AD-A070 204 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/2

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF VARIABLE METRI--ETC(U)

APR 79 K RITTER DAAG29-75-C-0024

UNCLASSIFIED MRC-TSR-1945 NL

| OF |

AD
A070204

END
DATE
FILMED
7-79
DDC

MRC Technical Summary Report # 1945

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF VARIABLE METRIC METHODS

Klaus Ritter

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

April 1979

(Received February 19, 1979)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF VARIABLE METRIC METHODS

Klaus Ritter

Technical Summary Report #1945

April 1979

ABSTRACT

This paper considers a class of variable metric methods for unconstrained minimization. Without requiring exact line searches it is shown that, under appropriate assumptions on the function to be minimized, each algorithm in this class converges globally and superlinearly.

AMS (MOS) Subject Classification: 90C30

Key Words: Unconstrained minimization, variable metric method, global convergence, superlinear convergence.

Work Unit Number 5 - Mathematical Programming and Operations Research.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

SIGNIFICANCE AND EXPLANATION

Many practical problems in operations research may be reduced to minimizing a function with or without contraints. By means of penalty functions and similar techniques a constrained minimization problem can be converted into a sequence of unconstrained minimization problems. In this paper we discuss a class of algorithms for unconstrained minimization problems which converge rapidly to the solution from a starting point which is not necessarily a good approximation to the solution of the given problem.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF VARIABLE METRIC METHODS

Klaus Ritter

1. Introduction

Variable metric methods have been used successfully for iteratively calculating an approximation to the least value of a function $F(x)$ of n variables. A variable metric method simultaneously generates a sequence of points $\{x_j\}$ and a sequence of matrices $\{H_j\}$. During each iteration a correction matrix of rank one or two is added to H_j with the intent to construct an approximation to the inverse Hessian matrix of $F(x)$.

A large class of such methods has been introduced by Huang [9]. This class contains symmetric and unsymmetric matrices H_j. A restriction of the Huang class to update formulas which are of rank two, satisfy the quasi-Newton equation and maintain the symmetry of H_j leads to a class of methods proposed by Broyden [1] and Fletcher [7]. Two well-know members of this class are the Davidon-Fletcher-Powell-method (DFP - method), [4], [6], and the Broyden-Fletcher-Goldfarb-Shanno-method (BFGS - method), [2], [7], [8], [16].

The first general global convergence result is due to Powell [12], [13] who proved that, if $F(x)$ satisfies certain assumptions and if the optimal step size is used, the DFP - method converges superlinearly to a global minimizer of $F(x)$. In [5] Dixon showed that under certain conditions the methods in the Huang class generate the same sequence $\{x_j\}$ if they are started with the same initial x_0, H_0 and if the optimal step size is used. Under the idealized assumption of an optimal step size these two results provide therefore a complete convergence theory. In practice, however, it is in general not possible to use an optimal step size. Therefore, it is important to establish global convergence for a non-optimal step size.

One such result was obtained by Lenard [10] who generalized Powell's convergence proof for the DFP - method. Another result is due to Powell [14] who proved that the BFGS - method converges superlinearly with a step size procedure that eventually results in a step size equal to one.

Using also a non-optimal step size Stoer [17] showed that every method in a subclass of the Broyden class, the so-called restricted Broyden-methods, converges n-step quadratically for

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

every positive definite starting matrix H_0 and every initial value x_0 sufficiently close to a minimizer z of $F(x)$.

If it is assumed that both x_0 and H_0 are sufficiently close to z and the inverse Hessian matrix of $F(x)$ at z, respectively, then it follows from results obtained by Broyden, Dennis and Moré [3] that the DFP - method and the BFGS - method converge superlinearly to z with step size one.

It is the purpose of this paper to show that with an appropriate non-optimal step size every method in the Broyden class converges globally and superlinearly provided $F(x)$ satisfies certain assumptions. In the next section we derive a representation of the matrix H_j as a sum of n matrices of rank 1. This representation allows us to study the dependence of H_{j+1} on the parameters used in the update formula for H_j and leads to a simple proof of Dixon's result. In Section 3 global convergence is established. The proof is based on a generalization of Powell's proof for the BFGS - method. In the final section it is shown that the sequence $\{x_j\}$ converges superlinearly and that the sequences $\{\|H_j\|\}$ and $\{\|H_j^{-1}\|\}$ are bounded.

2. Basic properties of variable metric methods

Let $x \in E^n$ and let $F(x)$ be a real-valued function. If $F(x)$ is twice differentiable at a point x_i, we denote the gradient and the Hessian matrix of $F(x)$ at x_i by $g_i = \nabla F(x_i)$ and $G_i = G(x_i)$, respectively. A prime is used for the transpose of a vector or a matrix. For any $x \in E^n$, $\|x\|$ denotes the Euclidean norm of x.

We consider the problem of determining a vector z such that

$$F(z) \leq F(x) \quad \text{for all} \quad x \in E^n \quad .$$

For later reference we formulate the following assumption.

Assumption 1.

$F(x)$ is a convex function. There exists an x_0 such that the set

$$S_0 = \{x | F(x) \leq F(x_0)\}$$

is bounded, and such that $F(x)$ is twice continuously differentiable on some convex open set containing S_0.

If a variable metric method is used to minimize $F(x)$, then at a given point x_j, a search direction s_j is determined by multiplying the gradient $g_j = \nabla F(x_j)$ by a appropriate matrix H_j, i.e.,

$$s_j = H_j g_j \quad ,$$

where H_j is an approximation to the inverse Hessian matrix of $F(x)$ at x_j. With a suitable step size σ_j a new point

$$x_{j+1} = x_j - \sigma_j s_j$$

is computed. If $g_{j+1} = \nabla F(x_{j+1}) \neq 0$, the matrix H_{j+1} is determined from H_j in such a way that the quasi-Newton equation is satisfied, i.e.,

(2.1) $$H'_{j+1}(g_j - g_{j+1}) = \sigma_j s_j \quad .$$

The various variable metric methods differ in the update procedure which is used to compute H_{j+1} from H_j. In many methods H_{j+1} is obtained by adding one or two matrices of rank one to H_j. A large class of such methods has been studied by Huang [9] and Dixon [5]. With

$$d_j = \frac{g_j - g_{j+1}}{\|\sigma_j s_j\|} \quad \text{and} \quad p_j = \frac{s_j}{\|s_j\|}$$

their update formula can be written as follows:

$$H'_{j+1} = H'_j + \rho\,\frac{p_j(\alpha_1 p'_j + \alpha_2 d'_j H_j)}{(\alpha_1 p'_j + \alpha_2 d'_j H_j)d_j} - \frac{H'_j d_j(\beta_1 p'_j + \beta_2 d'_j H_j)}{(\beta_1 p'_j + \beta_2 d'_j H_j)d_j}\,, \tag{2.2}$$

where ρ, α_1, α_2, β_1 and β_2 are parameters such that $\alpha_1^2 + \alpha_2^2 > 0$ and $\beta_1^2 + \beta_2^2 > 0$ and it is assumed that the denominators are not zero.

The equation (2.1) is satisfied if and only if $\rho = 1$. Therefore, we shall always assume that $\rho = 1$. Under suitable assumptions the inverse Hessian matrix of $F(x)$ is symmetric. Since H_j is intended to be an approximation to this matrix it is reasonable to restrict the parameters in such a way that H_j is symmetric for all j. With $\rho = 1$ we obtain from (2.2)

$$\begin{aligned} H'_{j+1} = H'_j &+ \frac{\alpha_1}{(\alpha_1 p'_j + \alpha_2 d'_j H_j)d_j}\,p_j p'_j + \frac{\alpha_2}{(\alpha_1 p'_j + \alpha_2 d'_j H_j)d_j}\,p_j d'_j H_j \\ &- \frac{\beta_1}{(\beta_1 p'_j + \beta_2 d'_j H_j)d_j}\,H'_j d_j p'_j - \frac{\beta_2}{(\beta_1 p'_j + \beta_2 d'_j H_j)d_j}\,H'_j d_j d'_j H_j\,. \end{aligned} \tag{2.3}$$

Thus, if H_j is symmetric then H_{j+1} is symmetric if and only if

$$\alpha_2(\beta_1 p'_j d_j + \beta_2 d'_j H_j d_j) = -\beta_1(\alpha_1 p'_j d_j + \alpha_2 d'_j H_j d_j)\,. \tag{2.4}$$

Assuming $\beta_1 \neq 0$ we can solve (2.4) for α_1. This gives

$$\alpha_1 = -\alpha_2\,\frac{\beta_1 p'_j d_j + (\beta_1 + \beta_2) d'_j H_j d_j}{\beta_1 p'_j d_j}\,.$$

Therefore,

$$\alpha_1 p'_j d_j + \alpha_2 d'_j H_j d_j = -\frac{\alpha_2}{\beta_1}(\beta_1 p'_j d_j + \beta_2 d'_j H_j d_j)$$

and

$$\frac{\alpha_1}{\alpha_1 p'_j d_j + \alpha_2 d'_j H_j d_j} = \frac{\beta_1(p'_j d_j + d'_j H_j d_j) + \beta_2 d'_j H_j d_j}{p'_j d_j(\beta_1 p'_j d_j + \beta_2 d'_j H_j d_j)}\,.$$

Substitution into (2.3) gives then the update formula for symmetric matrices H_j.

(2.5)
$$H_{j+1} = H_j + \frac{\beta_1(p_j'd_j+d_j'H_jd_j)+\beta_2d_j'H_jd_j}{p_j'd_j(\beta_1p_j'd_j+\beta_2d_j'H_jd_j)}\,p_jp_j'$$
$$- \beta_1\,\frac{p_jd_j'H_j+H_jd_jp_j'}{\beta_1p_j'd_j+\beta_2d_j'H_jd_j} - \beta_2\,\frac{H_jd_jd_j'H_j}{\beta_1p_j'd_j+\beta_2d_j'H_jd_j}\;.$$

The update formula (2.5) represents the subclass of the Huang class of update formulas with the property that all matrices H_j are symmetric and satisfy the quasi-Newton equations. This subclass is identical with a class of update formulas obtained by Broyden [1] and in different form by Fletcher [7].

First we consider three special case. If we choose $\beta_1 = 0$ then (2.4) implies $\beta_2 = 0$ and (2.3) reduces to

$$H_{j+1} = H_j + \frac{p_jp_j'}{d_j'p_j} - \frac{H_jd_jd_j'H_j}{d_j'H_jd_j}\;.$$

This is the update formula used in the Davidon-Fletcher-Powell - method [4], [6]. With $\beta_1 = 1$ and $\beta_2 = 0$ we obtain from (2.5)

(2.6)
$$H_{j+1} = H_j + \frac{p_j'd_j+d_j'H_jd_j}{(p_j'd_j)^2}\,p_jp_j' - \frac{p_jd_j'H_j+H_jd_jp_j'}{p_j'd_j}\;,$$

i.e., the update formula of the Broyden-Fletcher-Goldfarb-Shanno - method [2], [7], [8], [16]. Finally if we choose $\beta_1 = 1$ and $\beta_2 = -1$, then (2.5) becomes

(2.7)
$$H_{j+1} = H_j + \frac{p_jp_j'-p_jd_j'H_j-H_jd_jp_j'+H_jd_jd_j'H_j}{p_j'd_j-d_j'H_jd_j}$$
$$= H_j + \frac{(p_j-H_jd_j)(p_j'-d_j'H_j)}{(p_j'-d_j'H_j)\,d_j}\;.$$

This is a symmetric rank one update formula. Because the vectors $p_j - H_jd_j$ and d_j can become (nearly) orthogonal it is, however, known to be unstable and not recommended for use.

Returning to the general formula (2.5) we assume that H_j is positive definite. Because

$$p_j = \frac{H_j q_j}{\|H_j q_j\|} \quad \text{and} \quad H_j d_j = \frac{H_j q_j - H_j q_{j+1}}{\|\rho_j s_j\|}$$

we observe that with

$$T_j = \{x \mid (H_j q_j)'x = (H_j q_{j+1})'x = 0\}$$

we have

$$H_{j+1}x = H_j x \quad \text{for} \quad x \in T_j \quad . \tag{2.8}$$

Since H_j is positive definite, $q_j \notin T_j$ and $q_{j+1} \notin T_j$. Hence using (2.8) we can determine H_{j+1} completely by defining it on

$$S_j = \mathrm{span}\{q_j, q_{j+1}\} \quad .$$

For this purpose we write H_j as a sume of three matrices. Setting

$$p_j = \frac{H_j q_j}{\|H_j q_j\|} \quad , \quad \rho_j = \frac{1}{\|H_j q_j\|}$$

and choosing $w_j \in S_j$ such that $w_j' p_j = 0$ and $q_j = H_j w_j$ has norm one we have

$$H_j = \frac{p_j p_j'}{\rho_j q_j' p_j} + \frac{q_j q_j'}{w_j' q_j} + \hat{H}_j \quad , \tag{2.9}$$

where $\hat{H}_j$ is a symmetric matrix of rank $n - 2$ with

$$\hat{H}_j q_j = \hat{H}_j w_j = 0$$

and

$$\hat{H}_j x = H_j x \quad \text{for} \quad x \in T_j \quad .$$

Note that $\hat{H}_j$ can be written in the form

$$\hat{H}_j = \sum_{i=3}^{n} \frac{p_{ij} p_{ij}'}{d_{ij}' p_{ij}} \tag{2.10}$$

where $d_{3j}, \ldots, d_{nj}$ are vectors in T_j such that

$$d_{ij}' H_j d_{kj} = 0 \qquad i,k = 3,\ldots,n, \qquad i \neq k$$

and

$$H_j d_{ij} = p_{ij} \quad \text{with} \quad \|p_{ij}\| = 1,\ i = 3,\ldots,n \quad .$$

Let H_{j+1} be determined by (2.5). In order to define H_{j+1} on S_j we observe that d_j and w_j are in S_j and that the two vectors are linearly independent because $d_j'p_j \neq 0$ and $w_j'p_j = 0$. Since H_{j+1} satisfies the quasi-Newton equation we have

$$H_{j+1}d_j = p_j \quad . \tag{2.11}$$

Furthermore by (2.5),

$$H_{j+1}w_j = q_j - \frac{\beta_1 p_j d_j' q_j + \beta_2 H_j d_j d_j' q_j}{\beta_1 d_j' p_j + \beta_2 d_j' H_j d_j} \quad . \tag{2.12}$$

Thus $H_{j+1}w_j \in \text{span}\{q_j, p_j\}$. Since

$$d_j' H_{j+1} w_j = p_j' w_j = 0 \tag{2.13}$$

it follows that, for every choice of the parameters β_1 and β_2, $H_{j+1}w_j$ is a vector in $\text{span}\{q_j, p_j\}$ which is orthogonal to d_j.

Let u_j be a vector such that

$$u_j \in \text{span}\{q_j, p_j\},\ \|u_j\| = 1,\ d_j'u_j = 0,\ w_j u_j > 0 \quad .$$

Since $d_j'p_j \neq 0$ and $w_j p_j = 0$ it follows that u_j exists and is uniquely determined. Therefore, using (2.12) and (2.13) we have

$$H_{j+1}w_j = \omega_j u_j \tag{2.14}$$

where ω_j is a number that depends on the particular values of the parameters β_1 and β_2 used to determine H_{j+1}. Combining (2.8), (2.11), and (2.14) we see that

$$H_{j+1} = \frac{p_j p_j'}{d_j' p_j} + \omega_j \frac{u_j u_j'}{w_j' u_j} + \hat{H}_j \quad . \tag{2.15}$$

Thus all matrices H_{j+1} defined by (2.5) are of the form (2.15) and differ only in the factor ω_j. Furthermore, if H_j is positive definite and if $d_j'p_j > 0$, then H_{j+1} is positive definite if and only if $\omega_j > 0$.

In order to study the dependence of ω_j on the parameters β_1 and β_2 more closely we first determine ω_j for the BFGS - method. From (2.6) we obtain

$$H_{j+1}w_j = H_jw_j - \frac{p_jd_j'H_jw_j}{d_j'p_j} = q_j - \frac{d_j'q_j}{d_j'p_j}p_j$$

$$= q_j + \alpha_jp_j \quad ,$$

where $\alpha_j = - d_jq_j/d_j'p_j$. Thus

$$u_j = \frac{q_j+\alpha_jp_j}{\|q_j+\alpha_jp_j\|},\ \omega_j = \|q_j + \alpha_jp_j\| \quad . \tag{2.16}$$

Observing that by (2.9)

$$H_jd_j = p_j\frac{d_j'p_j}{\rho_jq_j'p_j} + q_j\frac{d_j'q_j}{w_j'q_j} \quad ,$$

$$d_j'H_jd_j = \frac{(d_j'p_j)^2}{\rho_jq_j'p_j} + \frac{(d_j'q_j)^2}{w_j'q_j} \tag{2.17}$$

we have for the general update formula (2.5)

$$H_{j+1}w_j = q_j - \frac{\beta_1p_jd_j'q_j+\beta_2H_jd_jd_j'q_j}{\beta_1d_j'p_j+\beta_2d_j'H_jd_j}$$

$$= (\beta_1d_j'p_j+\beta_2d_j'H_jd_j)^{-1}[(\beta_1d_j'p_j+\beta_2d_j'H_jd_j-\beta_2\frac{(d_j'q_j)^2}{w_j'q_j})q_j$$

$$- (\beta_1d_j'q_j+\beta_2\frac{d_j'p_jd_j'q_j}{\rho_jq_j'p_j})p_j]$$

$$= (q_j - \frac{d_j'q_j}{d_j'p_j}p_j)\frac{\beta_1d_j'p_j+\beta_2\frac{(d_j'p_j)^2}{\rho_jq_j'p_j}}{\beta_1d_j'p_j+\beta_2d_j'H_jd_j} \quad .$$

Thus

$$\omega_j = \gamma_j\|q_j+\alpha_jp_j\| \quad , \tag{2.18}$$

where

$$\gamma_j = \frac{\beta_1 d_j' p_j + \beta_2 \dfrac{(d_j' p_j)^2}{\rho_j g_j' p_j}}{\beta_1 d_j' p_j + \beta_2 d_j' H_j d_j} \tag{2.19}$$

and $\gamma_j = 1$ if $\beta_2 = 0$, i.e., for the BFGS - method. For the DFP - method we have $\beta_1 = 0$ and

$$\gamma_j = \frac{(d_j' p_j)^2}{\rho_j g_j' p_j d_j' H_j d_j} \quad . \tag{2.20}$$

Assuming that $d_j' p_j > 0$ and H_j is positive definite we see that the subset of the updating formulas (2.5) with

$$\beta_1 \beta_2 \geq 0, \quad \beta_1 + \beta_2 \neq 0$$

preserves the positive definiteness of H_j. More generally we have the following result.

Lemma 1

Let H_0 be a symmetric positive definite matrix and assume that, for every j, $d_j' p_j > 0$ and H_{j+1} is determined by (2.5). Then H_{j+1} is positive definite for every j if and only if at least one of the following two conditions is satisfied.

i) $\beta_1 \beta_2 \geq 0, \quad \beta_1 + \beta_2 \neq 0$

ii) $\left(\beta_1 + \beta_2 \dfrac{d_j' p_j}{\rho_j g_j' p_j}\right)\left(\beta_1 + \beta_2 \dfrac{d_j' H_j d_j}{d_j' p_j}\right) > 0$.

Proof:

Observing that by (2.15) H_{j+1} is positive definite if and only if $\omega_j > 0$ we immediately see that the lemma follows from (2.18) and (2.19).

From (2.17) and (2.19) we obtain

(2.21) $$\gamma_j = \frac{\beta_1 d_j'p_j+\beta_2 d_j'H_jd_j-\beta_2\dfrac{(d_j'q_j)^2}{w_j'q_j}}{\beta_1 d_j'p_j+\beta_2 d_j'H_jd_j}$$

$$= 1 - \frac{\beta_2\dfrac{(d_j'q_j)^2}{w_j'q_j}}{\beta_1 d_j'p_j+\beta_2 d_j'H_jd_j} .$$

If $d_j'q_j = 0$, then $\gamma_j = 1$ and, by (2.16), $u_j = q_j$ and $\omega_j = 1$. Therefore, it follows from (2.15) that in this case H_{j+1} is independent of the parameters β_1 and β_2. Since

$$d_j'q_j = \frac{(g_j'-g_{j+1})q_j}{\|\sigma_j s_j\|} = \frac{-g_{j+1}'q_j}{\|\sigma_j s_j\|}$$

this happens if and only if g_j and g_{j+1} are parallel. Excluding this case we have the following lemma.

Lemma 2

Let H_j be positive definite and suppose that $d_j'p_j > 0$ and $d_j'q_j \neq 0$. If $\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j \neq 0$, then

i) $\gamma_j = 1$ if and only if $\beta_2 = 0$

ii) $\gamma_j > 1$ if and only if $\dfrac{\beta_2}{\beta_1 d_j'p_j+\beta_2 d_j'H_jd_j} < 0$

iii) $0 < \gamma_j < 1$ if and only if $\beta_1 + \beta_2\dfrac{d_j'p_j}{\rho_j g_j'p_j} > 0$ and either $\beta_2 > 0$ or

$\beta_2 < 0$ and $\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j < 0$.

Proof:

The first statement of the lemma follows immediately from (2.21). Let $\beta_2 \neq 0$. Suppose first that $\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j > 0$. By (2.21) we have $\gamma_j > 1$ if $\beta_2 < 0$ and $\gamma_j < 1$ if $\beta_2 > 0$ in which case it follows from (2.19) that $\gamma_j > 0$ if and only if $\beta_1\rho_j g_j'p_j + \beta_2 d_j'p_j > 0$. Next let $\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j < 0$. Then $\beta_2 > 0$ implies $\gamma_j > 1$ and $\beta_2 < 0$ implies $\gamma_j < 1$

with $\gamma_j > 0$ if and only if $\beta_1\rho_j g_j' p_j + \beta_2 d_j' p_j > 0$. Since $\beta_1 d_j' p_j + \beta_2 d_j' H_j d_j > 0$ if $\beta_2 > 0$ and $\beta_1\rho_j g_j' p_j + \beta_2 d_j' p_j > 0$, this completes the proof of the lemma.

The above lemma shows that all update formulas (2.5) with

$$\beta_1\beta_2 \geq 0, \quad \beta_1 + \beta_2 \neq 0$$

in addition to preserving the positive definiteness of H_j produce a γ_j with $0 < \gamma_j \leq 1$. Let $\bar{\gamma}_j$ and $\hat{\gamma}_j$ denote the value of γ_j that corresponds to the DFP - method and the BFGS - method, respectively. It is interesting to observe that, if $d_j' p_j > 0$ and $d_j' q_j \neq 0$, (2.21) implies

$$0 < \bar{\gamma}_j < \gamma_j < \hat{\gamma}_j = 1$$

for every γ_j corresponding to an update formula (2.5) with

$$\beta_1\beta_2 > 0 \quad .$$

For the results obtained so far we have only assumed that σ_j is chosen in such a way that $d_j' p_j > 0$, i.e., $g_{j+1}' p_j < g_j' p_j$. Now we assume that σ_j is the optimal step size; more precisely let σ_j be the smallest value of σ such that

$$F(x_j - \sigma_j s_j) = \min\{F(x_j - \sigma s_j) \mid \sigma \geq 0\} \quad .$$

Then $g_{j+1}' p_j = 0$ and it follows from the definition of w_j that

(2.22) $$w_j = \lambda_{j+1} g_{j+1} \quad \text{where} \quad \lambda_{j+1} = \|H_j g_{j+1}\|^{-1} \quad .$$

Therefore, (2.15) becomes

$$H_{j+1} = \frac{p_j p_j'}{d_j' p_j} + \omega_j \frac{u_j u_j'}{\lambda_{j+1} g_{j+1}' u_j} + \hat{H}_j$$

and

(2.23) $$s_{j+1} = H_{j+1} g_{j+1} = \omega_j \frac{u_j}{\lambda_{j+1}} \quad ,$$

i.e. the search directions at x_{j+1} computed by any of the matrices (2.5) differ only in the factor ω_j. This observation suggests a simple proof for a theorem due to Dixon [5] which essentially states that, if the optimal step size is used, all members of the class (2.5) of update formulas produce the same sequence of points $\{x_j\}$.

Theorem 1

Let an initial point x_0 and a symmetric positive definite matrix H_0 be given. Suppose that for every j, σ_j is the optimal step size,

$$s_j = H_j g_j, \quad x_{j+1} = x_j - \sigma_j s_j$$

and H_{j+1} is determined by (2.5). Any choice of the parameters β_1 and β_2 for which $\omega_j > 0$, i.e., $g'_{j+1}s_{j+1} > 0$ for all j, results in the same sequence of points $\{x_j\}$.

Proof:

Suppose that, for some j, all matrices H_j in the class generated by the update formulas (2.5) have the form

$$(2.24) \qquad H_j = \omega_{j-1}\frac{p_j p'_j}{\lambda_j g'_j p_j} + \frac{p_{j-1}p'_{j-1}}{d'_{j-1}p_{j-1}} + \hat{H}_{j-1} \quad ,$$

where only ω_{j-1} depends on the particular values of β_1 and β_2. Since the optimal step size is used it follows that x_{j+1} and g_{j+1} are independent of ω_{j-1}. Thus span$\{H_j g_j, H_j g_{j+1}\}$ is independent of ω_{j-1} which implies that $p_{j+1} = u_j$ is independent of ω_{j-1}. Thus we can write

$$H_j = \omega_{j-1}\frac{p_j p'_j}{\lambda_j g'_j p_j} + \frac{p_{j+1}p'_{j+1}}{\lambda_{j+1}g'_{j+1}p_{j+1}} + \hat{H}_j$$

where $\lambda_{j+1} = \|H_j g_{j+1}\|^{-1}$ is independent of ω_{j-1} and the matrix $\hat{H}_j$ is as defined in (2.10) and independent of ω_{j-1}. Therefore (2.15) becomes

$$H_{j+1} = \frac{p_j p'_j}{d'_j p_j} + \omega_j \frac{p_{j+1}p'_{j+1}}{\lambda_{j+1}g'_{j+1}p_{j+1}} + \hat{H}_j \quad .$$

This representation of H_{j+1} is equivalent to the representation (2.24) of H_j. Since (2.24) holds for $j = 1$, this proves the theorem.

In practical computation σ_j differs from the optimal step size and numerical experience shows that the efficiency of a variable metric method depends very much on the particular update formula (2.5) which is being used. From (2.15) we obtain

$$s_{j+1} = H_{j+1}g_{j+1} = p_j \frac{p_j'g_{j+1}}{d_j'p_j} + \omega_j u_j \frac{u_j'g_{j+1}}{w_j'u_j} \ .$$

Thus depending on $p_j'g_{j+1}$ and γ_j, i.e., on the closeness of the step size used to the optimal step size and on the choice of β_1 and β_2, the directions s_{j+1} can differ considerably.

3. Convergence

For any initial point x_0 for which Assumption 1 is satisfied and any symmetric positive definite matrix H_0 let $\{x_j\}$ be a sequence with the following properties

i) $F(x_{j+1}) < F(x_j)$, $j = 0,1,\ldots$

ii) $x_{j+1} = x_j - \sigma_j s_j$, $s_j = H_j g_j$, $\sigma_j > 0$, $j = 0,1,\ldots$.

iii) H_{j+1} is obtained from H_j by (2.5) with arbitrary parameters β_1 and β_2.

Throughout the remainder of the paper we shall assume that, if necessary, the parameters β_1 and β_2 are adjusted in such a way that H_{j+1} is defined and positive definite, i.e., that the conditions of Lemma 1 are satisfied.

It is the purpose of this section to show that if Assumption 1 is satisfied and σ_j is chosen appropriately, then the sequence $\{g_j\}$ converges to zero and every cluster point of the sequence $\{x_j\}$ is a global minimizer of $F(x)$.

We shall prove this result by generalizing a proof due to Powell [14] for the case of the BFGS - method, i.e., $\beta_1 = 1$, $\beta_2 = 0$. Powell's proof uses the inverse of H_j rather than H_j. Setting

$$B_j = H_j^{-1}$$

we obtain from (2.9) and (2.10)

$$B_j = \frac{\rho_j q_j q_j'}{q_j' p_j} + \frac{w_j w_j'}{w_j' q_j} + \hat{B}_j \tag{3.1}$$

where

$$\hat{B}_j = \sum_{i=3}^{n} \frac{d_{ij} d_{ij}'}{d_{ij}' p_{ij}} .$$

Similarly, (2.15) implies

$$B_{j+1} = \frac{d_j d_j'}{d_j' p_j} + \frac{1}{\omega_j} \frac{w_j w_j'}{w_j' u_j} + \hat{B}_j \tag{3.2}$$

As a first step we derive a relation between the trace of B_j and the trace of B_{j+1}. By definition the trace of B_j is equal to the sum of eigenvalues of B_j which is equal to the sum of the diagonal elements of B_j. Since with H_j the matrix B_j is positive definite, too, the trace of B_j is positive. From (3.1) we obtain

$$\mathrm{tr}(B_j) = \frac{\rho_j \|g_j\|^2}{g_j' p_j} + \frac{\|w_j\|^2}{w_j' q_j} + \mathrm{tr}(\hat{B}_j) \quad .$$

Thus using (3.2) we have

$$\begin{aligned}(3.3)\qquad \mathrm{tr}(B_{j+1}) &= \mathrm{tr}(B_j) - \frac{\rho_j \|g_j\|^2}{g_j' p_j} + \frac{\|d_j\|^2}{d_j' p_j} - \frac{\|w_j\|^2}{w_j' q_j} + \frac{1}{\omega_j}\frac{\|w_j\|^2}{w_j' u_j} \\ &= \mathrm{tr}(B_j) - \frac{\rho_j \|g_j\|^2}{g_j' p_j} + \frac{\|d_j\|^2}{d_j' p_j} - (1 - \frac{1}{\gamma_j})\frac{\|w_j\|^2}{w_j' q_j} \quad ,\end{aligned}$$

where the last equality follows from the definition of ω_j (see (2.18)) and

$$(3.4)\qquad w_j' u_j = \frac{w_j'(q_j + \alpha_j p_j)}{\|q_j + \alpha_j p_j\|} = \frac{w_j' q_j}{\|q_j + \alpha_j p_j\|} \quad .$$

Since $g_j' p_j > 0$ for all j, we deduce from (3.3) the inequality

$$(3.5)\qquad \mathrm{tr}(B_{j+1}) \le \mathrm{tr}(B_0) + \sum_{i=0}^{j} \frac{\|d_i\|^2}{d_i' p_i} + \sum_{i=0}^{j} \frac{1-\gamma_i}{\gamma_i}\frac{\|w_i\|^2}{w_i' q_i} \quad .$$

Next we establish a relation between the determinants of B_{j+1} and B_j. For the special case of the BFGS - method, i.e., for $\gamma_j = 1$, the result has been obtained by Pearson [11].

Lemma 3

Let B_j and B_{j+1} be defined by (3.1) and (3.2), respectively. Then

$$(3.6)\qquad \det(B_{j+1}) = \frac{1}{\gamma_j}\frac{d_j' p_j}{\rho_j g_j' p_j}\det(B_j) \quad .$$

Proof:

Set

$$D_j^{-1} = \left(\frac{p_j}{\sqrt{\rho_j g_j' p_j}}, \frac{q_j}{\sqrt{w_j' q_j}}, \frac{p_{3j}}{\sqrt{d_{3j}' p_{3j}}}, \ldots, \frac{p_{nj}}{\sqrt{d_{nj}' p_{nj}}}\right)$$

and

$$D_{j+1} = \left(\frac{d_j}{\sqrt{d_j' p_j}}, \frac{w_j}{\sqrt{\gamma_j w_j' q_j}}, \frac{d_{3j}}{\sqrt{d_{3j}' p_{3j}}}, \ldots, \frac{d_{nj}}{\sqrt{d_{nj}' p_{nj}}} \right) .$$

Then it follows from (2.9) and (3.2) that

$$H_j = D_j^{-1} D_j'^{-1} \quad \text{and} \quad B_{j+1} = D_{j+1} D_{j+1}' \quad .$$

Therefore,

$$\det(B_{j+1} H_j) = \det(D_{j+1} D_{j+1}' D_j^{-1} D_j'^{-1})$$

$$= (\det(D_{j+1}' D_j^{-1}))^2 = \frac{d_j' p_j}{\gamma_j \rho_j g_j' p_j} \quad ,$$

which because of $B_j = H_j^{-1}$ implies

$$\det(B_{j+1}) = \frac{d_j' p_j}{\gamma_j \rho_j g_j' p_j} \, \frac{1}{\det(H_j)} = \frac{1}{\gamma_j} \, \frac{d_j' p_j}{\rho_j g_j' p_j} \det(B_j) \quad .$$

For the BFGS - method, ${}_j = 1$. Assuming that

(3.7) $$\frac{\|d_j\|^2}{d_j' p_j} \leq \delta_0 \quad \text{for some} \quad \delta_0 \quad \text{and all} \quad j$$

Powell [14] used (3.5) and (3.6) to prove that

$$\liminf_{j \to \infty} \|g_j\| = 0 \quad .$$

A review of Powell's proof shows that it can be adapted for a general update formula of type (2.5) if in addition to (3.7) we have

(3.8) $$\frac{1-\gamma_j}{\gamma_j} \, \frac{\|w_j\|^2}{w_j' q_j} \leq \delta_1 \quad \text{for some} \quad \delta_1 > 0 \quad \text{and all} \quad j$$

and

(3.9) $$\gamma_j \leq \delta_2 \quad \text{for some} \quad \delta_2 > 1 \quad \text{and all} \quad j \quad .$$

Unfortunately, it does not seem to be possible to determine any choice of the parameters β_1 and β_2, (other than $\beta_1 = 1$, $\beta_2 = 0$, resulting in $\gamma_j = 1$) for which (3.8) and (3.9) can be verified a priori: Indeed, if $\gamma_j > 1$, then (3.8) is satisfied. However, by Lemma 2, we have then

$$\frac{\beta_2}{\beta_1 d_j' p_j + \beta_2 d_j' H_j d_j} < 0$$

which by (2.21) could result in an arbitrary large γ_j. On the other hand, if we choose β_1 and β_2 such that $\gamma_j < 1$ it does not seem to be possible to find a positive lower bound for γ_j. Thus $(1-\gamma_j)/\gamma_j$ may become arbitrarily large. Even if these numbers are bounded $\|w_j\|^2/w_j' q_j$ could become large since we cannot show a priori that the sequence $\{\beta_j\}$ is bounded.

In order to overcome this difficulty we replace the matrix H_j by a matrix $H_j(\eta_j)$ which is defined as follows.

(3.10)
$$H_j(\eta_j) = H_j + \frac{\eta_j}{1-\eta_j} \frac{p_j p_j'}{\rho_j g_j' p_j}, \qquad \eta_j < 1$$

$$= \frac{1}{1-\eta_j} \frac{p_j p_j'}{\rho_j g_j' p_j} + \frac{q_j q_j'}{w_j' q_j} + \hat{H}_j \ .$$

Setting

$$\tilde{s}_j = H_j(\eta_j) g_j$$

we have

$$\tilde{s}_j = \frac{1}{1-\eta_j} s_j$$

and with a modified step size

$$\tilde{\sigma}_j = (1-\eta_j)\sigma_j$$

we obtain

$$x_{j+1} = x_j - \sigma_j s_j = x_j - \tilde{\sigma}_j \tilde{s}_j \ .$$

Furthermore (2.15) shows that H_{j+1} is not affected by the change in H_j.

Denoting the inverse matrix of $H_j(\eta_j)$ by $B_j(\eta_j)$ we see from (3.1) that

$$B_j(\eta_j) = (1-\eta_j)\frac{\rho_j q_j q_j'}{q_j' p_j} + \frac{w_j w_j'}{w_j' q_j} + \hat{B}_j$$

$$= B_j - \eta_j \frac{\rho_j q_j q_j'}{q_j' p_j} \quad .$$

Using the same argument as in the proof of Lemma 3 it is easy to verify that

$$\det(B(\eta_j)) = (1-\eta_j)\det(B_j) \quad .$$

Therefore, replacing B_j and B_{j+1} by $B_j(\eta_j)$ and $B_{j+1}(\eta_{j+1})$, respectively, in (3.3) and (3.6) we obtain

(3.11)
$$\mathrm{tr}(B_{j+1}(\eta_{j+1})) = \mathrm{tr}(B_j(\eta_j)) - \frac{(1-\eta_j)\rho_j \|q_j\|^2}{q_j' p_j} + \frac{\|d_j\|^2}{d_j' p_j}$$

$$- \left(1-\frac{1}{\gamma_j}\right)\frac{\|w_j\|^2}{w_j' q_j} - \eta_{j+1}\frac{\rho_{j+1}\|q_{j+1}\|^2}{q_{j+1}' p_{j+1}}$$

and

(3.12)
$$\det(B_{j+1}(\eta_{j+1})) = \frac{1-\eta_{j+1}}{\gamma_j}\,\frac{d_j' p_j}{(1-\eta_j)\rho_j q_j' p_j}\det(B_j(\eta_j)) \quad .$$

If we assume that σ_j is the optimal step size, then it follows from (2.22) and (2.23) that

$$w_j = \lambda_{j+1} q_{j+1},\ p_{j+1} = u_j,\ \rho_{j+1} = \frac{\lambda_{j+1}}{\omega_j}$$

which by (2.18) and (3.4) implies

$$\frac{\rho_{j+1}\|q_{j+1}\|^2}{q_{j+1}' p_{j+1}} = \frac{1}{\gamma_j}\,\frac{\|w_j\|^2}{w_j' q_j} \quad .$$

Thus if we set

$$\eta_{j+1} = 1-\gamma_j$$

then the sum of the last two terms in (3.11) is zero and

$$\frac{1-\eta_{j+1}}{\gamma_j} = 1 \quad .$$

Since it suffices to have the sum of the last two terms in (3.11) bounded from above, σ_j need only be an approximation to the optimal step size which satisfies the following condition.

Condition 1

The step size σ_j is determined such that, for all j,

$$(3.13)\qquad (1-\gamma_j)\left[\frac{\|g_{j+1}-\varepsilon_j d_j\|^2}{(g_{j+1}-\varepsilon_j d_j)' H_{j+1}(g_{j+1}-\varepsilon_j d_j)} - \frac{\|g_{j+1}\|^2}{g'_{j+1}H_{j+1}g_{j+1}}\right] \le \delta_3 \ ,$$

where δ_3 is an arbitrary positive constant and

$$\varepsilon_j = \frac{g'_{j+1}p_j}{d'_j p_j} \ .$$

For $\gamma_j = 1$ Condition 1 is trivially satisfied. If σ_j is the optimal step size, then $\varepsilon_j = 0$. Therefore, for every j, there is an interval, containing the optimal step size, such that every σ_j in this interval satisfies Condition 1.

Since $H_{j+1}(g_{j+1} - \varepsilon_j d_j) \in \mathrm{span}\{q_j, p_j\}$ and $d'_j H_{j+1}(g_{j+1} - \varepsilon_j d_j) = 0$ it follows that

$$(3.14)\qquad u_j = \frac{s_{j+1}-\varepsilon_j p_j}{\|s_{j+1}-\varepsilon_j p_j\|} \quad \text{and} \quad w_j = \omega_j \frac{g_{j+1}-\varepsilon_j d_j}{\|s_{j+1}-\varepsilon_j p_j\|} \ .$$

Thus

$$\frac{\|g_{j+1}-\varepsilon_j d_j\|^2}{(g_{j+1}-\varepsilon_j d_j)' H_{j+1}(g_{j+1}-\varepsilon_1 d_j)} = \frac{\|w_j\|^2}{\omega_j w'_j u_j} = \frac{1}{\gamma_j}\,\frac{\|w_j\|^2}{w'_j q_j} \ ,$$

and observing that $p_{j+1} = s_{j+1}/\|s_{j+1}\| = \rho_{j+1}s_{j+1}$ we deduce from (3.13) the inequality

$$\frac{1-\gamma_j}{\gamma_j}\,\frac{\|w_j\|^2}{w'_j q_j} - (1-\gamma_j)\,\frac{\rho_{j+1}\|g_{j+1}\|^2}{g'_{j+1}p_{j+1}} \le \delta_3 \ .$$

Choosing

$$\eta_{j+1} = 1-\gamma_j$$

and assuming that the inequalities (3.7) and (3.13) are satisfied we obtain from (3.11) the relation

$$\operatorname{tr}(B_{j+1}(\eta_{j+1})) \leq \operatorname{tr}(B_j(\eta_j)) - \frac{(1-\eta_j)\rho_j\|g_j\|^2}{g_j'p_j} + \delta_0 + \delta_3$$

which shows that, for every j,

$$\operatorname{tr} B_{j+1}(\eta_{j+1}) \leq \operatorname{tr}(B_0(\eta_0)) - \sum_{i=0}^{j} \frac{(1-\eta_i)\rho_i\|g_i\|^2}{g_i'p_i} + (j+1)(\delta_0+\delta_3) \tag{3.15}$$

where $B_0(\eta_0) = B_0$.

Using this upper bound for the trace of $B_{j+1}(\eta_{j+1})$ we can prove the following key lemma.

<u>Lemma 4</u>

Suppose the inequalities (3.7) and (3.13) are satisfied. Then there is $\delta_4 > 0$ such that for $j = 0,1,\ldots,$

$$\prod_{i=0}^{j} \frac{\|g_i\|^2}{g_i'p_i} \leq \delta_4^j \prod_{i=0}^{j} \frac{g_i'p_i}{d_i'p_i} \quad . \tag{3.16}$$

<u>Proof</u>:

Since $B_j(\eta_j)$ is positive definite for all j, *we obtain from* (3.15)

$$\sum_{i=0}^{j} \frac{(1-\eta_i)\rho_i\|g_i\|^2}{g_i'p_i} \leq \operatorname{tr}(B_0) + (j+1)(\delta_0+\delta_3) \leq (j+1)\delta_5$$

with $\delta_5 = \operatorname{tr}(B_0) + \delta_0 + \delta_3$. Applying the geometric/arithmetic mean inequality we obtain the relation

$$\prod_{i=0}^{j} \frac{(1-\eta_i)\rho_i\|g_i\|^2}{g_i'p_i} \leq \delta_5^{j+1} \quad \text{for } j = 0,1,\ldots \quad . \tag{3.17}$$

Observing that $1-\eta_{j+1} = \gamma_j$ *and using* (3.12) we find

$$\prod_{i=0}^{j} \frac{d_i'p_i}{(1-\eta_i)\rho_i g_i'p_i} = \frac{\det(B_{j+1}(\eta_{j+1}))}{\det(B_0)} \quad . \tag{3.18}$$

Next we deduce from (3.15) the inequality

$$\operatorname{tr}(B_{j+1}(\eta_{j+1})) \leq (j+1)\delta_5 \quad . \tag{3.19}$$

Since the determinant of $B_{j+1}(\eta_{j+1})$ is equal to the product of its eigenvalues we can use (3.19) and the geometric/arithmetic mean inequality to find the relation

$$\det(B_{j+1}(\eta_{j+1})) \leq \left(\frac{(j+1)\,\delta_5}{n}\right)^n .$$

Combining (3.17), (3.18) and the above inequality we obtain the expression

$$\prod_{i=0}^{j} \frac{\|g_i\|^2}{g_i'p_i} \leq \delta_5^{j+1}\left(\frac{(j+1)\,\delta_5}{n}\right)^n \frac{1}{\det(B_0)} \prod_{i=0}^{j} \frac{g_i'p_i}{d_i'p_i}$$

$$\leq \delta_4^j \prod_{i=0}^{j} \frac{g_i'p_i}{d_i'p_i}$$

where δ_4 is a suitable constant.

Since the inequality (3.13) is trivially satisfied if $\gamma_j = 1$, i.e. in the BFGS - method, we need an additional condition for the step size σ_j in order to be able to draw further conclusions from the inequality (3.16).

Condition 2.

Let γ and γ^* be constants satisfying the inequalities

$$0 < \gamma < \gamma^* < 1, \quad \gamma < \frac{1}{2}$$

and let σ_j be determined such that

i) $g_{j+1}'p_j \leq \gamma^* g_j'p_j$

ii) $F(x_{j+1}) \leq F(x_j) - \gamma\|\sigma_j s_j\| g_j'p_j$ or $\sigma_j \geq \sigma_j^+$ and $F(x_{j+1}) \leq F(x_j - \sigma_j^+ s_j)$ where σ_j^+ is the smallest positive number with

$$F(x_j - \sigma_j^+ s_j) = F(x_j) - \gamma\|\sigma_j^+ s_j\| g_j'p_j$$

iii) $\sigma_j = \sigma_j^*$ if possible with

$\sigma_j^* = 1$ if $\beta_2 = 0$

$$\sigma_j^* = \frac{g_j's_j}{2(F(x_j-s_j) - F(x_j) + g_j's_j)} \quad \text{if } \beta_2 \neq 0 \quad .$$

Let $\hat{\sigma}_j$ denote the optimal step size. Since $\hat{\sigma}_j$ could be greater than σ_j^+ and Condition 1 could force σ_j to be close to $\hat{\sigma}_j$ we cannot insist on the inequality $F(x_{j+1}) \leq F(x_j) - \gamma\|\sigma_j s_j\| g_j'p_j$.

Under suitable assumptions it can be shown [15] that with

$$\sigma_j^* = \frac{g_j's_j}{2(F(x_j-s_j) - F(x_j) + g_j's_j)}$$

we have

$$|\nabla F(x_j-\sigma_j^* s_j)'p_j| = O(\|g_j\|^2) \quad . \tag{3.20}$$

Furthermore, it will be shown in the next section that for every update formula (2.5) with $\beta_1\beta_2 \geq 0$, $\sigma_j = \sigma_j^*$ is an acceptable step size for j sufficiently large.

Using a step size which satisfies Conditions 1 and 2 we obtain the following result.

<u>Lemma 5</u>

Suppose the inequality (3.7) holds and σ_j satisfies the Conditions 1 and 2. Then

$$\lim_{j\to\infty} \inf \|g_j\| = 0 \quad .$$

<u>Proof:</u>

Since for all j,

$$\frac{g_j'p_j}{d_j'p_j} = \frac{\|\sigma_j s_j\| g_j'p_j}{g_j'p_j - g_{j+1}'p_j} \leq \frac{\|\sigma_j s_j\|}{1-j^*} \leq \delta_6$$

where δ_6 is a suitable constant, it follows from (3.16) that

$$\prod_{i=0}^{j} \frac{\|g_i\|^2}{g_i'p_i} \leq (\delta_4\delta_6)^j, \quad j = 0,1,\ldots \quad . \tag{3.21}$$

The sequence $\{F(x_j)\}$ is decreasing. Therefore, $\{x_j\} \subset S_0$. If there is an infinite subset $J \subset \{0,1,\ldots\}$ and an $\varepsilon > 0$ such that

$$p_j' g_j \geq \varepsilon \quad \text{for} \quad j \in J \ ,$$

then it follows from $p_j' g_{j+1} \leq \gamma^* p_j' g_j$ and the uniform continuity of $\nabla F(x)$ on S_0 that

$$\min\{\|\sigma_j s_j\|, \|\sigma_j^+ s_j\| \geq \varepsilon_1 > 0 \quad \text{for some} \quad \varepsilon_1 > 0 \quad \text{and} \quad j \in J \ .$$

Because $F(x)$ is bounded from below and

$$F(x_{j+1}) \leq F(x_j) - \gamma p_j' g_j \min\{\|\sigma_j s_j\|, \|\sigma_j^+ s_j\|\}$$

this implies that $p_j g_j \to 0$ as $j \to \infty$, which by (3.21) proves that $\{\|g_j\|\}$ is not bounded away from zero.

We are now ready to prove the main convergence theorem.

Theorem 2

Let Assumption 1 and Conditions 1 and 2 be satisfied. Then

$$g_j \to 0 \quad \text{as} \quad j \to \infty$$

and every cluster point of the sequence $\{x_j\}$ is a global minimizer of $F(x)$.

Proof:

It has been shown in [14] that if $F(x)$ is convex and twice continuously differentiable on S_0 then the inequality (3.7) holds for all j. Therefore, we deduce from Lemma 5 that there is an infinite subset $J \subset \{0,1,\ldots\}$ and a $z \in S_0$ such that

$$\nabla F(z) = 0 \quad \text{and} \quad x_j \to z \quad \text{as} \quad j \to \infty,\ j \in J \ .$$

If $\{g_j\}$ does not converge to zero, then the sequence $\{x_j\}$ has a cluster point z^*, say, such that $\nabla F(z^*) \neq 0$. Since $F(x)$ is convex this implies $F(z^*) > F(z)$. Because $F(x_{j+1}) < F(x_j)$, this contradiction shows that $g_j \to 0$ as $j \to \infty$. Therefore, it follows from the continuity of $\nabla F(x)$ and the convexity of $F(x)$ on S_0 that every cluster point of $\{x_j\}$ is a global minimizer of $F(x)$.

4. Superlinear convergence

In order to prove that the sequence $\{x_j\}$ converges superlinearly to a global minimizer of $F(x)$ we require that in addition to the assumptions stated in the previous sections the following assumption is satisfied.

Assumption 2.

The sequence $\{x_j\}$ converges to a point z. The Hessian matrix $G = G(z)$ is positive definite. There is a neighborhood $U_1(z)$ such that the Lipschitz condition

(4.1) $$\|G(x) - G(z)\| \leq L\|x-z\|$$

holds for all $x \in U_1(z)$, where L is a constant.

The above assumption implies that there are a neighborhood $U_2(z)$ and constants $0 < \mu < \eta$ such that, for every $x \in U_2(z)$,

(4.2) $$\mu\|y\|^2 \leq y'G(x)y \leq \eta\|y\|^2 \quad \text{for all} \quad y \in E^n \; .$$

Therefore there is a neighborhood $U(z)$ such that the inequalities (4.1) and (4.2) hold for every $x \in U(z)$. By deleting finitely many members of the sequence $\{x_j\}$ if necessary, we may assume without loss of generality that $\{x_j\} \subset U(z)$.

In proving that the sequence $\{x_j\}$ converges superlinearly we will use the weighted matrices

$$G^{1/2}H_j(\eta_j)G^{1/2} \quad , \quad G^{-1/2}B_j(\eta_j)G^{-1/2} \quad ,$$

where the symmetric positive definite matrix $G^{1/2}$ is the square root of G and $G^{-1/2} = (G^{1/2})^{-1}$. As a first result we will show that

$$\psi_j = \mathrm{tr}(G^{1/2}H_j(\eta_j)G^{1/2}) + \mathrm{tr}(G^{-1/2}B_j(\eta_j)G^{-1/2})$$

is bounded if we choose $\eta_j = 1-\gamma_{j-1}$ as before and impose an appropriate condition on the step size σ_j.

We observe that by (3.10) and (2.15)

$$(4.3) \qquad H_{j+1}(\eta_{j+1}) = H_j(\eta_j) - \frac{1}{1-\eta_j}\,\frac{p_jp_j'}{\rho_jq_j'p_j} + \frac{p_jp_j'}{d_j'p_j} - \frac{q_jq_j'}{w_j'q_j}$$

$$+ \omega_j\frac{u_ju_j'}{w_j'u_j} + \frac{\eta_{j+1}}{1-\eta_{j+1}}\,\frac{p_{j+1}p_{j+1}'}{\rho_{j+1}q_{j+1}'p_{j+1}} .$$

Therefore, choosing $\eta_j = 1-\gamma_{j-1}$ and setting

$$\tau_j = \frac{p_j'Gp_j + d_j'G^{-1}d_j}{d_j'p_j} ,$$

$$\varphi_j = \frac{1}{\omega_j}\frac{w_j'G^{-1}w_j}{w_j'u_j} - \frac{w_j'G^{-1}w_j}{w_j'q_j} - \eta_{j+1}\frac{\rho_{j+1}q_{j+1}'G^{-1}q_{j+1}}{q_{j+1}'p_{j+1}}$$

$$= (1-\gamma_j)\left[\frac{1}{\gamma_j}\frac{w_j'G^{-1}w_j}{w_j'q_j} - \frac{\rho_{j+1}q_{j+1}'G^{-1}q_{j+1}}{q_{j+1}'p_{j+1}}\right] ,$$

$$\mu_j = \frac{(q_j+\alpha_jp_j)'G(q_j+\alpha_jp_j)}{w_j'q_j} - \frac{q_j'Gq_j}{w_j'q_j} ,$$

$$\xi_j = \omega_j\frac{u_j'Gu_j}{w_j'u_j} - \frac{q_j'Gq_j}{w_j'q_j} + \frac{\eta_{j+1}}{1-\eta_{j+1}}\,\frac{p_{j+1}'Gp_{j+1}}{\rho_{j+1}q_{j+1}'p_{j+1}} - \mu_j$$

$$= (\gamma_j-1)\left[\frac{(q_j+\alpha_jp_j)'G(q_j+\alpha_jp_j)}{w_j'q_j} - \frac{1}{\gamma_j}\,\frac{p_{j+1}'Gp_{j+1}}{\rho_{j+1}q_{j+1}'p_{j+1}}\right]$$

we deduce from (4.3) and (3.11) that for every j

$$(4.4) \qquad \psi_{j+1} = \psi_j - \frac{p_j'Gp_j + (1-\eta_j)^2\rho_j^2q_j'G^{-1}q_j}{(1-\eta_j)\rho_jq_j'p_j} + \tau_j + \varphi_j + \xi_j + \mu_j .$$

In order to show that the sequence $\{\psi_j\}$ is bounded we have to derive upper bounds for the terms τ_j, φ_j, ξ_j, and μ_j. This will be done in the following few lemmas.

<u>Lemma 6</u>

Let G be a symmetric nonsingular (n,n) matrix and let $y,x \in E^n$ be such that $y'x \neq 0$. Then

$$\frac{x'Gx+y'G^{-1}y}{y'x} = 2 + \frac{v'G^{-1}v}{y'x}$$

where $v = y - Gx$.

Proof:

$$\frac{x'Gx+y'G^{-1}y}{y'x} = \frac{x'(y-v)+y'(x+G^{-1}v)}{y'x}$$

$$= 2 + \frac{v'(G^{-1}y-x)}{y'x} = 2 + \frac{v'G^{-1}v}{y'x} \quad .$$

Lemma 7

Under the assumptions stated the sequence

$$\frac{\|x_{j+1} - z\|}{\|x_j - z\|}$$

is bounded and the sum

$$\sum_{j=0}^{\infty} \|x_j - z\| \tag{4.5}$$

is finite.

Proof :

By Taylor's theorem there is a v_j on the line segment joining z and x_j such that

$$2(F(x_j)-F(z)) = (x_j-z)'G(v_j)(x_j-z) \quad .$$

Therefore,

$$\mu\|x_j-z\|^2 \leq 2(F(x_j)-F(z)) \leq \eta\|x_j-z\|^2 \tag{4.6}$$

which implies

$$\frac{\|x_{j+1} - z\|^2}{\|x_j - z\|^2} \leq \frac{\eta}{\mu} \frac{F(x_{j+1})-F(z)}{F(x_j)-F(z)} \leq \frac{\eta}{\mu} \quad .$$

By Taylor's theorem and Condition 2 we have

$$\gamma^* g_j'p_j \geq g_{j+1}'p_j \geq g_j'p_j - \eta\|\sigma_j s_j\|$$

and

$$F(x_j) - \gamma\sigma_j^+ g_j' s_j = F(x_j - \sigma_j^+ s_j) \le F(x_j) - \sigma_j^+ g_j' s_j + \frac{1}{2}\eta\|\sigma_j^+ s_j\|^2 \quad .$$

Therefore,

$$\text{(4.7)} \qquad \min\{\|\sigma_j s_j\|, \|\sigma_j^+ s_j\|\} \ge \frac{1}{\eta}\min\{1-\gamma^*, 2(1-\gamma)\} g_j' p_j$$

$$= \frac{g_j' p_j}{\eta}(1-\gamma^*) \quad .$$

Using Condition 2 once more we deduce from (4.6) and (4.7) the relation

$$\text{(4.8)} \qquad F(x_{j+1}) - F(z) \le F(x_j) - F(z) - \gamma g_j' p_j \min\{\|\sigma_j s_j\|, \|\sigma_j^+ s_j\|\}$$

$$\le F(x_j) - F(z) - \frac{\gamma(1-\gamma^*)}{\eta}(g_j' p_j)^2$$

$$\le (F(x_j) - F(z))\left(1 - \frac{\gamma(1-\gamma^*)}{\eta}\,\frac{2\|g_j\|^2}{\eta\|x_j - z\|^2}\,\frac{(g_j' p_j)^2}{\|g_j\|^2}\right)$$

$$\le (F(x_j) - F(z))\left(1 - \gamma(1-\gamma^*)\,\frac{2\mu^2}{\eta^2}\,\frac{(g_j' p_j)^2}{\|g_j\|^2}\right) ,$$

where the last inequality follows from the relation $\mu\|x_j - z\| \le \|g_j\|$, see [12] for instance. Setting

$$\zeta_j = 1 - \gamma(1-\gamma^*)\,\frac{2\mu^2}{\eta^2}\,\frac{(g_j' p_j)^2}{\|g_j\|^2}$$

we obtain from (4.8)

$$\text{(4.9)} \qquad F(x_{j+1}) - F(z) \le (F(x_0) - F(z)) \prod_{i=0}^{j} \zeta_i \quad .$$

Since $d_j' p_j \ge \mu$ it follows from (3.16) that there is $\delta_7 < 1$ such that

$$\prod_{i=0}^{j}\left(\frac{g_i' p_i}{\|g_i\|}\right)^2 \ge \delta_7^j \quad , \quad j = 1,2,\ldots \quad .$$

Observing that $g_i' p_i \le \|g_i\|$ we deduce from this inequality that for every j, at least half of the numbers

$$\frac{g_i' p_i}{\|g_i\|} \quad , \; i = 0,1,\ldots,j$$

are greater than or equal to δ_7. This implies that, for every j, at least half of the number ζ_i, $i = 0,1,\ldots,j$, are less than or equal to some $\delta_8^2 < 1$. Therefore, it follows from (4.9) that

$$F(x_{j+1}) - F(z) \le \delta_8^j (F(x_0) - F(z)) \quad , \quad j = 0,1,\ldots \quad ,$$

which by (4.6) implies that the sume (4.5) is finite.

<u>Lemma 8</u>

The assumptions stated imply that

i) $\|d_j - Gp_j\| = 0(\|x_j - z\|)$

ii) $\sum_{j=0}^{\infty} (\tau_j - 2)$ is finite.

<u>Proof</u>:

By Taylor's theorem

(4.10)
$$d_j = \frac{g_j - g_{j+1}}{\|\sigma_j s_j\|} = Gp_j + E_j p_j$$

where

$$E_j = \int_0^1 G(x_j + t(x_{j+1} - x_j))dt - G \quad .$$

Hence

(4.11)
$$\begin{aligned}\|E_j\| &\le \max_{0\le t\le 1} \|G(x_j + t(x_{j+1} - x_j)) - G\| \\ &\le \max_{0\le t\le 1} \|L(x_j + t(x_{j+1} - x_j) - z)\| \\ &\le L \max(\|x_j - z\|, \|x_{j+1} - z\|) = 0(\|x_j - z\|) \quad ,\end{aligned}$$

where the last relation follows from Lemma 7. Using the inequality $d_j' p_j \ge \mu$ and Lemma 6 we have therefore

(4.12)
$$0 \le \tau_j - 2 = 0(\|x_j - z\|^2) \quad ,$$

which by Lemma 7 implies that the sum

$$\sum_{j=0}^{\infty} (\tau_j - 2)$$

is finite.

<u>Lemma 9</u>

The assumptions stated imply that

$$\mu_j = O\left(\frac{\|x_j - z\|}{w_j' q_j}\right) .$$

<u>Proof</u>:

By definition $\alpha_j = -d_j' q_j / d_j' p_j$. Therefore, using (4.10) we have

$$\begin{aligned}
\mu_j &= \frac{1}{w_j' q_j} \left(2\, \alpha_j p_j' G q_j + \alpha_j^2 p_j' G p_j\right) \\
&= \frac{1}{w_j' q_j} \left(-\, 2\, \frac{d_j' q_j}{d_j' p_j} (d_j' - p_j' E_j) q_j + \left(\frac{d_j' q_j}{d_j' p_j}\right)^2 p_j' (d_j - E_j p_j)\right) \\
&= \frac{1}{w_j' q_j} \left(-\, \frac{(d_j' q_j)^2}{d_j' p_j} + 2\, \frac{d_j' q_j}{d_j' p_j} p_j' E_j q_j - \left(\frac{d_j' q_j}{d_j' p_j}\right)^2 p_j' E_j p_j\right) \\
&\leq \frac{\|E_j\|}{w_j' q_j} \left(\frac{2\|d_j\|}{d_j' p_j} + \frac{\|d_j\|^2}{(d_j' p_j)^2}\right) \\
&= O\left(\frac{\|x_j - z\|}{w_j' q_j}\right) ,
\end{aligned}$$

where the last relation follows from (4.11) and $\|d_j\| \leq \eta$, $d_j' p_j \geq \mu$.

In order to find an upper bound for the terms φ_j and ξ_j we observe that, by (3.4) and the definition of w_j, $\varphi_j = \xi_j = 0$ if σ_j is the optimal step size. Thus we can control $\varphi_j + \xi_j$ by imposing a condition on σ_j which ensures that σ_j is sufficiently close to the optimal step size.

<u>Condition 3.</u>

The step size σ_j is determined such that for all j

$$|1-\gamma_j| \max\left\{\frac{|\varepsilon_j|\ \|g_{j+1}\|}{(g_{j+1}-\varepsilon_j d_j)' H_{j+1}(g_{j+1}-\varepsilon_j d_j)},\ \left(\frac{|\varepsilon_j|\ \|g_{j+1}\|}{(g_{j+1}-\varepsilon_j d_j)' H_{j+1}(g_{j+1}-\varepsilon_j d_j)}\right)^2\right\} \le \delta_9 \nu_j$$

$$\frac{|1-\gamma_j|}{\gamma_j} \max\left\{\frac{|\varepsilon_j|(\|s_{j+1}\|+|\varepsilon_j|)}{(g_{j+1}-\varepsilon_j d_j)' H_{j+1}(g_{j+1}-\varepsilon_j d_j)},\ \left(\frac{|\varepsilon_j|(\|s_{j+1}\|+|\varepsilon_j|)}{(g_{j+1}-\varepsilon_j d_j)' H_{j+1}(g_{j+1}-\varepsilon_j d_j)}\right)^2\right\} \le \delta_9 \nu_j$$

where

$$\varepsilon_j = \frac{g'_{j+1} p_j}{d'_j p_j} \quad ,$$

δ_9 is a positive constant and $\{\nu_j\}$ is a sequence of positive numbers such that

$$\sum_{j=0}^{\infty} \nu_j$$

is finite and either $\sigma_j = \sigma_j^*$ or $|g'_{j+1} p_j| \le |\nabla F(x_j - \sigma_j^* s_j)' p_j|$.

Condition 3 is trivially satisfied if $\gamma_j = 1$, i.e., for the BFGS - method. If σ_j is the optimal step size then $\varepsilon_j = 0$. For every j, there is therefore an interval, containing the optimal step size, such that every σ_j in this interval satisfies Condition 3.

Since by Lemma 7 the sum

$$\sum_{j=0}^{\infty} \|x_j - z\|$$

is finite and $\|g_j\| = O(\|x_j - z\|)$ (see [12], for instance) it is possible to choose

$$\nu_j = \|g_j\| \quad \text{for} \quad j = 0,1,\ldots \quad .$$

In the next lemma it is shown that Condition 3 implies Condition 1.

Lemma 10

If σ_j satisfies Condition 3, then

i) σ_j satisfies Condition 1

ii) $\sum_{j=0}^{\infty} (\varphi_j + \xi_j)$ is finite.

Proof:

Observing that

$$(4.13)\qquad (g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j) = g'_{j+1}H_{j+1}g_{j+1} - \epsilon_j^2 d_j' p_j \le g'_{j+1}H_{j+1}g_{j+1}$$

we obtain the relation

$$(4.14)\qquad |1-\gamma_j| \left| \frac{(g_{j+1}-\epsilon_j d_j)'G^{-1}(g_{j+1}-\epsilon_j d_j)}{(g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j)} - \frac{g'_{j+1}G^{-1}g_{j+1}}{g'_{j+1}H_{j+1}g_{j+1}} \right| \le$$

$$|1-\gamma_j| \left[\frac{g'_{j+1}G^{-1}g_{j+1}(\epsilon_j^2 d_j' p_j)}{[(g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j)]g'_{j+1}H_{j+1}g_{j+1}} + \frac{|\epsilon_j^2 d_j' G^{-1} d_j - 2\epsilon_j d_j' G^{-1} g_{j+1}|}{(g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j)} \right] = O(\delta_9 \nu_j)$$

where the equality follows from (4.13), Condition 3 and the fact that $\|\epsilon_j d_j\| = O(|g'_{j+1}p_j|)$.

Replacing G^{-1} with the unit matrix we deduce from (4.14) that σ_j satisfies Condition 1. Moreover since it follows from (3.14) that

$$\frac{1}{\gamma_j}\frac{w_j' G^{-1} w_j}{w_j' q_j} = \frac{(g_{j+1}-\epsilon_j d_j)'G^{-1}(g_{j+1}-\epsilon_j d_j)}{(g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j)}$$

we obtain from (4.14) the inequality

$$(4.15)\qquad \varphi_j \le \delta_{10}\nu_j \ , \quad j = 0,1,\dots \ ,$$

for some constant δ_{10}. Since it follows from (2.16), (3.4), and (3.14) that

$$\frac{(g_j+\alpha_j p_j)'G(g_j+\alpha_j p_j)}{w_j' q_j} = \frac{1}{\gamma_j}\frac{(s_{j+1}-\epsilon_j p_j)'G(s_{j+1}-\epsilon_j p_j)}{(g_{j+1}-\epsilon_j d_j)'H_j(g_{j+1}-\epsilon_j d_j)}$$

we have similar to (4.14) the inequality

$$(4.16)\qquad \xi_j \le \frac{|\gamma_j - 1|}{\gamma_j}\left[\frac{s'_{j+1}Gs_{j+1}(\epsilon_j^2 d_j' p_j)}{[g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j)]g'_{j+1}H_{j+1}g_{j+1}} + \frac{|\epsilon_j^2 p_j' G p_j - 2\epsilon_j p_j' G s_{j+1}|}{(g_{j+1}-\epsilon_j d_j)'H_{j+1}(g_{j+1}-\epsilon_j d_j)} \right] \le \delta_{11}\nu_j \ .$$

for some constant δ_{11}.

Using Condition 3 we will now establish the boundedness of the sequence $\{\psi_j\}$ and two important consequences.

Lemma 11

Let Assumptions 1 and 2 be satisfied and suppose that the step size σ_j satisfies Conditions 2 and 3. Set $\eta_j = 1 - \gamma_{j-1}$. Then

i) The sequence $\{\psi_j\}$ is bounded.

ii) The sequences $\{\|H_j(\eta_j)\|\}$ and $\{\|H_j^{-1}(\eta_j)\|\}$ are bounded.

iii) $\|(1-\eta_j)\rho_j q_j - Gp_j\| \to 0$ as $j \to \infty$.

Proof:

i) By (4.4) and Lemma 6 we have for every j

$$\begin{aligned} \psi_{j+1} &\le \psi_j - 2 + \tau_j + \varphi_j + \xi_j + \mu_j \\ &\le \psi_j + \delta_{12}\|x_j - z\| + (\delta_{10}+\delta_{11})\nu_j + \delta_{13}\frac{\|x_j - z\|}{w_j' q_j} , \end{aligned} \tag{4.17}$$

where δ_{12} and δ_{13} are positive constants and the last inequality follows from (4.12), (4.15), (4.16), and Lemma 9. Because for every j

$$\psi_j \ge 1 \quad \text{and} \quad \frac{1}{w_j' q_j} \le \frac{\psi_j}{q_j' G q_j} \le \frac{\psi_j}{\mu} \tag{4.18}$$

we obtain from (4.17) the relation

$$\begin{aligned} \psi_{j+1} &\le \psi_j (1 + (\delta_{12} + \frac{\delta_{13}}{\mu})\|x_j - z\| + (\delta_{10} + \delta_{11})\nu_j) \\ &\le \psi_0 \prod_{i=0}^{j} (1 + \delta_{14}\|x_i - z\| + \delta_{15}\nu_i) , \end{aligned}$$

where $\delta_{14} = \delta_{12} + \frac{\delta_{13}}{\mu}$ and $\delta_{15} = \delta_{10} + \delta_{11}$. Therefore,

$\ln \psi_{j+1} \le \ln \psi_0 + \sum_{i=0}^{j} \ln(1 + \delta_{14}\|x_i - z\| + \delta_{15}\nu_i)$. Since by Lemma 7 and Condition 3 the two sums

$$\sum_{j=0}^{\infty} \|x_j - z\| \quad \text{and} \quad \sum_{j=0}^{\infty} \nu_j$$

are finite this shows that $\{\psi_j\}$ is bounded.

ii) Because $H_j(\eta_j)$ and $H_j^{-1}(\eta_j)$ are positive definite for every j and ψ_j is equal to the sum of the eigenvalues of $H_j(\eta_j)$ and $H_j^{-1}(\eta_j)$ the second statement of the theorem follows from the boundedness of $\{\psi_j\}$.

iii) By (4.4) we have for every j

$$\sum_{i=0}^{j} \left(\frac{p_i' G p_i + (1-\eta_i)^2 \rho_i^2 g_i' G^{-1} g_i}{(1-\eta_i)\rho_i g_i' p_i} - 2 \right) \leq \psi_0 + \sum_{i=0}^{j} (\tau_j - 2 + \varphi_j + \xi_j + \mu_j) \quad . \tag{4.19}$$

Since by Lemmas 7 through 10, inequality (4.18) and part i) of the theorem, we have

$$\sum_{j=0}^{\infty} (\tau_j - 2 + \varphi_j + \xi_j + \mu_j) < \infty \tag{4.20}$$

the inequality (4.19) implies that

$$\frac{p_j' G p_j + (1-\eta_j)^2 \rho_j^2 g_j' G^{-1} g_j}{(1-\eta_j)\rho_j g_j' p_j} \to 2 \quad \text{as} \quad j \to \infty \quad . \tag{4.21}$$

By the second part of the theorem

$$(1-\eta_j)\rho_j g_j' p_j = (1-\eta_j)^2 \rho_j^2 g_j' H_j(\eta_j) g_j$$

is bounded away from zero. Therefore it follows from (4.21) and Lemma 6 that

$$\|(1-\eta_j)\rho_j g_j - G p_j\| \to 0 \quad \text{as} \quad j \to \infty \quad .$$

Before we can use the above results to prove the superlinear convergence of the sequence $\{x_j\}$ to z we need some properties of the two sequences $\{\gamma_j\}$ and $\{\sigma_j^*\}$. These are established in the following two lemmas.

<u>Lemma 12</u>

Let Assumptions 1 and 2 and Conditions 2 and 3 be satisfied. Then for every update formula (2.5) with $\beta_1 + \beta_2 \neq 0$ the following statements hold.

i) If $\beta_1\beta_2 \geq 0$ then $\gamma_j \to 1$ as $j \to \infty$.

ii) If $\beta_1\beta_2 < 0$, then

$$\gamma_j \to 1 \text{ or } -\frac{\beta_1}{\beta_2} \text{ as } j \to \infty.$$

iii) If $\gamma_j \to 1$ as $j \to \infty$, then

$$|1-\gamma_j| = O(\min\{\left(\frac{\|q_{j+1}\|}{\|q_j\|}\right)^2, (d_j'q_j)^2\}) \quad .$$

__Proof__:

Since

$$d_j'q_j = p_j'Gq_j + (d_j - Gp_j)'q_j \tag{4.22}$$

$$= (1-\eta_j)\rho_j q_j'q_j + (Gp_j - (1-\eta_j)\rho_j q_j)'q_j + (d_j - Gp_j)'q_j$$

$$= (Gp_j - (1-\eta_j)\rho_j q_j)'q_j + (d_j - Gp_j)'q_j$$

it follows from Lemmas 8 and 11 that

$$d_j'q_j \to 0 \text{ as } j \to \infty \quad . \tag{4.23}$$

Let $\beta_1\beta_2 > 0$. Because

$$|\beta_1 d_j'p_j + \beta_2 d_j'H_j d_j| \geq |\beta_1| d_j'p_j \geq |\beta_1|\mu > 0$$

and, by part ii) of Lemma 11, $w_j'q_j$ is bounded away from zero it follows from (2.21) and (4.23) that $\gamma_j \to 1$ as $j \to \infty$. Now assume that $\beta_1 = 0$. By (2.20) and (2.17)

$$\frac{1}{\gamma_j} = \frac{\rho_j q_j'p_j}{(d_j'p_j)^2} d_j'H_j d_j = 1 + \frac{\rho_j q_j'p_j}{(d_j'p_j)^2} \frac{(d_j'q_j)^2}{w_j'q_j} \tag{4.24}$$

$$= 1 + \left(\frac{q_j'p_j}{\|q_j\|} \frac{(d_j'q_j)^2}{(d_j'p_j)^2 w_j'q_j}\right) \frac{\|q_j\|}{\|s_j\|}$$

$$\leq 1 + \delta_{16}\left(\frac{q_j'p_j}{\|q_j\|} \frac{(d_j'q_j)^2}{(d_j'p_j)^2 w_j'q_j}\right)\frac{1}{\gamma_{j-1}}$$

for some positive constant δ_{16}, where the inequality follows from the relation

$$H_j(\eta_j)q_j = \frac{1}{1-\eta_j}\frac{p_j}{\rho_j} = \frac{s_j}{\gamma_{j-1}}$$

which by part ii) of Lemma 11 implies

$$\frac{\|q_j\|}{\|s_j\|} = O\left(\frac{1}{\gamma_{j-1}}\right) \quad .$$

Since by Lemma 2, $1/\gamma_j \geq 1$ for $j = 0,1,2,\ldots$ we deduce from (4.23) and (4.24) that $\gamma_j \to 1$ as $j \to \infty$.

Finally let $\beta_1\beta_2 < 0$. By (2.17) we have

$$\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j = d_j'p_j\left(\beta_1 + \beta_2 \frac{d_j'p_j}{\rho_j q_j'p_j} + \beta_2 \frac{(d_j'q_j)^2}{w_j'q_j d_j'p_j}\right) \tag{4.25}$$

Furthermore, since

$$d_j'p_j = p_j'Gp_j + (d_j - Gp_j)'p_j, \quad \gamma_{j-1}\rho_j q_j'p_j = p_jGp_j + (\gamma_{j-1}\rho_j q_j - Gp_j)'p_j$$

it follows from Lemmas 8 and 11 that

$$\frac{d_j'p_j}{\rho_j q_j'p_j} = \gamma_{j-1}\frac{d_j'p_j}{\gamma_{j-1}\rho_j q_j'p_j} = \gamma_{j-1}(1+\varepsilon_j), \quad \varepsilon_j \to 0 \quad \text{as} \quad j \to \infty \quad . \tag{4.26}$$

Using (4.23, (4.25), and (4.26) we see that there is $\varepsilon > 0$ and j_0 such that $|1 - \gamma_{j-1}| < \varepsilon$ and $j > j_0$ imply

$$|\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j| \geq \frac{d_j'p_j}{2}|\beta_1 + \beta_2| > 0 \quad .$$

Therefore, it follows from (2.21) and (4.23) that the sequence $\{|1 - \gamma_j|\}$ either converges to zero or is bounded away from zero. In the latter case (2.21) and (4.23) imply that $\beta_1 d_j'p_j + \beta_2 d_j'H_jd_j \to 0$ as $j \to \infty$ which by (4.25) and (4.26) shows that $\beta_1 + \beta_2\gamma_{j-1} \to 0$ as $j \to \infty$.

Finally assume that $\gamma_j \to 1$ as $j \to \infty$. By (4.25) and (4.26) this implies that

(4.27) $$|\beta_1 d_j' p_j + \beta_2 d_j' H_j d_j| \geq \frac{1}{2} d_j' p_j |\beta_1 + \beta_2| \geq \frac{\mu}{2} |\beta_1 + \beta_2| > 0$$

for j sufficiently large. Since by definition

$$d_j' q_j = \frac{(q_j - q_{j+1})' q_j}{\|\sigma_j s_j\|} = \frac{-q_{j+1}' q_j}{\|\sigma_j s_j\|} = \frac{-q_{j+1}' q_j}{\|q_j\|} \frac{\|q_j\|}{\|\sigma_j s_j\|}$$

and, under Assumption 2, $\{\|q_j\| / \|\sigma_j s_j\|\}$ is bounded it follows from (2.21) and (4.27) that

$$|1 - \gamma_j| = O(\min\{\left(\frac{\|q_{j+1}\|}{\|q_j\|}\right)^2, (d_j' q_j)^2\}) .$$

Lemma 13

Let Assumptions 1 and 2 and Conditions 2 and 3 be satisfied. Then for every update formula (2.5) with $\beta_1 + \beta_2 \neq 0$ and for j sufficiently large

$$\nabla F(x_j - \sigma_j^* s_j)' p_j \leq \gamma^* q_j' p_j$$

$$F(x_j - \sigma_j^* s_j) \leq F(x_j) - \gamma \|\sigma_j^* s_j\| q_j' p_j .$$

Proof:

First assume that $\beta_2 = 0$, i.e. $\gamma_j = 1$. Then $\sigma_j^* = 1$. By Taylor's theorem there is v_j in the set

(4.28) $$\{x \mid x = x_j - t s_j,\ 0 \leq t \leq 1\}$$

such that

$$\nabla F(x_j - s_j)' p_j = q_j' p_j - p_j' G(v_j) s_j$$

$$= q_j' p_j \frac{\|s_j\|}{q_j' p_j} (p_j'(G - G(v_j)) p_j - p_j'(G p_j - \sigma_j q_j) p_j) .$$

Since by Lemma 11, $\|s_j\| = O(q_j' p_j)$ and $\|G p_j - \sigma_j q_j\| \to 0$ as $j \to \infty$ this implies that $\nabla F(x_j - s_j)' p_j \leq \gamma^* q_j' p_j$ for j sufficiently large. Furthermore, there is y_j in the set (4.28) such that

(4.29) $$F(x_j - s_j) - F(x_j) = -\|s_j\| q_j' p_j + \frac{\|s_j\|^2}{2} p_j' G(y_j) p_j .$$

Since

$$p_j'G(y_j)p_j = \rho_j g_j'p_j + (Gp_j - \rho_j g_j)'p_j + p_j'(G(y_j) - G)p_j$$

it follows from (4.29) and Lemma 11 that

$$F(x_j-s_j) - F(x_j) \le -\|s_j\|g_j'p_j(\frac{1}{2} - \frac{\|s_j\|}{2g_j'p_j}(\|Gp_j-\rho_j g_j\| - \|G(y_j)-G\|))$$

$$\le -\gamma\|s_j\|g_j'p_j$$

for j sufficiently large.

Now suppose that $\beta_2 \neq 0$. Then it follows from (3.20) that

$$\nabla F(x_j - \sigma_j^* s_j)'p_j \le \gamma^* g_j'p_j$$

for j sufficiently large. Finally by Taylor's theorem there is

$$v_j \in \{x \mid x = x_j - t(\sigma_j^* s_j),\quad 0 \le t \le 1\}$$

such that

$$F(x_j-\sigma_j^* s_j) - F(x_j) = -\|\sigma_j^* s_j\|g_j'p_j + \frac{\|\sigma_j^* s_j\|^2}{2} p_j'G(v_j)p_j \quad .$$

Since

$$\sigma_j^* = \frac{g_j's_j}{2(F(x_j-s_j)-F(x_j)+g_j's_j)} = \frac{g_j's_j}{s_j'G(y_j)s_j} \tag{4.30}$$

for some y_j in the set (4.28), the above equality and (4.30) imply that

$$F(x_j - \sigma_j^* s_j) - F(x_j) - \frac{1}{2}\|\sigma_j^* s_j\|g_j'p_j \to 0 \quad \text{as} \quad j \to \infty.$$

Because $\gamma < \frac{1}{2}$ this completes the proof.

We are now ready to prove the main result of this section.

<u>Theorem 3</u>

Let Assumptions 1 and 2 and Conditions 2 and 3 be satisfied. Then for every update formula (2.5) with $\beta_1 + \beta_2 \neq 0$ the following statements hold.

i) The sequences $\{\|H_j\|\}$ and $\{\|H_j^{-1}\|\}$ are bounded.

ii) $\dfrac{\|g_{j+1}\|}{\|g_j\|} \to 0$ as $j \to \infty$, $\dfrac{\|x_{j+1}-z\|}{\|x_j-z\|} \to 0$ as $j \to \infty$.

iii) The two sums

$$\sum_{j=0}^{\infty} \left(\frac{\|g_{j+1}\|}{\|g_j\|}\right)^2 \quad \text{and} \quad \sum_{j=0}^{\infty} \left(\frac{\|x_{j+1}-z\|}{\|x_j-z\|}\right)^2$$

are finite.

iv) If $\beta_1\beta_2 \geq 0$ or $\beta_1\beta_2 < 0$ and $\gamma_j \to 1$ as $j \to \infty$, then

$$\sigma_j \to 1 \text{ as } j \to \infty \text{ and } \hat{\sigma}_j \to 1 \text{ as } j \to \infty ,$$

where $\hat{\sigma}_j$ denotes the optimal step size.

v) If $\beta_1\beta_2 < 0$ and $\gamma_j \to -\dfrac{\beta_1}{\beta_2}$ as $j \to \infty$ then

$$\sigma_j \to -\frac{\beta_2}{\beta_1} \text{ as } j \to \infty \text{ and } \hat{\sigma}_j \to -\frac{\beta_2}{\beta_1} \text{ as } j \to \infty .$$

vi) If $\beta_1\beta_2 \geq 0$ or $\beta_1\beta_2 < 0$ and $\gamma_j \to 1$ as $j \to \infty$ then

$$\sigma_j = \sigma_j^*$$

for j sufficiently large, provided $\|g_j\| = O(\nu_j)$.

Proof:

The first statement of the theorem follows immediately from part ii) of Lemma 11 and parts i) and ii) of Lemma 12.

Since $1-\eta_j = \gamma_{j-1}$ it follows from (4.19) and (4.20) that the sum

$$\sum_{j=0}^{\infty} \left(\frac{p_j'Gp_j+\gamma_{j-1}^2\rho_j^2g_j'G^{-1}g_j}{\gamma_{j-1}\rho_jg_j'p_j} - 2\right)$$

is finite. By Lemmas 6 and 11 this implies

(4.31) $$\sum_{j=0}^{\infty} \|\gamma_{j-1}\rho_j g_j - Gp_j\|^2 < \infty .$$

Furthermore it follows from (4.10) and (4.11) that

(4.32) $$\frac{\|g_{j+1}\|}{\|g_j\|} \le \left\| \frac{g_j}{\|g_j\|} - G \frac{\sigma_j s_j}{\|g_j\|} \right\| + \|E_j\| \frac{\|\sigma_j s_j\|}{\|g_j\|}, \quad \|E_j\| = O(\|x_j - z\|) .$$

First assume that $\beta_2 = 0$, i.e., $\gamma_j = 1$, for $j = 0,1,2,\ldots$. Then Lemma 13 implies that

(4.33) $$\sigma_j = \sigma_j^* = 1 \text{ for } j \text{ sufficiently large} .$$

Since

$$\|\rho_j g_j - Gp_j\| = \frac{\|g_j\|}{\|s_j\|} \left\| \frac{g_j}{\|g_j\|} - G \frac{s_j}{\|g_j\|} \right\|$$

and, by the first part of the theorem, $\{\|g_j\|/\|s_j\|\}$ is bounded we deduce from (4.5), (4.31), (4.32), and (4.33) that

(4.34) $$\frac{\|g_{j+1}\|}{\|g_j\|} \to 0 \text{ as } j \to \infty \text{ and } \sum_{j=0}^{\infty} \left(\frac{\|g_{j+1}\|}{\|g_j\|} \right)^2 < \infty .$$

Now suppose that $\beta_2 \ne 0$. Then it follows from Condition 3 that either

$$\sigma_j = \sigma_j^* \quad \text{or} \quad |g_{j+1}' p_j| \le |\nabla F(x_j - \sigma_j^* s)' p_j| .$$

By (3.20) this implies

(4.35) $$\frac{|g_{j+1}' p_j|}{\|g_j\|} = O(\|g_j\|) .$$

Furthermore, because (see [12] for instance)

(4.36) $$\|g_j\| = O(\|x_j - z\|) \quad \text{and} \quad \|x_j - z\| = O(\|g_j\|)$$

we conclude from (4.35) and Lemma 7 that

(4.37) $$\sum_{j=0}^{\infty} \frac{|g_{j+1}' p_j|}{\|g_j\|} < \infty .$$

By (4.22), (4.31), and Lemmas 7 and 8 we have

$$\sum_{j=0}^{\infty}\left(\frac{\|q_j\|}{\|\sigma_j s_j\|}\ \frac{q'_{j+1}q_j}{\|q_j\|}\right)^2 = \sum_{j=0}^{\infty}(d'_j q_j)^2 < \infty \quad ,$$

which implies

(4.38)
$$\sum_{j=0}^{\infty}\left(\frac{q'_{j+1}q_j}{\|q_j\|}\right)^2 < \infty$$

since, under Assumption 2, $\{\|q_j\|/\|\sigma_j s_j\|\}$ is bounded away from zero.

Observing that $q_{j+1} \in \text{span}\{\rho_j q_j,\ w_j\}$ we deduce from the first part of the theorem that

(4.39)
$$\frac{\|q_{j+1}\|}{\|q_j\|} = O(\max\{\frac{|q'_{j+1}p_j|}{\|q_j\|},\ \frac{|q'_{j+1}q_j|}{\|q_j\|}\}) \quad .$$

Therefore, the second and third part of the theorem follows from (4.34) and (4.36) through (4.39).

In order to prove the next two parts of the theorem we use Taylor's theorem to show that there is

$$v_j \in \{x \mid x = x_j - t(\hat{\sigma}_j s_j),\ 0 \le t \le 1\}$$

such that

(4.40)
$$\gamma_{j-1}\hat{\sigma}_j = \frac{\gamma_{j-1}q'_j p_j}{p'_j G(v_j) s_j} = \frac{\gamma_{j-1}\rho_j q'_j p_j}{p'_j G(v_j) p_j}$$

$$= \frac{p'_j G p_j + (\gamma_{j-1}\rho_j q_j - G p_j)' p_j}{p'_j G p_j + p_j (G(v_j) - G) p_j} \quad .$$

Since $G(v_j) - G \to 0$ as $j \to \infty$, it follows from Lemma 11 that

(4.41)
$$\gamma_{j-1}\hat{\sigma}_j \to 1 \quad \text{as} \quad j \to \infty \quad .$$

If $\beta_2 = 0$, then $\gamma_j = 1$ and $\sigma^*_j = 1$ for all j and the parts iv) and v) of the theorem follow from (4.41) and Lemmas 12 and 13. Let $\beta_2 \neq 0$, then we obtain from (4.30) and (4.40) the relation

(4.42) $$|\sigma_j^* - \hat{\sigma}_j| = \left|\frac{g_j'p_j}{p_j'G(y_j)s_j} - \frac{g_j'p_j}{p_jG(v_j)s_j}\right|$$

$$= \frac{g_j'p_j}{\|g_j\|}\,\frac{\|g_j\|}{\|s_j\|}\left|\frac{p_j'(G(v_j)-G(y_j)p_j}{(p_j'G(y_j)p_j)\,(p_j'G(v_j)p_j)}\right|$$

$$\leq \frac{g_j'p_j}{\|g_j\|}\,\frac{\|g_j\|}{\|s_j\|}\,\frac{\|G(v_j)-G(y_j)\|}{\mu^2}$$

$$\to 0 \quad \text{as} \quad j \to \infty \ .$$

Finally we deduce from Condition 3 and Taylor's theorem the inequalities

(4.43) $$\mu\|s_j\|\ |\sigma_j-\hat{\sigma}_j| \leq |g_{j+1}'p_j| \leq |\nabla F(x_j-\sigma_j^*s_j)'p_j| \leq \eta\|s_j\|\ |\sigma_j^*-\hat{\sigma}_j| \ .$$

Therefore the parts iv) and v) of the theorem are a consequence of Lemmas 12, 13 and (4.41) through (4.43).

To complete the proof of the theorem we observe that in view of Lemma 13 it suffices to show that if we set $\sigma_j = \sigma_j^*$, then the resulting γ_j and g_{j+1} satisfy the two inequalities of Condition 3 for j sufficiently large.

Because $(g_{j+1} - \epsilon_j d_j)'p_j = 0$, $d_j'p_j \geq \mu$ and $\|d_j\| \leq \eta$ it follows that the sequence

$$\left\{\left\|\frac{g_{j+1}}{\|g_{j+1}\|} - \frac{\epsilon_j}{\|g_{j+1}\|}\,d_j\right\|\right\}$$

is bounded away from zero. By the first part of the theorem this implies that the sequence

$$\left\{\frac{\|g_{j+1}\|^2}{(g_{j+1}-\epsilon_jd_j)'H_{j+1}(g_{j+1}-\epsilon_jd_j)}\right\}$$

is bounded. Therefore we obtain from part iii) of Lemma 12 the relation

(4.44) $$|1-\gamma_j|\,\frac{|\epsilon_j|\ \|g_{j+1}\|}{(g_{j+1}-\epsilon_jd_j)'H_{j+1}(g_{j+1}-\epsilon_jd_j)} = 0\left(\frac{\|g_{j+1}\|^2}{\|g_j\|^2}\,\frac{|g_{j+1}'p_j|}{\|g_{j+1}\|}\right) \ .$$

Since $\|g_j\| = 0(\nu_j)$, $\|g_{j+1}\|/\|g_j\| \to 0$ as $j \to \infty$ and, by (3.20), $|g_{j+1}'p_j| = 0(\|g_j\|^2)$ it follows from (4.44) that the first inequality of Condition 3 is satisfied for j sufficiently large. A completely anologuous argument shows that the second inequality is satisfied too, if j if sufficiently large.

REFERENCES

[1] C. G. Broyden, "Quasi-Newton methods and their application to function minimization", Mathematics of Computation 21 (1967), pp. 368-381.

[2] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms", Parts 1 and 2, Journal of the Institute of Mathematics and its Applications 6 (1970), pp. 76-90, pp. 222-231.

[3] C. G. Broyden, J. E. Dennis, Jr. and J. J. Moré, "On the local and superlinear convergence of quasi-Newton methods", Journal of the Institute of Mathematics and its Applications 12 (1973), pp. 223-245.

[4] W. C. Davidon, "Variable metric methods for minimization", Argonne National Laboratories rept. ANL-5990 (1959).

[5] L. C. W. Dixon, "Variable metric algorithms: necessary and sufficient conditions for identical behaviour on non-quadratic functions", Journal of Optimization Theory and Applications 10 (1972), pp. 34-40.

[6] R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization", The Computer Journal 6 (1963), pp. 163-168.

[7] R. Fletcher, "A new approach to variable metric algorithms", The Computer Journal 13 (1970), pp. 317-322.

[8] D. Goldfarb, "A family of variable metric methods derived by variational means", Mathematics of Computation 24 (1970), pp. 23-26.

[9] H. Y. Huang, "Unified approach to quadratically convergent algorithms for function minimization", Journal of Optimization Theory and Applications 5 (1970), pp. 405-423.

[10] M. L. Lenard, "Practical convergence conditions for the Davidon-Fletcher-Powell-method", Mathematical Programming 9 (1975), pp. 69-86.

[11] J. D. Pearson, "Variable metric methods of minimization", The Computer Journal 12 (1969), pp. 171-178.

[12] M. J. D. Powell, "On the convergence of the variable metric algorithm", Journal of the Institute of Mathematics and its Applications 7 (1971), pp. 21-36.

[13] M. J. D. Powell, "Some properties of the variable metric algorithm", Numerical Methods for Non-linear Optimization, F. A. Lootsma, ed., Academic Press, London, 1972.

[14] M. J. D. Powell, "Some global convergence properties of a variable metric algorithm for minimization without exact line searches, Nonlinear Programming, SIAM-AMS Proceedings, Vol. 9, American Mathematical Society, Providence, R.I., 1976.

[15] K. Ritter, "A method of conjugate directions for linearily constrained nonlinear programming problems", SIAM Journal on Numerical Analysis 12 (1975), pp. 273-303.

[16] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization", Mathematics of Computation 24 (1970), pp. 647-656.

[17] J. Stoer, "On the convergence rate of imperfect minimization algorithms in Broyden's β-class", Mathematical Programming 9 (1975), pp. 313-335.

KR/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #1945	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9 Technical
4. TITLE (and Subtitle) 6 GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF VARIABLE METRIC METHODS.		5. TYPE OF REPORT & PERIOD COVERED Summary Report - ~~no specific~~ [redacted]
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Klaus Ritter		8. CONTRACT OR GRANT NUMBER(s) 15 DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE 11 Apr[redacted]79
		13. NUMBER OF PAGES 43
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 12 47p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

14 MRC-TSR-1945

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Unconstrained minimization, variable metric method, global convergence, superlinear convergence.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper considers a class of variable metric methods for unconstrained minimization. Without requiring exact line searches it is shown that, under appropriate assumptions on the function to be minimized, each algorithm in this class converges globally and superlinearly.

221200 Gu

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)