix A is a survey of all the derivative-free methods for unconstrained optimization. Appendix B contains a FORTRAN subroutine for calculating the finite difference approximation to the gradient and some numerical results which are given by the subroutine MINFA with this subroutine.

II. THE ALGORITHMS

Definitions

Let R^n be the real n-dimensional linear space of column vectors and $L(R^n)$ be the linear space of real matrices of order n. And $||\cdot||$ will denote an arbitrary norm in R^n or the operator norm it induces in $L(R^n)$.

Definition 1: The Frobenious norm is defined by

$$||A||_F = (\text{trace } (A^T A))^{\frac{1}{2}} = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}}$$

for all $A \in L(R^n)$.

Definition 2: Let $P\{L(R^n)\}$ be the collection of all non-empty subsets of $L(R^n)$. Then the update function U is a mapping from $R^n \times L(R^n) \times R^n$ to $P\{L(R^n)\}$.

Definition 3: A function $g: R^n \to R^n$ is (Gateaux) differentiable if and only if its Jacobian matrix $G(x)$ satisfies

$$\lim_{t \to 0} \frac{g(x+th) - g(x)}{t} = G(x)\, h$$

for all $x \in R^n$, $h \in R^n$.

Definition 4: For a differentiable function $f: R^n \to R$, the finite difference derivative $\bar{g}(x, h)$ of f is a vectorvalued function defined on $R^n \to R^n$ such that

$$(\bar{g}(x,\ h))_i = \begin{cases} \dfrac{f(x + (h^T e_i)e_i) - f(x)}{h^T e_i} & \text{if } h^T e_i \neq 0, \\ \dfrac{\partial f(x)}{\partial x_i} = (g(x))_i & \text{if } h^T e_i = 0. \end{cases} \qquad (2.1)$$

for all $x \in R^n$ and $h \in R^n$. Where e_i's are unit coordinate vectors, i.e., $e_k = (\delta_{k1}, \delta_{k2}, \cdot \cdot \cdot, \delta_{kn})$ where δ_{kj} is the kronecker delta.

Remark 1: Suppose f is twice differentiable in an open set $D \subset R^n$, g is the gradient of f which satisfies a Lipschitz condition on D and G is the Hessian matrix of f. Then we have

$$|\bar{g}(x,h)_i - (g(x))_i| = \left|\frac{f(x+h_i e_i) - f(x)}{h_i} - \frac{\partial f}{\partial x_i}(x)\right|$$

$$= \left|\frac{\partial f}{\partial x_i}(x+\theta_i h_i e_i) - \frac{\partial f}{\partial x_i}(x)\right|$$

$$= |(g(x+\theta_i h_i e_i))_i - (g(x))_i| \leq ||g(x+\theta_i h_i e_i) - g(x)||$$

$$\leq C_0 \; ||h||$$

where $0 \leq \theta_i \leq 1$, C_0 is the Lipschtz constant. Hence, we have

$$||\bar{g}(x,h) - g(x)|| \leq C_0 \; ||h||.$$

In particular, this is true if $||G(x)||$ is bounded above by C_0.

Definition 5: The matrix norm $||\cdot||_m$ is consistent with the vector norm $||\cdot||_v$ if $||Ax||_v \leq ||A||_m \, ||x||_v$ where $A \in L(R^n)$, and $x \in R^n$.

Remark 2: If a matrix norm $||\cdot||_m$ is induced by the vector norm $||\cdot||_v$ then we have $||Ax||_v \leq ||A||_m \, ||x||_v$ for all $A \in L(R^n)$ and $x \in R^n$. Therefore, $||\cdot||_m$ is consistent with $||\cdot||_v$.

Description of the Algorithms

The algorithms are based on modifying some superlinearly convergent methods for the unconstrained minimization of a function of several variables. Let the function be $f:R^n \to R$ and let $g:R^n \to R^n$ be the gradient. We seek an algorithm which can solve $g(x) = 0$. It is assumed that $f(x)$ is twice differentiable in an open convex set D, that is $g(x)$ is differentiable on D. Moreover, we will assume that for constant $p > 0$ and $K > 0$

$$||G(x) - G(x^*)|| \leq K||x - x^*||^p$$

for all $x \in D$, and x^* satisfies $g(x^*) = 0$.

The methods which we will modify are quasi-Newton methods.

They take the form $x^{k+1} = x^k - B_k^{-1} g(x^k)$ where $\{B_k\}$ is generated by some update to satisfy the quasi-Newton equation, i.e.,

$$B_{k+1} (x^{k+1} - x^k) = g(x^{k+1}) - g(x^k).$$

There are two classes of quasi-Newton methods that are included in the Broyden-Dennis-Moré's paper [4].

The first class of methods are single rank update methods, in which the update matrices are of the form

$$\overline{B} = B + (y - Bs) \frac{c}{c^T s}$$

where $s = \overline{x} - x$ with $\overline{x} = x - B^{-1} g(x)$, $y = g(\overline{x}) - g(x)$, $c^T s \neq 0$. Here the identification of x, B and $\overline{x}$, $\overline{B}$ with x^k, B_k and x^{k+1}, B_{k+1} is intended. We can choose to work with $H_k = B_k^{-1}$ where $\{H_k\}$ are generated by the following formula

$$\overline{H} = H + (s - Hy) \frac{d}{d^T y}$$

where $s = \overline{x} - x$ with $\overline{x} = x - Hg(x)$, $y = g(\overline{x}) - g(x)$ and $d = H^T c$. The value of c, d are different from one method to another. They include Broyden's first method [3] ($c = s$, $d = H^T s$), Pearson's method [19] ($c = y$, $d = H^T y$), McCormick's method [19] ($c = B^T s$, $d = s$) and Broyden's second method [3] ($c = B^T y$, $d = y$).

For the double rank methods, the update matrices are

$$\bar{B} = B + \frac{(y - Bs)c^T + c(y - Bs)^T}{c^T s} - \frac{s^T(y - Bs)\ cc^T}{(c^T s)^2}$$

where $s = \bar{x} - x$ with $\bar{x} = x - B^{-1} g(x)$, $y = g(\bar{x}) - g(x)$, $c^T s \neq 0$ or update H:

$$\bar{H} = H + \frac{(s - Hy)d^T + d(s - Hy)^T}{d^T y} - \frac{y^T(s - Hy)\ dd^T}{(d^T y)^2}$$

where $s = \bar{x} - x$ with $\bar{x} = x - Hg(x)$, $y = g(\bar{x}) - g(x)$. They include the Powell symmetric Broyden algorithm [21] $(c = s)$, Davidon-Fletcher-Powell algorithm [12] $(c = y)$, Greenstadts method [13] $(d = y)$, Broyden-Fletcher-Goldfarb-Shanno algorithm [27] $(d = s)$.

We will modify all these methods by using the finite difference derivative $\bar{g}(x, h)$ of $f(x)$ to substitute for $g(x)$. Therefore, in our case $\bar{x} = x - B^{-1}\ \bar{g}(x, h)$, $y = \bar{g}(\bar{x}, \bar{h}) - \bar{g}(x, h)$ where the step-sizes h, $\bar{h}$ satisfy the following two conditions:

(i) $||\bar{h}|| \leq \min(C_1\ ||\bar{x} - x||^2, ||h||)$ for some $0 < C_1 < \infty$. (2.2)

(ii) $||h|| - ||\bar{h}|| \leq C_2\ ||\bar{x} - x||^2$ for some $0 < C_2 < \infty$. (2.3)

And we further require h, $\bar{h}$ to be properly chosen that

$\bar{g}(x, h) = 0$, $\bar{g}(\bar{x}, \bar{h}) = 0$ will not happen when $g(x) \neq 0$, $g(\bar{x}) \neq 0$. The capability of choosing such h, $\bar{h}$ is proved by the following lemma:

Lemma 2.1: Suppose $f:R^n \to R$ is differentiable in the open convex set D. Let g(x) be the gradient of f(x) and assume that $g(x^*) = 0$ and $g(x) \neq 0$ for all $x \in D - \{x^*\}$. Let $\bar{g}(x, h)$ be the finite difference derivative defined by (2.1). Then for all $x \neq x^*$ in D, the set

$$H_1(x) = \{h \in R^n : \bar{g}(x, h) \text{ is well-defined and nonzero}\}$$

is nonempty and there is $\delta_x > 0$ such that

$$H_1(x) \supset \{h \in R^n : ||h|| \leq \delta_x\}$$

Proof: For all $x \neq x^*$ in D, there is at least one component of g(x), say $(g(x))_i$, which is nonzero. Since

$$(g(x))_i = \lim_{\lambda \to 0} \frac{f(x + \lambda e_i) - f(x)}{\lambda}$$

this implies that for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $|\lambda| < \delta$

$$-\varepsilon < \frac{f(x + \lambda e_i) - f(x)}{\lambda} - (g(x))_i < \varepsilon$$

Therefore, let $\varepsilon_0 = |\frac{(g(x))_i}{2}| > 0$ and there is $\delta_0(x) > 0$ such that for all λ with $|\lambda| < \delta_0(x)$ we have

$$\left|\frac{f(x + \lambda e_i) - f(x)}{\lambda}\right| > \left|\frac{(g(x))_i}{2}\right| > 0$$

Hence, for any $x \in D$, if we choose $\delta_x > 0$ so small that for $h = (h_1, h_2, \cdot \cdot \cdot, h_n) \in R^n$ with $||h|| < \delta_x$ we have $|h_i| < \delta_0(x)$, then we have $(\overline{g}(x, h))_i \neq 0$, that is $\overline{g}(x, h)$ is nonzero. So, $h \in R^n$ with $||h|| < \delta_x$ implies that $h \in H_1(x)$. This concludes the proof.

Lemma 2.2: Let f, g and $\overline{g}(x, h)$ satisfy all the assumptions of Lemma 2.1 and assume that for some $x \in D$ and nonsingular $B \in L(R^n)$ the vector $x - B^{-1} g(x)$ belongs to D. Then the set

$$H_2(x) = \{h \in H_1(x) : \overline{x} = x - B^{-1} \overline{g}(x, h) \in D\}$$

is nonempty and there exists $\eta_x > 0$ such that

$$H_2(x) \supset \{h \in R^n : ||h|| < \eta_x\}.$$

Proof: Since D is an open set, there exists $\delta > 0$ such that for all $y \in D$ with $||y - (x - B^{-1} g(x))|| < \delta$, $y \in D$. Let $\gamma_0 > 0$ be an upper bound for $||B^{-1}||$. By the definition of $g(x)$, there exists $\eta_0(x) > 0$ such that for all $h \in R^n$ with $||h|| < \eta_0(x)$, we have

$$||\overline{g}(x, h) - g(x)|| < \delta/\gamma_0$$

Hence, if $h \in R^n$ with $||h|| < \eta_x = \min(\eta_0(x), \delta_x)$ then $h \in H_1(x)$, and $||x - B^{-1}\,\bar{g}(x, h) - (x - B^{-1}\,g(x))|| \leq ||B^{-1}||\,||\bar{g}(x, h) - g(x)|| \leq \gamma_0\,\delta/\gamma_0 = \delta$. Therefore, $x - B^{-1}\,\bar{g}(x, h) \in D$ which implies $h \in H_2(x)$, and this completes the proof.

Discussion of the Algorithms

The consistency of condition (2.2) and (2.3) can be verified by the following theorem which proves that for any $C_1 > 0$, as long as we take precautions in choosing the h which defines x, (2.3) can be satisfied by any $\bar{h}$ which satisfies (2.2).

Theorem 2.3: Let f(x), g(x) and $\bar{g}(x, h)$ satisfy the assumptions of Lemma 2.1 and assume that for some $x \in D$ and nonsingular $B \in L(R^n)$ the vector $x - B^{-1}\,g(x)$ belongs to D. Then for any nonzero $h \in H_2(x)$ the set

$$H(h) = \{\bar{h} \in H_1(x) : ||\bar{h}|| \leq \min(C_1\,||\bar{x} - x||^2,\ ||h||)\} \tag{2.4}$$

is nonempty, provided $x \neq x^*$. Moreover, if f is twice continuously differentiable, let G be the Hessian of f. Suppose for any $x \in D$, G(y) is nonsingular for all $y = rx + (1-r)x^*$, $r \in [0, 1]$, then for any $C_2 > 0$ there exists $h \in H_2(x)$ such that for all $\bar{h} \in H(h)$ the following holds

$$||h|| - ||\bar{h}|| \leq C_2\,||\bar{x} - x||^2 \tag{2.5}$$

<u>Proof</u>: Since $x - B^{-1} g(x) \in D$, it follows from Lemma 2.2 that $H_2(x) \neq \phi$. Therefore, we can find an $h \in H_2(x)$ and $\bar{x} = x - B^{-1} \bar{g}(x, h) \in D$. If $x \neq x^*$, Lemma 2.1 implies that there exists $\delta_{\bar{x}} > 0$ such that for all $\bar{h} \in R^n$ with $||\bar{h}|| < \delta_{\bar{x}}$ we have $\bar{h} \in H_1(x)$. So, for any $\bar{h} \in R^n$ which satisfies $||\bar{h}|| \leq \min(C_1 ||\bar{x} - x||^2, ||h||, \delta_{\bar{x}})$ we obtain $\bar{h} \in H(h)$. Hence $H(h)$ is well-defined and nonempty.

The second part of this theorem follows if we can prove that there exists an $h \in H_2(x)$ such that $||h|| \leq C_2 ||\bar{x} - x||^2$. Since

$$||\bar{x} - x|| = ||B^{-1} g(x, h)|| = ||B^{-1} g(x) + B^{-1} \bar{g}(x, h) - B^{-1} g(x)||$$

$$\geq ||B^{-1} g(x)|| - ||B^{-1}|| \, ||\bar{g}(x, h) - g(x)||.$$

Let λ_0 be an upper bound for $||B^{-1}||$, λ_1^{-1} be an upper bound for $||B||$ and γ_0 be an upper bound for $||G(y)^{-1}||$ for $y \in [x, x^*]$,

$$||\bar{x} - x|| \geq \lambda_1 ||g(x)|| - \lambda_0 ||\bar{g}(x, h) - g(x)||. \qquad (2.6)$$

By a well-known result, for all $x \in D$, $y \in D$

$$||g(x) - g(x^*) - G(y)(x - x^*)|| \leq \sup_{t \in (0,1)} ||G(tx + (1-t)x^*) - G(y)|| \, ||x - x^*||$$

Since G is continuous, there exists $t_0 \in (0, 1)$ such that $||g(x) - g(x^*) - G(y)(x - x^*)|| \leq ||G(t_0 x + (1 - t_0)x^*) - G(y)|| \, ||x - x^*||$. And there exists $y_0 = rx + (1 - r) x^*$

for some $r \in (0, 1)$ such that $||G(t_0 x + (1 - t_0)x^*) - G(rx + (1 - r)x^*)|| \leq 1/2\gamma_0$.

Hence with this y_0, we have $G(y_0)$ nonsingular, $||G(y_0)^{-1}|| < \gamma_0$ and

$$||g(x)|| \geq (||G(y_0)^{-1}||^{-1} - 1/(2\gamma_0))\ ||x - x^*|| \geq 1/2\gamma_0\ ||x - x^*||. \tag{2.7}$$

for all $x \in D$. Therefore, from (2.6) and (2.7) $||\overline{x} - x|| \geq \frac{\lambda_1}{2\gamma_0}\ ||x - x^*|| - \lambda_0\ ||\overline{g}(x, h) - g(x)||$.

By the definition of $g(x)$, there exists $\delta_0 > 0$ such that for all $h \in R^n$ with $||h|| < \delta_0$,

$$||\overline{g}(x, h) - g(x)|| < \frac{\lambda_1\ ||x - x^*||}{4\ \gamma_0 \lambda_0}$$

So, we choose $h \in R^n$ with $||h|| \leq \min(\delta_0, \frac{C_2\ \lambda_1^2\ ||x - x^*||^2}{16\ \gamma_0^2})$.

Then, $||h|| \leq C_2\ (\frac{\lambda_1\ ||x - x^*||}{4\ \gamma_0})^2 \leq C_2\ (\frac{\lambda_1}{2\gamma_0}\ ||x - x^*|| - \lambda_0\ ||\overline{g}(x, h) - g(x)||)^2 \leq C_2\ ||\overline{x} - x||^2$.

Thus, we know with this h, (2.5) holds for all $\overline{h} \in H(h)$.

Theorem 2.3 shows that there are nonzero h's and $\overline{h}$'s that will satisfy both (2.2) and (2.3). This implies (2.2) and (2.3) are consistent. From the proof of Theorem 2.3 we can see that the existence of h does not depend on the particular B we chose. It depends on upper bounds of $||B||$, $||B^{-1}||$ and the constant L such that $||g(x)|| \geq L\ ||x - x^*||$. Hence we have the following corollary:

Corollary 2.4: Let $f(x)$, $g(x)$ and $\overline{g}(x, h)$ satisfy all the assumptions of Lemma 2.1. For any $x \in D$, suppose

$$||g(x)|| \geq L\, ||x - x^*|| \text{ for some constant } L > 0.$$

Let γ_0 γ_1 be two positive constants. Then for any $C_2 > 0$, there exists $h \in H_2(x)$ such that $||h|| \leq C_2\, ||B^{-1}\, \overline{g}(x, h)||^2$ for every nonsingular $B \in L(R^n)$ with $||B|| \leq \gamma_0$, $||B^{-1}|| \leq \gamma_1$. Moreover, this h will make (2.5) hold for all $\overline{h} \in H(h)$ provided $\overline{x} = x - B^{-1}\, \overline{g}(x, h) \neq x^*$.

Proof: Immediately follows from the proof of Theorem 2.3.

Now we can state the modified algorithm (call it Algorithm MQ) more specifically in the following way:

1. Start from $x^0 \in D$ and nonsingular $B_0 \in L(R^n)$. Choose $h^0 \in H_2(x^0)$. Then get $x^1 = x^0 - B_0{}^{-1}\, \overline{g}(x^0\ \ h^0)$. Choose $h^1 \in H_2(x^1)$ such that

$$||h^0|| - ||h^1|| \leq C_2\, ||x^1 - x^0||^2 \tag{2.8}$$

and

$$||h^1|| \leq \min(C_1\, ||x^1 - x^0||^2,\ ||h^0||) \tag{2.9}$$

2. For any interer $k \geq 1$, suppose x^k, B_{k-1}, h^k are known and $x^k \neq x^*$. Get B_k by updating B_{k-1} and get $x^{k+1} = x^k - B_k{}^{-1}\, \overline{g}(x^k, h^k)$. If $x^{k+1} \neq x^k$, choose $h^{k+1} \in H_2\ (x^{k+1})$ such that

$$||h^{k+1}|| \le \min(C_1\ ||x^{k+1} - x^k||^2,\ ||h^k||) \tag{2.10}$$

$$||h^k|| - ||h^{k+1}|| \le C_2\ ||x^{k+1} - x^k||^2 \tag{2.11}$$

If x^0, B_0 are chosen properly such that $x^0 - B_0^{-1}\, g(x^0)$ is in the region where f is defined then Lemma 2.2 and Theorem 2.3 imply that there exists $h^0 \varepsilon H_2(x^0)$ such that $x^1 = x^0 - B_0^{-1}\, \overline{g}(x^0, h^0)$ is in the region where f is defined and we can choose $h^1 \varepsilon H_2(x^1)$ such that (2.8), (2.9) hold. For iteration k, $k \ge 1$, Theroem 2.3 guarantees that we can find an $h^{k+1} \varepsilon H_2(x^{k+1})$ which satisfies (2.10) (2.11). And it follows from Lemma 2.2 that whenever $x^k - B_k^{-1}\, g(x^k)$ is in the region where f is defined we will have x^{k+1} in this region too. Hence Algorithm MQ is a well-defined algorithm.

There are several ways of choosing the appropriate step-size h in implementing the Algorithm MQ. The straightforward way is try to pick an h which satisfies both (2.2) and (2.3). If in any iteration, say after computing x^k, we cannot find an h^k that satisfies both (2.10) and (2.11), this means we did not get an h^{k-1} which was small enough. Hence we back up one iteration and get a smaller h^{k-1} and then continue the algorithm from the (k-1)th iteration. There are plenty of choices of h^{k-1} that will enable us to get a proper step-size h^k. Clearly the smaller the ratio C_1/C_2 is, the easier we can get a suitable h^{k-1}. In practice, we can either choose

C_1, C_2 in such a way that C_1/C_2 is so small that $\frac{C_1}{C_1 + C_2}$ $\leq \frac{||x^k - x^{k-1}||^2}{||x^{k-1} - x^{k-2}||^2}$ will always be true when $\frac{||x^k - x^{k-1}||^2}{||x^{k-1} - x^{k-2}||^2}$ is not approximately zero or we can pick h^k as a solution to the inequality $||h^k|| \leq C_2 \, ||B_k^{-1} \, \overline{g}(x^k, h^k)||^2$ as suggested in Corollary 2.4. Since $\overline{g}(x^k, h^k)$ is just a vector with component equivalent to $(f(x^k + h_i^k e_i) - f(x^k))/h_i^k$ and B_k, x^k are known, we can continue to reduce the value of h^k and finally obtain a solution. Both ways can help us to avoid back up. The algorithm sounds complicated, but it has been successfully implemented in such a way that it is quite easy to use. The reader is invited to consult Appendix B in this context.

III. GENERAL CONVERGENCE THEOREM

First we will prove a general local convergence theorem that will apply to Algorithm MQ with all the update methods discussed in Chapter II. For convenience we use $||\cdot||$ to denote the ℓ_2-norm and $||A||$ with $A \in L(R^n)$ will be any matrix norm consistent with the vector norm $||\cdot||$.

The following is a well-known lemma that we will use.

Lemma 3.1.

Assume

(1) $f:R^n \to R$ is twice differentiable in the open convex set D

(2) G(x) is the Hessian of f(x) for $p > 0$, $K > 0$

$$||G(x) - G(x^*)|| \leq K \;\; ||x - x^*||^p \tag{3.1}$$

Then, for each u and v in D, g, the gradient of f, satisfies

$$\begin{aligned} ||g(v) - g(u) - G(x^*)(v - u)|| &\leq K \max \{||v - x^*||^p, \\ &||u - x^*||^p\}||v - u||. \end{aligned} \tag{3.2}$$

Moreover, if $G(x^*)$ is invertible, there is an $\varepsilon > 0$ and a $\rho > 0$ such that $\max \{||v - x^*||, ||u - x^*||\} < \varepsilon$ implies that u and v belong to D and

$$(1/\rho)||v - u|| \leq ||g(v) - g(u)|| \leq \rho \;||v - u|| \tag{3.3}$$

<u>Proof</u>: See [4]

We will use $\lfloor\cdot\rfloor$ to denote the floor function. Hence $\lfloor t\rfloor$ means the greatest integer smaller than or equal to t, for all $t \in R$.

<u>Theorem 3.2</u>: Assume

(1) $f:R^n \to R$ is twice differentiable in the open convex set D,

(2) g(x) is the gradient of f(x), G(x) is the Hessian of f and for some x^* in D, $g(x^*) = 0$ and $G(x^*)$ is nonsingular. For $p > 0$ $K > 0$ (3.1) is satisfied,

(3) $\overline{g}(x, h)$ is the finite difference deviative of f(x) defined by (2.1),

(4) $U:R^n \times L(R^n) \times R^n \to p(L(R^n))$ is defined in a neighborhood $N = N_1 \times N_2 \times N_3$ of $(x^*, G(x^*), 0)$ where N_1 is contained in D and $\forall\, x \in N_1$, $g(x) \neq 0$ if $x = x^*$. N_2 contains only nonsingular matrices and N_3 is a neighborhood of $0 \in R^n$ such that for all $h \in N_3$, $\overline{g}(x, h)$ is defined for all $x \in N_1$.

(5) There are nonnegative constants α_1, α_2, α_3 such that for each (x, B, h) in N with $\overline{g}(x, h) \neq 0$ and for $\overline{x} = x - B^{-1}\overline{g}(x, h)$ the function U satisfies

$$||\overline{B} - G(x^*)|| \leq [1 + \alpha_1 \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\}]$$

$$||B - G(x^*)|| + \alpha_2 \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\}$$

$$+ \alpha_3 \frac{||h|| + ||\overline{h}||}{||\overline{x} - x||} \tag{3.4}$$

for all $\overline{B} \in U(x, B, h)$ and for all $\overline{h} \in H_2(\bar{x})$.
then, for each r in (0, 1), there exists constants $\varepsilon(r)$, $\delta(r)$ and $d(r)$ such that for $||x^0 - x^*|| < \varepsilon(r)$, $||B_0 - G(x^*)|| < \delta(r)$, $||h^0|| < d(r)$ the sequence $\{x^*\}$ generated by Algorithm MQ with $B_{k+1} \in U(x_k, B_k, h_k)$ is well-defined and convergent to x^*. Moreover,

$$||x^{k+1} - x^*|| < r^{\lfloor\frac{k+1}{2}\rfloor} ||x^0 - x^*|| \tag{3.5}$$

for each $k \geq 0$ and $\{||B_k||\}$ $\{||B_k^{-1}||\}$ are uniformly bounded.

<u>Proof</u>: Let $r \in (0, 1)$ be given. Set $\gamma \geq ||G(x^*)^{-1}||$, $C \geq ||G(x^*)||$. Now choose $\varepsilon(r) = \varepsilon$, $\delta(r) = \delta$ such that

$$4\gamma(1 + r)(C + k\,\varepsilon^p)\,C_1\,\varepsilon \leq r/2 \tag{3.6}$$

$$\gamma(1 + r)(2\delta + k\,\varepsilon^p) < r/2 \tag{3.7}$$

If necessary, in addition we restrict ε, δ, $d = d(r)$ such that $||x - x^*|| < \varepsilon$, $||h|| < d$ and $||B - G(x^*)|| < 2\delta$ imply $(x, B, h) \in N$. Then we further restrict ε, d such that

$$\frac{2(2\alpha_1\,\delta\,\varepsilon^p + \alpha_2\,\varepsilon^p + 2\alpha_3\,\varepsilon(C_2 + 2C_1))}{1 - r^p} < \delta/3 \tag{3.8}$$

$$k\,\varepsilon^r \leq \frac{1}{2\gamma} \tag{3.9}$$

and

$$\gamma(1 + r)(C + k\,\varepsilon^p)\,d \leq \varepsilon/2 \tag{3.10}$$

Suppose $||x^0 - x^*|| < \varepsilon$, $||B_0 - G(x^*)|| < \delta$, $||h^0|| < d$ and h^0 is chosen according to Algorithm MQ (so, $\overline{g}(x^0, h^0) \neq 0$). By (3.2) of Lemma 3.1, $||g(x^0) - G(x^*)(x^0 - x^*)|| \leq k\ \varepsilon^p\ ||x^0 - x^*||$. Hence, $||g(x^0)|| \geq (||G(x^*)^{-1}||^{-1} - k\ \varepsilon^p)\ ||x^0 - x^*||$. And condition (3.9) implies that $||g(x^0)|| \geq \frac{1}{2\gamma}\ ||x^0 - x^*||$. Apply Corollary 2.4 with $L = \frac{1}{2\gamma}$, $\gamma_1 = C + 2\delta$ and $\gamma_2 = (1 + r)\gamma$, then we know we can further restrict h^0 such that $||h^0|| \leq C_2\ ||B^{-1}\ \overline{g}(x^0, h^0)||^2$ for any nonsingular matrix B with $||B|| \leq C + 2\delta$ and $||B^{-1}|| \leq (1 + r)\gamma$. Since $||B_0 - G(x^*)|| < 2\delta$, with the use of (3.7) we have $2\gamma(1 + r)\delta \leq r/2 \leq r$, it follows from the Banach Lemma that $||B_0^{-1}|| \leq (1 + r)\gamma$ and $||B_0|| < 2\delta + C$. By assumption (2), we have $||G(x)|| \leq k\ ||x - x^*||^p + ||G(x^*)|| \leq k\ \varepsilon^p + C$. With Remark 1 on page 5, we have $||g(x^0) - \overline{g}(x^0, h^0)|| \leq (k\ \varepsilon^p + C)\ ||h^0||$. By Lemma 3.1 and Equations (3.7), (3.10)

$$
\begin{aligned}
||x^1 - x^*|| &\leq ||B_0^{-1}||\ (||B_0 - G(x^*)||\ ||x^0 - x^*|| \\
&\quad + ||g(x^0) - g(x^*) - G(x^*)(x^0 - x^*)|| \\
&\quad + ||g(x^0) - \overline{g}(x^0, h^0)||) \\
&\leq ||B_0^{-1}||\ ((2\delta + k\ \varepsilon^p)\ ||x^0 - x^*|| + (k\ \varepsilon^p + C)\ ||h^0||) \\
&\leq r/2\ \varepsilon + \gamma\ (1 + r)(k\ \varepsilon^p + C)\ d \leq \varepsilon
\end{aligned}
$$

Therefore, $x^1 \in N_1 \subset D$. Let us choose h^1 according to Algorithm MQ. That is, $h^1 \in H_2(x^0)$ and (2.8), (2.9) hold. This h^1 exists because by the way we close h^0 and $||B_0|| \leq \gamma_1$,

$||B_0^{-1}|| \leq \gamma_2$, we have $||h^0|| < C_2\ ||B_0^{-1}\overline{g}(x^0, h^0)||^2$ then Corollary 2.4 with $L = \frac{1}{2\gamma}$, $\gamma_1 = C + 2\delta$, $\gamma_2 = (1 + r)\ \gamma$ implies that for all h^1 that satisfy (2.8) will satisfy (2.9). Use the same argument we used in choosing h^0, we know we can further restrict h^1 such that $||h^1|| < C_2\ ||B^{-1}\overline{g}(x^1, h^1)||^2$ for all nonsingular B with $||B|| \leq C + 2\delta$ and $||B^{-1}|| \leq (1 + r)\ \gamma$. Since $||B_0 - G(x^*)|| < \delta$ and h^1 satisfies (2.8) and (2.9), it follows from (3.4) that

$$
\begin{aligned}
&||B_1 - G(x^*)|| - ||B_0 - G(x^*)|| \leq \\
&\quad \leq \alpha_1 \max \{||x^1 - x^*||^p, ||x^0 - x^*||^p\}\ ||B_0 - G(x^*)|| \\
&\quad + \alpha_2 \max \{||x^1 - x^*||^p, ||x^0 - x^*||^p\} \\
&\quad + \alpha_3 \frac{2||h^1|| + (||h^0|| - ||h^1||)}{||x^1 - x^0||} \\
&\quad \leq \alpha_1 \max \{||x^1 - x^*||^p, ||x^0 - x^*||^p\}\ ||B_0 - G(x^*)|| \\
&\quad + \alpha_2 \max \{||x^1 - x^*||^p, ||x^0 - x^*||^p\} \\
&\quad + \alpha_3 \frac{2C_1\ ||x^1 - x^0||^2 + C_2\ ||x^1 - x^0||^2}{||x^1 - x^0||}
\end{aligned}
$$

With (3.8) we have

$$||B_1 - G(x^*)|| \leq \delta + \alpha_1 \delta \varepsilon^p + \alpha_2 \varepsilon^p + (\alpha_3 C_2 + 2\alpha_3 C_1)\ 2\varepsilon \leq 4\delta/3 < 2\delta$$

This implies that $(x^1, B_1) \varepsilon N_1 \times N_2$ and $||B_1|| \leq C + 2\delta$, $||B_1^{-1}|| \leq (1+r)\gamma$. Thus we are all set to get x^2, $x^2 \neq x^1$ (x^2

$= x^1$ iff $x^1 = x^*$, in which case we are done). We are going to use induction to complete the proof. Suppose for all $k = 0, \ldots, m$, $||B_k - G(x^*)|| < 2\delta$, h^k is chosen from $H(x^k)$ such that (2.10) (2.11) hold and $||h^k|| \le C_2 \, ||B^{-1}\overline{g}\,(x^k, h^k)||^2$ for all nonsingular B with $||B|| \le C + 2\delta$. $||B^{-1}|| \le \gamma(1 + r)$, x^{k+1} is well-defined and $||x^{k+1} - x^*|| \le r^{\lfloor\frac{k+1}{2}\rfloor} \, ||x^0 - x^*||$ with $x^{k+1} \ne x^*$. The proof is completed if we can prove there is $h^{m+1} \varepsilon H(x^{m+1})$ such that (2.10) (2.11) are satisfied and $||h^{m+1}|| \le C_2 \, ||B^{-1}\overline{g}(x^{m+1}, h^{m+1})||^2$ for all nonsingular B with $||B|| \le C + 2\delta$. $||B^{-1}|| \le \gamma(1 + r)$, $||B_m - G(x^*)|| < 2\delta$ and x^{m+2} is well-defined with $||x^{m+2} - x^*|| \le r^{\lfloor\frac{m+2}{2}\rfloor} \, ||x^0 - x^*||$. Since $||B_m - G(x^*)|| < 2\delta$, we have $||B_m|| \le C + 2\delta$, $||B_m{}^{-1}|| \le \gamma(1+r)$. By the induction hypothesis on h^m, we have $||h^m|| \le C_2 \, ||B_m{}^{-1}\overline{g}(x^m, h^m)||^2$. Since $||x^{m+1} - x^*|| < \varepsilon$, similar to the case $k = 0$, we have $||g(x^m)|| \ge \frac{1}{2\gamma} \, ||x^m - x^*||$ and it follows from Corollary 2.4 that there exists an $h \varepsilon H(h^m)$ such that $||h|| \le C_2 \, ||B^{-1}\overline{g}(x^{m+1}, h^{m+1})||^2$ for all nonsingular B with $||B|| \le C + 2\delta$, $||B^{-1}|| \le \gamma(1 + r)$. We will let h^{m+1} be this h. By the way we chose h^m and by Corollary 2.4 with $L = \frac{1}{2\gamma}$, $\gamma_1 = C + 2\delta$, $\gamma_2 = (1 + r)\gamma$ we have that for all $h \varepsilon H(h^m)$ (2.11) holds. Hence the h^{m+1} we chose satisfies (2.10) and (2.11). From (3.4) $||B_{k+1} - G(x^*)|| - ||B_k - G(x^*)||$ $\le \alpha_1 \max \{||x^{k+1} - x^*||^p, ||x^k - x^*||^p\} \, 2\delta$ $+ \alpha_2 \max \{||x^{k+1} - x^*||^p, ||x^k - X^*||^p\}$

$$+ \alpha_3 \frac{2C_1 ||x^{k+1} - x^k||^2 + C_2 ||x^{k+1} - x^k||^2}{||x^{k+1} - x^k||}$$

for $k = 0, \ldots, m$. The induction hypothesis on h^k and x^k give us: $||B_{k+1} - G(x^*)|| - ||B_k - G(x^*)||$ $\leq 2\alpha_1 r^{\lfloor \frac{k}{2} \rfloor} p_\varepsilon p_\delta + \alpha_2 r^{\lfloor \frac{h}{2} \rfloor} p_\varepsilon p + 2\alpha_3 (C_2 + 2C_1) r^{\lfloor \frac{k}{2} \rfloor} \varepsilon$. Summing both sides of the above inequality from $k = 0$ to $k = m$

$$||B_{m+1} - G(x^*)|| - ||B_0 - G(x^*)|| \leq (2\alpha_1 \delta \varepsilon^p + \alpha_2 \varepsilon^p) \frac{2}{1-r^p} + \frac{4\alpha_3 \varepsilon (C_2 + 2C_1)}{1-r} \leq 2(2\alpha_1 \delta \varepsilon^p + \alpha_2 \varepsilon^p + 4\alpha_3 \varepsilon (C_2 + 2C_1)) \frac{1}{1-r^p}$$

for $p \leq 1$ which by (3.8) is less than $\delta/3$. Hence x^{m+2} is well-defined and again $x^{m+2} \neq x^{m+1}$. Now let us prove that $||x^{m+2} - x^*|| \leq r^{\lfloor \frac{m+2}{2} \rfloor} ||x^0 - x^*||$. Similar to the case for $k = 0$, we have $||x^{m+2} - x^*|| \leq r/2 \; ||x^{m+1} - x^*|| + \gamma(1 + r)(C + \varepsilon^p k) \; ||h^{m+1}||$. It follows from (2.10) and (3.6)

$$||x^{m+2} - x^*|| \leq r/2 \; ||x^{m+1} - x^*|| + 4\gamma(1 + r)(C + k\varepsilon^p) C_1 \max \{||x^{m+1} - x^*||^2, ||x^m - x^*||^2\}$$
$$\leq ((r/2) + 4\gamma(1 + r)(C + k\varepsilon^p) C_1 \varepsilon) \max \{||x^{m+1} - x^*||, ||x^m - x^*||\} \leq r^{\lfloor \frac{m+3}{2} \rfloor} ||x^0 - x^*||.$$

Thus we have completed the proof.

This Theorem shows that a quasi-Newton method using a proper finite difference derivative instead of the analytic derivative will still be locally convergent as long as its update function satisfies (2.4). However, Theorem 3.2 does not give an estimate of its rate of convergence. It can be proved easily by a generalization of the Characterization Theorem of Dennis and Moré [9] that if there is some subsequence of $\{||B_k - G(x^*)||\}$ convergent to zero, then $\{x^k\}$ convergent Q-superlinearly. Let us state the generalization of the Characterization Theorem first.

Theorem 3.3: Assume

(1) $g: R^n \to R^n$ be differentiable in the open convex at in R^n,

(2) for some x^* in D, g' is continuous at x^* and $g'(x^*)$ is nonsingular,

(3) $\{B_k\}$ in $L(R^n)$ is a sequence of nonsingular matrices,

(4) $\bar{g}(x, h)$ is an approximation rule for $g(x)$ and suppose for some $x^0 \in D$ the sequence $\{x^k\}$ where $x^k = x^{k-1} - B_{k-1}{}^{-1}\, \bar{g}(x^{k-1}, h^{k-1})$ remains in D and converges to x^*,

then $\{x^k\}$ converges Q-superlinearly to x^* and $g(x^*) = 0$

$$\text{iff} \lim_{k\to\infty} \frac{||(B_k - G(x^*))(x^{k+1} - x^k) + \bar{g}(x^k, h^k) - g(x^k)||}{||x^{k+1} - x^k||} = 0 \tag{3.11}$$

Proof: Since $(B_k - g'(x^*))(x^{k+1} - x^k) + \bar{g}(x^k, h^k) - g(x^k)$

$$= -\bar{g}(x^k, h^k) - g'(x^*)(x^{k+1} - x^k) + \bar{g}(x^k, h^k) - g(x^k)$$

$$= g(x^{k+1}) - g(x^k) - g'(x^*)(x^{k+1} - x^k) - g(x^{k+1})$$

Then following the proof of the Characterization Theorem 2.2

of Dennis-Moré [9] will complete this proof.

And we need the following Lemma.

Lemma 3.4: Let $\{\phi_k\}$ and $\{\delta_k\}$ be sequences of nonnegative numbers such that

$$\phi_{k+1} \le (1 + \delta_k{}^p)\ \phi_k + \delta_k{}^p + \delta_k \tag{3.12}$$

and

$$\sum_{k=0}^{\infty} \delta_k{}^p < +\infty. \qquad \sum_{k=0}^{\infty} \delta_k < +\infty \tag{3.13}$$

then $\{\phi_k\}$ converges.

Proof: Let $U_k = \prod_{j=0}^{k-1} (1 + \delta_j{}^p)$, by (3.13) U_k is bounded above, say by $U > 1$ and $U_k > 1$ for all $k \ge 1$

It follows that

$$\frac{\phi_{k+1}}{U_{k+1}} \le \frac{\phi_k}{U_k} + \frac{\delta_k{}^p}{U_{k+1}} + \frac{\delta_k}{U_{k+1}} \le \frac{\phi_k}{U_k} + \delta_k{}^p + \delta_k$$

so, $\dfrac{\phi_{k+1}}{U_{k+1}} - \dfrac{\phi_k}{U_k} \le \delta_k{}^p + \delta_k.$

Summing from $k = 0$ to $k = m$, we have

$$\frac{\phi_{m+1}}{U_{m+1}} \le \frac{\phi_0}{U_0} + \sum_{k=0}^{m} \delta_k{}^p + \sum_{k=0}^{m} \delta_k$$

Therefore, $\{\phi_k\}$ is bounded above and so it has at least one limit point. To prove it has only one limit point, we assume that there are two subsequence $\{\phi_{k_n}\}$ $\{\phi_{k_m}\}$ convergent respectively to ϕ', ϕ''. Let ϕ be the bound of $\{\phi_k\}$. With the use of (3.13) we have $\phi_{k+1} \leq \phi_k + \delta_k{}^p (\phi + 1) + \delta_k$. For $k_n \geq k_m$, $\phi_{k_n} - \phi_{k_m} \leq (\phi + 1) \sum_{j=k_m}^{\infty} \delta_k{}^p + \sum_{j=k_m}^{\infty} \delta_k$. Let $k_m \to \infty$, $\phi_{k_n} - \phi''$ $\leq (\phi+1) \sum_{j=k_m}^{\infty} \delta_k^p + \sum_{j=k_m}^{\infty} \delta_k$. Let $k_n \to \infty$, $\phi' \leq \phi''$. For $k_m \geq k_n$, $\phi_{k_m} - \phi_{k_n}$ $\leq (\phi + 1) \sum_{j=k_n}^{\infty} \delta_k{}^p + \sum_{j=k_n}^{\infty} \delta_k$. Then let $k_m \to \infty$, $\phi'' - \phi_{k_n}$ $\leq (\phi + 1) \sum_{j=k_n}^{\infty} \delta_k{}^p + \sum_{j=k_n}^{\infty} \delta_k$. Let $k_n \to \infty$, we have $\phi'' - \phi' \leq 0$. Hence, $\phi' = \phi''$. $\{\phi_k\}$ converges.

Corollary 3.5: Assume the hypotheses of Theorem 3.2 hold. If some subsequences of $\{||B - G(x^*)||\}$ converges to zero, than the sequence $\{x^k\}$ of Theorem 3.2 converges Q-superlinearly to x^*.

Proof: It follows from the proof of Theorem 3.2 that (3.6) implies that

$$
\begin{aligned}
||B_{k+1} - G(x^*)|| \leq {} & [1 + \alpha_1 \max \{||x^k - x^*||^p, \\
& ||x^{k-1} - x^*||^p\}] \; ||B_k - G(x^*)|| \\
& + \alpha_2 \max \{||x^k - x^*||^p, ||x^{k-1} - x^*||^p\} \\
& + \hat{\alpha}_3 \max \{||x^k - x^*||, ||x^{k-1} - x^*||\}
\end{aligned}
\tag{3.14}
$$

where $\hat{\alpha}_3 = \alpha_3 (C_2 + 2C_1)$. Let $\alpha = \max \{\alpha_1, \alpha_2, \hat{\alpha}_3{}^p\}$, and $\delta_k = \alpha^{1/p} \max \{||x^k - x^*||, ||x^{k-1} - x^*||\}$. Hence (3.14) becomes $||B_{k+1} - G(x^*)|| \le (1 + \delta_k{}^p) ||B_k - G(x^*)|| + \delta_k{}^p + \delta_k$.

By Theorem 3.2, since $0 < r < 1$, $0 \le p \le 1$

$$\sum_{k=0}^{\infty} \delta_k{}^p \le \alpha ||x^0 - x^*||^p \sum_{k=0}^{\infty} r^{p\lfloor\frac{k}{2}\rfloor} \le \frac{2\alpha ||x^0 - x^*||^p}{1 - r^p} < +\infty$$

$$\sum_{k=0}^{\infty} \delta_k \le \alpha^{1/p} ||x^0 - x^*|| \sum_{k=0}^{\infty} r^{\lfloor\frac{k}{2}\rfloor} \le \frac{2\alpha^{1/p} ||x^0 - x^*||}{1 - r} < +\infty.$$

Apply Lemma 3.4 with $\phi_k = ||B_k - G(x^*)||$, since (3.12)(3.13) are satisfied we have $\{||B_k - G(x^*)||\}$ converges to zero.

Then,

$$0 \le \lim_{k\to\infty} \frac{||(B_k - G(x^*))(x^{k+1} - x^k)||}{||x^{k+1} - x^k||}$$

$$\le \lim_{k\to\infty} \frac{||B_k - G(x^*)||\ ||x^{k+1} - x^k||}{||x^{k+1} - x^k||} = 0$$

By Remark 1 on page 5, there exists $C_0 > 0$ such that

$$\lim_{k\to\infty} \frac{||\bar{g}(x^k, h^k) - g(x^k)||}{||x^{k+1} - x^k||} \le \lim_{k\to\infty} \frac{C_0 ||h^k||}{||x^{k+1} - x^k||}$$

And, h^k is chosen according to Algoithm MQ. Hence it satisfies (2.10)(2.11).

Then, $0 \le \lim_{k\to\infty} \dfrac{||\bar{g}(x^k, h^k) - g(x^k)||}{||x^{k+1} - x^k||}$

$$\le C_0 \lim_{k\to\infty} \frac{||h^k|| - ||h^{k+1}|| + ||h^{k+1}||}{||x^{k+1} - x^k||}$$

$$\le C_0 \left(\lim_{k\to\infty} \frac{C_2 ||x^{k+1} - x^k||^2 + C_1 ||x^{k+1} - x^k||^2}{||x^{k+1} - x^k||}\right) = 0$$

Therefore,

$$\lim_{k\to\infty} \frac{||(B_k - G(x^*))(x^{k+1} - x^k) - \bar{g}(x^k, h^k) - g(x^k)||}{||x^{k+1} - x^k||} = 0$$

Hence (3.11) of Theorem 3.3 is satisfied, which implies that $\{x^k\}$ converges Q-superlinearly to x^*.

Sometimes in the Algorithm, we use H_k, the approximation to the inverse of the Hassian matrix, instead of B_k. The following theorem is analoguous to Theorem 3.2.

Theorem 3.6: Let the assumption (1) to (4) of Theorem 3.2 hold. In addition assume that:

(5) for some $C_1 > 0$, there are nonegative constants $\alpha_1, \alpha_2, \alpha_3$ such that for each (x, H, h) in N with $\bar{g}(x, h) \neq 0$ and for $\bar{x} = x - H\bar{g}(x, h)$ the function U satisfies

$$\begin{aligned} ||\bar{H} - G(x^*)|| \le 1 &+ \alpha_1 \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}\, ||H - G(x^*)|| \\ &+ \alpha_2 \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\} \\ &+ \alpha_3 \frac{||\bar{h}|| + ||h||}{||\bar{x} - x||} \end{aligned} \tag{3.15}$$

for each $\bar{H}$ in $U(x, H, h)$ and $\bar{h} \in H_2(x)$ such that $\bar{h}$ satisfies (2.2) and (2.3), then, for each r in $(0, 1)$, there exists constants $\varepsilon(r)$, $\delta(r)$ and $d(r)$ such that for $||x^0 - x^*|| < \varepsilon(r)$, $||H_0 - G(x^*)|| < \delta(r)$, $||h^0|| < d(r)$ the sequence $\{x_k\}$ generated by Algorithm MQ with $H_{k+1} \in U(x^k, H_k, h^k)$ is well-defined and convergent to x^*. Moreover,

$$||x^{k+1} - x^k|| \le r^{\lfloor \frac{k+1}{2} \rfloor} ||x^0 - x^*|| \tag{3.16}$$

for each $k \geq 0$ and $\{||H_k||\}$, $\{||H_k^{-1}||\}$ are uniformly bounded.

<u>Proof</u>: Let $r \in (0, 1)$ be given. Set $\gamma \geq ||G(x^*)^{-1}||$ and $C \geq ||G(x^*)||$. Now, choose $\varepsilon(r) = \varepsilon$, $\delta(r) = \delta$, such that

$$(2\delta + \gamma)\ k\ \varepsilon^p + C\ \delta \leq r/2 \tag{3.17}$$

$$2(2\delta + \gamma)(k\ \varepsilon^p + C)\ C_1\ \varepsilon \leq r/2, \tag{3.18}$$

If necessary, in addition we restrict ε, δ and $d = d(r)$ such that $||x - x^*|| < \varepsilon$, $||h|| < d$, $||H - G(x^*)^{-1}|| < 2\delta$ imply $(x, H, h) \in N$. And we further restrict ε, d such that

$$2\left(\frac{\alpha_1\ \delta\ \varepsilon^p + \alpha_2\ \varepsilon^p + 2\alpha_3\ (2C_1 + C_2)\ \varepsilon}{1 - r^p}\right) \leq \delta/3 \tag{3.19}$$

$$k\ \varepsilon^p \leq \frac{1}{2\gamma} \tag{3.20}$$

$$(2\delta + \gamma)(k\ \varepsilon^p + C)\ d \leq \varepsilon/2 \tag{3.21}$$

Suppose $||x^0 - x^*|| < \varepsilon$, $||H_0 - G(x^*)^{-1}|| < \delta$, $||h^0|| < d$ and h^0 is chosen according to Algorithm MQ (so $\overline{g}(x^0, h^0) \neq 0$). By (3.2) of Lemma 3.1, $||g(x^0) - G(x^*)(x^0 - x^*)|| \leq k\ \varepsilon^p\ ||x^0 - x^*||$. Hence $||g(x^0)|| \geq (||G(x^*)^{-1}||^{-1} - k\ \varepsilon^p)\ ||x^0 - x^*||$. And condition (3.20) implies that $||g(x^0)|| \geq \frac{1}{2\gamma}\ ||x^0 - x^*||$. Apply Corollary 2.4 with $L = \frac{1}{2\gamma}$,

$\gamma_1 = \frac{C}{1-\gamma}$, $\gamma_2 = 2\delta + \gamma$, then we know we can further restrict h^0 such that $||h^0|| \leq C_2\ ||B^{-1}\bar{g}(x^0, h^0)||^2$ for any nonsingular matrix B with $||B|| \leq \frac{C}{1-\gamma}$ and $||B^{-1}|| \leq 2\delta + \gamma$. Since $||H_0 - G(x^*)^{-1}|| < \delta$, with the use of the Banach Lemma, we have $||H_0^{-1}|| \leq \frac{C}{1-\gamma}$. And, $||H_0|| \leq 2\delta + \gamma$. Hence h^0 satisfies $||h^0|| \leq C_2\ ||H_0\bar{g}(x^0, h^0)||^2$. Since $x - x^* = x^0 - H_0\bar{g}(x^0, h^0) - x^* = x^0 - H_0 g(x^0) - x^* + H_0 g(x^0) - H_0\bar{g}(x^0, h^0)$ $= -H_0[g(x^0) - g(x^*) - G(x^*)(x^0 - x^*)] + [I - H_0 G(x^*)](x^0 - x^*)$ $+ H_0[g(x^0) - \bar{g}(x^0, h^0)]$.

By the proof of Theorem 3.2, we have $||G(x)|| \leq k\varepsilon^p + C$ and $||g(x^0) - \bar{g}(x^0, h^0)|| \leq (k\ \varepsilon^p + C)\ ||h^0||$. By Lemma 3.1 and Equations (3.17), (3.21), $||x^1 - x^*||$

$\leq (||H_0||\ k\ \varepsilon^p + C\delta)\ ||x^0 - x^*|| + ||H_0||\ (k\ \varepsilon^p + C)\ ||h^0||$

$\leq ((2\delta + \gamma)\ k\ \varepsilon^p + C\delta)\ ||x^0 - x^*|| + (2\delta + \gamma)(k\ \varepsilon^p + C)d$

$\leq r/2\ \varepsilon + (2\delta + \gamma)(k\ \varepsilon^p + C)\ d \leq \varepsilon$.

Therefore, $x^1 \in N_1 \subset D$. Let us choose h^1 according to Algorithm MQ. That is, $h^1 \in H_2(h^0)$ and (2.8) (2.9) hold. This h^1 exists because by the way we choose h^0 and $||H_0|| \leq 2\delta + \gamma$, $||H_0^{-1}|| \leq \frac{C}{1-\gamma}$, the Corollary 2.4 with $L = \frac{1}{2\gamma}$, $\gamma_1 = \frac{C}{1-\gamma}$, $\gamma_2 = 2\delta + \gamma$ implies that all h that satisfy (2.8) will satisfy (2.9). Use the same argument we used in choosing h^0, we know we can further restrict h^1 such that $||h^1||$ $\leq C_2\ ||B^{-1}\ \bar{g}(x^1, h^1)||^2$ for all nonsingular B with $||B|| \leq \frac{C}{1-\gamma}$, $||B^{-1}|| \leq 2\delta + \gamma$. Since $||H_0 - G(x^*)^{-1}|| < \delta$ and h^1 satisfies (2.8) and (2.9) it follows from (3.15) that

$$||H_1 - G(x^*)^{-1}|| - ||H_0 - G(x^*)^{-1}||$$

$$\le \alpha_1 \max\{||x^1 - x^*||^p, ||x^0 - x^*||^p\} ||H_0 - G(x^*)^{-1}||$$

$$+ \alpha_2 \max\{||x^1 - x^*||^p, ||x^0 - x^*||^p\}$$

$$+ \alpha_3 \frac{||h^0|| - ||h^1|| + 2||h^1||}{||x^1 - x^0||}$$

$$\le \alpha_1 \delta \varepsilon^p + \alpha_2 \varepsilon^p + \alpha_3 \frac{C_2||x^1 - x^0||^2 + 2C_1 ||x^1 - x^0||^2}{||x^1 - x^0||}$$

With (3.19) we

$$||H_1 - G(x^*)^{-1}|| \le \delta + \alpha_1 \delta \varepsilon^p + \alpha_2 \varepsilon^p + \alpha_3(2C_1 + C_2) 2\varepsilon$$

$$\le 4\delta/3 < 2\delta.$$

This implies that $(x^1, H^1) \in N_1 \times N_2$ and $||H_1|| \le \gamma + \delta$, $||H_1{}^{-1}|| \le \frac{C}{1 - \gamma}$. Thus we are all set to get x^2, $x^2 \neq x^1$ ($x^2 = x^1$ iff $x^1 = x^*$ in which case we all done). We are going to use induction to complete the proof. Suppose for all $k = 0, \ldots, m$, $||H_k - G(x^*)^{-1}|| < 2\delta$, h^k is chosen from $H(x^k)$ such that (2.10) (2.11) hold and $||h^k|| \le C_2 ||B^{-1}\bar{g}(x^k, h^k)||^2$ for all nonsingular B with $||B|| \le \frac{C}{1 - \gamma}$, $||B^{-1}|| \le \gamma + 2\delta$, then x^{k+1} is well-defined and $||x^{k+1} - x^*|| \le r^{\lfloor\frac{k+1}{2}\rfloor} ||x^0 - x^*||$ and $x^{k+1} \neq x^*$. The proof is completed if we can prove there is $h^{m+1} \in H(x^{m+1})$ such that (2.10) (2.11) are satisfied and $||h^{m+1}|| \le C_2 ||B^{-1}\bar{g}(x^m, h^m)||^2$ for all nonsingular B with $||B|| \le \frac{C}{1 - \gamma}$, $||B^{-1}|| \le \gamma + \delta$, $||H_m - G(x^*)^{-1}|| < 2\delta$ and x^{m+2} is well-defined with $||x^{m+2} - x^*|| \le r^{\lfloor\frac{m+2}{2}\rfloor} ||x^0 - x^*||$.

Since $||H_m - G(x^*)^{-1}|| < 2\delta$, we have $||H_m|| \leq 2\delta + \gamma$. By the induction hypothesis on h^m, we have $||h^m|| \leq C_2\, ||H_m \bar{g}(x^m, h^m)||^2$. Since $||x^{m+1} - x^*|| < \varepsilon$, similar to the case $k = 0$, we have $||\, g(x^m)|| \leq \frac{1}{2\gamma}\, ||x^m - x^*||$ and it follows from Corollary 2.4 that we can get $h^{m+1} \varepsilon H(h^m)$ such that $||h^{m+1}|| \leq C_2\, ||B^{-1}\bar{g}(x^{m+1}, h^{m+1})||^2$ for all nonsingular B with $||B|| \leq \frac{C}{1 - \gamma}, ||B^{-1}|| \leq 2\delta + \gamma$. By the way we chose h^m and by Corollary 2.4 with $L = \frac{1}{2\gamma}$, $\gamma_2 = \gamma + 2\delta$, $\gamma_1 = \frac{C}{1 - \gamma}$, we have that for all $h \varepsilon H(h^m)$ (2.11) holds. Hence the h^{m+1} we chose satisfies (2.10) and (2.11).

From (3.15),

$$||H_{k+1} - G(x^*)^{-1}|| - ||H_k - G(x^*)||$$

$$\leq \alpha_1 \max \{||x^{k+1} - x^*||^p, ||x^k - x^*||^p\}\, 2\delta$$

$$+ \alpha_2 \max \{||x^{k+1} - x^*||^p, ||x^k - x^*||^p\}$$

$$+ \alpha_3 \frac{2C_1\, ||x^{k+1} - x^k||^2 + C_2\, ||x^{k+1} - x^*||^2}{||x^{k+1} - x^k||}$$

for $h = 0, \ldots, m$. The induction hypothesis on h^k and x^k gives us $||H_{k+1} - G(x^*)^{-1}|| - ||H_k - G(x^*)^{-1}||$

$$\leq 2\,\alpha_1\, r^{\lfloor\frac{k}{2}\rfloor p}\, \varepsilon^p\, \delta + \alpha_2\, r^{\lfloor\frac{k}{2}\rfloor p}\, \varepsilon^p + 2\,\alpha_3 (2C_1 + C_2)\, r^{\lfloor\frac{k}{2}\rfloor}\, \varepsilon.$$

Summing both sides of the above inequality from $k = 0$ to $k = m$,

$$||H_{m+1} - G(x^*)^{-1}|| - ||H_0 - G(x^*)^{-1}||$$

$$\leq (2\alpha_1\, \delta\, \varepsilon^p + \alpha_2\, \varepsilon^p)\, \frac{2}{1 - r^p} + \frac{4\, \alpha_3\, \varepsilon\, (C_2 + 2C_1)}{1 - r}$$

$$\le \frac{2(\alpha_1 \delta \varepsilon^p + \alpha_2 \varepsilon^p + 2\alpha_3 \varepsilon (2C_1 + C_2))}{1 - r^p} \quad \text{for } p \le 1$$

which by (3.19) is less than $\delta/2$. Hence $||H_{m+1} - G(x^*)^{-1}||$ $\le 2\delta$.

And similar to the case k = 0, we know x^{m+2} is well-defined and again $x^{m+2} \neq x^{m+1}$ and we have

$$||x^{m+2} - x^*|| \le r/2 \, ||x^{m+1} - x^*|| + ||H_{m+1}|| \, (k \, \varepsilon^p + C) \, ||h^{m+1}||$$

$$\le r/2 \, ||x^{m+1} - x^*|| + (2\delta + \gamma)(k \, \varepsilon^p + C) \, C_1 \, ||x^{m+1} - x^m||^2$$

$$\le [r/2 + (2\delta + \gamma)(k \, \varepsilon^p + C) \, C_1 \, 2 \max \{||x^{m+1} - x^*||, ||x^m - x^*||\}] \, r^{\lfloor \frac{m}{2} \rfloor} \, ||x^0 - x^*||$$

$$\le (r/2 + 2(2\delta + \gamma) \, (k \, \varepsilon^p + C) \, C_1 \, \varepsilon) \, r^{\lfloor \frac{m}{2} \rfloor} \, ||x^0 - x^*||.$$

By (3.18), $||x^{m+2} - x^*|| \le r^{\lfloor \frac{m+2}{2} \rfloor} \, ||x^0 - x^*||$.

Thus we have completed the proof.

Dennis and Moré [9] have proved that for the case of H,

$$\lim_{k\to\infty} \frac{||(H_k - G(x^*)^{-1}) \, y^k||}{||y^k||} = 0 \text{ where } y^k = g(x^{k+1}) - g(x^k) \qquad (3.22)$$

is a sufficient condition for $\overline{x} = x - Hg(x)$ to converge Q-superlinearly. Use the same technique as we used in Theorem 3.3 and Corollary 3.5 we have the following theorem and corollary in terms of H_k.

Theorem 3.7: Let the assumptions (1) (2) of Theorem 3.3 hold. And assume that

(3) $\{H_k\}$ in $L(R^n)$ is a sequence of nonsingular matrices and $||H_k^{-1}||$ is uniformly bounded,

(4) $\bar{g}(x, h)$ is an approximation rule for $g(x)$ and for some $x^0 \in D$ the sequence $\{x^k\}$ where $x^k = x^{k-1} - H_{k-1}\, \bar{g}(x^{k-1}, h^{k-1})$ remains in D and converges to x^*, then $\{x^k\}$ converges Q-superlinearly to x^* and $g(x^*) = 0$

$$\text{iff } \lim_{k\to\infty} \frac{||(H_k - G(x^*)^{-1})[g(x^{k+1}) - g(x^k)] - H_k(\bar{g}(x^k, h^k) - g(x^k))||}{||g(x^{k+1}) - g(x^k)||} = 0 \tag{3.23}$$

Proof:

$$(H_k - G(x^*)^{-1})[g(x^{k+1}) - g(x^k)] - H_k[\bar{g}(x^k, h^k) - g(x^k)]$$

$$= H_k\, g(x^{k+1}) - G(x^*)^{-1}\, [g(x^{k+1}) - g(x^k) - G(x^*)(x^{k+1} - x^k)]$$

By the assumptions and $\{x^k\}$ converges to x^*, (3.23) is equivalent to

$$\lim_{k\to\infty} \frac{||H_k\, g(x^{k+1})||}{||g(x^{k+1}) - g(x^k)||} = 0$$

And there exists $\gamma > 0$ such that $||g(x^{k+1}) - g(x^k)|| \geq \gamma\, ||x^{k+1} - x^k||$.

Hence $\lim_{k\to\infty} \frac{||H_k\, g(x^{k+1})||}{||x^{k+1} - x^k||} = 0$. The rest of the proof is the same as the proof of Theorem 2.2 of Dennis and Moré [9].

Corollary 3.8: Assume that the hypotheses of Theorem 3.6 hold. If some subsequence of $\{||H_k - G(x^*)^{-1}||\}$ converges to zero, then $\{x^k\}$ converges Q-superlinearly.

Proof: Similar to Corollary 3.5, it follows from (3.15) and Lemma 3.4, that $\{||H_k - G(x^*)^{-1}||\}$ converges to zero. And,

$$\lim_{k\to\infty} \frac{||H_k\ [\bar{g}(x^k,\ h^k) - g(x^k)]\,||}{||x^{k+1} - x^k||}$$

$$\le \lim_{k\to\infty} ||H_k|| \lim_{k\to\infty} \frac{||\bar{g}(x^k,\ h^k) - g(x^k)||}{||x^{k+1} - x^k||}\ .$$

By $\{||H_k||\}$ is uniformly bounded and Remark 1 on page 5, there exists constant $L > 0$ such that

$$\lim_{k\to\infty} \frac{||H_k\ [\bar{g}(x^k,\ h^k) - g(x^k)]\,||}{||x^{k+1} - x^k||} \le L \lim_{k\to\infty} \frac{||h^k||}{||x^{k+1} - x^k||}$$

Follow the proof of Corollary 2.5, we will get that (3.23) is satisfied. Hence, apply Theorem 3.7, then we know $\{x^k\}$ converges Q-superlinearly.

IV. CONVERGENCE OF THE MODIFIED SINGLE-RANK METHODS

Theorem 3.2 and Theorem 3.6 give two sufficient conditions for a modified quasi-Newton method to converge locally. In order to use these results for the modified single-rank methods, we need two lemmas which have been proved by Broyden-Dennis-Moré.

Lemma 4.1: Let $B \in L(R^n)$, y, c in R^n with $c^T s \neq 0$, define

$$\bar{B} = B + \frac{(y - Bs)c^T}{c^T s} \tag{4.1}$$

then for any $A \in L(R^n)$ and any symmetric nonsingluar $M \in L(R^n)$ it follows that

$$\bar{E} = E[I - \frac{M^{-1}s(Mc)^T}{c^T s}] + \frac{M(y - As)(Mc)^T}{c^T s}$$

where $\bar{E} = M(\bar{B} - A)M$ and $E = M(B - A)M$

Proof: See [4].

Lemma 4.2: Let $M \in L(R^n)$ be a nonsingular symmetric matrix such that

$$||Mc - M^{-1} s|| \leq \beta \; ||M^{-1} s|| \tag{4.3}$$

for some $\beta \in [0, 1/3]$ and vector c and s in R^n with $s \neq 0$.

Then

(a) $$(1 - \beta) \, ||M^{-1} s||^2 \leq c^T s \leq (1 + \beta) \, ||M^{-1} s||^2 \qquad (4.4)$$

and for some nonzero $E \in L(R^n)$

(b) $$\left|\left|E\left[I - \frac{M^{-1} s (Ms)^T}{c^T s}\right]\right|\right|_F \leq \sqrt{1-\alpha\theta^2} \, ||E||_F.$$

(c) $$\left|\left|E\left[I - \frac{M^{-1} s (Mc)^T}{c^T s}\right]\right|\right|_F \leq \left[\sqrt{1-\alpha\theta^2} + (1-\beta)^{-1} \frac{||Mc - M^{-1}s||}{||M^{-1}s||}\right] ||E||_F$$

where

$$\alpha = \frac{1 - 2\beta}{1 - \beta^2} \in [3/8, 1] \qquad (4.5)$$

and

$$\theta = \frac{||E \, M^{-1} \, s||}{||E||_F \, ||M^{-1} \, s||} \in [0, 1] \qquad (4.6)$$

Moreover, for any $y \in R^n$

(d) $$\left|\left|\frac{(y - As)(Mc)^T}{c^T s}\right|\right|_F \leq 2 \, \frac{||y - As||}{||M^{-1} \, s||}$$

<u>Proof</u>: See [4].

Let us define an M-norm $||\cdot||_M$ by $||A||_M = ||MAM||_F$ for any nonsingular matrix $M \in L(R^n)$ before we start the following lemma.

<u>Lemma 4.3</u>:

(1) Let $f: R^n \to R$ be twice differentiable in the open convex set D and g be the gradient of f.

(2) Let G be the Hessian of f and suppose that for some $p > 0$ (3.1) is satisfied for some x^* where $g(x^*) = 0$ and $G(x^*)$ is nonsingular, $||G(x^*)|| \le C$ for some $C > 0$.

(3) Let $N = N_1 \times N_2 \times N_3$ be a neighborhood of $(x^*, G(x^*), 0)$ such that N_1 is contained in D and for all $x \in N_1$, $g(x) \ne 0$, N_2 contains all the nonsingular matrices, N_3 is a neighborhood of the vector 0 such that for all $x \in N_1$, $\bar{g}(x, h)$ is well-defined.

(4) If

$$\frac{||Mc - M^{-1} s||}{||M^{-1} s||} \le \mu_1 ||s||^p \tag{4.8}$$

where $s = B^{-1} \bar{g}(x, h)$ with $(x, B, h) \in N$ and $x \ne x^*$, $\bar{g}(x, h) \ne 0$, μ_1 is a nonegative constant, M is a nonsingular symmetric matrix in $L(R^n)$.

(5) The update function U is defined by

$$U(x, B, h) = \{\bar{B} : \bar{B} = B + \frac{(y - s)\, c^T}{c^T s}$$

where

$$y = \bar{g}(\bar{x}, \bar{h}) - \bar{g}(x, h)$$

with

$$\bar{x} = x + s,\ s = B^{-1} \bar{g}(x, h),\ \bar{h} \in H_2(\bar{x}) \cap N_3.\}$$

for all $(x, B, h) \in N$ with $\bar{g}(x, h) \ne 0$, $x \ne x^*$.

Then there exists a neighborhood $\bar{N}$ of $(x^*, G(x^*), 0)$ such that $\bar{B} \in U(x, B, h)$ is well-defined and for all $\bar{x} = x - B^{-1} \bar{g}(x, h)$,

$$||\bar{B} - G(x^*)||_M \le (1 + \alpha_1 \max \{||x - x^*||^p,$$

$||x - x^*||^p\} \; ||B - G(x^*)||_M +$

$$\alpha_2 \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\} + \alpha_3 \frac{||\bar{h}|| + ||h||}{||\bar{x} - x||} \tag{4.9}$$

where α_1, α_2, α_3 are nonnegative constants, for all $(x, B, h) \in \bar{N}$ with $\bar{g}(x, h) \neq 0$, $x \neq x^*$.

<u>Proof</u>: Let γ be the constant such that $\gamma \geq ||G(x^*)^{-1}||$. Choose $\varepsilon_1 > 0$, $d_1 > 0$ such that

$$2\,\gamma\,(c + k\,\varepsilon_1^p)\,\varepsilon_1 < \frac{1}{2(3\mu_1)^{1/p}} \tag{4.10}$$

$$(c + k\,\varepsilon_1^p)\,d_1 < \frac{1}{2(3\mu_1)^{1/p}} \tag{4.11}$$

Let $\bar{N} = \bar{N}_1 \times \bar{N}_2 \times \bar{N}_3$ where

$\bar{N}_1 = \{x \in N_1 \text{ and } ||x - x^*|| < \varepsilon_1\}$

$\bar{N}_2 = \{B \in N_2, \; ||B - G(x^*)|| < 1/2\gamma\}$

$\bar{N}_3 = \{h \in N_3, \; ||h|| < d_1\}$

Now suppose $(x, B, h) \in \bar{N}$, by the Banach Lemma, since $G(x^*)$ is nonsingular, for all $B \in \bar{N}_2$, B is nonsingular and $||B^{-1}||$

$\leq \dfrac{||G(x^*)||}{1 - ||B - G(x^*)|| \; ||G(x^*)^{-1}||} \leq \dfrac{\gamma}{1 - \gamma/(2\gamma)} = 2\gamma$. Therefore,

by the Remark 1 on page 5

$$||s|| \le ||B^{-1}|| \; ||\bar{g}(x, h)|| \le ||B^{-1}|| \; (||g(x)||$$
$$+ (c + k\,\varepsilon_1^{p}) \; ||h||).$$

And Lemma 3.1 implies that $||g(x)|| \le ||G(x^*)|| \; ||x - x^*||$ $+ K\,\varepsilon_1^{p} \; ||x - x^*||$. Hence $||s|| \le 2\gamma(c + K\,\varepsilon_1^{p})\,\varepsilon_1 + (c + K\,\varepsilon_1^{p})\,d_1$, which by (4.10) (4.11) is less than $(1/3\mu_1)^{1/p}$. Hence, it follows from (4.8) that for all $(x, B, h) \in \bar{N}$, $\frac{||Mc - M^{-1}s||}{||M^{-1}\,s||} \le 1/3$. For all $(s, B, h) \in \bar{N}$ with $\bar{g}(x, h) \ne 0$, $x \ne x^*$, $B^{-1}\,\bar{g}(x, h) \ne 0$. Hence, $s \ne 0$. Now we can apply Lemma 4.2(a) with $s = B^{-1}\,\bar{g}(x, h)$ then we have $c^T s \ge 2/3\,||M^{-1}\,s||^2 > 0$. Therefore, $\bar{B} \in U\,(x, B, h)$ with U defined as in assumption (5) is well-defined. Apply Lemma 4.1 with $A = G(x^*)$, y, s as these defined in this lemma, we have

$$||\bar{B} - G(x^*)||_M \le ||(M(B - G(x^*))M)\,[I - \frac{M^{-1}s\;(M^{-1}c)^T}{c^T s}]\,||_F$$
$$+ \,||\frac{M(y - G(x^*))\;Mc^T}{c^T s}||_F$$

Apply Lemma 4.2(c) (d) with the same A, y, s we have

$$||\bar{B} - G(x^*)||_M \le [\sqrt{1-\alpha\theta^2} + (1-\beta)^{-1}\,\frac{||Mc - M^{-1}s||}{||M^{-1}s||}]\;||B - G(x^*)||_M$$
$$+ \,2\,||M||\;\frac{||y - G(x^*)s||}{||M^{-1}s||}$$

where

$$\alpha = \frac{1 - 2\beta}{1 - \beta^2} = \frac{3}{8},$$

$$\theta = \frac{||M(B - G(x^*))M\ M^{-1}s||}{||M(B - G(x^*))M||\ ||M(B - G(x^*))s||}$$

Let C_0 be an upper bound of $||G(x)||$ for all $x \in \bar{N}$, Lemma 3.1 and the Remark 1 on page 5 imply that

$$||y - G(x^*)s|| = ||g(\bar{x}) - g(x) - G(x^*)(\bar{x} - x)||$$

$$+ ||\bar{g}(x, h) - g(x)|| + ||\bar{g}(\bar{x}, h) - g(\bar{x})||$$

$$\leq K \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}\ ||\bar{x} - x||$$

$$+ C_0\ (||\bar{h}|| + ||h||)$$

Hence

$$||\bar{B} - G(x^*)||_M \leq (1 + \bar{\alpha}_1\ ||s||^p)\ ||B - G(x^*)||_M + 2\ ||M||\ ||M||$$

$$\frac{K \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}\ ||\bar{x} - x|| + C_0\ (||\bar{h}|| + ||h||)}{||\bar{x} - x||}$$

where $\bar{\alpha}_1 = \frac{3}{2}\mu_1$. So we have

$$||\bar{B} - G(x^*)||_M \leq (1 + \alpha_1 \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}\ ||B - G(x^*)||_M$$

$$+ \alpha_2 \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\} + \alpha_3\ \frac{||\bar{h}|| + ||h||}{||\bar{x} - x||}$$

where $\alpha_1 = \frac{3}{2}\mu_1 2^p$, $\alpha_2 = 2K||M||^2$, $\alpha_3 = 2||M||^2 C_0$.
This concludes the proof.

If we can prove that (3.4) can be satisfied by this update then we can have the sequence generated by the Algorithm MQ with this update is convergent. We need the following lemma.

Lemma 4.4: Let $||\cdot||_M$ be the M-norm we define above, $||\cdot||$ be the ℓ_2-norm then there exists a constant $\gamma > 0$ such that the norm $||\cdot||_*$ defined by $||A||_* = \gamma ||A||_M$ is the matrix norm consistent with the ℓ_2-norm.

Proof: We want to prove that $||Ax||_2 \le ||A||_* ||x||_2$ holds for all $A \in L(R^n)$, $x \in R^n$. By the definition of Frobenious norm,

$$||Ax||_2 \le ||A||_2 \; ||x||_2 \le ||A||_F^2 \; ||x||_2 \le ||M^{-1}MAMM^{-1}||_F^2 \; ||x||_2$$

$$\le ||M^{-1}||_F^2 \; ||MAM||_F ||x||_2 \le ||M^{-1}||_F^2 \; ||A||_M \; ||x||_2.$$

If we choose $\gamma = ||M^{-1}||_F^2$ then $||Ax||_2 \le ||A||_* ||x||_2$. That means, the matrix form $||\cdot||_*$ defined by using this γ is consistent with the ℓ_2-norm.

Now we are ready to prove the following theorem about the convergence of the single-rank methods.

Theorem 4.5: With the assumptions of Lemma 4.3, there is a neighborhood N_0 of $(x^*, G(x^*), 0)$ such that the iteration generated by Algorithm MQ with the update function $U(x, B, h)$ is well-defined and convergent to x^*. Furthermore, it converges Q-superlinearly.

Proof: Let $N_0 = N \cap \tilde{N} \cap \overline{N}$ where $\tilde{N}$ in Lemma 4.3, $\overline{N} = \{(x, B, h)$ such that $||x - x^*|| < \varepsilon$, $||B - G(x^*)|| < \delta$, $||h|| < d$ where ε, δ, d are defined in Theorem 3.2$\}$.

It follows from Lemma 4.3 that (4.9) exists. Let $\gamma = ||M^{-1}||_F^2$ multiply (4.9) by γ, then (4.9) becomes

$$\gamma\ ||\overline{B} - G(x^*)||_M \le (1 + \alpha_1 \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\})\ \gamma\ ||B - G(x^*)||_M + \alpha_2\ \gamma \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\} + \gamma\ \alpha_3 \frac{||h|| + ||\overline{h}||}{||\overline{x} - x||}$$

By Lemma 4.4, we know the matrix norm $||A|| = \gamma\ ||A||_M$ is consistent with the ℓ_2-norm. And it follows from the above inequality that,

$$||\overline{B} - G(x^*)|| \le (1 + \alpha_1 \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\})\ ||B - G(x^*)|| + \alpha_2\gamma \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\} + \gamma\ \alpha_3 \frac{||h|| + ||\overline{h}||}{||\overline{x} - x||}$$

where α_1, $\alpha_2\ \gamma$, $\alpha_3\ \gamma$ are positive constants. Hence, (3.4) is satisfied. So by applying Theorem 3.2 with the particular norm, we conclude that the iteration is well-defined and convergent.

To prove $\{x^k\}$ is superlinearly convergent. There are two cases:

Case I: There is a subsequence of $\{B_k\}$ convergent to $G(x^*)$. Then it follows directly from Corollary 3.5 that $\{x^k\}$ converges Q-superlinearly.

Case II: If $\{||B_k - G(x^*)||\}$ is bounded away from zero. The proof of Theorem 4.3 shows that the following inequality

$$||B_{k+1} - G(x^*)|| \le \sqrt{1 - 3/8\ \theta_k{}^2}\ ||B_k - G(x^*)||$$
$$+ \max\{||x - x^*||^p, ||\bar{x} - x^*||^p\}$$
$$[\gamma\ \alpha_1\ ||B_k - G(x^*)|| + \alpha_2\ \gamma] + \alpha_3\ \gamma\ \max\{||x - x^*||, ||\bar{x} - x^*||\}$$

where

$$\theta_k = \frac{\gamma\ ||M(B_k - G(x^*))\ s_k||}{||B_k - G(x^*)||\ ||M^{-1}\ s_k||} \tag{4.13}$$

holds with $(x^0, B_0, h^0) \in N_0$. Since $\sqrt{1 - \alpha} \le 1 - \frac{\alpha}{2}$ we have

$$\frac{3}{16}\ \theta_k{}^2\ ||B_k - G(x^*)|| \le ||B_k - G(x^*)|| - ||B_{k+1} - G(x^*)||$$
$$+ \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}\ [\alpha_1\ \gamma\ ||B_k - G(x^*)|| + \alpha_2\ \gamma]$$
$$+ \alpha_3\ \gamma\ \max\{||\bar{x} - x^*||, ||x - x^*||\}$$

Therefore, $\Sigma\ \theta_k^2\ ||B_k - G(x^*)|| < +\infty$ because $||x^k - x^*|| \leq \gamma^{\lfloor \frac{k}{2} \rfloor}\ ||x^0 - x^*||$ and $p < 1$. Since $||B_k - G(x^*)||$ is bounded away from zero, we conclude θ_k tends to zero. Now $||B_k - G(x^*)||$ is bounded above, hence (4.13) implies that $\frac{||(B_k - G(x^*))s^k||}{||s^k||}$ tends to zero. Similar to the proof of Corollary 3.5 we can get

$$\frac{||\bar{g}(x^k, h^k) - g(x^k)||}{||s^k||} \to 0$$

Then apply Theorem 3.3, $\{x^k\}$ converges to x^* locally and Q-superlinearly.

The update of Broyden's first method and Pearson's method are two special cases of the update defined in Theorem 4.5. Therefore, we have the following corollary:

Corollary 4.6: Let the hypotheses (1) to (3) of Theorem 4.5 hold. Then the modified form of Broyden's first method ($c = s$) is locally and Q-superlinearly convergent. If, in addition, $G(x^*)$ is symmetric and positive definite, then the modified form of Pearson's method ($c = y$) is also locally and Q-superlinearly convergent at x^*.

Proof: For Broyden's method, follow the proof of Corollary 4.4 in [4] by Broyden-Dennis-Moré. For Pearson's method, put $c = y$. Since $G(x^*)$ is positive definite, there is a symmetric $M \in L(R^n)$ such that $G(x^*)^{-1} = M^2$. Then

$$My - M^{-1}s = M[y - G(x^*)s]$$

By Lemma 3.1 and the Remark 1 on page 5:

$$||My - M^{-1}s|| \leq ||M||\ ||y - G(x^*)s|| \leq ||M||$$
$$[K \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\}\ ||\overline{x} - x||$$
$$+ C_0\ (||h|| + ||\overline{h}||)]$$

where C_0 is an upper bound for $||G(x)||$ for all $x \in D$. Since **h**, $\overline{h}$ satisfy (2.2) and (2.3), we have

$$||My - M^{-1}s|| \leq ||M||\ [K \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\}$$
$$+ C_0\ (C_2\ ||\overline{x} - x|| + 2C_1\ ||\overline{x} - x||)]\ ||\overline{x} - x|| \tag{4.14}$$

The result will then follow from Theorem 4.5 if we can show that there is a neighborhood N of $(x^*, G(x^*), 0)$ such that $\max\{||x - x^*||^p, ||x - x^*||^p\} \leq \lambda\ ||s||^p$ for some constant $\lambda > 0$ and all (x, B, h) in N' with $x \neq x^*$, $\overline{g}(x, h) \neq 0$. Now, by Lemma 3.1, there is an $\varepsilon > 0$ and a $\rho > 0$ such that (3.3) holds if $\max\{||x - x^*||, ||x - x^*||\} < \varepsilon$. Set $N_1 =$ $\{x \in R^n: ||x - x^*|| < \min(\frac{\varepsilon}{2}, 1), 2\rho\ ||G(x^*)^{-1}||\ ||x - x^*|| < \min(\frac{\varepsilon}{4}, \frac{1}{2})$ and $G(x)$ is nonsingular.$\}$

$N_2 = \{B \varepsilon L(R^n) : ||G(x^*)^{-1}|| \; ||B - G(x^*)|| < \frac{1}{2}\}$

$N_3 = \{h \varepsilon R^n : \overline{g}(x, h)$ is well-defined for all $x \varepsilon N_1$

and $||h|| \leq \dfrac{\min(\frac{\varepsilon}{4}, \frac{1}{2})}{2C_0}\}$

Then, $N' = N_1 \times N_2 \times N_3$ is a neighborhood of $(x^*, G(x^*), 0)$. If $(x, B, h) \varepsilon N'$, with $x \neq x^*$, $\overline{g}(x, h) \neq 0$, then by the Banach Lemma, the matrix B is invertible and

$$||B^{-1}|| \leq 2 \; ||G(x^*)^{-1}||$$

Equation (3.3) and the Remark 1 on page 5 then imply that

$$||s|| = ||B^{-1} \; \overline{g}(x, h)|| \leq 2 \; ||G(x^*)^{-1}|| \; (\rho||x - x^*|| + C_0 \; ||h||)$$

$$\leq \min(\tfrac{\varepsilon}{2}, 1)$$

and therefore $\overline{x} \varepsilon D$ and $||s|| < 1$. Moreover, (3.3) and the Remark 1 on page 5 also imply that

$$||B|| \; ||s|| \geq ||B \; s|| \geq (\tfrac{1}{\rho}) \; ||x - x^*|| - C_0 \; ||h||$$

Let $\overline{h}$ be the step-size corresponding to $\overline{x}$, hence according to Algorithm MQ, $\overline{h}$, h should satisfy (2.3).

Then

$$||B||\ ||s|| \geq (\frac{1}{\rho})\ ||x - x^*|| - C_0\ (||h|| + ||\bar{h}|| - ||\bar{h}||)$$

$$\geq (\frac{1}{\rho})\ ||x - x^*|| - C_0\ (C_1\ ||s|| + C_2\ ||s||)\ ||s||$$

Therefore $||x - x^*|| \leq (\rho||B|| + \rho\ C_0\ (C_1 + C_2)\ ||s||)\ ||s||$

Thus,

$$\max\{||\bar{x} - x^*||^p,\ ||x - x^*||^p\} \leq (||s|| + ||x - x^*||)^p$$

$$\leq \lambda\ ||s||^p,$$

with

$$\lambda = (\rho\ ||B|| + \rho\ C_0\ (C_1 + C_2)\ ||s|| + 1)^p > 0.$$

Hence (4.14) becomes

$$||My - M^{-1}s|| \leq ||M||\ (K\ \lambda\ ||s||^p + C_0\ (C_2 + 2C_1)\ ||s||)\ ||s||$$

Since $B^{-1}\ \bar{g}(x, h) \neq 0$, we have $1 \geq ||s|| > 0$. So by $0 < p < 1$, we obtain $||s||^p > ||s||$. Hence,

$$\frac{||My - M^{-1}s||}{||M^{-1}s||} \leq \bar{\lambda}\ ||s||^p$$

with

$$\bar{\lambda} = ||M||^2\ (K\ \lambda + C_0\ (2C_1 + C_2)) > 0$$

Therefore (4.8) is satisfied. We conclude the proof by applying Theorem 4.5 with $c = y$.

In order to prove the convergence of some methods which use the H-update form, we will need a theorem similar to Theorem 4.5 but in terms of the H-formula. We prove the following lemma first.

<u>Lemma 4.7</u>: Let the hypotheses (1) and (2) of Lemma 4.3 hold.

(3) Let $N = N_1 \times N_2 \times N_3$ be a neighborhood of $(x^*, G(x^*), 0)$ such that N_1 is contained in D and for all $x \in N_1$, $g(x) \neq 0$, N_2 contains all the nonsingular matrices, N_3 is a neighborhood of vector 0 such that for all $x \in N_1$, $\overline{g}(x, h)$ is well-defined.

(4) If $$\frac{||Md - M^{-1}y||}{||M^{-1}y||} \leq \mu_2 ||y||^p, \; y \neq 0 \qquad (4.15)$$

where $y = \overline{g}(\overline{x}, \overline{h}) - g(x, h)$ with $(x, H, h) \in N$, $x \neq x^*$, $\overline{g}(x, h) \neq 0$, $\overline{x} = x - H\overline{g}(x, h)$ and $\overline{h} \in H_2(x)$, μ_2 is a nonnegative constant, M is a nonsingular symmetric matrix in $L(R^n)$.

(5) The update function U is defined by

$$U(x, H, h) = \{\overline{H}:\overline{H} = \frac{(s - Hy)\, d^T}{d^Ty} \text{ where } y = \overline{g}(\overline{x}, \overline{h}) - \overline{g}(x, h)$$

$\overline{x} = x + H\overline{g}(x, h)$ with $\overline{h} \in H_2(x) \cap N_3$ such that $y \neq 0\}$ for all $(x, H, h) \in N$ with $\overline{g}(x, h) \neq 0$, $x \neq x^*$.

Then there exists a neighborhood $\bar{N}$ of $(x^*, G(x^*)^{-1}, 0)$ such that $\overline{H} \in U(x, H, h)$ is well-defined and for all $\overline{x} = x - H\overline{g}(x, h)$

$$\begin{aligned} ||\overline{H} - G(x^*)^{-1}||_M \leq\; & (1 + \alpha_1 \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\}) \\ & ||H - G(x^*)^{-1}||_M \\ & + \alpha_2 \max\{||\overline{x} - x^*||^p, ||x - x^*||^p\} \\ & + \alpha_3 \frac{||h|| + ||\overline{h}||}{||\overline{x} - x||} \end{aligned} \qquad (4.16)$$

where α_1, α_2, α_3 are nonnegative constants, for all $(x, H, h) \varepsilon N$ with $x \neq x^*$, $\overline{g}(x, h) \neq 0$ and $h \varepsilon H_2(x)$ such that (2.2) (2.3) are satisfied.

Proof: Let γ, C_0 be the constant such that $\gamma \geq ||G(x^*)^{-1}||$, $C_0 \geq ||G(x)||$ for all $x \varepsilon D$. By Lemma 3.1, there exists $\varepsilon > 0$, $\rho > 0$ such that $\max\{||x - x^*||, ||x - x^*||\} < \varepsilon$ implies that (3.3) holds. Now choose $\varepsilon_1 > 0$, $d > 0$ such that

$$\frac{\varepsilon_1}{4} + (\gamma + \frac{1}{2\gamma})(C + K(\frac{\varepsilon_1}{4})^p)\ d_1 < \frac{\varepsilon_1}{2} \tag{4.17}$$

$$\rho\ \frac{\varepsilon_1}{2} + 2C_0\ d_1 < (\frac{1}{3\mu_2})^{1/p} \tag{4.18}$$

$$\varepsilon_1 < \varepsilon \tag{4.19}$$

$$C_0\ (C_2 + 2C_1)\ (\frac{\varepsilon_1}{4}) < \frac{1}{\rho} \tag{4.20}$$

where C_1, C_2 are two constants used in (2.2) (2.3). Let $N = \bar{N}_1 \times \bar{N}_2 \times \bar{N}_3$ where

$$\bar{N}_1 = \{\mathbf{x}: ||x - x^*|| < \min(\frac{\varepsilon_1}{4}, \frac{\varepsilon_1}{4(\gamma + \frac{1}{2\gamma})\ \rho})\}$$

$$\bar{N}_2 = \{H\ \varepsilon\ N_2: ||H - G(x^*)^{-1}|| < \frac{1}{2\gamma}\}$$

$$\bar{N}_3 = \{h\ \varepsilon\ N_3: ||h|| < d_1\}$$

By (3.1), $||G(x)|| < C + K(\frac{\varepsilon_1}{4})^p$ for all $x\ \varepsilon\ \bar{N}_1$. Now suppose

$(x, H, h) \in \bar{N}$, note that H is nonsingular for all $H \in \bar{N}_2$ and $||H|| < \frac{1}{2\gamma} + \gamma$. It follows from (3.3), the Remark 1 on page 5 and (4.17) that

$$||s|| \leq ||H||\ ||\bar{g}(x, h)|| \leq (\gamma + \frac{1}{2\gamma})(||g(x)|| + (C + K\ (\frac{\varepsilon_1}{4})^p)\ ||h||)$$
$$\leq (\gamma + \frac{1}{2\gamma})(\rho\ ||x - x^*|| + (C + K\ (\frac{\varepsilon_1}{4})^p)\ d_1)$$
$$\leq \frac{\varepsilon_1}{4} + (\gamma + \frac{1}{2\gamma})(C + K\ (\frac{\varepsilon_1}{4})^p)\ d_1 < \frac{\varepsilon_1}{2}$$

Thus, $||\bar{x} - x^*|| \leq ||s|| + ||x - x^*|| < \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{2} < \varepsilon_1$
Hence, we proved $\bar{x} \in D$. We apply the Remark 1 on page 5 again,

$$||y|| = ||\bar{g}(\bar{x}, \bar{h}) - \bar{g}(x, h)|| \leq ||g(\bar{x}) - g(x)|| + C_0\ (||\bar{h}|| + ||h||) \tag{4.21}$$

Since (4.19) implies that $||\bar{x} - x^*|| < \varepsilon$, it follows from (3.3) with x, $\bar{x}$ that

$$||y|| \leq \rho\ \frac{\varepsilon_1}{2} + 2C_0\ d_1$$

which by (4.18) is less than $(\frac{1}{3\mu_2})^{1/p}$. Then we proved

$$\mu_2\ ||y||^p \leq \mu_2\ \frac{1}{3\mu_2} \leq \frac{1}{3}$$

Furthermore, we can prove that there exists $\bar{h} \in H_2\ (x) \cap N_3$ such that $y \neq 0$. Because if $\bar{h}$ is chosen according to Algorithm MQ, that is, if $\bar{h}$ satisfies (2.2) (2.3), then (3.3) and (4.21) imply that

$$\rho \; ||\bar{x} - x|| + C_0(C_2 + 2C_1) \; ||\bar{x} - x||^2 \geq ||\bar{g}(\bar{x}, \bar{h}) - \bar{g}(x, h)||$$

$$\geq \frac{1}{\rho} \; ||\bar{x} - x|| - C_0(C_2 + 2C_1) \; ||\bar{x} - x||^2 \tag{4.22}$$

By (4.22) we know $y = 0$ if and only if $x = x^*$. Hence it shows that there are $\bar{h}$'s such that $y \neq 0$. Now we are ready to prove that the update function is well-defined. Apply Lemma 4.2 with $c = d$, $s = y$, then we know condition (4.3) is satisfied. Hence by (4.4) we have, $d^T y = 0$ iff $y = 0$. But, in the update function $\bar{h}$ is chosen in such a way that $y \neq 0$. Hence, $d^T y \neq 0$. And, hence $\bar{H} \in U(x, H, h)$ with $(x, H, h) \in \bar{N}$ is well-defined. The inequality (4.5) can be proved by following the proof of the inequality (4.9) but replace B by H, interchange s and y, replace c by d and replace $G(x^*)$ by $G(x^*)^{-1}$. Hence, Lemma 4.1 yields that

$$\bar{E} = E + [I - \frac{M^{-1} y (Md)^T}{d^T y}] + \frac{M(s - Ay)(Md)^T}{d^T y}$$

for any $A \in L(R^n)$ with $\bar{E} = M(\bar{H} - A)M$, $E = M(H - A)M$, $A = G(x^*)^{-1}$. Then, it follows from (c) and (d) of Lemma 4.2 that

$$||\bar{H} - G(x^*)^{-1}||_M \leq (\sqrt{1 - 3/8\,\theta^2} + \frac{3}{2}\,\mu_2\; ||y||^p)\; ||H - G(x^*)^{-1}||_M$$

$$+ \frac{2\; ||M||\; ||s - G(x^*)^{-1} y||}{||M^{-1} y||}$$

$$\theta = \frac{||E\, M^{-1}\, y||}{||E||_F\; ||M^{-1} y||}$$

But for any $x \in N_1$, $||s - G(x^*)^{-1} y|| \leq ||G(x^*)^{-1}||\ ||y - G(x^*) s||$ so by the proof of Lemma 4.3,

$$||s - G(x^*)^{-1} y|| \leq ||G(x^*)^{-1}|| K \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}$$
$$||\bar{x} - x|| + C_0 \ (||h|| + ||\bar{h}||)$$

If $\bar{h}$ satisfies (2.2) (2.3), we know from (4.20) and (4.22) that there exists a positive constant ρ' such that for $||\bar{x} - x^*|| < \varepsilon_1$, $||x - x^*|| < \varepsilon_1$

$$\rho' \ ||\bar{x} - x|| \geq ||\bar{g}(\bar{x}, \bar{h}) - \bar{g}(x, h)|| \geq \frac{1}{\rho'} ||\bar{x} - x|| \qquad (4.23)$$

Therefore, we obtain

$$||\bar{H} - G(x^*)^{-1}||_M \leq \sqrt{1 - 3/8\ \theta^2}\ ||H - G(x^*)^{-1}||_M$$
$$+ \max\{||\bar{x} - x||^p, ||x - x^*||^p\} \ (\alpha_1 \ ||H - G(x^*)||_M + \alpha_2)$$
$$+ \alpha_3 \frac{||h|| + ||\bar{h}||}{||\bar{x} - x||} \qquad (4.24)$$

where

$$\alpha_1 = (\tfrac{3}{2})\ \mu_2\ (2\rho')^p$$
$$\alpha_2 = 2\rho'\ K\ ||G(x^*)^{-1}||\ ||M^2||$$
$$\alpha_3 = 2\ ||M||^2\ \rho'$$
$$\theta = \frac{||M - (H - G(x^*)^{-1})\ y||}{||H - G(x^*)^{-1}||_M\ ||M^{-1}\ y||}$$

This concludes the proof.

Theorem 4.8: With the assumptions of Lemma 4.7, there is a neighborhood N_0 of $(x^*, G(x^*)^{-1}, 0)$ such that the iteration generated by Algorithm MQ with the update $U(x, H, h)$ is well-defined and convergent to x^*. Furthermore, it converges Q-superlinearly.

Proof: Let $N_0 = N \cap \tilde{N} \cap \bar{N}$ where $\tilde{N}$ is the $\tilde{N}$ in Theorem 4.7, $\bar{N} = \{(x, H, h)$ such that $||x - x^*|| < \varepsilon$, $||H - G(x^*)^{-1}|| < \delta$, $||h|| < d$ where ε, δ, d are defined in Theorem 3.6.$\}$
It follows from Lemma 4.7 that (4.16) exists for all $(x, H, h) \in N$. Similar to the proof of Theorem 4.5, we can prove that there is a matrix norm $||\cdot||$ consistent with ℓ_2-norm such that

$$||H - G(x^*)^{-1}|| \le (1 + \alpha_1 \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\})$$
$$||H - G(x^*)||$$
$$+ \alpha_2 \gamma \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\} + \gamma \alpha_3 \frac{||h|| + ||\bar{h}||}{||\bar{x} - x||} \tag{4.25}$$

where α_1, α_2, α_3, γ are positive constants for all $(x, H, h) \in N_0$, $\bar{x} = x - H\bar{g}(x, h)$ and $\bar{h} \in H_2(x)$ such that $\bar{h}$ satisfies (2.2), (2.3). So it follows from Theorem 3.6 that the iteration is convergent.

To prove the superlinearly convergence we again follow the proof of Theorem 4.5. We discuss it in two cases.

Case I: There is a sequence $\{H_k\}$ converges to $G(x^*)^{-1}$. Then the result follows directly from Corollary 3.8.

Case II: The sequence $||H_k - G(x^*)^{-1}||$ is bounded away from zero. From (4.24) and (4.25) we have

$$\frac{3}{16}\theta_k^2 \, ||H_k - G(x^*)^{-1}|| \le ||H_k - G(x^*)^{-1}|| - ||H_{k+1} - G(x^*)^{-1}||$$

$$+ \max\{||x^k - x^*||^p, ||x^{k+1} - x^*||^p\}$$

$$[\alpha_1 \, ||H_k - G(x^*)^{-1}|| + \alpha_2 \, \gamma]$$

$$+ \hat{\alpha}_3 \, \gamma \max\{||\bar{x} - x^*||, ||x - x^*||\}$$

where $\theta_k = \dfrac{||M[H_k - G(x^*)^{-1}] \, y^k||}{\gamma \, ||H_k - G(x^*)^{-1}|| \, ||M^{-1} \, y^k||}$

Therefore, $\Sigma \, \theta_k^2 \, ||H_k - G(x^*)^{-1}|| < +\infty$ by $||x^k - x^*|| \le r^{\lfloor \frac{k}{2} \rfloor} ||x^0 - x^*||$ and $p < 1$. Since $\{||H_k - G(x^*)^{-1}||\}$ is bounded away from zero, we conclude that $\theta_k \to 0$ as $k \to \infty$. Because $||H_k - G(x^*)^{-1}||$ is bounded above, we have

$$\frac{||[H_k - G(x^*)^{-1}] \, y^k||}{||y^k||} \to 0 \text{ as } k \to \infty$$

Since

$$||(H_k - G(x^*)^{-1}) \, y^k|| \le ||(H_k - G(x^*)^{-1})(g(x^{k+1}) - g(x^k))||$$

$$- ||H_k - G(x^*)^{-1}|| \; (||\bar{g}(x^{k+1}, h^{k+1}) - g(x^{k+1})||$$
$$+ ||\bar{g}(x^k, h^k) - g(x^k)||)$$

it follows from the fact that $||H_k - G(x^*)^{-1}||$ is bounded above and the Remark 1 on page 5 that there exists $L > 0$ such that

$$||(H_k - G(x^*)^{-1})\, y^k|| \geq ||(H_k - G(x^*)^{-1})(g(x^{k+1}) - g(x^k))||$$
$$- L(||h^{k+1}|| + ||h^k||)$$

Now, h^k, h^{k+1} are chosen according to Algorithm MQ that means they satisfy (2.8) and (2.9), therefore, (4.23) will be true with $\bar{x} = x^{k+1}$ and $x = x^k$. So,

$$\frac{||(H_k - G(x^*)^{-1})\, y^k||}{||y^k||} \geq \frac{||(H_k - G(x^*)^{-1})(g(x^{k+1}) - g(x^k))||}{\rho' \; ||s^k||}$$
$$- \frac{L(2C_1 + C_2)\; ||s^k||^2}{\rho' \; ||s^k||}$$

When $k \to 0$, $||s^k|| \to 0$, therefore,

$$\frac{||[H_k - G(x^*)^{-1}]\, y^k||}{||y^k||} \geq \frac{||[H_k - G(x^*)^{-1}](g(x^{k+1}) - g(x^k))||}{\rho' \; ||s^k||}$$

Apply (3.3) with $u = x^{k+1}$, $v = x^k$, then

$$||g(x^{k+1}) - g(x^k)|| \geq \frac{1}{\rho} \; ||s^k||$$

Hence,

$$\frac{||[H_k - G(x^*)^{-1}]y^k||}{||y^k||} \geq \frac{||(H_k - G(x^*)^{-1})(g(x^{k+1}) - g(x^k))||}{\rho\rho' \; ||g(x^{k+1}) - g(x^k)||}$$

when $k \to 0$. Thus,

$$\frac{||(H_k - G(x^*)^{-1})(g(x^{k+1}) - g(x^k))||}{||g(x^{k+1}) - g(x^k)||} \to 0 \quad \text{as } k \to \infty.$$

The proof of Corollary 3.8 shows that

$$\frac{||H_k[\overline{g}(x^k, h^k) - g(x^k)]||}{||g(x^{k+1}) - g(x^k)||} \to 0 \qquad \text{as } k \to \infty$$

Hence, (3.23) of Theorem 3.7 is satisfied. And we conclude the proof.

Corollary 4.9: Let the hypotheses (1) to (3) of Theorem 4.8 hold. Then the modified form of Broyden's second method ($d = y$) is locally and Q-superlinearly convergent at x^*. If, in addition, $G(x^*)$ is symmetric and positive definite, then the modified form of McCormick's method ($d = s$) is also locally and Q-superlinearly convergent at x^*.

Proof: For Broyden's method, (4.15) can be satisfied easily by using $M = I$.

For McCormick's method, since $G(x^*)$ is positive definite and symmetric, let $M^2 = G(x^*)$, and C_0 be an upper bound for $||G(x)||$ for all $x \in D$. Then,

$$M^{-1} y - Ms = M^{-1} (y - G(x^*) s)$$

It is proved in the proof of Lemma 4.3 that

$$\frac{||M^{-1}y - Ms||}{||M^{-1}y||} \le \frac{||y - G(x^*)s||}{||y||} \le K \max\{||\bar{x} - x^*||^p,$$

$$||x - x^*||^p\} \frac{||\bar{x} - x||}{||y||} + C_0 \frac{(||\bar{h}|| + ||h||)}{||y||}$$

Since $\bar{h}$ is chosen according to Algorithm MQ, by (4.23)′ we obtain

$$\frac{||M^{-1}y - Ms||}{||M^{-1}y||} \le \rho' K \max\{||\bar{x} - x^*||^p, ||x - x^*||^p\}$$

$$+ C_0(2C_1 + C_2)\ ||s||$$

Then, similar to the proof of Corollary 4.6, we can prove that there is a neighborhood N' of $(x^*, G(x^*)^{-1}, 0)$ such that $\max\{||x - x||^p, ||x - x^*||^p\} \le \lambda\ ||s||^p$, and $||s|| < 1$. Hence, $\frac{||M^{-1}y - Ms||}{||M^{-1}y||} \le \rho' (K + C_0 (C_2 + 2C_1))\ ||s||^p$. Apply (4.23) again then

$$\frac{||M^{-1}y - Ms||}{||M^{-1}y||} \le (\rho')^{p+1} (K + C_0 (2C_1 + C_2))\ ||y||^p$$

Therefore, the result follows form Theorem 4.8.

V. CONVERGENCE OF THE MODIFIED DOUBLE-RANK METHOD

Similar to the modified single-rank methods, the convergence of the modified double-rank methods can be proved by using Theorem 3.2 or Theorem 3.6. We start with the double-rank analogue of Lemma 4.1.

Lemma 5.1: Let $B \in L(R^n)$ be symmetric, and set

$$\bar{B} = B + \frac{(y - Bs)\, c^T + c(y - Bs)^T}{c^T s} - \frac{s^T (y - Bs)\, cc^T}{(c^T s)^2} \tag{5.1}$$

where y, c and s are vectors in R^n with $c^T s \neq 0$. If $A \in L(R^n)$ is symmetric, and $M \in L(R^n)$ is nonsingular and symmetric, then

$$\bar{E} = [I - \frac{(Mc)(M^{-1} s)^T}{c^T s}]\, E\, [I - \frac{(M^{-1} s)(Mc)^T}{c^T s}] + \frac{M(y - As)(Mc)^T}{c^T s} + \frac{(Mc)(y - As)^T M}{c^T s}\, [I - \frac{(M^{-1} s)(Mc)^T}{c^T s}] \tag{5.2}$$

when $\bar{E} = M(\bar{B} - A)M$ and $E = M(B - A)M$

Proof: See [4].

Lemma 5.2: Let $M \in L(R^n)$ be a nonsingular symmetric matrix such that inequality (4.3) holds for some $\beta \in [0, 1/3]$ and vectors c and s in R^n with $s \neq 0$. Let $B \in L(R^n)$ be symmetric and define B by (5.1) where $y \in R^n$. If $||\cdot||_M$ is the matrix norm defined by $||Q||_M = ||MQM||_F$, then for any symmetric $A \in L(R^n)$ with $B \neq A$,

$$||\overline{B} - A||_M \leq [\sqrt{1 - \alpha\theta^2} + (\tfrac{5}{2})(1 - \beta)^{-1} \frac{||Mc - M^{-1} s||}{||M^{-1} s||}] \; ||B - A||_M]$$

$$+ 2(1 + 2\sqrt{n}) \; ||M||_F \frac{||y - As||}{||M^{-1} s||} \qquad (5.3)$$

where

$$\alpha = \frac{1 - 2\beta}{1 - \beta^2} \in [\tfrac{3}{8}, 1]$$

and

$$\theta = \frac{||M[B - A] s||}{||B - A||_M \; ||M^{-1} s||}$$

Proof: See [4].

The following is a result analogous to Theorem 4.5 but for the update (5.1).

Theorem 5.3: Suppose the assumptions (1) to (4) of Lemma 4.3 hold and, in addition assume that:

(5) the update function U is defined by

$$U(x, B, h) = \{\overline{B}:\overline{B} = B + \frac{(y - Bs)c^T + C(y - Bs)^T}{c^T s} - \frac{s^T(y - Bs) \; c \; c^T}{(c^T s)^2} .$$

where $\overline{x} = x + s$ with $s = B^{-1} \; \overline{g}(x, h)$

$y = \overline{g}(\overline{x}, \overline{h}) - \overline{g}(x, h)$ with $\overline{h} \in H_2 \; (x)\}$

for all $(x, B, h) \in N$ with $\bar{g}(x, h) \neq 0$, $x \neq x^*$, B is symmetric. Then there exists a neighborhood $\bar{N}$ of $(x^*, G(x^*), 0)$ such that $\bar{B}$ is well-defined and the corresponding iteration generated by Algorithm MQ with the update U is well-defined, locally and Q-superlinearly convergent at x^*.

Proof: Similar to the proof of Lemma 4.3, we could find a neighborhood $\bar{N}$ of $(x^*, G(x^*), 0)$ such that $0<\mu_1||s||^p \leq 1/3$ for all $(x, B, h) \in \bar{N}$. Apply Lemma 5.2 with $A = G(x^*)$ and follow the proof of Lemma 4.3 and the proof of Theorem 5.3 in [4] we get (4.9) with $\alpha_1 = (\frac{5}{2})(1 - \beta)^{-1} \mu_1 2^p$, $\alpha_2 = 2K(1 + 2\sqrt{n})\, ||M||_F\, ||M||$, $\alpha_3 = 2C_0 (1 + 2\sqrt{n})\, ||M||_F\, ||M||$ where C_0 is an upper bound for $||G(x)||$ for all $x \in D$. Thus the proof can be completed by following the proof of Theorem 4.5.

The results that held for the single-rank updates will hold for the double-rank updates provided the relevant matrices are symmetric. Hence, we will just state the results, the proofs are the same as those for the corresponding results in Chapter IV.

Corollary 5.4: Let the assumptions of Theorem 5.3 hold. Then the modified Powell symmetric Broyden algorithm $(c = s)$ is locally and Q-superlinearly convergent at x^*. If, in addition $G(x^*)$ is positive definite, then the modified Davidon-Fletcher-Powell $(c = y)$ method is also locally and Q-superlinearly convergent at x^*.

Theorem 5.5: Let the assumptions (1) to (4) of Lemma 4.7 hold and in addition assume:

(5) the update function U is defined by

$$U(x, H, h) = \{\bar{H} : \bar{H} = H + \frac{(s - Hy)\, d^T + d(s - Hy)^T}{d^T y} - \frac{y^T (s - Hy)\, dd^T}{(d^T y)^2}$$

where $x = x + H\bar{g}(x, h)$, $y = \bar{g}(\bar{x}, \bar{h}) - \bar{g}(x, h)$ with $h \in H_2(x)$ such that $y \neq 0.\}$

for all $(x, H, h) \in N$ with $\bar{g}(x, h) \neq 0$, $x \neq x^*$.
Then there exists a neighborhood $\tilde{N}$ of $(x^*\ G(x^*)^{-1}, 0)$ such that $\bar{H} \in U(x, H, h)$ is well-defined and the corresponding iteration by Algorithm MQ with this update function U is locally and Q-superlinearly convergent at x^*.

Corollary 5.6: Let the hypotheses of Theorem 5.5 hold. Then the modified Greenstadt method ($d = y$) is locally and Q-superlinearly convergent at x^*. If, in addition, $G(x^*)$ is positive definite, then the modified Broyden-Fletcher-Goldfarb-Shanno method ($d = s$) is also locally and Q-superlinearly convergent at x^*.

VI. CONVERGENCE OF THE MODIFIED MINFA

The program MINFA, which is an implementation of the dogleg algorithm, suggested by Powell [21] in 1970, uses function and first derivative values to solve unconstrained minimization problems. This algorithm has been proved to have some desirable superlinearly convergent properties.

The modified MINFA (we call it Algorithm MD) uses the finite difference derivative $\bar{g}(x, h)$ instead of the real derivative $g(x)$. The algorithm is iterative. Given a starting vector x^1, it generates a sequence of vectors x^k ($k = 1, 2, 3, \ldots$) which is intended to converge to a point at which the objective function $f(x)$ has a local minimum.

At the beginning of each iteration a point x^k is available, with a matrix B_k and a step bound Δ^k.

To define x^{k+1} we first define the quadratic approximation:

$$\phi(s^k) = f(x^k) + s^{k^T} \bar{g}(x^k, h^k) + \tfrac{1}{2} s^{k^T} B_k s^k \qquad (6.1)$$

then we calculate a value s^k such that

$$\phi(s^k) \leq f(x^k). \qquad (6.2)$$

Thus x^{k+1} is defined by

$$x^{k+1} = \begin{cases} x^k & \text{if } f(x^k) \le f(x^k + s^k) \qquad (6.3a) \\ x^k + s^k & \text{otherwise.} \qquad (6.3b) \end{cases}$$

At the end of the iteration we test the convergence criterion:

$$||\bar{g}(x^k, h^k)|| < \varepsilon \qquad (6.4)$$

where h^{k+1} is choosen according to Chapter II if $x^{k+1} \neq x^k$.

There are two ways of defining s^k in the algorithm. For ordinary iteration, s^k is defined like the one used by Powell [21] (1970).

$$s^k = \begin{cases} -(B_k^{-1}\, \bar{g}(x^k, h^k)) & \text{if } B_k \text{ is positive definite and} \\ & \quad ||B_k^{-1}\, \bar{g}^k|| < \Delta^k, \\ -\Delta^k \dfrac{\bar{g}(x^k, h^k)}{||\bar{g}(x^k, h^k)||} & \text{if } (\bar{g}^{k^T} B_k\, \bar{g}^k)\, \Delta^k \le ||\bar{g}^k||^2, \\ (1-\alpha) \dfrac{||\bar{g}^k||^2\, \bar{g}^k}{\bar{g}^{k^T} B_k\, \bar{g}^k} + \alpha\, B_k^{-1}\, \bar{g}^k = s^k(\alpha) & \text{otherwise} \\ \text{where } \alpha \text{ is choosen so that } \phi(s^k(\alpha)) \text{ is the least} & \\ \text{for } ||s^k(\alpha)|| \le \Delta^k & \end{cases}$$

where $\bar{g}^k = \bar{g}(x^k, h^k)$.

The step bound Δ^{k+1} is calculated according to the success of the kth iteration,

$$\Delta^k = \begin{cases} ||s^k|| \text{ or } 2||s^k|| & \text{if } f(x^k + s^k) - f(x^k) \le 0.1(\phi(s^k) - f(x^k)), \\ \tfrac{1}{2}\, ||s^k|| & \text{otherwise.} \end{cases}$$

Let $y = \bar{g}(x + s, h_s) - \bar{g}(x, h)$ where h_s is one of those $\bar{h}$'s which satisfies (2.2), (2.3) with $\bar{x} = x + s$, and the update function of the approximation to the Hessian is

$$\bar{B} = B + \theta \frac{(y - Bs)s^T + s(y - Bs)^T}{||s||^2} - \theta^2 \frac{s^T(y - Bs)\ ss^T}{||s||^4} \tag{6.5}$$

where θ is the number closest to 1 such that $|\det B^{k+1}| > 0.1\ |\det B^k|$. Every third iteration is a special iteration, for which s^k is defined in such a way that for some constant $E < 1$

$$||\prod_{j=k}^{k+\ell} (I - \theta^j \frac{s^i\ s^{j^T}}{||s^j||^2})||_2 \le E$$

will be true for all k and a fixed ℓ. Unlike an ordinary iteration, a special iteration always sets $\Delta^{k+1} = \Delta^k$, but B_k, x^{k+1}, h^{k+1} are calculated in the usual way.

Recently, Powell [24] has proved that MINFA actually belongs to a class of minimization algorithms which have the properties that there is no need for the starting point x^1 to be close to the solution, the function f(x) need not be convex and the superlinearly convergence can be proved even though the second derivative approximation may not converge to the real Hessian at the solution. We will first prove that similar results still hold after we modify this class of algorithms. Then we will prove that the Algorithm

MD belongs to the class of algorithms with real gradient replaced by the finite difference quotients and we can conclude the convergence properties of the Algorithm MD by applying those results.

The methods under consideration here will generate a sequence $\{x^k\}$ exactly the same way Algorithm MD does. This implies that equations (6.1), (6.2), (6.3) will all be true for any one of the methods in the class. However, s^k, B_k, Δ^k will be defined more generally in the following way:

$$s^k = \begin{cases} B_k^{-1}\,\bar{g}(x^k, h^k) & \text{if } B_k \text{ is positive-definite and } ||B_k^{-1}\,\bar{g}^k|| \le \Delta^k \\ ||s^k|| = ||\Delta^k|| & \text{otherwise.} \end{cases}$$

Furthermore, on every iteration the choice of s^k must satisfy the inequality:

$$f(x^k) - \phi(s^k) \ge C_3\,||\bar{g}^k||\,\min[||s^k||, \frac{||\bar{g}^k||}{||B_k||}] \tag{6.6}$$

where the step-size h^k is chosen according to Chapter II, provided $x^{k+1} \ne x^k$, $C_3 > 0$, B_1 may be any symmetric matrix and B_k will be generated by any method which provides that the condition

$$||B_k|| \le C_{12} + C_{13} \sum_{i=1}^{k} ||s^i|| \tag{6.7}$$

is satisfied, where $C_{12} > 0$, $C_{13} > 0$. The step bound Δ^k is defined in the following way. If

$$f(x^k) - f(x^k + s^k) \geq C_4 \, (f(x^k) - \phi(s^k)) \tag{6.8}$$

is obtained with $0 < C_4 < 1$, then Δ^{k+1} satisfies the bound

$$||s^k|| \leq \Delta^{k+1} \leq C_5 \, ||s^k|| \tag{6.9}$$

where $C_5 \geq 1$. If (6.8) fails, Δ^{k+1} satisfies

$$C_6 \, ||s^k|| \leq \Delta^{k+1} \leq C_7 \, ||s^k|| \tag{6.10}$$

where $0 < C_6 \leq C_7 < 1$. Moreover, we impose a fixed upper bound on the step lengths, so we have $\Delta^k \leq \bar{\Delta}$ for $k = 1, 2, 3, \ldots$

The class of algorithms that is analyzed consists of the algorithm that meet all the conditions given so far. We will prove that the algorithms after the modification still have all the properties they originally had. In an attempt to increase the readability of this material we have used notation allowing intermediate results occuring in one proof to be used in subsequent proofs. For example, if k_6 is chosen greater or equal to k_5 in a proof, k_5 may have been chosen in a previous proof.

<u>Theorem 6.1</u>: Suppose f(x) is bounded below and differentiable,

$g(x)$ is uniformly continuous on a convex hull of the level set $L(x^1)$ of the starting point x^1. Then the vectors $\overline{g}(x^k, h^k)$ $(k = 1, 2, 3, \ldots)$ are not bounded away from zero.

<u>Proof</u>: Although most of the proof follows the proof of Theorem 1 in [24], for the sake of completeness we will not omit any part of it.

Let Σ' denote the sum over the iterations for which condition (6.8) is satisfied. Suppose (6.8) holds for $k = p$ and fails for $k = p + 1, \ldots, q$ then expression (6.9), (6.10) and the fact that $||s^k|| \le \Delta^k$ imply the bound

$$\sum_{i=p}^{q} ||s^i|| \le ||s^p|| \, [1 + C_5 + C_5 C_7 + C_5 C_7{}^2 + \cdots + C_5 C_7{}^{q-p-1}] \le ||s^p|| \, [1 + \frac{C_5}{1 - C_7}].$$

Therefore, the following inequality

$$\sum_{i=1}^{k} ||s^i|| \le [1 + \frac{C_5}{1 - C_7}] \, [||s^1|| + \sum_{i=2}^{k}{}' \, ||s^i||] \tag{6.11}$$

holds. Thus we deduce from inequality (6.7) that there exists constants $C_8 > 0$ and $C_9 > 0$ such that $\{B_k\}$ satisfy the condition

$$||B_k|| \le C_8 + C_9 \sum_{k=1}^{k}{}' \, ||s^i|| \tag{6.12}$$

The fact that $f(x)$ is bounded below and the inequality (6.3) imply that sum $\sum_{k=1}^{\infty}{}' \, [f(x^k) - f(x^{k+1})]$ is convergent. Because Σ' denotes the sum over iterations for which condition (6.8) is satisfied, the sum $\sum_{k=1}^{\infty}{}' \, (f(x^k) - \phi(s^k))$ is convergent. Thus by applying the elementary inequality

$$\min[|a|, |b|] \geq \frac{|ab|}{|a| + |b|} \tag{6.13}$$

we deduce from expression (6.6) that the sum

$$\sum_{k=1}^{\infty}{}' \frac{||s^k|| \; ||\bar{g}^k||^2}{||\bar{g}^k|| + ||s^k|| \; ||B_k||} \tag{6.14}$$

is also convergent. The theorem is proved by obtaining a contridiction if $\bar{g}^k$ satisfies the bound $||\bar{g}^k|| \geq C_{10}$ where $C_{10} > 0$. In this case (6.14) and $\Delta^k \leq \bar{\Delta}$ imply that

$$\sum_{k=1}^{\infty} \frac{||s^k||}{1 + (\bar{\Delta}/C_{10})(C_8 + C_9 \sum_{i=1}^{k}{}' \; ||s^i||)} .$$

is finite. It follows from the fact if $\sum_{k=1}^{\infty}{}' \frac{a_k}{\sum_{i=1}^{k} a_i}$ is finite then $\sum_{k=1}^{\infty}{}' \, a_k$ is finite. Thus (6.12) shows that there exists a constant $C_{17} > 0$ such that $||B_k|| \leq C_{17}$ for all k.

Moreover, from (6.11) we find the limit $||s^k|| \to 0$. Let k_1 be so large that for all $k \geq k_1$, we have $||s^k|| \leq C_{10}/C_{17}$. Hence it follows from (6.6) that

$$f(x^k) - \phi(s^k) \geq C_3 \; ||\bar{g}^k|| \; ||s^k|| \tag{6.15}$$

for all $k \geq k_1$. Thus the inequality (6.1) gives

$$|1 + \frac{s^{k^T} \bar{g}^k}{f(x^k) - \phi(s^k)}| = |\frac{\frac{1}{2} s^{k^T} B_k s^k}{f(x^k) - \phi(s^k)}| \leq \frac{\frac{1}{2} |s^{k^T} B_k s^k|}{C_3 ||\bar{g}^k|| \, ||s^k||}$$

Since B_k is uniformly bounded and s^k tends to zero, the right hand side tends to zero, hence

$$\lim_{k\to\infty} \frac{s^{k^T} \bar{g}^k}{\phi(s^k) - f(x^k)} = 1 \tag{6.16}$$

Let $k_2 \geq k_1$ such that for all $k \geq k_2$ the left hand side of (6.16) is at least $\frac{1}{2}$. By (6.2) and (6.16) we have

$$-s^{k^T} \bar{g}^k \geq \tfrac{1}{2} (f(x^k) - \phi(s^k)) \geq \frac{C_3}{2} ||\bar{g}^k|| \, ||s^k|| \tag{6.18}$$

for all $k \geq k_2$. Since, for all $i = 1, \ldots, n$

$$|\bar{g}_i(x^k, h^k) - g_i(x^k)| = |g_i(x^k + \theta_i h^{k^T} e_i) - g_i(x^k)|$$
$$\leq ||g(x^k + \theta_i h^{k^T} e_i) - g(x^k)||$$

where

$$0 \leq \theta_i \leq 1,$$

we have the following inequality:

$$|f(x^k + s^k) - f(x^k) - s^{k^T} \bar{g}^k| \leq |\int_{\theta=0}^{1} s^{k^T} (g(x^k + \theta s^k)$$
$$-g(x^k)) \, d\theta| + ||s^k|| \, ||\bar{g}^k - g^k||$$

$$\le ||s^k|| \{\omega(||s^k||) + \sqrt{n} \sup_{1\le i\le n}$$

$$\{||g(x^k + \theta_i h^{k^T} e_i) - g(x^k)||\}\}$$

$$\le ||s^k|| (\omega(||s^k||) + \sqrt{n}\, \omega(||h^k||)).$$

Here, $w(\cdot)$ is the modulus of continuity of $g(x)$ which is finite by the fact that $g(x)$ is uniformly continuous. Thus, since $||s^k|| \to 0$ and $||h^k|| \le C_1 ||s^{k-1}||^2$,

$$\lim_{k\to\infty} \frac{f(x^k+s^k) - f(x^k)}{||s^k||} = \lim_{k\to\infty} \frac{s^{k^T} \bar{\sigma}^k}{||s^k||} \tag{6.19}$$

Since (6.18) and $||\bar{g}^k||$ is bounded away from zero show that this right hand side is bounded away from zero, the ratio of the left hand side to the right hand side of equation (6.19) tends to 1. Therefore equation (6.16) gives the limit

$$\lim_{k\to\infty} \frac{f(x^k) - f(x^k + s^k)}{f(x^k) - \phi(s^k)} = 1 \tag{6.20}$$

showing that the test (6.8) holds for all sufficiently large k. Thus (6.9) implies that $\Delta^{k+1} \ge ||s^k||$ for $k \ge k_3 > 0$. Since $||s^k||$ is either Δ^k or $||B_k^{-1} g^k||$, and $||B_k^{-1} g^k|| \ge C_{10}/C_{17}$ with the fact that $||s^k|| \to 0$, there exists $k_4 > 0$ such that

$$||s^k|| = \Delta^k \text{ for all } k \ge k_4.$$

Hence, $||s^{k+1}|| = \Delta^{k+1} \geq ||s^k||$ for $k \geq \max(k_4, k_3)$. In other words, after a finite number of iterations, $||s^k||$ stops decreasing. Since $||s^k||$ is always positive, we cannot obtain $||s^k|| \to 0$. This is a contradiction. Therefore $||\bar{g}^k||$ cannot be bounded away from zero.

Lemma 6.2: Assume

(1) the sequence x^k converges to a limit point x^*,

(2) the Hessian $G(x)$ of $f(x)$ exists and is continuous in a neighborhood N_0 of x^*, $G(x^*)$ is positive-definite, then there exists an integer $k_5 > 0$ and positive constants m, M, C_{18} such that for all $k \geq k_5$

(i) $m\ ||y||^2 \leq y^T\ G(x^k)\ y \leq M\ ||y||^2$ for $y \in R^n$,

(ii) $m\ ||x^k - x^*|| \leq ||g(x^k)|| \leq M\ ||x^k - x^*||$,

(iii) $m/2\ ||x^k - x^*||^2 \leq f(x) - x(x^*) \leq M/2\ ||x^k - x^*||^2$,

(iv) $\dfrac{||g(x^k)||}{||s^k||} \geq C_{18}$ if $x^{k+1} \neq x^k$.

Proof: Since $G(x^*)$ is positive definite and $G(x)$ is continuous in a neighborhood N_0 of x^*, we can find another neighborhood N_1 of x^* such that for all $x \in N_1$, $G(x)$ is positive definite. Let $M \geq ||G(x)||$ for all $x \in N_1$ and $\bar{m}$ be a lower bound for the eigenvalues of $G(x)$ for all $x \in N_1$, then we have

$$\bar{m}\ ||y||^2 \leq ||y^T\ G(x)\ y|| \leq M\ ||y||^2 \text{ for all } x \in N_1.$$

Let $\bar{m}$ be an upper bound for $||G(x)^{-1}||$ for all $x \in N_1$ and $\varepsilon > 0$ be so small that for all $||x - x^*|| < \varepsilon$ we have

$$||G(x) - G(x^*)|| < \frac{1}{2\bar{m}}$$

Since x^k converges to x^* there exists an integer $k_5 > 0$ such that for all $k \geq k_5$, $x^k \in N_1$ and $||x^k - x^*|| < \varepsilon$. It follows from a well-known result that

$$||g(x)|| \leq \sup_{t \in [0,1]} ||G(tx + (1 - t)\, x^*)(x - x^*)||$$

and

$$||g(x) - g(x^*) - G(y)(x - x^*)|| \leq \sup_{t \in [0,1]} ||G(tx + (1 - t)\, x^*) - G(y)||\ ||x - x^*||$$

where $y = rx + (1 - r)\, x^*$ for any $r \in [0,1]$. Therefore, we have

$$||g(x^k)|| \leq M\, ||x^k - x^*|| \text{ for } k \geq k_5$$

and $$||g(x^k)|| \geq \left(\frac{1}{||G(y)^{-1}||} - \sup_{t \in [0,1]} ||G(tx^k + (1 - t)\, x^k - G(y)||\right) ||x^k - x^*||,$$ because

$$||tx^k + (1 - t)\, x^* - (rx^k + (1 - r)\, x^*)|| < \varepsilon$$

for all $t \in [0,1]$, we have

$$\sup_{t \in [0,1]} ||G(tx^k + (1 - t)\, x^*) - G(y)|| < \frac{1}{2\bar{m}}$$

Hence,

$$||g(x^k)|| \geq \frac{1}{2\bar{m}}\, ||x^k - x^*||$$

for all $k \geq k_5$. Choose $m = \min(\frac{1}{2\bar{m}}, \hat{m})$, then we have proved (i) and (ii). From the identity

$$f(x^k) - f(x^*) = \int_0^1 (1 - \theta)(x^k - x^*)^T G(x^* + \theta(x^k - x^*))$$
$$(x^k - x^*)\, d\theta$$

and inequality (i), we deduce the bound (iii). Finally, by applying (6.3b) and (iii), we obtain

$$\begin{aligned} ||s^k|| = ||x^{k+1} - x^k|| &\leq ||x^{k+1} - x^*|| + ||x^k - x^*|| \\ &\leq ||x^k - x^*|| + (2[f(x^{k+1}) - f(x^*)]/m)^{\frac{1}{2}} \\ &\leq ||x^k - x^*|| + (2[f(x^k) - f(x^*)]/m)^{\frac{1}{2}} \\ &\leq (1 + \sqrt{M/m})\, ||x^k - x^*|| \end{aligned} \tag{6.21}$$

provided $x^k \neq x^{k+1}$

Thus it follows from (ii) that $||s^k|| \leq (1 + \sqrt{M/m}) \frac{||g^k||}{m}$

Let $C_{18} = \frac{m}{1 + \sqrt{M/m}}$, then $\frac{||g^k||}{||s^k||} \geq C_{18}$ for all $k \geq k_5$ and $x^{k+1} \neq x^k$.

<u>Lemma 6.3</u>: Let the assumptions of Lemma 6.2 hold. Then the sum $\Sigma||s^k||$ is convergent.

<u>Proof</u>: Because $(f(x^k) - f(x^*))$ is a monotonically decreasing sequence, we obtain

$$\sum_{k=1}^{m} [(f(x^k) - f(x^*)) - (f(x^{k+1}) - f(x^*))]/\sqrt{f(x^k) - f(x^*)}$$

$$\leq 2 \sum_{k=1}^{m} (\sqrt{f(x^k) - f(x^*)} - \sqrt{f(x^{k+1}) - f(x^*)})$$

$$< 2(\sqrt{f(x^1) - f(x^*)} - \sqrt{f(x^{m+1}) - f(x^*)} < 2\sqrt{f(x^1) - f(x^*)}$$

Hence, the sum

$$\sum_{k=1}^{\infty} [f(x^k) - f(x^{k+1})]/\sqrt{f(x^k) - f(x^*)} \tag{6.22}$$

is convergent. Suppose $K' = \{k : k \geq k_5$ and x^{k+1} is defined by (6.3b)$\}$. Then by applying Lemma 6.2(ii), (6.21), the Remark 1 on page 5 and the fact the step sizes $\{h^k\}_{k \in K'}$ satisfy (2.2) and (2.3), we get

$$\begin{aligned} ||\bar{g}(x^k, h^k)|| &\leq ||g(x^k)|| + M\,||h^k|| \\ &\leq M\,||x^k - x^*|| + M\,(||h^k|| - ||h^{k+1}|| + ||h^{k+1}||) \\ &\leq M\,||x^k - x^*|| + M\,(C_1 + C_2)\,||s^k||^2 \\ &\leq (M + M(C_1 + C_2)(1 + \sqrt{M/m})\,||s^k||)\,||x^k - x^*|| \\ &< \hat{M}\,||x^k - x^*|| \end{aligned} \tag{6.23}$$

where $\hat{M} > 0$. Choose ε_1 so small that

$$\hat{m} = m - M(C_1 + C_2)(1 + \sqrt{M/m})\,\varepsilon_1 > 0.$$

Then there exists $k_6 \geq k_5 > 0$ such that for all $k \geq k_6$ $||x^k - x^*|| < \varepsilon$. Thus again with Lemma 6.2(ii) and (6.21) we have

$$||\bar{g}(x^k, h^k)|| \geq ||g(x^k)|| - M\, ||h^k|| \geq m\, ||x^k - x^*||$$

$$- M(||h^k|| - ||h^{k+1}|| + ||h^{k+1}||)$$

$$\geq m\, ||x^k - x^*|| - M\,(C_1 + C_2)\, ||s^k||^2$$

$$\geq (m - M(C_1 + C_2)(1 + \sqrt{M/m})^2\, ||x^k - x^*||)\, ||x^k - x^*||$$

Then $||g(x^k, h^k)|| \geq \hat{m}\, ||x^k - x^*||$ (6.24)

for $k \geq k_6$ and $k \in K'$. It follows from (6.24) and Lemma 6.2(iii) that

$$\sqrt{f(x^k) - f(x^*)} \leq \sqrt{M/2\hat{m}^2}\ \ ||\bar{g}(x^k, h^k)||$$

for $k \geq k_6$ and $k \in K'$. Therefore, the expression (6.22) implies that

$$\sum_{k \epsilon k'} \frac{f(x^k) - f(x^{k+1})}{||\bar{g}^k||}$$

is convergent. Since K' includes all those k's included in Σ' with equation (6.6) we obtain

$$\Sigma' \min[||s^k||,\ ||\bar{g}^k||/||B_k||] < +\infty\ .$$

Thus by applying (6.13), we deduce the sum

$$\Sigma' \frac{||s^k||\ ||\bar{g}^k||}{||\bar{g}^k|| + ||s^k||\ ||B_k||} \tag{6.25}$$

is also convergent. Remember that it followed form (6.21) that

$$||s^k|| \le (1 + \sqrt{M/m}\,) \frac{||\bar{g}^k||}{\hat{m}}$$

Hence, there is a constant $C_{14} = \dfrac{1 + \sqrt{M/m}}{\hat{m}} > 0$ such that

$$\frac{||s^k||}{||\bar{g}^k||} \le C_{14}$$

for all $k \ge k_6$ and $k \in K'$. So, (6.12) and (6.25) show that

$$\infty > \Sigma' \frac{s^k}{1 + C_{14}\, ||B_k||} \ge \Sigma' \frac{||s^k||}{1 + C_8 + C_9 \sum\limits_{i=1}^{k}{}' \, ||s^i||}$$

By the fact that if $\sum\limits_{k=1}^{\infty} \dfrac{a_k}{\sum\limits_{i=1}^{k} a_i}$ is finite then $\Sigma\, a_k$ is also finite, we have $\Sigma' \, ||s^k||$ is finite. Therefore, because of inequality (6.11) the proof is completed.

<u>Theorem 6.4</u>: Let the hypotheses of Lemma 6.2 hold and in addition assume

(4) the matrices B_k satisfy the following condition

$$\lim_{k\to\infty} \frac{||(B_k - G(x^*))\, s^k + \bar{g}(x^k, h^k) - g(x^k)||}{||s^k||} = 0 \qquad (6.26)$$

Then $\{x^k\}$ converges Q-superlinearly to x^*.

<u>Proof</u>: The assumption (4) implies that there exists $k_7 \ge k_6$ such that for all $k \ge k_7$

$$||(B_k - G(x^*))\, s^k + \bar{g}(x^k, h^k) - g(x^k)|| < \tfrac{1}{2} m\, ||s^k|| \tag{6.27}$$

where m is the constant in Lemma 6.2. Therefore, by (6.1) (6.27) and Lemma 6.2(i), we obtain

$$\begin{aligned} 0 < f(x^k) - \phi(s^k) &= -s^{k^T} \bar{g}^k - \tfrac{1}{2} s^{k^T} B_k\, s^k \\ &\le - s^{k^T} \bar{g}^k - \tfrac{1}{2} s^{k^T} [G(x^*)\, s^k - (\bar{g}^k - g^k)] + \tfrac{m}{4} ||s^k||^2 \\ &\le - s^{k^T} \bar{g}^k - \tfrac{1}{2} s^{k^T} G(x^*)\, s^k + \tfrac{1}{2} s^{k^T} [\bar{g}^k - g^k] + \tfrac{m}{4} ||s^k||^2 \\ &\le - s^{k^T} \bar{g}^k - \tfrac{1}{4} m\, ||s^k||^2 + \tfrac{1}{2} ||s^k||\; ||\bar{g}^k - g^k|| \end{aligned} \tag{6.28}$$

If $x^{k+1} \neq x^k$, then it follows from the Remark 1 on page 5 and h^k satisfy (2.2)(2.3) that

$$\begin{aligned} ||\bar{g}^k - g^k|| &\le M\, ||h^k|| \le M\, (||h^k|| - ||h^{k+1}|| + ||h^{k+1}||) \\ &\le M\, (C_1 + C_2)\, ||s^k||^2 \end{aligned} \tag{6.29}$$

However, since x^k is convergent to x^*, if $x^{k+1} = x^k$ is defined by (6.3a), we can always find a x^{k+i}, $i > 1$ such that $x^{k+i} \neq x^k$ and $x^{k+i-j} = x^k$ for all $1 \le j \le i - 1$. In this case we have $h^k = h^{k+1} = \ldots, = h^{k+i-1}$ and

$$||h^{k+i-1}|| - ||h^{k+i}|| \le C_2\, ||s^{k+i-1}||^2$$

$$||h^{k+i}|| \le C_1\, ||s^{k+i-1}||^2$$

Because $x^k = x^{k+i-j}$, $1 \le j \le i - 1$, the inequality (6.8) must be failed for x^{k+i-j} with $1 \le j \le i - 1$. Therefore, by (6.10) we obtain

$$||s^{k+i-1}|| \le \Delta^{k+i-1} \le C_7 \, ||s^{k+i-2}|| \le C_7 \, \Delta^{k+i-2} \le \ldots$$
$$\le C_7^{\,i-1} \, ||s^k||$$

Since $C_7 < 1$, the above inequalities imply that

$$||s^{k+i-1}|| \le ||s^k||$$

which gives

$$||h^k|| - ||h^{k+i}|| \le C_2 \, ||s^k||^2$$
$$||h^{k+i}|| \le C_1 \, ||s^k||^2$$

Hence, (6.29) can be proved even if s^k is not used to define x^{k+1}, i.e., $x^k = x^{k+1}$ is defined by (6.3a). Therefore, (6.28) gives the following inequality

$$0 < -s^{k^T} \bar{g}^k - \tfrac{1}{4} m \, ||s^k||^2 + \tfrac{1}{2} M(C_1 + C_2) \, ||s^k||^3.$$

If we further choose $k_8 \ge k_7$ so large that for all $k \ge k_1$

$$M \, (C_1 + C_2) \, ||s^k|| \le m/8,$$

then we have

$$0 < -s^{k^T} \bar{g}^k - \tfrac{1}{4} m \, ||s^k||^2 + \tfrac{1}{8} m \, ||s^k||^2 < -s^{k^T} \bar{g}^k - \tfrac{m}{8} \, ||s^k||^2$$

for all $k \ge k_8$, which gives the inequality

$$||\bar{g}^k|| \geq \frac{1}{8} m \; ||s^k|| \qquad k \geq k_8$$

Therefore, from (6.6) we obtain

$$f(x^k) - \phi(s^k) \geq C_3 \min[\frac{1}{8} m \; ||s^k||^2, \frac{(m \; ||s^k||)^2}{64 \; ||B_k||}].$$

It follows from (6.12) and Theorem 6.3 that there exists $C_{15} \geq 0$ such that $||B_k|| \leq C_{15}$ for $k \geq k_4$. Hence, there exists a positive constants C_{16} such that

$$f(x^k) - \phi(s^k) \geq C_{16} \; ||s^k||^2 \tag{6.30}$$

for $k \geq k_8$.

From Taylor's Theorem we can deduce

$$f(x^k + s^k) - f(x^k) = g^{k^T} s^k + \int_0^1 s^{k^T} G(x^k + ts^k) \; s^k \; (1 - t) \; dt.$$

Since $\int_0^1 s^{k^T} G(x^*) \; s^k \; (1 - t) \; dt = \frac{1}{2} s^{k^T} G(x^*) \; s^k$ this provides

$$f(x^k + s^k) - f(x^k) = g^{k^T} s^k + \int_0^1 s^k [G(x^k + ts^k) - G(x^*)] \; s^k \; (1 - t) \; dt + \frac{1}{2} s^{k^T} G(x^*) \; s^k$$

From (6.1) we have

$$f(x^k) - \phi(s^k) = -\bar{g}^{k^T} s^k - \frac{1}{2} s^{k^T} B_k \; s^k$$

Let us assume that $f(x^k + s^k) - \phi(s^k)$ is positive. Add the above two identities, we obtain

$$f(x^k + s^k) - \phi(s^k) = (g^k - \bar{g}^k)^T s^k + \tfrac{1}{2} s^{k^T} [G(x^*) - B_k] s^k$$

$$+ \int_0^1 s^{k^T} [G(x^k + ts^k) - G(x^*)] s^k (1 - t)\, dt$$

$$\le \tfrac{1}{2} ||g^k - \bar{g}^k|| \, ||s^k|| + \tfrac{1}{2} ||(B_k - G(x^*)) s^k$$

$$+ \bar{g}^k - g^k|| \, ||s^k||$$

$$+ \tfrac{1}{2} ||s^k|| \max_{t \in [0,1]} ||G(x^k + ts^k) - G(x^*)||$$

By (6.29),

$$0 < \frac{f(x^k + s^k) - \phi(s^k)}{||s^k||^2} \le M(C_1 + C_2) \, ||s^k||$$

$$+ \tfrac{1}{2} \frac{||(B_k - G(x^*)) s^k + \bar{g}^k - g^k||}{||s^k||}$$

$$+ \tfrac{1}{2} \max_{t \in [0,1]} ||G(x^k + ts^k) - G(x^*)||$$

As $||s^k|| \to 0$, $x^k \to x^*$ so $x^k + ts^k \to x^*$, by the continuity of $G(x)$,

$$\frac{f(x^k + s^k) - \phi(s^k)}{||s^k||^2} \to 0 \qquad (6.31)$$

Since

$$\frac{f(x^k) - \phi(s^k)}{f(x^k) - f(x^k + s^k)} = \frac{f(x^k) - f(x^k + s^k) + f(x^k + s^k) - \phi(s^k)}{f(x^k) - f(x^k + s^k)}$$

$$= 1 + \frac{(f(x^k + s^k) - \phi(s^k))/||s^k||^2}{\dfrac{f(x^k) - \phi(s^k)}{||s^k||^2} - \dfrac{f(x^k + s^k) - \phi(s^k)}{||s^k||^2}}$$

with (6.30) and (6.31) we have

$$\frac{f(x^k) - \phi(s^k)}{f(x^k) - f(x^k + s^k)} \to 1 \qquad \text{as } k \to +\infty$$

Thus there exists an $k_9 \geq k_8$ such that for all $k \geq k_9$ inequality (6.8) is satisfied if $f(x^k + s^k) > \phi(s^k)$. However, if $f(x^k + s^k) \leq \phi(s^k)$, (6.8) is by all means true. So for all $k \geq k_9$, (6.8) is satisfied, and the conditions

$$\begin{cases} \Delta^{k+1} \geq ||s^k|| \\ x^{k+1} = x^k + s^k \end{cases} \qquad \text{for all } k \geq k_9$$

hold. Therefore, if an iteration gives the reduction $||s^{k+1}|| \leq ||s^k||$, then the rule governing the definition of s^{k+1} implies that $s^{k+1} = -B_{k+1}{}^{-1}\,\bar{g}^{k+1}$. Since $||s^k||$ tends to zero as k tend to infinity, it follows that the Newton's formula is applied an infinite many times. Suppose, $k_{10} \geq k_9$ be so large that $s^{k_{10}}$ is defined by Newton's formula, and such that for $k \geq k_{10}$

$$||G(x^*)\, s^k + g(x^k) - g(x^k + s^k)|| < \frac{C_{18}\, ||s^k||}{2}$$

and

$$||(B_k - G(x^*))\, s^k + \bar{g}^k - g^k|| \leq \frac{C_{18}}{2}\, ||s^k||$$

Add the above two inequalities and we have for all $k \geq k_{10}$

$$||B_k\, s^k + \bar{g}^k - g(x^k + s^k)|| < C_{18}\, ||s^k|| \tag{6.32}$$

Let $k \geq k_{10}$ be an integer such that $s^k = -B_k{}^{-1}\,\bar{g}^k$ then $x^{k+1} = x^k + s^k$ since $k \geq k_9$. Furthermore, (6.32) implies that

$$||g(x^{k+1})|| < C_{18}\ ||s^k||$$

By Lemma 6.2(iv), we obtain

$$\frac{||g^{k+1}||}{||s^{k+1}||} > C_{18} > \frac{||g^{k+1}||}{||s^k||}$$

Hence $||s^{k+1}|| \leq ||s^k||$. It follows by induction that $s^k = -B_k^{-1}\ \bar{g}^k$ for all $k \geq k_{10}$. Hence apply Theorem 3.3 to the subsequence $\{x^k\}_{k \geq k_{10}}$ we have x^k converges to x^* Q-superlinearly.

The above Theorems prove all the convergence properties of the class of algorithms we discussed in the beginning of this section. We will now prove that Algorithm MD belongs to this class and hence also has these properties. First we need the following theorem.

Theorem 6.5: Let $\eta \in R^n$ be the value of s that minimizes $\phi(s)$ subject to the inequality

$$||\eta^k|| \leq ||s^k|| \qquad (6.33)$$

and subject to the condition that has the form

$$\eta^k = -\alpha\ \bar{g}^k \qquad (6.34)$$

Then the bound

$$f(x^k) - \phi(\eta^k) \geq \tfrac{1}{2} \, ||\bar{g}^k|| \, \min[||s^k||, \frac{||\bar{g}^k||}{||B_k||}] \tag{6.35}$$

is obtained

<u>Proof</u>: By (6.1) and (6.34) we have

$$f(x^k) - \phi(\eta^k) = \alpha \, ||\bar{g}^k||^2 - \tfrac{1}{2} \alpha^2 \, \bar{g}^{k^T} B_k \, \bar{g}^k$$

Therefore, if $\bar{g}^{k^T} B_k \bar{g}^k \leq 0$, then the required vector η^k is obtained when α has the value $\alpha = \frac{s^k}{\bar{g}^k}$ in which case

$$f(x^k) - \phi(\eta^k) \geq ||s^k|| \, ||\bar{g}^k||$$

which is consistent with the bound (6.35). However, if $\bar{g}^{k^T} B_k \bar{g}^k \geq 0$, then α is the number which will make the derivate of $(f(x^k) - \phi(\eta^k))$ w. r. t. α become 0, if $||\eta^k|| \leq ||s^k||$ still holds. Hence, $\alpha = \min(\frac{||s^k||}{||\bar{g}^k||}, ||\bar{g}^k||^2/\bar{g}^{k^T} B_k \bar{g}^k)$. Thus,

$$\alpha \, \bar{g}^{k^T} B_k \, \bar{g}^k \leq ||\bar{g}^k||^2 \tag{6.36}$$

and so

$$\alpha \geq \min[||s^k||/||\bar{g}^k||, 1/||B_k||] \tag{6.37}$$

Thus, in this case, it follows from (6.36) and (6.37) that

$$\begin{aligned} f(x^k) - \phi(\eta^k) &\geq \tfrac{1}{2} \alpha \, ||\bar{g}^k||^2 \\ &\geq \tfrac{1}{2} \, ||\bar{g}^k|| \, \min[||s^k||, ||\bar{g}^k||/||B_k||] \end{aligned}$$

The theorem is proved.

This theorem shows that requiring condition (6.6) is equivalent to requiring the difference $f(x^k) - \phi(s^k)$ to be no less than a positive constant multiple of the greatest value of the difference $f(x^k) - \phi(\eta^k)$ that can be obtained when η^k is subject to (6.33) and (6.34). Algorithm MD defines s^k in three ways. If s^k is defined either by Newton's formula or as a gradient step, (6.6) is clearly satisfied. If s^k is defined by the combination of Newton's formula and the gradient step such that $\phi(x^k + s^k)$ will have the least value, then $\phi(s^k) \leq \phi(\eta^k)$ where η^k is defined by (6.33), (6.34). Thus (6.6) is satisfied by this s^k too. Since Dennis [8] has proved that under the assumption $||G(x) - G(y)|| \leq L\ ||x - y||$ the update of the Powell symmetric Broyden's method satisfies (6.7), by applying his proof replacing $g(x)$ by $\overline{g}(x, h)$ we can prove that the matrices B_k of Algorithm MD satisfy (6.7). Now, we are ready for the following theorem which is a global convergence theorem for Algorithm MD.

Theorem 6.6: Suppose $f(x)$ is bounded below and twice differentiable and $g(x)$ is uniformly continuous and there exists $L > 0$ such that

$$||G(x) - G(y)|| \leq L\ ||x - y||$$

for all x, y in a convex hull of the level ste $L(x^1)$, where x^1 is the initial point. Then the vector sequences $\overline{g}(x^k, h^k)$ where $\{x^k\}$ is generated by Algorithm MD are not bounded away from the zero vector.

Proof: Since the s^k's of the ordinary iterations satisfy (6.6) and the B_k's satisfy (6.7), this proof is concluded if we can prove that the special iterations do not affect the proof of Theorem 6.1. Let us redefine Σ' to denote the sum over the ordinary iterations for which (6.8) is satisfied. Because in a special iteration, $||s^k|| \leq \Delta^k$ and $\Delta^{k+1} = \Delta^k$, the inequality (6.11) will still hold for the new definition of Σ'. Hence the rest of the proof of Theorem 6.1 applies unchanged to this theorem.

The previous theorem provides the limit

$$\lim\inf ||g^k|| = 0$$

Therefore, if one of the points x^k falls into a region where $f(x)$ is locally convex, if $f(x)$ has a local minimum in this region and if the step bounds Δ^k and the inequality $f(x^{k+1}) \leq f(x^k)$ keep the latter points of the sequence x^k $(k = 1, 2, \ldots)$ in the region then convergence is obtained to the local minimum. This is the following convergence theorem:

Theorem 6.7: Let the hypotheses of Theorem 6.6 hold, and assume

(3) $f(x)$ is strictly convex in a closed neighborhood S of the local minimum x^*,

(4) there exists an integer $\sigma > 0$ such that for all $k \geq \sigma$,

the iterate points x^k generated by Algorithm MD all lie in S. Then $\{x^k\}$ converges to x^*.

<u>Proof</u>: Let $\rho_1 = \inf ||x^k - x^*||$, $k \geq \sigma$. If $\rho_1 > 0$, then we define $\rho_2 > 0$ so large that $||x - x^*|| < \rho_2$ for all $x \in S$. Hence for all $k \geq \sigma$, $\rho_1 \leq ||x^k - x^*|| \leq \rho_2$. Set $\bar{S} = \{x : \rho_1 \leq ||x - x^*|| \leq \rho_2\}$ and $\bar{f} = \min_{x \in \bar{S}} f(x)$. Since f is strictly convex on S and $\rho_1 > 0$, we have $\bar{f} > f(x^*)$ and

$$\begin{aligned} f(x^*) &\geq f(x^k) + (x^k - x^*)^T g^k \\ &\geq f(x^k) - ||x^k - x^*|| \; ||g^k|| \\ &\geq \bar{f} - \rho_2 \; ||g^k|| \end{aligned}$$

Hence, we deduce the bound

$$||g^k|| \geq (\bar{f} - f(x^*))/\rho_2 \qquad\qquad k \geq \sigma_1$$

By (6.11) we know $||s^k|| \to 0$ as k tends to infinity. Since $||h^k|| \leq C_1 \; ||s^{k-1}||^2$ there exists $\sigma_1 \geq \sigma \geq 0$ such that for all $k \geq \sigma_1$

$$||h^k|| \leq \frac{\bar{f} - f(x^*)}{2 \rho_2 M} \qquad .$$

Therefore, with Remark 1 on page 5, we obtain

$$||\bar{g}(x^k, h^k)|| \geq ||g(x^k)|| - M \; ||h^k|| \geq \frac{\bar{f} - f(x^*)}{2 \rho_2} > 0$$

Then we have a contridiction to Theorem 6.6. That implies

$\inf ||x^k - x^*|| = 0$ for $k \geq \sigma$. Because $f(x)$ is continuous and $f(x^k)$ decreases monotonically, we deduce the limit

$$\lim_{k\to\infty} f(x^k) = f(x^*) \tag{6.38}$$

Now we want to prove that for any $\varepsilon > 0$ there exists $\sigma_3(\varepsilon)$ such that for all $k \geq \sigma_3(\varepsilon)$, $||x^k - x^*|| < \varepsilon$. Let us define $\hat{f}(\varepsilon) = \min\limits_{\substack{||x-x^*||>\varepsilon \\ x\varepsilon S}} f(x)$, ε is any positive number, then $\hat{f}(\varepsilon) > f(x^*)$ and (6.38) implies that there exists $\sigma_3(\varepsilon) > 0$ such that for all $k \geq \sigma_3(\varepsilon)$, $f(x^k) < \hat{f}(\varepsilon)$. Hence for all $k \geq \sigma_3(\varepsilon)$, $||x^k - x^*|| < \varepsilon$ since $x^k \varepsilon S$. This concludes the proof.

<u>Theorem 6.8</u>: If the hypotheses of Theorem 6.6 hold and assume the sequence $\{x^k\}$ generated by Algorithm MD converges to x^*. Then $\{x^k\}$ is two-step Q-superlinearly convergent to x^*.

<u>Proof</u>: Let us consider the ordinary iterations first. Since B_k is just the Powell symmetric Broyden update, if we let $M = I$, (4.8) can be easily verified. Following the proofs of Lemma 4.3 and Theorem 4.5 we can get (3.11) is satisfied by the B_k's of Algorithm MD. Then we can apply Theorem 6.4 to the ordinary iterations to have that x^k's generated only by ordinary iterations converge Q-superlinearly. The proof is completed if we can show the superlinear convergence is not damaged by the special iterations. It follows from Lemma 6.2(iii) that since $\{f(x^k) - f(x^*)\}$ is monotonically decreasing

$$||x^{k+1} - x^*|| \le \sqrt{M/m}\ ||x^k - x^*||$$

Because every special iteration is followed by two ordinary iterations, if kth iteration is a special iteration, (k-1)th, (k-2)th, (k+1)th, (k+2)th are ordinary iterations. Hence, the ratio

$$0 \le \frac{||x^k - x^*||}{||x^{k-2} - x^*||} \le \sqrt{M/m}\ \frac{||x^{k-1} - x^*||}{||x^{k-2} - x^*||}$$

tends to zero when k tends to infinity. And

$$0 \le \frac{||x^{k+2} - x^*||}{||x^k - x^*||} \le \frac{||x^{k+2} - x^*||}{\sqrt{m/M}\ ||x^{k+1} - x^*||}$$

tends to zero when k tends to infinity. Therefore, $\{x^k\}$ generated by Algorithm MD is two-step Q-superlinearly convergent.

From the above theorems we know the special iterations do not affect any convergence property of the algorithm. In fact it can be proved that the special iterations contribute in the convergence of the B_k to $G(x^*)$. However, Theorem 6.4 and Theorem 6.8 prove that it is not necessary to have B_k converge to $G(x^*)$. As a matter of fact the Algorithm MD converges Q-superlinearly without the special iterations in Algorithm MD. But sometimes it is desirable if we could know some properties of $G(x^*)$, in this case we do need B^k to converge to $G(x^*)$. For example, in implementing our algorithm MD, we use Stewart's scheme to start out step-sizes $\{h^k\}$, it does need the condition B^k converges to $G(x^*)$. Furthermore, numerical results show that Powell's algorithm has the tendency of converging faster with the special iterations. Therefore, we still include the special iterations in our Algorithm MD.

APPENDIX A

A SURVEY OF UNCONSTRAINED MINIMIZATION METHODS WITHOUT DERIVATIVES

There are many derivative-free algorithms for solving the unconstrained minimization problems. Kowalik and Osborne [16] divided these algorithms into two classes called 'Direct search methods' and 'Descent methods'.

The direct search methods are based on a sequential examination of trial solutions which by simple comparisons give an indication for a further searching procedure. Most of the earlier algorithms for the unconstrained minimization problems belong to this class. The earliest algorithm is a classical method of optimization which adjust each variable separately. Specifically, the initial estimate $(x_1, x_2, \ldots, x_n)$ is altered to $(x_1 + \lambda_1, x_2, \ldots, x_n)$, and so on until the nth stage of the process replaces $(x_1 + \lambda_1, x_2 + \lambda_2, \ldots, x_{n-1} + \lambda_{n-1}, x_n)$ by $(x_1 + \lambda_1, \ldots, x_n + \lambda_n)$. These n stages are repeated, usually until the corrections $(\lambda_1, \ldots, \lambda_n)$ becomes very small. This process can always reduce the value of the objective function unless all the gradient components are zero. The convergence properties of this method are like those of the steepest descent algorithm, in particular the rate of convergence is usually too slow to use.

The introduction of computer as well as stimulating the extension to classical methods led to many direct search methods. Rosenbrock (1960) [26] provided an algorithm which is an extension of the above classical methods. The basis of Rosenbrock's method is that it tries to identify the direction of the fastest direction of the function in order to use it as a search direction. Initially the variables are changed one at a time, as in the classical methods, so on the first iteration the initial estimate $(x_1, x_2, \ldots, x_n)$ is changed to $(x_1, \ldots, x_n) + \lambda_1 d_1$, this estimate is then changed to $(x_1, \ldots, x_n) + \lambda_1 d_1 + \lambda_2 d_2$ and so on, until the complete iteration replaces the initial guess of the solution by $(x_1, \ldots, x_n) + \sum_{i=1}^{n} \lambda_i d_i$ where $\{d_i\}_{i=1, \ldots, n}$ is the initial set of mutually orthogonal search directions. Before we start a new iteration, the set of n search directions is changed, and the first search direction is replaced by $d_1^* = \sum_{i=1}^{n} \lambda_i d_i$ which is the total step achieved by the previous stage. The remaining new search directions are obtained by an orthogonalization over the set $\{ \sum_{i=1}^{n-k} \lambda_i d_i \}$ $k=0, 1, \ldots, n-1$ and then the iterative process is repeated. It is essential that the first new search direction is defined in this way. Because it could be expected that this direction is oriented toward the direction of the fastest decrease of the function, if this is the case then the rate of convergence becomes faster. A further development of this idea is given by Davies, Swann and Campey [30] which is in essence an application of linear minimization to the Rosenbrock method. Hooke and Jeeves (1961) [15] published another direct search method which is composed of two parts: pattern move and exploratory move. The algorithm applies the pattern move to change the current estimate x^k of the solution by the total change made in the last iteration (except on the first iteration there is no pattern move). From the resultant point y an exploratory move is made, which in fact is a fine adjustment of the values of the variables. Specifically, small steps are taken along each of the coordinate directions in order to decrease the objective function. Let the resultant point be z. If f(z) is smaller than $f(x^k)$ then z becomes the starting point for the next iteration, otherwise the iteration is treated as a failure, and an exploratory move is made from $(x_1^k, \ldots, x_n^k)$. A special feature in Hooke and Jeeves method is it permits f(y) to be larger than f(x). Since the exploratory move following the pattern move can always pull the estimate point back to the minimum point, this method is particular suitable for objective functions that have curve ridges. The simplex method is another valuable addition to the derivative-free optimization algorithms made by Spendley, Hext and Himsworth (1962) [28]. The algorithm starts with a given simplex in the space of the variable, then an iterative procedure is applied to move the given simplex to a position that is near to the optimal point. Each iteration moves the given simplex by reflecting it in one of the faces, so the position of only one vertex is changed by an iteration. The vertex is changed is usually the one at which the value of the objective function is least, but there are exceptions to this rule to prevent permature cycling. Once the simplex is close to the optimal point, it can be contracted, in order to determine the required vector of variables more precisely. The special property of this method is that it depends only on the comparison of function values at the vertices of a simplex. Nelder and Mead (1965) [17] then suggested a more flexible approach in which the simplex can be altered both in size and in geometry so that it moves toward an optimum and be ultimately shrinks at this point. However, Box (1966) [2] gave some experiment results which indicate Nelder and Mead's method is a good algorithm for optimizing an exact function only if the number of variables is small.

Generally speaking, direct search methods can be applied to functions that are not very smooth because of its action is depending only on comparisons of function values. However, with the same reason it does not take full advantage of this smoothness whenever the objective function has a continuous derivative. Hence, their final rate of convergence is always very disappointing if a very accurate approximation to the minimum is needed. The descent methods were developed to remedy this difficiency.

The descent methods are a class of methods in which the solution of the general optimization problem is found by solving a sequence of one-dimensional problems. Powell (1964) [20] constructed a conjugate direction method. It requires an estimate x^k of the position of the optimum and n linear independent search directions, say $d_1^k, \ldots, d_n^k$ which initially are chosen to be the coordinate vectors. Each iteration can change both the estimate and the search direction. It begins by searching along each directions in turn, then x^k is changed by the amount $s = \sum_{i=1}^{n} \lambda_i^k d_i^k$ where λ_i^k is calculated to minimize the objective function $f(x^k + \sum_{i=1}^{n} \lambda_i d_i^k)$ then the resultant point x^{k+1} is the new starting estimate for the next iteration and the new search directions are $d_i^{k+1} = d_{i+1}^k$, for $i = 1, 2, \ldots, n - 1$ and $d_n^{k+1} = s$. And the set $\{d^{k+1}\}_{i=1, \ldots, n}$ has the property that the last k of them are mutually conjugate unless linear dependence occurs. Fletcher (1965) [11] has compared this method with the method of Davies, Swann and Campey. In terms of number of function evaluations, this method emerges as definitely superior. Regrettably, Powell's method has a tendency of generating directions which are linear dependent. Because of this tendency Powell modified his procedure by updating the search directions only if it is predicted that the updating will improve the linear independence of the search directions. An unfortunate consequence of this decision is that for large value of n, the search directions are generally changed less frequently, which gives slower convergence, so it is seldom practical to use this method to minimize functions of more than twenty variables. Zangwill (1967) [31] developed a promising extension to Powell's method. It adds frequent searches along coordinate directions into the algorithm, for then linear dependent cannot occur, and therefore the test that sometimes prevents the updating of the search directions can be removed. No numerical results are available to tell if it is worthwhile. However, theoretical convergence is established for a strictly convex continuously differentiable function. And Daniel [7] proved its convergence for a uniformly quasi-convex, strictly psedo-convex, continuously differentiable function. Powell (1972) [23] gave another idea of revising the search directions to avoid the possibility of linear dependence. He recommended the set of search directions $\{d_i^k\}$ be revised by first construct $\bar{d}_i = \sum_{j=1}^{n} \Omega_{ij} d_j$ $i=1, . . ., n$. Then let the new search directions be $\{d_i^*\}$, where $d_i^* = \dfrac{\bar{d}_i}{(\bar{d}_i, G\bar{d}_i)}$ $i = 1, . . ., n$, where Ω is any orthogonal matrix and G is the Hessian of the objective function. It was proved that the conjugacy properties of the new directions are at least as good as those of the old directions. But good methods for choosing the orthogonal matrix Ω where there are three or more variables are less obvious. More research can be done on this subject to prove some useful algorithms.

Another descent method which is based on searching along some fixed directions is called the lattice approximation algorithm due to Berman (1969) [1]. The algorithm seeks the least value of f(x) subject to the constraint that x is to have the form $x = x_0 + \lambda h$ where x_0 is fixed, h is the fixed size of the lattice, and where the components of λ are forced to be integers. By refining the lattice, (i.e., dividing h by integer) sequence of points are obtained by Berman that the convergance rate is fairly uniform but there is no other proof or details about the rate of convergence. And unfortunately Powell [21] pointed out that the amount of computation required by the algorithm increases exponentially with the number of variables, so it is not suitable for large problems. There is another kind of descent method which fits quadratic functions to calculated values of f(x). The first algorithm of this kind is proposed by Fiaco and McCormick [10]. It begins by searching for the least value of f(x) along each of the coordinate direction in turn, and thus the coefficients of the quadratic $\phi(x) = c + (b, x) + \frac{1}{2}(x, Gx)$ and the diagonal elements of G are obtained. Then another $\frac{1}{2} n(n - 1)$ searches are made along $e_i + e_j$, $i \neq j$ lead to an estimate of the elements G_{ij}, $i \neq j$. If G is positive definite, then $-G^{-1}b$ is the minimum of $\phi(x)$. Therefore $-G^{-1}b$ is used to be an estimate of the vector of variables that minimizes f(x). But they did not mention what should we do if G is not positive definite. Apparently this method needs to be modified. Another similar work is given by Winfiel (1967)[31] which does not require any linear searches. Instead it uses $\frac{1}{2}(n + 1)(n + 2)$ values of the objective function, and it forces $\phi(x) = f(x)$ for each of these function values. Therefore the calculation of the coefficients of $\phi(x)$ is equivalent to the solution of a system of $\frac{1}{2}(n + 1)(n + 2)$ linear equations. Then solve the quadratic programming problem min $\phi(x)$, $||x - x_0||_\infty < \alpha$ where $f(x_0)$ is the smallest value of all the values of $f(x)$ that has been calculated so far and where α is a parameter. Let ξ be the solution, then include $f(\xi)$ in the list of $\frac{1}{2}(n + 1)(n + 2)$ function values, in order to repeat the process. The value of α is also adjusted by each iteration. And α is the factor that makes this algorithm work without G necessary positive definite. Winfield notes two deficienties of this algorithm: it takes a large amount of work on each iteration and the matrix of the linear equations may happen to be singular.

All the methods we considered above do not make any apparent approximations to the gradient of f. Since there are a lot of successful unconstrained optimization algorithms which use gradient values, it would be useful if we cound find a proper way to estimate the gradient directly so that all these successful algorithms could become derivative-free. We have not mentioned any work that deals with the question, "Can one construct better algorithm that do not require the computation of derivatives by using function values to approximate derivatives by differences?" Stewart [29] (1967) is the first one who partly answer this question by giving a modification of Davidon-Fletcher-Powell algorithm which uses difference approximation to estimate first derivatives of the objective function. A typical estimate is the expression

$$\frac{\partial f(x)}{\partial x_i} \approx \frac{f(x + h_i\, e_i) - f(x)}{h_i} \tag{1}$$

where e_i is the vector whose ith component is one and whose other components are zero. The h_i is the step-size, and it is important to choose its value well because too small a value gives excessive loss of accuracy due to cancellation and too large a value will generate a large approximation error. Therefore, Stewart's paper concentrated in finding a proper way to choose step-size. The basis is try to choose the step-size that balances trunction error against cancellation error. A technique is used that depends on an estimate of the curvature of f(x) along the direction e_i. By this means it is sometimes predicated that formula (1) is not sufficiently accurate for any h_i within reason, then he replaces (1) by the central difference approximation

$$\frac{\partial f(x)}{\partial x_i} \approx \frac{f(x + h_i\, e_i) - f(x - h_i\, e_i)}{h_i} \quad . \tag{2}$$

Although tests run by Stewart (1967) [29] and Fletcher (1969) [11] have shown the algorithm to be superior to all previously known methods using only values, there is no theoretical justification for the method. Since in Stewart's algorithm the truncation error is estimated by employing the DFP approximation to the second derivatives, if the approximation G does not converge to $\nabla^2 f(x^*)$ then a poor h_i may be obtained. Gill and Murray [11] suggested another way to define the step-size. In their algorithm the value h is specified initially and is the same for every iteration. They found out in practice that if using a machine with a mantissa of t binary digits, $h = 2^{-t/2}$ gives a reasonable balance between truncation error and cancellation error. They observed that their algorithm needed less function evaluations then Stewart's method when applied to Wood's function. However, Gill and Murray's algorithm is far more sophisticated in other ways and is based on recurring the factorization of an approximation to the Hessian. They and Stewart did not apply the finite difference estimate scheme on exactly the same algorithm, without theoretical proofs it is hard to tell which way of choosing h will give a better result. Cullum (1972) [6] noticed that if the gradient approximations are inaccurate, Stewart's scheme may converge very slowly or even terminate prematurely. To avoid this difficiency, she developed two algorithms: modified rank 1 algorithm (MR1) and modified Davidon-Fletcher-Powell algorithm [12]. They are just a combination of modified Stewart scheme and Zangwill's searching along the coordinate line periodically technique. The insertion of the periodic cyclic linear miniminzation along the coordinate line prevented the algorithm from terminating prematurely and it contributes in the proof that both algorithms are convergent if f is strictly convex and continuously differentiable. Since Stewart's algorithm is improved by searching along the coordinate line periodically not by choosing a more reasonable step-size, the step-size does not affect the convergence. It may be a good idea to modify the step-size so that it can make some contribution in the convergence of the algorithm or the rate of convergence.

Recently (1974), Miffin [17] established a new derivative-free hybrid method which combines the Polak's quasi-Newton method [25] and Chernous'kos' [5] local variation method. The algorithm uses (2) to estimate the first derivatives. The step-size h is restricted by the condition $||\bar{g}|| > \alpha||h||$ where $\bar{g}$ is the estimation of the gradient and α is a constant. Unlike the method we mentioned above this one does not require any exact one variable minimization. The algorithm is supported by a strong theoretical result which is that the algorithm converge Q-superlinearly for a twice continuously differentiable strongly convex function. However, the detailed computational results have not published yet, this algorithm requires order n^2 function evaluations and order n^3 arithmetic operations per iteration due to the fact that the second partial derivatives are approximated by finite differences and an $n \times n$ linear system is solved at each iteration. There are a lot of well-known quasi-Newton methods which require only order n^2 operations at each iteration. The question now is "Is there a way of defining the step-size which we could apply to the quasi-Newton methods and still retain the superlinearly convergence property?" In this paper, we find out a positive answer to this question. And the theoretical proofs show that the modified quasi-Newton methods still converge superlinearly. They formed a new class of derivative-free algorithms for unconstrained optimization.

APPENDIX B

A FORTRAN SUBROUTINE FOR CALCULATING THE GRADIENT

The Parameters of the Subroutine

The name of the subroutine and its parameter are:

```
SUBROUTINE GRAD(N,X,FUN,G,GG,IT,NF,ERF,EPSI)
```

X,G,GG are one-dimensional arrays, whose length must not be less than N. FUN,ERF,EPSI are real variables. N,IT,NF are integer variables. The user must give values to N,X(I),GG(I) (I = 1, . . ., N), ERF, EPSI and FUN. The subroutine calculates the value of G(I) (I = 1, . . ., N).

N is the number of variables of the objective function, and it must be in the range 2 ≤ N ≤ an upper bound. The upper bound is imposed by the dimension of the arrays private to the subroutine, so it is easy to change.

X(I) (i = 1, . . ., N) is the argument vector for which the user wants to find the gradient.

FUN is the value of the objective function corresponding to the vector X(I) (I = 1, . . ., N).

GG(I) (I = 1, . . ., N) is an array containing values of the diagonal elements of the approximate Hessian of the objective function evaluated at X(I) (I = 1, . . ., N).

IT indicates how many times the subroutine has been called.

NF indicates the number of function evaluations we have made up to this point.

ERF is the relative error of the function evaluation.

EPSI is the machine epsilon of the machine this subroutine applied to,

The user must provide a function subroutine with the name

```
FUNCTION U(X)
```

subroutine GRAD calls U(X) everytime it wants to evaluate the objective function at a certain point x.

Initial Gradient of the Algorithm

This subroutine calculated the gradient by the forward difference approximation:

$$g_j = \frac{f(x + \delta_j e_j) - f(x)}{\delta_j}$$

where e_j is the vector with its jth component unity and its other component zero. The main duty of this subroutine is to get a proper step-size δ_j such that the error between the estimated gradient and the real gradient will not destroy the convergence of the algorithm to which this subroutine is applied. However, for the initial gradient there is not enough information to calculate a proper step-size. We use $10^{-6} \times |x_j|$ for $j = 1, \ldots, N$ which we think is a proper step-size. The user of course could supply his own appropriate step-size for the calculation of the initial gradient.

The subroutine tests of this is the call for the initial gradient in Line 29. If it is the case, it goes down to Line 30, 31 and 32 which set the step-size to $\delta_i = 10^{-6} \times |x_i|$ $i = 1, \ldots, N$. But if $x_i = 0$, δ_i is then set to 10^{-6} (in double precision). Line 33 to Line 35 calculate the norm of the step-size which will be useful for the next iteration. Then it goes down to Line 104. From Line 104 to Line 125, the subroutine calculates the initial gradient gradient by the simple difference formula (1) and store the value of X, δ and $||\delta||$ for the next iteration. This wraps up the operations of finding the initial gradient. If Line 29 implies that this is not a call for the initial gradient, it goes down to Line 37. From there it begins the general process to find the gradient approximation.

General Operations to Calculate an Approximated Gradient

Let us use the superscript k to denote the number of the iteration. The main purpose of the operation is to get a good step-size δ^k (the step-size for the kth iteration) for all $k \geq 1$. We have proved that the following two conditions:

$$\text{(i)} \quad ||\delta^k|| \leq C_1 \, ||x^k - x^{k-1}||^2 \qquad \text{for } 0 < C_1 < \infty \qquad (2)$$

$$\text{(ii)} \quad ||\delta^k|| \leq C_2 \, ||x^k - x^{k-1}||^2 \qquad \text{for } 0 < C_2 < \infty \qquad (3)$$

define a class of prospective step-sizes. However, we need one which is hopefully the best in the class for each iteration. Since the step-size calculated by Stewart's scheme [29] should work very well if the information it requires is accurate, it will naturally be our best choice provided it satisfies (2) and (3). The procedure goes as follows:

First we will compute a δ^k by Stewart's scheme which is based on balancing the truncation error against cancallation error. The scheme includes three steps:

1. Set $\eta = \max\{\eta_\phi, \frac{|g^k|}{|f^k|} \; |x_i{}^k| \; \eta_m\}$ where η_ϕ is an estimate of the relative error in the function evaluation. This value is a parameter of the subroutine and is supplied by the user. The value of η_m is the relative error due to cancellation of significant figures on the current machine. The number η will be taken as a bound for the error caused by cancellation when the value of f is calculated. The value of η is computed in Line 43 and 44 and put in the working space WK.

2. By balancing the cancellation error and truncation error, δ^k can be deduced from the following equations:

if $(g_i{}^{k-1})^2 \geq |f^k| \; |G_{ii}{}^{k-1}|\eta$, $\delta_p = \delta_i(1 - \frac{|G_{ii}{}^{k-1}| \; \delta_i}{3|G_{ii}{}^{k-1}| \; \delta_i + 4|g_i{}^k|})$

where $\delta_i' = 2(\frac{|f^k|\eta}{|G_{ii}{}^{k-1}|})$

otherwise,

$$\delta_p = \delta_i' \left(1 - \frac{2|g_i^{k-1}|}{3|G_{ii}^{k-1}|\ \delta_i' + |g_i^k|}\right)$$

where

$$\delta_i' = 2 \left(\frac{|f^k|\ |g_i^k|}{|G_{ii}^{k-1}|}\right)$$

Let $\delta_i^k = \text{sign}\ (G_{ii}^{k-1},\ g_i^{k-1})\ \delta_p$. In this way, the block of instructions from Line 45 to Line 55 define the step-size δ^k and then it is stored in the array DELTA.

3. If the δ^k we got from step 2 makes the value $\frac{|G_{ii}^{k-1}\ k|}{2|g_i^{k-1}|}$ greater than some prescribed upper bound, say 10^{-m}, the derivative will not be sufficiently accurate. So Stewart suggests that in this case we let δ^k be the positive solution of

$$\tfrac{1}{2}\ |G_{ii}^{k-1}|\ \delta^2 + |g_i^{k-1}|\ \delta - 10^m\ |f^k|\ \eta = 0 \qquad (4)$$

And use the central difference formula instead of the s mple difference:

$$g_i = \frac{f(x + \delta_i e_i) - f(x - \delta_i e_i)}{2\ \delta_i} \qquad (5)$$

to compute the gradient approximation. Hence, Line 56 to

Line 58 check if δ^k should be recomputed. If this is the case, Line 59 and 60 give a solution to equation (4) and reset identifier ID to indicate later that it requires the central difference equation (5).

Thus after Line 62 is reached, we have a δ^k which is supposed to balance the effect of truncation error and cancellation error. However, whole Stewart's theory depends on the assumptions that we have a pretty good approximation to the second derivative at the current point and that the estimate of the error bound on the calculation of the function value is good. Actually, we estimate the truncation error by employing the approximation to the second derivative which may not be very close to the real Hessian. As to η_ϕ, it is really not easy to get a good estimation. Since it involves specific properties of the x^k's the algorithm generated as well as f(x), if we try to calculate it carefully, we may get into more trouble than to compute the gradient directly. Inaccuracy of these two quantities may lead us to a poor δ^k. However, as long as we have some condition to test whether the Stewart's δ^k is good enough or not, it is worthwhile to compute it. Condition (2) and (3) will determine the applicability of Stewart's δ^k. In the subroutine, we first check $||\delta^k|| \le ||\delta^{k-1}||$ in Line 75. If this is not true,we set $\delta^k = \delta^{k-1}$. Because it is hard to tell what values constant C_1, C_2 should be, we use the second iteration to help determining starting values of C_1 and C_2. That is, in the second iteration we suppose δ^k

satisfies (2) and (3). Hence, set

$$C_1 = \frac{||\delta^k||}{||x^2 - x^1||^2} \text{ and } C_2 = \frac{||\delta^1|| - ||\delta^2||}{||x^2 - x^1||^2}$$

which are done in Line 82,83. Moreover, it can be proved that the condition (2) and (3) will be more reasonable if we require $C_1 < C_2$. So Line 85 is set for this purpose. However, we will always make some adjustment to C_1, C_2 latter whenever it is appropriate. We start applying condition (2) and (3) from the third iteration. Line 88 checks if Stewart's δ^k satisfies (2). If it does, we go down to Line 91 which is for condition (3). Otherwise, we discard Stewart's δ^k and reset $\delta^k = C_1 \, ||x^k - x^{k-1}||^2$ in Line 89. Since a δ^k which is less than machine epsilon would cause $\bar{g}(x^k, \delta^k) = 0$ even if x^k is very far from the solution point. So Line 91 resets $\delta^k = \delta^{k-1}$ whenever it finds out $||\delta^k||$ is smaller than machine epsilon. In fact any value that lies between machine epsilon and $||\delta^{k-1}||$ can be used as $||\delta^k||$, but we choose $\delta^k = \delta^{k-1}$ in order to reduce the truncation error in y^k and f^k. Condition (2) is used to prevent δ^k being too large, but if the case δ^k less than machine epsilon happens, this implies that condition (2) is over restricting the step-size. Hence, to properly enlarge C_1 is necessary. Line 94 adjusts C_1 by $\frac{||\delta^k||}{||x^k - x^{k-1}||^2}$ with $\delta^k = \delta^{k-1}$. Condition (3) is checked by Line 98. If δ^k satisfies condition (3), then we are set to compute the gradient. If even this δ^k does not satisfy (3), it means all the step-size in the class defined by condition (2) will not satisfy (3). Hence we will have to either enlarge C_1 or enlarge C_2 properly. A larger C_2 may only cause the problem that a too small δ^k pass the condition (3). But since we have set up a statement that will not allow δ^k to be smaller than the machine epsilon, this problem will not happen. It is better to adjust C_2 instead of C_1. Line 103 enlarges C_2 to $\frac{||\delta^{k-1}|| - ||\delta^k||}{||x^k - x^{k-1}||^2}$. Hence, after Line 103 we get our final step-size δ^k. Then the gradient will be computed by the statement in the block from Line 107 to Line 122. Whether the subroutine will use the forward difference or the central difference formula is depending on the array ID whose value was set during the process of finding δ^k. Finally we store the value of x^k, δ^k and $||\delta^k||$ in Line 123, 124, 125 for the sake of the next call. This gives an end to the subroutine.

Numerical Results

We coded the subroutine GRAD and applied it to Powell's MINFA. The new algorithm generated in this way is our Algorithm MD. It has been tested in FORTRAN on Cornell University's IBM360/65 on four functions. The numerical results on these four functions and a table which has the comparisons with Powell's MINFA and Stewart's DFP are listed.

Four testing functions are:

(1) Rosenbrock function (Table 1)

$$f(x) = 100\,(x_x - x_1^2)^2 + (x_1 - 1)^2$$

starting point = (1.2, 1.0)

(2) Four variable function (Table 2)

$$f(x) = x_1^2 + 2x_2^3 + 4x_4^2 + (x_1 + x_2 + x_3 + x_4)^4$$

starting point = (1, -1, -1, 1)

(3) Wood-Coville function (Table 3)

$$f(x) = 100\,(x_2 - x_1^2)^2 + (1 - x_1)^2 + 90(x_3 - x_3^2)^2 + (1 - x_3)^2$$
$$+ 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8\,(x_2 - 1)(x_4 - 1)$$

starting point = (-3, -1, -3, -1)

(4) Powell's function (Table 4)

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4$$
$$+ 10(x_1 - x_4)^4$$

starting point = (3, -1, 0, 1)

```
      SUBROUTINE GRAD(X,N,FUN,G,GG,IT,NF,EPSI,ERF)
C     FUN: THE FUNCTION VALUE
C     G: GRADIENT OF THE FUNCTION
C     GG; DIAGONAL ELEMENTS OF THE APPROXIMATE HESSIAN MATRIX
C     IT: THE NUMBER OF ITERATIONS
C     NF: THE NUMBER OF THE FUNCTION EVALUATIONS
C     OLDG: THE GRADIENT OF THE PREVIOUS ITERATION
C     DELTA: THE STEP-SIZE FOR THE FINITE DIFFERENCE
C     OLDEL: THE STEP-SIZE OF THE PREVIOUS ITERATION
C     STEPN: THE NORM OF THE STEP-SIZE
C     OLDX; ITERATE POINT OF THE PREVIOUS ITERATION
C     OLSIN: THE STEPN OF THE PREVIOUS ITERATION
C     DELX: THE SQURE OF THE NORM OF THE DIFFERENCE OF THE TWO
C     SUCCESSIVE POINTS
C     ERF: AN ESTIMATE OF THE RELATIVE ERROR IN THE FUNCTION EVALUATION
C     EPSI: MACHINE EPSILON
C     XD: THE DIFFERNECE OF TWO SUCCESSIVE POINTS
C     FD: THE DIFFERENCE OF THE FUNCTION VALUES OF TWO SUCCESSIVE POINTS
C     ID:AN ARRAY FOR IDENTIFYING PURPOSES
      IMPLICIT REAL*8 (A-H,O-Z), INTEGER (I-N)
      DIMENSION X(N),G(N),GG(N),DELTA(9),XD(9),OLDG(9),OLDEL(9),OLDX(9)
     1,ID(10)
C     IF THE CORRESPONDING COMPONENT OF THE INITIAL POINT IS 0, LET
C     THE INITIAL STEP-SIZE BE 0,000001 FOR DOUBLE PRECISION,
C     0,001 FOR SINGAL PRECISION
      DATA DELTA,ID/9*1.0D-6,10*1/
      M=N+1
C     INITIAL GRADIENT
      IF ( IT-1) 11,11,21
   11 DO 12 I=1,N
      ID(I)=1
   12 IF ( X(I) .NE. 0.0) DELTA(I)=DABS(X(I))*0,000001
      STEPN= DELTA(1)
      DO 14 J=2,N
   14 STEPN= DMAX1(DELTA(J), STEPN)
      GO TO 41
C     GRADIENT FOR THE ORDINARY ITERATIONS
C     FIRST GET THE STEP-SIZE BY STEWART'S SCHEME
   21 DO 31 I=1,N
      ID(I)=1
      IF ( G(I) ,EQ, 0,0D0) GO TO 31
      OLDG(I)=G(I)
      WK= DABS(OLDG(I)*X(I)*EPSI/FUN)
      WK= DMAX1( ERF, WK)
      WL=OLDG(I)*OLDG(I)
      WR=WK*DABS(GG(I)*FUN)
      IF ( WL- WR) 22,22,23
   22 WR=2,0*(DABS(FUN*OLDG(I))*WK/(GG(I)*GG(I)))**0,33333333333333
      WR= WR*(1.0-2.0*DABS(OLDG(I))/(3.0*DABS(GG(I))*WR+4.0
     1*DABS(OLDG(I))))
      GO TO 24
   23 WR= 2.0*DSQRT(DABS(FUN/GG(I))*WK)
      WR=WR *(1,0 -DABS(GG(I))*WR/(3,0*DABS(GG(I))*WR+4,0
     1*DABS(OLDG(I))))
   24 DELTA(I)=DSIGN( WR, OLDG(I)*GG(I))
```

```
         WL     =0.5*DABS(GG(I)*DELTA(I)/OLDG(I))
         WR=1.D0-2
         IF (WL-WR) 31,25,25
   25    ID(I)=2
         DELTA(I)=(-DABS(OLDG(I))+DSQRT(DABS(OLDG(I))**2.0+200.0*
        1DABS(GG(I)*FUN)*WK))/DABS(GG(I))
   31    CONTINUE
C        CHECK TWO CONDITIONS:
C        (1) ||STEP-SIZE||<= CONS1*||X - OLDX||**2
C        (2) ||OLD STEP-SIZE|| - ||STEP-SIZE||<=CONS2*||X-OLDX||**2
C        FIRST GET ||X - OLDX||**2 AND NORM OF THE STEP-SIZE
         WK=DABS(X(1)-OLDX(1))
         DO 32 J=2,N
   32    WK=DMAX1(DABS(X(J)-OLDX(J)),WK)
         DELX= WK* WK
         IF ( DELX .EQ. 0.0) GO TO 45
         STEPN= DABS(DELTA(1))
         DO 33 J=2,N
   33    STEPN = DMAX1(STEPN, DABS(DELTA(J)))
         IF ( STEPN - OLSTN) 36,36,35
   35    STEPN=OLSTN
         DO 351 I=1,N
         ID(I)=1
  351    DELTA(I)=OLDEL(I)
   36    IF( IT .NE. 2) GO TO 37
C        USE THE 2ND ITERATION TO GET THE INITIAL GUESS OF CONS1, CONS2
         CONS1=STEPN/DELX
         CONS2=(OLSTN - STEPN)/DELX
C        WE WANT CONS1<<CONS2
         IF ( CONS2 .EQ. 0.0 .OR. CONS2 .LE. CONS1) CONS2=CONS1*10.0
         GO TO 41
C        FOR THE ITERATIONS OTHER THAN THE 1ST AND 2ND ITERATION
   37     IF ( STEPN - CONS1*DELX) 39,39,38
   38    STEPN=CONS1*DELX
         ID(M)=2
   39    IF ( STEPN .GE. EPSI) GO TO 391
         STEPN=OLSTN
         ID(M)=1
         CONS1=STEPN/DELX
         DO 381 I=1,N
         ID(I)=1
  381    DELTA(I)=OLDEL(I)
  391    IF (( OLSTN - STEPN) - CONS2*DELX) 41,41,40
   40    IF (( OLSTN - CONS1*DELX) - CONS2*DELX) 46,46,47
   46    ID(M)=2
         STEPN=CONS1*DELX
         GO TO 41
   47    CONS2 = (OLSTN -STEPN )/DELX
   41    DO 42 J=1,N
         IF ( IT .EQ. 1 .OR. IT .EQ. 2) GO TO 42
         IF ( ID(M) .EQ. 2) DELTA(J)=STEPN
   42    XD(J)=X(J)
         DO 44 I=1,N
         IF ( IT .NE. 1 .AND. G(I) .EQ. 0.0) GO TO 44
         IF ( I .NE. 1) XD(I-1)=X(I-1)
```

```
      XD(I)=X(I)+DELTA(I)
      FO=U(XD)
      NF=NF+1
C     SIMPLE DIFFERENCE FOR THE GRADIENT
      G(I)=(FO-FUN)/DELTA(I)
C     SAVE THE ITERATE POINT & STEP-SIZE & THE NORM OF THE STEP SIZE
      IF ( ID(I) .EQ. 1) GO TO 43
C     CENTRAL DIFFERENCE FOR THE GRADIENT
      XD(I)=X(I)-DELTA(I)
      FE=U(XD)
      NF=NF+1
      G(I)=(FO-FE)/(2.0*DELTA(I))
   43 OLDEL(I)= DELTA(I)
      OLSTN=STEPN
      OLDX(I)= X(I)
   44 CONTINUE
   45 RETURN
      END
```

REFERENCES

[1] Berman, G. "Lattice approximation to the minima of functions of several variables." Journal of A.C.M. 16 (1969) 286-294.

[2] Box, M.J. "A comparison of several current optimization methods and the use of transformations in constrained problems." Comput. J. 9 (1966) 67-77.

[3] Broyden, C.G. "A class of methods for solving nonlinear simultaneous equations." Math. Comp. 19 (1965) 577-593.

[4] Broyden, C.G., J.E. Dennis, Jr. and J.J. Moré. "On the local and superlinear convergence of quasi-Newton methods." J. Inst. Maths. Applics. (1973) 12, 223-245.

[5] Chernous'ko, F.L. "A local variation method for the numerical solution of variational problems." U.S.S.R. Computational Math. and Math. Phys. 5 (1965) 234-242.

[6] Cullum, J. "Unconstrained minimization of functions without explicit use of their derivatives." IBM Waston Research Center, Yorktown Heights, New York, 1971.

[7] Daniel, J.W. "The approximate minimization of functionals (Prentice-Hall, Englewood Cliffs, New Jersey, 1971), 205-206.

[8] Dennis, J.E., Jr. "On some methods based on Broyden's secant approximation to the Hessian in numerical methods for nonlinear optimization." ed. F.A. Lootsma, Academic Press, 1972.

[9] Dennis, J.E.,and Jorge J. Moré. "A characterization of superlinear convergence and its applications to quasi-Newton methods." Cornell Computer Science Technical Report 73-157.

[10] Fiacco, A.V. and G.P. McCormick. "Nonlinear programming: sequential unconstrained minimization techniques." (John Wiley and Son, Inc., New York, 1968).

[11] Fletcher, R. "A review of methods for unconstrained minimization." in Optimization ed. R. Fletcher (Academic Press, London, 1969). 1-12.

[12] Fletch, R. and M.J.D. Powell. "A rapidly convergent descent method for minimization." Compt. J. 6 (1963) 163-168.

[13] Gill, P.E. and W. Murray. "Quasi-Newton methods for unconstrained optimization." J. Inst. Maths. Applies (1972) 9, 91-108.

[14] Greenstadt, J. "Variations on variable-metric methods." Math. Comp. 24 (1970) 1-18.

[15] Hooke, R. and T.A. Jeeves. "Direct search solution of numerical and statistical problems." J. Assoc. Compt. Mach., 8 (1961) 212-221.

[16] Kowalik, J. and M.R. Osborne. "Methods for unconstrained optimization problems." American Elsevier Publishing Company, Inc., New York, 1968.

[17] Mufflin, R. "A superlinear convergent algorithm for minimization without evaluating derivatives." Yale University Technical Report No. 65, 1974.

[18] Nelder, J.A. and R. Mead. "A simplex method for function minimization." Compt. J. 7 (1965) 308-313.

[19] Pearson, J.D. "Variable metric methods of minimization." Comp. J. 12 (1969) 171-178.

[20] Powell, M.J.D. "An efficient method for finding the minimum of a function of several variables without calculating derivatives," Ibid. 7 (1964) 155-162.

[21] Powell, M.J.D. "A new algorithm for unconstrained optimization." Nonlinear Programming edited by J.B. Rosen, O.L. Manfasarian, K. Ritter, Academic Press, New York, (1970).

[22] Powell, M.J.D. "Recent advances in unconstrained optimization." Mathematical Programming, v. 1, 26-57.

[23] Powell, M.J.D. "Unconstrained minimization algorithms without computation of derivatives." Theoretical Physics Division, U.K.A.E.A. Research Group, Atomic Energy Research Establishment, Harwell, 1972.

[24] Powell, M.J.D. "Convergence properties of a class of minimization algorithms." SIGMAP Nonlinear Programming Symposium, Univ. of Wisconsin, 1974.

[25] Polak, E. "Computational methods in optimization." Academic Press, New York, 1971.

[26] Rosenbrock, H.H. "An automatic method for finding the greatest or the least value of a function." Compt. J. 3 (1960) 175-184.

[27] Shanno, D.F. "Conditioning of quasi-Newton methods for function minimization." Math. Comp. 24 (1970) 647-656.

[28] Spendley, W., G.R. Hext and F.R. Himsworth. "Sequential application of simplex design in optimization and evaluationary operation technometrics," 4 (1962) 441.

[29] Stewart, G.W. "A modification of Davidon's minimization method to accept difference approximation of derivative." J.A.C.M. 14 (1967) 72-83.

[30] Swann, W.H. "Report on the development of a new direct search method of optimization." Research Note 64/3, Central Instrument Laboratory, I.C.I. Ltd. 1964.

[31] Winfield, D.H. "Function minimization without derivation by a sequence of quadratic programming problems." Technical Report No. 537. Engineering and Applied Physics Division, Harvard Univ. (1967).

[32] Zangwill, W.I. "Minimizing a function without calculating derivatives." Compt. J. 10 (1967) 293-296.