Applied and
Computational
Mathematics
Division

NISTIR 4554

Computing and Applied Mathematics Laboratory

Convergence Properties of a Class of Rank-Two Updates

P.T. Boggs, J.W. Tolle

April 1991

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards and Technology
Gaithersburg, MD 20899

NISTIR 4554

Convergence Properties of a Class of Rank-Two Updates

P. T. Boggs*

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards and Technology
Computing and Applied Mathematics Laboratory
Applied and Computational Mathematics Division
Gaithersburg, MD 20899

J. W. Tolle**

Department of Mathematics and Operations Research,
University of North Carolina
Chapel Hill, NC 27599

***Research supported by the Air Force Office of Scientific Research, #AFOSR-ISSA-90-0004**

****Research supported by the Air Force Office of Scientific Research, #AFOSR-88-0267**

April 1991

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
John W. Lyons, Director

ABSTRACT

Many optimization algorithms generate, at each iteration, a pair (x_k, H_k) consisting of an approximation to the solution, x_k, and a Hessian matrix approximation, H_k, which contains local second order information about the problem. Much is known about the convergence of the x_k to the solution of the problem but relatively little about the behavior of the sequence of matrix approximations. We analyze the sequence $\{H_k\}$ generated by the extended Broyden class of updating schemes independently of the optimization setting in which they are used, deriving various conditions under which convergence is assured and delineating the structure of the limits. Rates of convergence are also obtained. Our results extend and clarify those already in the literature.

Key words: optimization, quasi-Newton algorithms, matrix updates, convergence analysis

1. Introduction

In this paper we investigate a general class of matrix updating schemes that are of major interest in optimization problems, both constrained and unconstrained. We consider the possible convergence of sequences generated by these schemes, attempting to isolate the crucial factors that determine convergence independent of any particular optimization setting. The results presented thus subsume and clarify some recent results on convergence of Hessian approximations in specific optimization problems.

In many iterative numerical optimization algorithms each iteration requires the solution (or partial solution) of an approximate quadratic model of the problem under consideration. Each surrogate model is usually characterized by the current iterate, x_k, and a symmetric, often positive definite, matrix, H_k. The solution of the quadratic model yields a step, s_k, from which the next iterate, $x_{k+1} = x_k + s_k$ is obtained. Once the new iterate is located a new quadratic model, i.e., a new matrix, H_{k+1}, is chosen. The choice is determined by a particular formula involving the information on hand, including the preceding matrix H_k, the step s_k, and the values of the problem functions at x_k and x_{k+1}. The success of the algorithm is known to be sensitive in many cases to the method by which the new matrix is chosen; in particular, the best results are to be expected when the quadratic model closely reflects the important features of the underlying problem as the solution is approached. For this reason the convergence properties of the sequence $\{H_k\}$ are of significant interest. This paper is devoted to the study of these properties.

We examine an iteration scheme of the form

$$H_{k+1} = U(\phi_k, H_k, s_k, y_k) \tag{1.1}$$

where the H_k are N x N symmetric matrices, the ϕ_k are scalar parameters, the s_k are given vectors in $\Re^N$, the y_k are vectors which depend on the s_k in some way to be specified, and the function U is the update formula. Of particular interest will be how the convergence (and rate of convergence) of the sequence $\{H_k\}$ depends on the choice of the ϕ_k, the initial matrix H_0, and the sequence $\{s_k\}$ for a

given U.

The primary motivation for this study is the use of these updates in those optimization problems where the function U adds a rank-two matrix to the current matrix H_k in such a way that the secant condition

$$H_{k+1} s_k = y_k \tag{1.2}$$

is satisfied for each k. The best known example of the use of this updating method takes place in quasi-Newton algorithms for solving unconstrained optimization problems where the matrix H_k is an approximation to the Hessian of the objective function. In this case s_k is the step generated by

$$s_k = - \alpha_k H_k^{-1} \nabla f(x_k) \tag{1.3}$$

and

$$y_k = \nabla f(x_k + s_k) - \nabla f(x_k). \tag{1.4}$$

Here f is the objective function, x_k is the current iterate and α_k is a step length parameter. Similar updating schemes are also used in trust region methods for the unconstrained problem and for the sequential quadratic programming algorithms designed to solve constrained optimization problems.

The most commonly used class of updating functions in these cases is the so-called "extended Broyden" class where

$$U(\phi_k, H_k, s_k, y_k) = (I - \frac{y_k s_k^t}{y_k^t s_k}) H_k (I - \frac{s_k y_k^t}{y_k^t s_k}) + \frac{y_k y_k^t}{y_k^t s_k} \tag{1.5}$$

$$- \phi_k (s_k^t H_k s_k) (\frac{y_k}{y_k^t s_k} - \frac{H_k s_k}{s_k^t H_k s_k}) (\frac{y_k}{y_k^t s_k} - \frac{H_k s_k}{s_k^t H_k s_k})^t.$$

The special cases $\phi_k = 0$ and $\phi_k = 1$ correspond to the well-known DFP and BFGS update formulas respectively while the values of $\phi_k \in [0,1]$ give rise to the "restricted Broyden class" of updates. There is also a similar set of updates in which s_k and y_k are interchanged; these updating schemes yield approximations to the inverse Hessian. (See Dennis and Schnabel [1] or Fletcher [2].)

There has been a significant amount of research on the convergence of the matrices generated by these updating schemes, most generally for the case described above in which each s_k is a step in an optimization algorithm. For example, Schuller [3], Ge and Powell [4], and Stoer [5] have determined conditions under which the sequence of matrices converge in the setting of variable metric algorithms where each s_k is given by (1.3). More recently, Byrd, Nocedal, and Yuan [6] have studied the Broyden class of updates seeking to explain the observed superiority of the BFGS update over the DFP update in quasi-Newton algorithms. Conn, Gould, and Toint [7] have analyzed the special case of (1.5) where $U(\theta_k, H_k, s_k, y_k)$ is the symmetric rank one update. More significantly, they do not require that s_k be generated by (1.3) but only that it represent the step $x_{k+1} - x_k$. (However, they do take the vector y_k to be the difference in the gradients at x_{k+1} and x_k.) As they point out, this more general formulation allows the analysis to include the important trust region methods as well as large scale methods that use partitioned matrices.

The results presented herein are in the spirit of the approach in [7] in that the s_k are not required to satisfy (1.3); however, we go a step further and disassociate them from any particular optimization problem, although the setting is motivated by the unconstrained optimization problem. We will establish the convergence of $\{H_k\}$ under only minimal restrictions on $\{s_k\}$ that are not related directly to its being generated by an optimization process. Moreover, we will not require the y_k to satisfy (1.4) but rather that they be related to the s_k in a certain way that includes (1.4) as a special case. In spite of the less restrictive assumptions made on the parameters, some of the properties of the limit of the sequence $\{H_k\}$ are derived under various conditions on the sequence $\{s_k\}$, thus clarifying the results in [3] – [7]. Finally, our results can be applied to most of the common rank-two updating schemes that have been used in unconstrained optimization.

In Section 2 the nuances of the problem to be solved are discussed and a simplified framework in which to analyze the convergence is introduced. Specifically, we consider the class of updates as nonlinear perturbations of a single linear update which derives from setting $\phi_k = 0$ in (1.5). Section 3 contains some basic notation and terminology as well as some fundamental results on difference equations that are used in the remainder of the paper. In Section 4 the convergence of the linear update is established for the case when the sequence $\{s_k\}$ repeatedly spans $\Re^N$ in a uniform manner and $y_k = s_k$ for each k. This latter condition occurs in the unconstrained optimization of a convex quadratic function. In Section 5 this result is generalized to allow the s_k to approach a proper subspace of $\Re^N$. In particular, the counterexample in [4] to general convergence of the sequence $\{H_k\}$ is clarified. In Section 6 the convergence results are extended to nonlinear updating schemes, i.e., choices of ϕ_k other than zero.

Finally, in Section 7 the condition $y_k = s_k$ is relaxed, permitting the previously obtained results to be applied to the updates of the form (1.5). The basic requirement is only that the difference $y_k - s_k$ converge to zero at a specified rate. Section 8 contains a summary of the major conclusions in the paper, observations on their relevance, and suggestions for further avenues of research.

2. Problem formulation

In the quasi-Newton algorithm for unconstrained optimization problems $s_k = x_{k+1} - x_k$ and y_k is given by (1.4). Assuming that the sequence $\{x_k\}$ converges to the solution x^*, we can write

$$y_k = F(x_k)s_k + O(|s_k|^2) = F(x^*)s_k + (F(x_k) - F(x^*))s_k + O(|s_k|^2).$$

where $F(x)$ is the Hessian of f at x. Thus

$$y_k = F(x^*)s_k + \xi_k$$

where $\frac{\xi_k}{|s_k|} \to 0$. If $F = F(x^*)$ is positive definite then, since the form of the update (1.5) is invariant under the transformations,

$$(T) \qquad s \leftrightarrow F^{-1/2}s, \quad y \leftrightarrow F^{1/2}y, \quad \text{and} \quad H \leftrightarrow F^{1/2}HF^{1/2},$$

$F(x^*)$ can, without loss of generality, be assumed to be the identity matrix. Using this case as motivation, the underlying assumption for our analysis of (1.5) will be that

$$(2.1) \qquad y_k = s_k + \xi_k$$

where $\frac{\xi_k}{|s_k|} \to 0$ at a rate to be specified. Note that if the objective function is quadratic then $\xi_k = 0$. Initially, we will analyze this special case with $y_k = s_k$. Then we will allow the more general formula (2.1) as a perturbation of this simpler situation. Further comments on this assumption are made in the course of the paper.

With the assumption that $y_k = s_k$, the update scheme generated by (1.5) no longer depends on the length of s_k and can be written:

$$(2.2) \qquad \hat{H} = U(\phi, H, s, s)$$

$$= (I - ss^t)H(I - ss^t) + ss^t - \phi(s^tHs)\left(s - \frac{Hs}{s^tHs}\right)\left(s - \frac{Hs}{s^tHs}\right)^t$$

where the subscript k has been deleted, $\hat{H}$ has replaced H_{k+1}, and $|s| = 1$. It is observed that both the Broyden class of update formulas for the approximations to the inverse Hessian and the Hessian satisfy this equation when $y = s$. In fact, (2.2) represents the most general rank-two symmetric update formula that uses the vectors s and Hs and satisfies the secant equation (1.2) (which is now $\hat{H}s = s$). Thus (2.2) represents a rather general starting point for analyzing the convergence properties of rank-two updates. It should be noted that other rank-two update formulas, such as the well-known PSB update, that are not members of the Broyden class may not be able to be placed in this form because they are not invariant under the transformations (T); that is, it cannot be assumed that the Hessian of the objective function is the identity matrix. As is noted in Section 7, however, in certain cases the convergence of the updates may still be analyzed by the methods we develop.

The two special cases identified in Section 1 are singled out for emphasis: $\phi = 0$, which yields

$$(L) \qquad \hat{H} = (I - ss^t)H(I - ss^t) + ss^t,$$

and $\phi = 1$, which gives

$$(C) \qquad \hat{H} = H + ss^t - \frac{Hss^tH}{s^tHs}.$$

These two updating formulas, which we will hereafter refer to as the "L update" (L for linear) and the "C update" (for complementary), can be thought of as versions of the well known DFP and BFGS updating formulas. The reader should be careful to note that the BFGS, DFP, and other updates are not uniquely identified in the formula (2.2) because the assumption that $y = s$ removes the reference to the particular optimization setting that might generate s. For example, the L update is derived from the DFP update for the approximation of the Hessian matrix or, alternatively, the BFGS update for approximating the inverse Hessian matrix. Similarly, the C update derives from the direct BFGS or inverse DFP update.

We will first be concerned with the convergence of the sequence $\{H_k\}$ generated by the iterative scheme (2.2) for arbitrary positive definite starting matrices H_0 (although many of our results will apply for arbitrary symmetric matrices) and specified sequences of unit vectors $\{s_k\}$. Two types of methods of choosing the parameter ϕ_k can be considered: a "static" procedure where ϕ_k is constant over all iterations and a "dynamic" method where the choice of ϕ_k will vary from iteration to iteration and may depend on the current values of s_k and H_k. (If the dynamic method is derived from the original update formula (1.5) then it must be invariant under the transformations given above if it is to be considered here.) The static method is typified by the L ($\phi_k = 0$ for all k) and the C ($\phi_k = 1$ for all k) updates and the dynamic method by the "symmetric rank one" update discussed below. In theory, the dynamic approach can be used to take better advantage of the current information and can therefore lead to more rapid convergence of the sequence $\{H_k\}$. This is illustrated by the aforementioned symmetric rank one scheme in which the choice of ϕ_S leads to convergence of the $\{H_k\}$ in a finite number ($\leq N$) of steps whenever the vectors $s_0,\ldots,s_{N-1}$ are linearly independent.

It is useful to consider the inverse corresponding to the update (2.2). Assuming that H and $\hat{H}$ are invertible, an application of the Sherman-Morrison formula yields

$$\hat{H}^{-1} = U(\mu(\phi), H^{-1}, s, s) \tag{2.3}$$

where

$$\mu(\phi) = \frac{(1-\phi)(s^tHs)(s^tH^{-1}s)}{\phi + (1-\phi)(s^tHs)(s^tH^{-1}s)}.$$

It is observed that $\mu(0) = 1$ and $\mu(1) = 0$, illustrating the property that the inverse of the L update is the C update of the inverse and vice versa. Moreover, if H is positive definite then $(s^tHs)(s^tH^{-1}s) \geq 1$, and it is seen that $0 < \phi < 1$ implies that $\mu(\phi)$ falls between 0 and 1. Thus the inverse of a restricted Broyden update of a positive definite matrix can be obtained by a (dynamic) restricted Broyden update of the inverse of the matrix.

It should also be observed that if H is positive definite then so is H^{-1} and thus, since the L and C updates preserve positive definiteness, (2.2) and (2.3) imply, by the interleaving eigenvalues theorem (see [8]), that $\hat{H}$ and $\hat{H}^{-1}$ are positive definite if either ϕ or $\mu(\phi)$ is less than or equal to zero. More specifically we have

Proposition 2.1 (Fletcher [2]): Let H be positive definite. Then the iterate, $\hat{H}$, given by (2.2) is positive definite if and only if $\phi < \phi_C$ where

$$\phi_C = - \frac{(s^t Hs)(s^t H^{-1}s)}{1-(s^t Hs)(s^t H^{-1}s)}.$$

Proof: Since $(s^t Hs)(s^t H^{-1}s) > 1$ clearly $\phi_C > 1$. If the inequality holds then either $\phi < 1$ or $\mu(\phi) < 0$ and by the remarks above $\hat{H}$ is positive definite. If $\phi = \phi_C$ then from (2.3) it is seen that $\hat{H}$ is singular. From (2.2) it follows that if $\hat{H}$ is positive definite for $\phi = \phi_1$ and for $\phi = \phi_2 > \phi_1$ then it is positive definite in (ϕ_1, ϕ_2). Thus $\hat{H}$ cannot be positive definite for any $\phi > \phi_C$. •

The case where $\phi = \phi_S = - \frac{s^t Hs}{1 - s^t Hs}$ is of special interest because the rank-two update formula (2.2) reduces to the rank one update

$$\text{(SR1)} \qquad \hat{H} = H + \frac{(s - Hs)(s - Hs)^t}{s^t(s - Hs)}$$

known as the "symmetric rank one" (SR1) update. Note that the update formula for $\hat{H}^{-1}$ is identical to that of $\hat{H}$ for the SR1 update (with H^{-1} replacing H). If H is positive definite the SR1 update is never a restricted Broyden update because ϕ_S is either greater than one (if $s^t Hs > 1$) or less than zero (if $s^t Hs < 1$). Moreover, $\hat{H}$ may not be positive definite when H is positive definite if $\phi_S > 1$; in fact, the SR1 update will not exist if $(I - H)s$ is orthogonal to s. It follows from Proposition 2.1 and the definition of ϕ_S that the SR1 update is positive definite if and only if either $s^t Hs < 1$ $(\phi_S < 0)$ or $s^t H^{-1}s < 1$ $(1 < \phi_S < \phi_C)$.

The convergence of the static schemes corresponding to the L and C updates are interrelated as shown in the following proposition.

Proposition 2.2: Let P be any subset of the symmetric matrices for which $H \in P$ implies $H^{-1} \in P$. If for a given sequence $\{s_k\}$ the sequence $\{H_k\}$ generated by the L update converges for every initial matrix chosen from P, then the same is true for the C update, and conversely.

Proof: Suppose that the sequence generated by the L updating formula converges to $L(H_0)$ for a starting matrix $H_0 \in P$. Then the sequence $\{H_k^{-1}\}$ converges to $L(H_0)^{-1}$. But the sequence $\{H_k^{-1}\}$ is generated by the C updating formula starting from H_0^{-1} and so the C formula yields convergent sequences starting in P. •

It is well-known that for any choice of linearly independent $\{s_k\}$, $k = 0,\dots,N-1$, and any initial H_0 the H_k generated by the SR1 update satisfy the "hereditary" secant condition

$$H_{k+1} s_j = s_j, \quad j = 0,\dots,k. \tag{2.4}$$

It is easily shown by induction that this condition is also satisfied by all updates of the form (2.2) if the $\{s_k\}$ are orthogonal (corresponding to conjugate directions in the quadratic optimization setting). The condition (2.4) implies that the H_k converge in at most N steps and if the full N steps are required then the limiting matrix is the identity. Thus the iteration matrices converge finitely for any sequence $\{\phi_k\}$ when the $\{s_k\}$ are orthogonal and for the SR1 updates when the $\{s_k\}$ are linearly independent. In both cases the limiting matrix is the identity matrix and so it is natural to expect that, generally, if the matrices converge nonfinitely they will also converge to the identity matrix (other possibilities are considered in Section 5). Accordingly, it is somewhat simpler to consider the convergence of the sequence of matrices $\{E_k\}$ where $E_k = H_k - I$. Using the relations $s_k^t s_k = 1$ and $H_k s_k = E_k s_k + s_k$ in (2.2), the E_k can be seen to satisfy the difference equation

$$\begin{aligned} E_{k+1} &= (I - s_k s_k^t) E_k (I - s_k s_k^t) \\ &\quad - \phi_k \left(\frac{1}{1 + s_k^t E_k s_k} \right) [(s_k^t E_k s_k) s_k - E_k s_k] [(s_k^t E_k s_k) s_k - E_k s_k]^t. \end{aligned} \tag{2.5}$$

It is this form of the iteration scheme that we shall analyze first. Note that the quasi-Newton condition for the E_k becomes $E_{k+1} s_k = 0$ so that the matrices E_k, $k \geq 1$, are always singular and hence never positive definite; however, they are symmetric if E_0 is symmetric. The eigenvectors of H_k and E_k are identical and the eigenvalues of H_k are exactly one unit greater than the eigenvalues of E_k. Thus if the H_k are positive definite, the eigenvalues of the E_k are greater than negative one and conversely. As a result, (2.5) is well-defined provided positive definiteness of the $\{H_k\}$ is

maintained, as in the L and C updates. However, if the matrix H_k is not positive definite then $s_k{}^t E_k s_k + 1$ can be zero and the resulting E_{k+1} undefined.

It is a well-known fact that both the L and C updating formulas (and therefore those of the restricted Broyden class) lead to sequences $\{ E_k \}$ whose Frobenius norms are nonincreasing and, in fact, are strictly decreasing if $E_k s_k \neq 0$. (See, for instance, [4] or [5]). However, as illustrated in [4], this does not ensure convergence of the matrices $\{ E_k \}$ nor does it yield any limiting matrix when the sequence does converge. It should be noted that not all updates of the form (2.5) lead to sequences $\{ E_k \}$ that decrease in the Frobenius norm. For example, the Frobenius norms of the matrices E_k generated by the SR1 update can have large jumps in value – although convergence within N steps is assured. The analysis of the next sections is intended to shed light on the convergence of these matrices and their limits.

3. Notation and preliminary results

Throughout this paper the notation $|x|$ will refer to the Euclidean norm of the vector x. For a matrix A, $||A||$ will denote the operator norm of A, i.e.,

$$||A|| = \sup \{ \frac{|Ax|}{|x|} \},$$

while $||A||_F$ will denote the Frobenius norm,

$$||A||_F = (\sum (a_{ij})^2)^{1/2}.$$

In referring to rates of convergence the term "m-step linear convergence" will mean Q-order convergence; that is, the sequence $\{ x_k \}$ converges to zero with an m-step linear rate if there is a β, $0 \leq \beta < 1$, independent of k such that

$$|x_{k+m} - x^*| \leq \beta |x_k - x^*|$$

for all k sufficiently large. One-step linear convergence will be called simply "linear convergence". We will also refer to R-(order) linear convergence, which requires

$$(|x_k - x^*|)^{1/k} \leq \beta < 1$$

for all k sufficiently large. The important implications for our work are that m-step linear convergence implies R-linear convergence and that the components of a vector converge R-linearly if and only if the vector itself converges R-linearly (which does not hold for m-step linear convergence). It will also be useful to note that a sequence converges R-linearly if and only if it is majorized by a sequence converging (one-step) linearly. That is, $\{x_k\}$ converges R-linearly if and only if there are constants ν and τ, $0 \leq \tau < 1$, such that for all k

$$|x_k - x^*| \leq \nu \cdot \tau^k.$$

More complete information on the definitions and properties of Q and R order convergence can be found in [9].

Many of the proofs in the following sections will depend upon the properties of solutions of difference equations. To simplify the presentations in Sections 4-7 we provide in this section some useful theorems on such systems.

Consider the system of nonlinear difference equations

$$\begin{aligned} z_{k+1} &= A_k z_k + Z_k(z_k, w_k, t_k) \\ w_{k+1} &= w_k + W_k(z_k, w_k, t_k) \end{aligned} \tag{3.1}$$

where $z_k \in \Re^p$, $w_k \in \Re^q$, $t_k \in \Re^n$, and A_k is a p x p matrix. Denote, for any fixed positive integer m,

$$\beta_{k,m} = \| \prod_{i=0}^{m-1} (A_{k+i}) \| . \tag{3.2}$$

The following give convergence results for the solution to (3.1) under a variety of conditions on the matrix A_k and the functions Z_k and W_k.

Theorem 3.1: Suppose that $\{z_k\}$ and $\{w_k\}$ satisfy (3.1) and, in addition,

(i) $\beta_{k,1} \leq 1$, i.e., $\|A_k\| \leq 1$, for every k;

(ii) for some fixed integer m, there is a constant $\beta < 1$ and an infinite sequence of integers $\{k_j\}$ such that $\beta_{k_j,m} \leq \beta$ for each j;

(iii) there exist constants κ_1 and κ_2 independent of k such that $|Z_k(z,w,t)| \leq \kappa_1 \cdot |t|$ and $|W_k(z,w,t)| \leq \kappa_2 \cdot |t|$ for every k; and

(iv) the sequence $\{t_k\}$ converges to zero R-linearly.

Then for any w_0 there exists a vector w^* such that $\{w_k\} \rightarrow w^*$ R-linearly and for any z_0 the sequence $\{z_k\}$ converges to the zero vector. If condition (ii) is replaced by

(ii)$'$ for some fixed integer m, there is a constant $\beta < 1$ such that $\beta_{k,m} \leq \beta$ for every k sufficiently large;

then the convergence rate of $\{z_k\}$ is R-linear and if $Z_k = 0$ for each k the rate is m-step linear.

Theorem 3.2: Suppose that conditions (i), (ii)$'$, and (iv) of Theorem 3.1 hold and, in addition:

(iii)$'$ there are positive constants $\kappa_1, \kappa_2, \kappa_3, \kappa_4$, and δ such that for each k

$$|Z_k(z,w,t)| \leq \kappa_1 |z|^2 + \kappa_2 (|z| + |w|) \cdot |t|$$

and

$$|W_k(z,w,t)| \leq \kappa_3 |z|^2 + \kappa_4 (|z| + |w|) \cdot |t|$$

whenever $|z| < \delta$.

Then there exist positive constants η_0 and τ_0 and a $w^* \in \Re^q$ such that if $|z_0| < \eta_0$ and $|t_k| < \tau_0$ for every k, then the solutions to (3.1) have the properties that $\{z_k\} \rightarrow 0$ and $\{w_k\} \rightarrow w^*$ and the convergence rates are R-linear. If, in particular, $\kappa_2 = 0$ then the convergence rate of $\{z_k\}$ is m-step linear.

Sketches of the proofs of these two theorems can be found in the Appendix.

4. Convergence of the L updating method

When $\phi_k = 0$ (which corresponds to the L update) the iterations (2.5) reduce to

$$E_{k+1} = (I - s_k s_k^t) E_k (I - s_k s_k^t). \tag{4.1}$$

Note that for fixed s_k (4.1) defines a linear transformation from the space of NxN matrices into itself. The other updates of the Broyden class ($\phi_k \neq 0$) can in fact be considered to be nonlinear perturbations of this linear mapping. We will exploit this fact in Section 6 to derive convergence results for the cases where $\phi_k \neq 0$. Denoting the mapping defined by (4.1) as

$$G(s_k ; \cdot): L(\mathfrak{R}^N, \mathfrak{R}^N) \rightarrow L(\mathfrak{R}^N, \mathfrak{R}^N)$$

we have that

$$E_{k+1} = G(s_k ; E_k). \tag{4.2}$$

Since the mapping preserves symmetry, S_m, the set of symmetric NxN matrices, is a $\frac{1}{2}N(N+1)$-dimensional invariant subspace for $G(s_k; \cdot)$. If we assign the Frobenius norm to $L(\mathfrak{R}^N,\mathfrak{R}^N)$ so that it is an inner product space then, as the following proposition shows, the mapping $G(s; \cdot)$ is a projection.

Proposition 4.1: Let s be a given unit vector. The mapping $G(s; \cdot)$ is a projection onto an $(N-1)^2$-dimensional subspace of $L(\mathfrak{R}^N, \mathfrak{R}^N)$. An orthonormal set of eigenvectors corresponding to the eigenvalue zero is

$$ss^t,\ s(v^2)^t,\ s(v^3)^t, \ldots,\ s(v^N)^t,\ v^2 s^t, \ldots,\ v^N s^t,$$

and an orthonormal set of eigenvectors corresponding to the eigenvalue one is

$$v^2(v^2)^t,\ v^2(v^3)^t,\ldots,\ v^2(v^N)^t,\ v^3(v^2)^t,\ldots,\ v^3(v^N)^t,\ldots,\ v^N(v^2)^t,\ldots,\ v^N(v^N)^t.$$

where $s, v^2, \ldots, v^N$ is an orthonormal basis for $\Re^N$.

Proof: Since $G^t = G$ and $G \circ G = G$, G is a projection. The orthonormality follows from the facts that for N-vectors u, v, w, and y ;

$$\| uv^t \|_F = |u||v|$$

and

$$(uv^t)^T(wy^t) = (u^t w)(v^t y)$$

where the superscript "T" represents the inner product in $L(\Re^N, \Re^N)$. The eigenvalue properties are easily checked. •

Since S_m is an invariant subspace of $G(s; \cdot)$, it is of interest to provide an orthonormal basis for S_m consisting of eigenvectors of $G(s; \cdot)$. The following is a direct consequence of Proposition 4.1.

Proposition 4.2: The following symmetric matrices are orthonormal eigenvectors of $G(s; \cdot)$ corresponding to the eigenvalue zero:

$$ss^t,\ \frac{s(v^2)^t + v^2 s^t}{\sqrt{2}},\ldots,\frac{s(v^N)^t + v^N s^t}{\sqrt{2}}\ ;$$

and the following are orthonormal eigenvectors corresponding to the eigenvalue one:

$$v^k(v^k)^t,\ \frac{v^j(v^k)^t + v^k(v^j)^t}{\sqrt{2}},\quad \text{for } k = 2,\ldots,N \text{ and } j = k+1,\ldots,N.$$

Since $G(s;\cdot)$ has eigenvalues equal to one it is not a contraction mapping, even when restricted to S_m, and so linear convergence of the E_k is ruled out in general. Since $G(s;\cdot)$ is a projection matrix it is possible, however, that a sequence of applications will reduce the norm of the E_k, i.e., that some multi-step linear convergence can occur. In particular, we consider the sequence of N applications of G,

$$G_N(\cdot) = G(s_{N-1}, \cdot) \circ G(s_{N-2}, \cdot) \circ \cdots \circ G(s_0, \cdot).$$

We will show that G^N is a contraction mapping provided $s_0, s_1, \ldots, s_{N-1}$ are linearly independent. To do that we must show that the operator norm of G_N,

$$\| G_N \| = \max \{ \| G_N(W) \|_F : W \in L(\Re^N, \Re^N) \text{ and } \| W \|_F = 1 \}$$

is less than one.

By Proposition 4.1 it is seen that there is an orthonormal basis for $L(\Re^N, \Re^N)$ consisting of rank one matrices. If these basis matrices are denoted by $W^1, W^2, \ldots, W^K$ with $K = N^2$, then for any $W \in L(\Re^N, \Re^N)$ with $\| W \|_F = 1$

$$W = \sum_{j=1}^{K} \gamma_j W^j \qquad \text{and} \qquad \sum_{j=1}^{K} (\gamma_j)^2 = 1 .$$

To obtain the desired results we define the set of vectors that generate the basis matrices W^j in a special manner. If we assume that the set of vectors $\{s_0, s_1, \ldots, s_{N-1}\}$ is linearly independent and for each j, $j = 0, \ldots, N\text{-}1$, let $S_j = \text{span}\{s_0, \ldots, s_j\}$, then we may choose the vectors v^j as follows:

(i) $v^0 = s_0$;
(ii) $v^j \in S_j$ and $v^j \in (S_{j-1})^{\perp}$;
(iii) $|v^j| = 1$.

This set of vectors exists (but is not unique) because of the linear independence of the s_j, and it forms an orthonormal basis for $\Re^N$. With this definition we have

$$s_j = \sum_{k=0}^{j} (s_j^t v^k)\, v^k \qquad \text{where} \qquad \sum_{k=0}^{j} (s_j^t v^k)^2 = 1. \tag{4.3}$$

Defining

$$u_0 = \sum_{k=0}^{N-1} \gamma_{0,k}\, v^k \qquad \text{where} \qquad \sum_{k=0}^{N-1} (\gamma_{0,k})^2 = 1 \tag{4.4}$$

and, for $i = 0, \ldots, N\text{-}1$,

$$u_{i+1} = \sum_{k=1}^{N-1} \gamma_{i+1,k}\, v^k = (I - s_i s_i^t) \cdots\cdots (I - s_0 s_0^t)\, u_0 \tag{4.5}$$

we have

Proposition 4.3: The $\gamma_{i,k}$ in (4.5) satisfy the recurrences

$$\gamma_{i+1,k} = \begin{cases} \gamma_{i,k} - (s_i^t v^k)(s_i^t u_i) & k \le i \\ \gamma_{i,k} & k > i \end{cases} \tag{4.6}$$

and

$$\sum_{k=0}^{N-1} (\gamma_{i+1,k})^2 = \sum_{k=0}^{N-1} (\gamma_{i,k})^2 - (s_i^t u_i)^2. \tag{4.7}$$

Proof: From (4.3) and (4.5) we have

$$u_{i+1} = (I - s_i s_i^t)u_i = u_i - (s_i^t u_i)s_i = u_i - (s_i^t u_i)\sum_{k=0}^{i} (s_i^t v^k) v^k$$

and (4.6) follows from equating coefficients of v^k. Then

$$\begin{aligned} \sum_{k=0}^{N-1}(\gamma_{i+1,k})^2 &= \sum_{k=i+1}^{N-1} (\gamma_{i,k})^2 + \sum_{k=0}^{i} (\gamma_{i,k} - (s_i^t v^k)(s_i^t u_i))^2 \\ &= \sum_{k=0}^{N-1} (\gamma_{i,k})^2 - (s_i^t u_i)^2 \end{aligned}$$

where we have used the fact (from (4.3)) that $\sum_{k=0}^{i} (s_i^t v^k)^2 = 1$. •

It now follows from (4.7) that

$$|u_{i+1}|^2 = (1 - \cos^2(\theta_i))\,|u_i|^2$$

where θ_i is the angle between s_i and u_i. Therefore,

$$|u_N|^2 = \prod_{i=0}^{N-1} (1 - \cos^2(\theta_i))\,|u_0|^2. \tag{4.8}$$

Setting

$$R = \prod_{j=0}^{N-1} (I - s_j s_j^t) = (I - s_{N-1}s_{N-1}^t)\cdots(I - s_0 s_0^t), \tag{4.9}$$

it is seen $|R u_0| = |u_N| < |u_0|$ unless $\theta_i = \pi/2$ for all i. However θ_i depends on u_i and hence u_0 so that, in principle, $|u_N|$ could be arbitrarily close to $|u_0|$. We would like to show that the factor in (4.8) actually is bounded away from one for all u_0. This is the content of the next proposition.

Proposition 4.4: Given linearly independent $s_0, \ldots, s_{N-1}$ there is a constant α, $0 < \alpha < 1$, such that for all $u_0 \in \Re^N$

$$|u_N|^2 \leq \alpha^2 |u_0|^2.$$

Proof: We prove the result holds for all u_0 with $|u_0| = 1$; the general case then follows in a straightforward manner. The proof is by contradiction. If the proposition is not true then there exists a sequence of vectors $\{u_0^r\}_{r=1}^{\infty}$ with $|u_0^r| = 1$ and

$$|u_N^r|^2 = |R u_0^r| \geq (1 - \tfrac{1}{r}) |u_0^r|^2.$$

Let $\hat{u}_0$ be any limit point of the sequence $\{u_0^r\}$ and $\hat{u}_N = R \hat{u}_0$. Then $|\hat{u}_N|^2 = |\hat{u}_0|^2 = 1$. Clearly, by (4.8) this implies that $\theta_i = \pi/2$ (or $(s_i^t \hat{u}_i) = 0$) for all i. Therefore, by (4.6), it follows that for all i

$$\hat{u}_i = \sum_{k=0}^{N-1} \hat{\gamma}_{0,k} v^k = \hat{u}_0$$

and hence $s_i^t \hat{u}_0 = 0$. The linear independence of the s_i implies that $\hat{u}_0 = 0$ contradicting the fact that $|\hat{u}_0| = 1$. •

Using the formula $\|V\|_F^2 = \text{trace}(V^t V)$ and the fact that $G_N(u v^t) = R u v^t R^t$ where R is given by (4.9), it is seen that Proposition 4.4 yields

$$\|G_N(u v^t)\|_F = |Ru||Rv| \leq \alpha^4 |u||v| = \alpha^4 \|u v^t\|_F.$$

In order to show that G_N is a contraction mapping it is necessary to extend this inequality to all matrices with unit Frobenius norm.

It now follows from Proposition 4.1 that each basis matrix W^j for $L(\Re^N, \Re^N)$ can be written in the form $v^i(v^k)^t$ for some i and k, $0 \le i \le N\text{-}1$, $0 \le k \le N\text{-}1$. Thus an arbitrary matrix W (with unit Frobenius norm) can be expressed as

$$W = \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} \zeta_{i,k}\, v^i (v^k)^t$$

where

$$\sum_{k=0}^{N-1} \sum_{i=0}^{N-1} (\zeta_{i,k})^2 = 1.$$

For any such matrix W, $G_N(W) = R W R^t$ where R is as given in (4.9). Thus

$$G_N(W) = R \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} \zeta_{i,k}\, v^i (v^k)^t R^t$$

$$= R \sum_{i=0}^{N-1} \beta_i v^i [R \sum_{k=0}^{N-1} \rho_{i,k}\, v^k]^t$$

where

$$\rho_{i,k} = \frac{\zeta_{i,k}}{\beta_i} \qquad \text{and} \qquad (\beta_i)^2 = \sum_{k=0}^{N-1} (\zeta_{i,k})^2.$$

Thus

$$\sum_{k=0}^{N-1} (\rho_{i,\,k})^2 = 1 \qquad \text{and} \qquad \sum_{i=0}^{N-1} (\beta_i)^2 = 1 .$$

Then by Proposition 4.4

$$R \sum_{k=0}^{N-1} \rho_{i,k}\, v^k = \sum_{k=0}^{N-1} \hat{\rho}_{i,k}\, v^k$$

where

$$\sum_{k=0}^{N-1} (\hat{\rho}_{i,k})^2 \le \alpha^2 \sum_{k=0}^{N-1} (\rho_{i,k})^2$$

for each i . Now

$$G_N(W) = R \sum_{i=0}^{N-1} \beta_i\, v^i \sum_{k=0}^{N-1} \hat{\rho}_{i,k}\, (v^k)^t$$

$$= [R \sum_{i=0}^{N-1} \beta_i \, (\sum_{k=0}^{N-1} \hat{\rho}_{i,k})\, v^i] (v^k)^t$$

$$= \sum_{k=0}^{N-1} [R \sum_{i=0}^{N-1} \mu_{i,k}\, v^i] (v^k)^t$$

where $\mu_{i,k} = \beta_i \hat{\rho}_{i,k}$. Using Proposition 4.4 again gives

$$G_N(W) = \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} \hat{\mu}_{i,k} v^i (v^k)^t$$

where for each k

$$\sum_{i=0}^{N-1} (\hat{\mu}_{i,k})^2 \leq \alpha^2 \sum_{i=0}^{N-1} (\mu_{i,k})^2.$$

Therefore,

$$\begin{aligned}(\| G_N(W) \|_F)^2 &= \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} (\hat{\mu}_{i,k})^2 \leq \alpha^2 \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} (u_{i,k})^2 \\ &\leq \alpha^2 \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} (\hat{\rho}_{i,k})^2 (\beta_i)^2 \leq \alpha^4 \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} (\rho_{i,k})^2 (\beta_i)^2 \\ &\leq \alpha^4 \sum_{k=0}^{N-1} \sum_{i=0}^{N-1} (\zeta_{i,k})^2.\end{aligned}$$

Thus we are lead to the following theorem.

Theorem 4.5: If $s_0, \cdots, s_{N-1}$ are linearly independent, then there exists an α, $0 \leq \alpha < 1$, and depending only on $(s_0, \cdots, s_{N-1})$ such that $\| G_N \| \leq \alpha^2 < 1$. Thus the linear transformation G_N is a contraction mapping on the set of NxN matrices.

Before stating our result on the convergence of the matrices generated by (4.1) we provide the following definitions (See also [7]).

Definition 4.1: Let $\{s_k\}$ be a sequence of unit N-vectors. For each k , let

$$G_{N,k} = G(s_{k+N-1}; \cdot) \circ G(s_{k+N-2}; \cdot) \circ \cdots \circ G(s_k ; \cdot)$$

and define

$$\alpha(k) = \| G_{N,k} \|.$$

Definition 4.2: Let $\{s_k\}$ be a sequence of unit $\mathfrak{R}^n$-vectors. If there is an infinite subsequence of the integers, $\{k_j\}$, and a constant α, $0 \leq \alpha < 1$, such that $\alpha(k_j) \leq \alpha$ for each j then the sequence $\{s_k\}$ is said to be subsequentially linearly independent. In this case, if the subsequence $\{k_j\}$ can be chosen so that

$$\chi = \limsup_{j \to \infty} \{k_{j+1} - k_j - 1\}$$

is finite, then the sequence $\{s_k\}$ is said to be uniformly linearly independent with gap χ.

The $\alpha(k)$ of the definition are well-defined by Theorem 4.5. Note that if there is an $\alpha < 1$ such that $\alpha(k) \leq \alpha$ for all k sufficiently large then the gap is zero, in which case we call the sequence $\{s_k\}$ uniformly linearly independent.

To obtain the basic convergence theorem, we put the system (4.2) into the form of the system (3.1) by identifying each E_k with an N^2-dimensional vector z_k and each $G(s_k;\cdot)$ with an $N^2 \times N^2$ matrix A_k. (Z_k and W_k are both zero for all k.) Since we are employing the Frobenius norm on the E_k, we have $|z_k| = \|E_k\|_F$ and also $\|A_k\| = \|G(s_k;\cdot)\|$. The following proposition will help establish the rate of convergence of the $\{E_k\}$.

Proposition 4.6: Let the sequence $\{s_k\}$ be given and for each k let A_k be the $N^2 \times N^2$ matrix representing $G(s_k;\cdot)$. Let $\{\beta_{k,m}\}$ be the constants defined by (3.2). Suppose $\{s_k\}$ is subsequentially linearly independent and let the constant $\alpha < 1$ and the subsequence $\{k_j\}$ be as specified in Definition 4.2. Then

$$\beta_{k,1} \leq 1 \tag{4.10}$$

for all k, and

$$\beta_{k_j,N} \leq \alpha \tag{4.11}$$

for each j. If $\{s_k\}$ is uniformly linearly independent with gap χ then

$$\beta_{k,(N+\chi)} \leq \alpha \tag{4.12}$$

for all k sufficiently large.

Proof: (4.10) and (4.11) follow directly from Definition 4.2. Assuming $\{s_k\}$ is uniformly linearly independent with gap χ, let $\{k_j\}$ be the sequence given by Definition 4.2. Then for every k sufficiently large there is a k_j such that $k_j - \chi \leq k \leq k_j$. Therefore

$$\| \prod_{i=0}^{N+\chi-1} A_{k+i} \| \leq \left(\prod_{i=k-k_j}^{-1} \| A_{k_j+i} \| \right) \cdot \| \left(\prod_{i=0}^{N-1} A_{k_j+i} \right) \| \cdot \left(\prod_{i=0}^{k-k_j+\chi-1} \| A_{k_j+N+i} \| \right)$$

$$\leq \alpha < 1$$

which yields (4.12). •

The main result of this section is now just a straightforward application of Theorem 3.1 (with Z_k and W_k both zero) and Proposition 4.6.

Theorem 4.7: Suppose the sequence $\{s_k\}$ is subsequentially linearly independent. Then the sequence $\{E_k\}$ defined by (4.1) converges to the zero matrix from any symmetric starting matrix E_0. If the sequence $\{s_k\}$ is uniformly linearly independent with gap χ then the convergence rate is $(N+\chi)$-step linear.

5. Extensions of convergence results for the L-update

In the previous section we proved that under relatively weak conditions on the sequence $\{s_k\}$ the L update leads to the convergence of the sequence $\{E_k\}$ regardless of the starting point. In this section we extend that result by relaxing the conditions on the sequence $\{s_k\}$.

As was pointed out in Section 1, sequences $\{s_k\}$ can be constructed for which the corresponding

$\{E_k\}$ will not converge. This was demonstrated by Ge and Powell [4] for the two-dimensional case. In their example, the vectors $\{s_k\}$ are pairwise linearly independent but become more nearly identical while not converging to a single vector. Thus the vectors are not subsequentially linearly independent and they do not converge to a proper subspace of $\Re^2$. This latter property is also necessary for nonconvergence of the $\{E_k\}$ in a sense that will be made clear in the following.

Theorem 4.7 is predicated on the (subsequential) uniform linear independence of the sequence $\{s_k\}$ which requires that they span $\Re^N$. We now relax this restriction and assume instead that they span a proper M-dimensional subspace, S, of $\Re^N$. In order to analyze this case we let P be an $N \times M$ matrix whose columns are an orthonormal basis for S and let Q be an $N \times (N-M)$ matrix whose columns are an orthonormal basis for $S^{\perp}$. For such P and Q any $N \times N$ symmetric matrix E can be written

$$E = PP^tEPP^t + QQ^tEQQ^t + PP^tEQQ^t + QQ^tEPP^t. \tag{5.1}$$

Note that PP^t is the projection of $\Re^N$ onto S and $QQ^t = I - PP^t$ is the projection onto $S^{\perp}$. For each k we let $r_k = P^t s_k$ or, equivalently, $s_k = Pr_k$, where $r_k \in \Re^M$ and $|r_k| = 1$. For a sequence $\{E_k\}$ satisfying (4.1), we set

$$V_k = P^t E_k P, \tag{5.2a}$$

$$U_k = Q^t E_k Q, \tag{5.2b}$$

$$Y_k = P^t E_k Q, \tag{5.2c}$$

so that

$$E_k = PV_kP^t + QU_kQ^t + PY_kQ^t + QY_k{}^tP^t. \tag{5.3}$$

Note that $Y_k = 0$ if and only if S is an eigenspace of E_k. Since $P^tP = I_M$ (the $M \times M$ identity matrix) and $Q^t s_k = 0$, we see from (4.1) that these matrices satisfy the following uncoupled difference equations

$$V_{k+1} = (I_M - r_k r_k{}^t)\, V_k\, (I_M - r_k r_k{}^t) \tag{5.4a}$$

$$U_{k+1} = U_k \tag{5.4b}$$

(5.4c) $$Y_{k+1} = (I_M - r_k r_k^t)\, Y_k .$$

Equation (5.4a) shows that the MxM matrices, V_k, satisfy the updating scheme (4.1); (5.4b) that the $(N-M)x(N-M)$ matrices, U_k, are constant with respect to k; and (5.4c) that the $Mx(N-M)$ matrices, Y_k, satisfy a one-sided type of (4.1) update. Thus if the sequence $\{r_k\}$ is uniformly linearly independent with gap χ it follows from Theorem 4.7 that $\{V_k\}$ converges to the MxM zero matrix and that the rate of convergence is $(M+\chi)$-step linear. A modification of the proofs of the propositions and theorems in Section 4 can be used to show that the Y_k also converge to a zero matrix at the $(M+\chi)$-linear rate. It is perhaps easier, however, to note that the symmetric positive semi-definite matrices $Z_k = Y_k Y_k^t$ satisfy the same equation as the V_k and hence tend to zero; the convergence of the Y_k to the zero matrix R-linearly then follows from the convergence of the Z_k. If the sequence $\{r_k\}$ is only subsequentially linearly independent then the convergence of the V_k and Y_k still holds but there is no multi-step linear rate. As a result of these observations we have the following theorem which shows that the restriction of the $\{s_k\}$ to a subspace does not prevent the convergence of the $\{E_k\}$.

Theorem 5.1: Let S be an M-dimensional subspace of $\Re^N$ and let the matrices P and Q be defined as above. Suppose the sequence of N-vectors $\{s_k\}$ is contained in S and that the sequence of M-vectors, $\{r_k\}$, defined by $r_k = P^t s_k$ is subsequentially linearly independent. Then for any initial symmetric matrix E_0, the sequence defined by (4.1) converges to the matrix $QQ^t E_0 QQ^t$. If the sequence $\{r_k\}$ is uniformly linearly independent with gap χ then the convergence rate is at least R-linear.

Proof: By the remarks above, $V_k \to 0$, $Y_k \to 0$, and $U_k = U_0$ for all k. Thus the convergence result follows from (5.3). The R-linear convergence follows from the fact that the components of E_k have $(M+\chi)$-step linear, and hence R-linear convergence, which implies (by the remarks in Section 3) the R-linear convergence of the E_k. •

The next result is somewhat surprising in that it demonstrates that if $\{s_k\}$ converges at a reasonable rate to a subspace, convergence of the $\{E_k\}$ can still be obtained. The importance of this type of convergence in constrained optimization will be discussed in Section 8.

Definition 5.1: Let S be an M-dimensional subspace of $\Re^N$ and suppose that $\{s_k\}$ is a sequence of unit vectors that satisfy $s_k = u_k + v_k$ where $u_k \in S$, $v_k \in S^{\perp}$, and $v_k \to 0$ as $k \to \infty$. Then we say that $\{s_k\}$ converges to S. If $|v_k|$ tends to zero R-linearly, we say that $\{s_k\}$ converges to S at an R-linear rate.

Theorem 5.2: Suppose that the $\{s_k\}$ converge to an M-dimensional subspace S at an R-linear rate. Let matrices P and Q be defined as above, let $\{r_k\}$ be the sequence of unit M-vectors given by

$$r_k = \frac{P^t s_k}{|P^t s_k|}, \tag{5.5}$$

and assume that $\{r_k\}$ is subsequentially linearly independent. Then for any initial matrix E_0 there is an $(N-M)x(N-M)$ matrix U^* such that the sequence $\{E_k\}$ defined by (4.1) converges to QU^*Q^t. If $\{r_k\}$ is uniformly independent with gap χ then the convergence rate is R-linear.

Proof: Let $s_k = u_k + v_k$ as in Definition 5.1. Then

$$s_k = \frac{u_k}{|u_k|} + \alpha_k u_k + v_k$$

where $\alpha_k = \dfrac{|u_k| - 1}{|u_k|}$ and from (5.5)

$$r_k = \frac{P^t u_k}{|u_k|}$$

and hence

$$s_k = P r_k + \alpha_k u_k + v_k.$$

It can be shown that the conditions $|s_k| = 1$ and $|v_k| \to 0$ R-linearly imply that $\alpha_k \to 0$ R-linearly. Thus we can write

$$s_k = P r_k + t_k \tag{5.6}$$

where $t_k \to 0$ R-linearly. Now using (5.1) and letting V_k, U_k, and Y_k be defined as in (5.2) we

obtain the equations

$$(5.7a) \qquad V_{k+1} = (I - r_k r_k^t) V_k (I - r_k r_k^t) + V(E_k, r_k, t_k)$$

$$(5.7b) \qquad U_{k+1} = U_k + U(E_k, r_k, t_k)$$

$$(5.7c) \qquad Y_{k+1} Y_{k+1}^t = (I - r_k r_k^t) Y_k Y_k^t (I - r_k r_k^t) + Y(E_k, r_k, t_k)$$

where, since the E_k and r_k are uniformly bounded, the nonlinear terms V, U, and Y are $O(|t_k|)$. Thus the system (5.7) can be put in the framework of (3.1) by identifying V_k and $Y_k Y_k^t$ with z_k, and U_k with w_k. The conclusion then follows from the application of Theorem 3.1. •

As a result of this theorem, it is seen that nonconvergence of the $\{E_k\}$ can occur only if the sequence $\{s_k\}$ does not approach any fixed subspace R-linearly in such a way that the projections onto the subspace are subsequentially linearly independent. As noted above, in the Ge and Powell example, the vectors approach each other but not a fixed subspace.

6. Convergence results for other updates

In this section we extend the convergence results obtained in Sections 4 and 5 to the update process (2.5) for ϕ_k other than zero. This will be accomplished by considering the general update as a perturbation of the L update. That is, from (2.5)

$$(6.1) \qquad E_{k+1} = G(s_k; E_k) - \phi_k H(s_k; E_k)$$

where H: $L(\mathfrak{R}^N, \mathfrak{R}^N) \to L(\mathfrak{R}^N, \mathfrak{R}^N)$ is a nonlinear mapping. If $||E_k||_F$ is sufficiently small then

$$|(1 + s_k^t E_k s_k)^{-1}| = 1 + O(||E_k||_F)$$

and hence

(6.2) $$\| H(s_k; E_k)\|_F \leq \sigma (\| E_k \|_F)^2$$

where σ is a constant independent of k. This fact leads directly to the next theorem.

Theorem 6.1: Let $\{s_k\}$ be uniformly linearly independent with gap χ and let the ϕ_k satisfy

(6.3) $$|\phi_k| \leq \kappa \cdot (1 + \|E_k\|_F)$$

for some constant κ independent of k and $\|E_k\|_F$ sufficiently small. Then there exists a postive constant δ such that if E_0 is any symmetric matrix satisfying $\|E_0\|_F < \delta$ the sequence defined by (2.5) converges to the zero matrix and the rate of convergence is $(N+\chi)$-step linear.

Proof: As in the proof of Theorem 4.7 we identify E_k with z_k and $G(s_k;\cdot)$ with A_k. Now identifying the term $\phi_k H(s_k;\cdot)$ with Z_k in (3.1) and applying (6.2) and (6.3) assures the hypotheses of Theorem 3.2 are satisfied. (No w_k is present is this application.) The result follows. •

The boundedness requirement on the ϕ_k, (6.3), generally rules out dynamic processes such as the SR1 in which the ϕ_k can be arbitrarily large for small E_k (although the SR1 update itself will converge finitely in this case). There are two significant differences between Theorem 6.1 and the corresponding theorem for the L update, Theorem 4.7. First, it appears that uniform linear independence is needed as a hypothesis for Theorem 6.1; otherwise, the gap between successive values of the sequence $\{k_j\}$ in Definition 4.2 could become large enough to allow the quadratic terms in $H(s_k;\cdot)$ to grow so rapidly as to preclude convergence of the solution of (3.1) to zero. More importantly, Theorem 6.1 yields local convergence only; inequality (6.2) is not valid without the assumption that $\|E_0\|$ is small. However, as will be observed in Section 7, for certain starting matrices the convergence of the C updating process can be obtained without these two restrictions.

To obtain the convergence of the general updating scheme (2.5) when the $\{s_k\}$ lie on a subspace we apply the perturbation technique of Theorem 6.1 to the analysis in Theorem 5.1. Defining the matrices P, Q, V_k, U_k, and Y_k as in Section 5 and setting $r_k = P^t s_k$ permit us to decompose (2.5) into the system

$$V_{k+1} = (I - r_k r_k^t) V_k (I - r_k r_k^t)$$

(6.4a)

$$- \phi_k \left(\frac{1}{1 + r_k^t V_k r_k} \right) [(r_k^t V_k r_k) r_k - V_k r_k] [(r_k^t V_k r_k) r_k - V_k r_k]^t$$

(6.4b) $$Y_{k+1} = (I - r_k r_k^t) Y_k - \phi_k \left(\frac{1}{1 + r_k^t V_k r_k} \right) [(r_k^t V_k r_k) r_k - V_k r_k] [r_k^t Y_k]$$

(6.4c) $$U_{k+1} = U_k - \phi_k \left(\frac{1}{1 + r_k^t V_k r_k} \right) Y_k^t r_k r_k^t Y_k.$$

From these equations we obtain the following result.

Theorem 6.2: Let S be an M-dimensional subspace of $\mathfrak{R}^N$ and let the matrices P, Q, V_k, Y_k, and U_k be defined as in Section 5. Suppose the sequence of N-vectors $\{s_k\}$ is contained in S and that the sequence of M-vectors, $\{r_k\}$, defined by $r_k = P^t s_k$ is uniformly linearly independent with gap χ. In addition suppose that the ϕ_k satisfy

(6.5) $$|\phi_k| \leq \kappa \cdot (1 + \|P^t E_k P\|_F)$$

for $\|P^t E_k P\|_F$ sufficiently small, where κ is independent of k. Then there exists a positive constant δ such that for any initial symmetric matrix E_0 satisfying $\|P^t E_0 P\|_F \leq \delta$ the sequence defined by (2.5) converges to a matrix of the form $Q U^* Q^t$ and the convergence rate is at least R-linear.

Proof: The equation (6.4a) for V_k is identical in form to that of (2.5). Therefore, by Theorem 6.1 $V_k \to 0$ at an $(N+\chi)$-step rate and hence an R-linear rate. Then by identifying V_k with t_k, $Y_k Y_k^t$ with z_k, and U_k with w_k it is seen that the last two equations in (6.4) have the form of the system (3.1) and satisfy the hypotheses of Theorem 3.2. The result follows. •

The matrix U^* is not specified by the theorem. In particular it is not $Q^t E_0 Q$ as in Theorem 5.1. It should also be observed that the convergence depends only on the initial value of the $P P^t E_0 P P^t$ component of E_0.

If the s_k are not on the subspace S but converge to it in a R-linear manner, then by using the combination of the analysis of Theorem 5.2 with that of the preceding theorem the following result can be obtained.

Theorem 6.3: Let the subspace S be given and let the matrices P and Q be defined as above. For each k let $s_k = P r_k + t_k$ where $\{r_k\}$ is the sequence of M-vectors defined by (5.5) and t_k is defined by (5.6). Assume that the $\{r_k\}$ are uniformly linearly independent with gap χ and that

$$|t_k| \leq |t_0| \tau^k \tag{6.6}$$

for some constant τ, $0 < \tau < 1$. Also assume that the sequence $\{\phi_k\}$ satisfies

$$|\phi_k| \leq \kappa(1 + \|P^t E_k P\|_F + |t_k| \cdot \|E_k\|_F) \tag{6.7}$$

for some constant κ independent of k. Let E_0 be a given symmetric matrix. Then there exist positive constants $\tau_0 = \tau(E_0)$ and δ such that if $\|P^t E_0 P\|_F + \|P^t E_0 Q\| < \delta$ and $|t_0| \leq \tau_0$, the sequence $\{E_k\}$ defined by (2.5) converges R-linearly to a matrix of the form $Q U^* Q^t$.

Proof: (6.6) implies that $t_k \to 0$ at an R-linear rate (and hence $\{s_k\}$ converges to S at an R-linear rate). Using (5.2) we obtain, from (2.5),

$$\begin{aligned} V_{k+1} &= (I - r_k r_k^t) V_k (I - r_k r_k^t) \\ &\quad - \phi_k \left(\frac{1}{1 + r_k^t V_k r_k} \right) [(r_k^t V_k r_k) r_k - V_k r_k][(r_k^t V_k r_k) r_k - V_k r_k]^t \\ &\quad + V_k(r_k, V_k, U_k, Y_k, t_k) \end{aligned}$$

$$Y_{k+1} = (I - r_k r_k^t) Y_k - \phi_k \left(\frac{1}{1 + r_k^t V_k r_k} \right) [(r_k^t V_k r_k) r_k - V_k r_k][r_k^t Y_k]$$

$$+ \; Y_k(r_k, V_k, k, Y_k, t_k)$$

$$U_{k+1} = U_k - \phi_k\left(\frac{1}{1 + r_k{}^t V_k r_k}\right) Y_k{}^t r_k r_k{}^t Y_k + U_k(r_k, V_k, U_k, Y_k, t_k).$$

Since the r_k are uniformly bounded, V_k, U_k, and Y_k are $O((\|V_k\|_F + \|U_k\|_F + \|Y_k\|_F)\cdot|t_k|)$. Using (6.7) and identifying V_k and $Y_k Y_k{}^t$ with z_k and U_k with w_k in (3.1) we see that the hypotheses of Theorem 3.2 are satisfied and hence the results follow. •

The hypotheses of this theorem do not require that the component of E_0 relative to $S^\perp$ be small however they do require that the components of $\{s_k\}$ start and remain close to S. An analysis of the proof of Theorem 3.2 will show that if the component of E_0 relative to $S^\perp$ is sufficiently small then the requirement that the $\{s_k\}$ start close to S can be relaxed, although the sequence still must converge to S at an R-linear rate.

7. Application to updates of H_k

In the previous three sections we have established the convergence of the sequence of matrices $\{E_k\}$ generated by the system (2.5) under certain restrictions on the sequences $\{s_k\}$ and $\{\phi_k\}$. In this section we apply those results to the sequence $\{H_k\}$ generated by (1.5). While the results obtained so far may be of interest in their own right, the updating schemes (1.5) are of practical interest in the optimization algorithms that motivated this study and are, consequently, of more import. For the simplest (quadratic) case, when $y_k = s_k$, the translation of the major theorems of Sections 4, 5, and 6 into results for the system (2.2) can be carried out by a straightforward substitution of $H_k - I$ for E_k. The results of this transformation are given in the next two theorems in condensed form; the first theorem deals with the case where the sequence $\{s_k\}$ remains on the subspace S (which may be $\Re^N$) and the second with the case where the $\{s_k\}$ converge to a proper subspace S.

Theorem 7.1: Let S be an M-dimensional subspace of $\Re^N$ with $M \leq N$ and let the matrices P and Q be defined relative to S as in Section 5 ($P = I$ and $Q = 0$ if $M = N$). Suppose each vector in the sequence of unit N-vectors, $\{s_k\}$, is in S and suppose that the sequence $\{r_k\}$, defined by $r_k = P^t s_k$, is uniformly independent with gap χ in S. Let H_0 be a symmetric matrix and generate the sequence $\{H_k\}$ by (2.2). Then

(i) if $\phi_k = 0$ for each k,

$$\lim_{k\to\infty} H_k = QQ^t H_0 QQ^t + PP^t;$$

or,

(ii) if the ϕ_k satisfy $|\phi_k| \leq \kappa (1 + || P^t H_k P ||_F)$ for some κ independent of k and

$|| P^t H_0 P - PP^t ||_F$ is sufficiently small, for some NxN matrix H^*

$$\lim_{k\to\infty} H_k = QH^*Q^t + PP^t.$$

The convergence rate of the matrices is $(N+\chi)$-step linear if $S = \Re^N$ and R-linear if S is a proper subspace of $\Re^N$. The uniform linear independence can be replaced in case (i) by subsequential linear independence, but the R-linear rate convergence is no longer guaranteed.

Theorem 7.2: Let S be a subspace of $\Re^N$ and assume that the sequence $\{s_k\}$ satisfies the hypotheses of Theorem 6.3. Assume that the ϕ_k satisfy

$$|\phi_k| \leq \kappa (1 + || P^t H_k P ||_F + |t_k| \cdot || H_k ||_F)$$

for some κ independent of k. Let H_0 be a given symmetric matrix. Then there exist constants $\tau_0 = \tau(H_0)$ and δ such that if $|| P^t H_0 P + P^t H_0 Q - PP^t ||_F \leq \delta$ and $\tau \leq \tau_0$ then the sequence $\{H_k\}$ generated by (1.5) satisfies

$$\lim_{k\to\infty} H_k = QH^*Q^t + PP^t$$

for some (N-M)x(N-M) matrix H^* and the convergence rate is R-linear. If $\phi_k = 0$ for all k then

the restriction on the initial matrix H_0 is not required.

Thus, generally speaking, the sequence of updates defined by (2.2) converges to a matrix which is the identity on the subspace spanned by the sequence $\{s_k\}$ (possibly in a limiting sense) provided the initial matrix is nearly the projection PP^t. In particular, the subspaces S and $S^{\perp}$ are eigenspaces of the limiting matrix.

We now drop the assumptions that each s_k is a unit vector and $y_k = s_k$ and require instead that at each iteration

$$y_k = s_k + \xi_k \tag{7.1}$$

where $\xi_k / |s_k| \to 0$ as $k \to \infty$. As was discussed in Section 2 this assumption is motivated by the unconstrained optimization problem where s_k represents the difference in successive iterates. We then have the following theorem.

Theorem 7.3: Let $\{s_k\}$ satisfy the hypotheses of Theorem 6.3. Let the sequence $\{y_k\}$ satisfy (7.1) and assume that $\xi_k / |s_k| \to 0$ at an R-linear rate and $|\xi_k| / |s_k| \leq \eta_0 \cdot \eta^k$ for all k and some η_0, where $0 \leq \eta < 1$. Assume that the sequence $\{\phi_k\}$ satisfies

$$|\phi_k| \leq \kappa \left(1 + \| P^t H_k P \|_F + (|t_k| + |\xi_k| / |s_k|) \cdot \| H_k \|_F\right)$$

for some constant κ independent of k. Let H_0 be a symmetric matrix. Then there exist positive constants δ and $\tau_0 = \tau(H_0)$ such that for $\| P^t H_0 P + P^t H_0 Q - PP^t \|_F < \delta$, $\eta_0 < \tau_0$, and $|t_0| < \tau_0$ the sequence $\{H_k\}$ generated by (1.5), satisfies

$$\lim_{k \to \infty} H_k = Q H^* Q^t + PP^t$$

for some $(N-M)x(N-M)$ matrix H^* and the convergence rate is R-linear. If $\phi_k = 0$ for all k then no restriction is required on the initial matrix H_0.

Proof: Let (7.1) hold. Then

$$y_k{}^t s_k = (s_k{}^t s_k)\left(1 + \frac{s_k{}^t \xi_k}{s_k{}^t s_k}\right)$$

and hence if $|\xi_k| / |s_k|$ is sufficiently small

$$(7.2) \qquad (y_k{}^t s_k)^{-1} = (s_k{}^t s_k)^{-1}\,(1 + O(\xi_k/|s_k|)).$$

Substituting (7.1) and (7.2) into the update formula (1.5), making the changes of variables

$$\hat{s}_k = \frac{s_k}{|s_k|} \quad \text{and} \quad \hat{\xi}_k = \frac{\xi_k}{|s_k|},$$

and replacing, as before, H_k by $E_k + I$ yield an equation of the form

$$(7.3) \qquad E_{k+1} = G(\hat{s}_k; E_k) - \phi_k\, H(\hat{s}_k; E_k) + W(\phi_k, \hat{s}_k, E_k, \hat{\xi}_k)$$

where now $|\hat{s}_k| = 1$, and $|\hat{\xi}_k| \le |\eta_0| \cdot \eta^k$. Using the bound on the ϕ_k it is seen that

$$(7.4) \qquad \| W(\phi, \hat{s}, E, \hat{\xi}) \|_F \le \kappa(\|E_k\|_F) \cdot |\hat{\xi}|.$$

Now the $\hat{s}_k$ can be decomposed as in Theorem 6.3 and the results of Section 3 (in particular Theorem 3.2) can be applied just as before. •

It may be disconcerting to observe that Theorems 7.1-7.3 yield global convergence (i.e., unrestricted initial matrix) of the $\{H_k\}$ only in the case where ϕ_k is zero, i.e., the L update. Proposition 2.2 allows us to extend these results to the C updating formula for positive definite starting matrices.

Theorem 7.4: If H_0 is positive definite and $\phi_k = 1$ for all k, then in each of Theorems 7.1-7.3 the convergence and rate of convergence of the sequence $\{H_k\}$ are independent of the initial matrix. In the case where the s_k all lie on a subspace S and $y_k = s_k$ for every k the uniform linear independence assumption can be relaxed to subsequential linear independence and the limit has the form

$$Q H^* Q^t + P P^t = [\, Q Q^t H_0^{-1} Q Q^t + P P^t \,]^{-1}.$$

Proof: The class of positive definite matrices satisfies the conditions of Proposition 2.2 so the sequence converges when the C update is used. Since the sequence is converging the iterates will eventually get close enough to their limit to assure that the initial matrix requirements of Theorems 7.1-7.3 are satisfied. The desired convergence rate is then achieved . The final form in the special case derives from the fact that the inverse of the C update of a matrix is always the L update of the inverse of the matrix. If only subsequential linear independence of $\{s_k\}$ is assumed in this case then the iterates will converge by the argument of Proposition 2.2 but no rate of convergence is specified. •

In addition to the class of "direct" Broyden updates generated by (1.5) there is the class of "inverse" Broyden updates that are obtained from (1.5) by interchanging s and y. In the unconstrained optimization setting this class of updates generates sequences $\{H_k\}$ that represent approximations to the inverse of the Hessian matrix of the objective function. When the $y_k = s_k$ the resulting class reduces to (2.2) and the theorems of Sections 4, 5, and 6 apply. When (7.1) holds, Theorem 7.3 can be repeated for the "inverse" class with no essential change. Only the function W is slightly different, although it still satisfies (7.4).

When the theorems above are applied to the unconstrained optimization problem, it is under the assumption that the Hessian matrix of the objective function at the optimal solution is the identity. Since the form of the Broyden class of updates (1.5), as well as the form of the "inverse" class, is invariant under the transformations (T) in Section 2, there is no loss in generality in making this assumption. The limiting value of the Hessian approximations in the untransformed case will be the Hessian matrix F if the vectors $F^{1/2}s_k$, equivalently the vectors s_k, do not approach a proper subspace of $\mathfrak{R}^N$. If $\{F^{1/2}s_k\}$ approaches a subspace S which is an eigenspace of F, the matrices P and Q can be chosen so that the columns are eigenvectors of F (and, hence, $F^{1/2}$). Then

$$F^{1/2}[P,Q] = [P,Q]\,D = [P,Q]\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

where D_1 and D_2 are the diagonal matrices whose diagonal elements are the square roots of the eigenvalues of F on the subspaces S and $S^{\perp}$ respectively. Now the general form for the limiting matrix gives

$$\lim_{k\to\infty} F^{-1/2}H_k F^{-1/2} = PP^t + QU^*Q^t$$

so

$$\lim_{k\to\infty} H_k = F^{1/2}[P,Q]\begin{bmatrix} I & 0 \\ 0 & U^* \end{bmatrix}[P,Q]^t F^{1/2}$$

$$= [P,Q]\, D\begin{bmatrix} I & 0 \\ 0 & U^* \end{bmatrix} D\,[P,Q]^t$$

$$= PD_1{}^2P^t + Q\hat{U}Q^t$$

for some matrix $\hat{U}$. If S is not an eigenspace of F then the columns of P and Q cannot be chosen as eigenvectors of F and the limit is $F^{1/2}[PP^t + QU^*Q^t]F^{1/2}$.

Finally, it should be pointed out that other rank two updates not of the Broyden class may be able to be analyzed by the methods presented here. For example, consider the well-known PSB update which has the form:

$$U(H,s,y) = \frac{(y - Hs)s^t + s(y - Hs)^t}{s^ts} - \frac{(y - Hs)^ts}{(s^ts)^2}\, ss^t.$$

This update is not invariant under the transformations given above; however, we can analyze it by letting $y = Fs + \xi$ where $\xi \to 0$ as above. Then letting $E = H - F$ and $\xi = 0$, we obtain the L updating scheme (4.1). Thus we can prove, under the appropriate restrictions on the $\{s_k\}$, that the sequence $\{H_k\}$ converges to F if the $\{s_k\}$ do not converge to a subspace. If the $\{s_k\}$ converge to an eigenspace, S, of F, then we obtain

$$\lim_{k\to\infty} (H_k - F) = Q\hat{U}Q^t$$

for some matrix $\hat{U}$. Now since the limiting subspace of the $\{s_k\}$ is an eigenspace of F,

$$F = PD_1P^t + QD_2Q^t$$

where D_1 and D_2 are diagonal matrices. (As above, P and Q are chosen to have eigenvectors of F as columns). Thus

$$\lim_{k\to\infty} H_k = PD_1P^t + QU^*Q^t$$

for some $(N-M)x(N-M)$ matrix U^*. This is a somewhat more general result than is available for the Broyden class of updates because there is no restriction that the limiting matrix F be positive definite; it only need be symmetric. In the Broyden class of updates, the term $y^t s$ that appears in the denominator can be zero when $y = Fs$ and F is not positive definite. Therefore, unless the s_k lie on, or converge to, a subspace on which F is positive definite, convergence results cannot be obtained for the Broyden class of updates when F is indefinite. The implications of these remarks for constrained optimization algorithms are discussed in the next section.

8. Final comments

In this paper we have analyzed the convergence of a sequence of symmetric matrices, $\{H_k\}$, satisfying rank two updating formulas, (1.5), of the type typically encountered in optimization algorithms. This analysis was carried out independently of the particular algorithm and, indeed, independently of any optimization problem. The results indicate that the convergence of the matrices follows from the secant condition (1.2), the requirement that the sequence $\{s_k\}$ satisfy certain linear independence conditions, and the condition that the sequence $\{y_k - s_k\}$ approach zero at a specified rate. It does not postulate that the sequence $\{s_k\}$ be related to any optimization iteration scheme. In addition to establishing convergence of the sequence of matrices the results presented show that the limiting matrix is the identity matrix when restricted to the subspace to which $\{s_k\}$ converges. In optimization problems where $y_k - F(x^*)\, s_k$ approaches zero and the Hessian matrix $F(x^*)$ is positive definite with the subspace as an eigenspace, the limit of the matrices is identical to $F(x^*)$ on the subspace.

These results do not subsume those of earlier research, e.g., those contained in references [3] – [7], because the conditions on the sequences $\{s_k\}$ and $\{y_k\}$ are assumed here and not derived as a consequence of their generation in optimization problems. They do, however, provide characterizations of the limiting matrix not given in those works and, in particular, illustrate how its structure depends upon the subspace that is spanned by the sequence $\{s_k\}$, an important consideration in, for example, constrained optimization. Another distinction of the results presented here is that some analysis of convergence rates is given. The convergence rate of these matrices, under the assumptions used, is essentially R-linear; thus it is unlikely to be recognized in an algorithm where the iterates are converging much more rapidly, such as in a Q-superlinearly convergent algorithm.

One of the original motives for undertaking this study was to try to find alternative proofs of the superlinear convergence of the BFGS and DFP algorithms for unconstrained optimization. Our efforts in this regard will appear in another report [10]. In particular, we wanted to try to simplify the proofs' dependence on the bounded deterioration estimates and the use of different weighted norms [1]. The relation between the superlinear convergence of these algorithms and the convergence of the Hessian approximations is best understood by recalling that if the iterates $\{x_k\}$ converge and $x_{k+1} - x_k = s_k$, then the iterates converge superlinearly if and only if

$$\lim_{k\to\infty} (H_k - I)\frac{s_k}{|s_k|} = 0. \tag{8.1}$$

(We assume the Hessian is the identity at the solution.) Now if the $\{s_k\}$ converge to a subspace S R-linearly and the projections of the s_k span a subspace S in the uniformly independent manner defined in this paper, then

$$(H_k - I)\frac{s_k}{|s_k|} = P(P^t(H_k - I)Pr_k + O(\xi_k)$$

where $\xi_k \to 0$. Therefore by the results of this paper the convergence is superlinear if and only if $P^t H_k P \to PP^t$.

Another motivation of this research was the desire to understand the question of superlinear convergence for sequential quadratic programming (SQP) methods for constrained optimization. The necessary and sufficient condition analogous to (8.1) for superlinear convergence is

$$\lim_{k\to\infty} PP^t(H_k - L)\frac{s_k}{|s_k|} = 0 \tag{8.2}$$

where L is the Hessian of the Lagrangian function with respect to the decision variable at the solution and PP^t is the projection onto the null space, S, of the active gradients at the solution. (See [11] or [12].) Note that the restriction of L to the subspace S is positive definite by the second order sufficient conditions. Thus, letting

$$s_k = PP^t s_k + QQ^t s_k = r_k + v_k$$

(8.2) becomes

$$\lim_{k\to\infty} \{ PP^t(H_k - L)PP^t \frac{r_k}{|s_k|} + (PP^t(H_k - L)QQ^t)\frac{v_k}{|s_k|} \} \to 0. \tag{8.3}$$

It can be seen that if the sequence $\{s_k\}$ spans $\Re^N$ in a uniformly linearly independent manner then the matrices $P^t(H_k - L)P$ and $P^t(H_k - L)Q$ must converge to zero if (8.3) is to hold. But L need not be positive definite except on the subspace S. Thus a BFGS or DFP updating scheme initiated with a positive definite matrix cannot be expected to lead to a superlinearly convergent algorithm unless the problem is convex. On the other hand, it is seen that if the sequence $\{r_k\}$ spans S in a uniformly linearly independent manner and $\{v_k\}$ tends to zero R-linearly then the results of this paper imply that $PP^t(H_k - L)PP^t \to 0$ and hence superlinear convergence occurs. The sequence of vectors $\{v_k\}$ tending to zero is implied by the sequence of iterates $\{x_k\}$ tending to the solution in a "tangential" sense relative to S and so this result reinforces and explains previous observations (see [13]) that tangential convergence of the iterates in an SQP algorithm implies superlinear convergence. The remarks in Section 7 concerning the PSB update suggest that superlinear convergence can be obtained by an SQP algorithm employing that update, since under the assumption of uniform linear independence of the sequence $\{s_k\}$ the PSB updates converge to the Hessian matrix without the assumption of positive definiteness. Indeed, Han in his early work on the SQP method [14] was able to prove superlinear convergence for this update.

It would be ideal to be able use to the convergence analysis of this paper to deduce the "best" update scheme for a given optimization algorithm. This is unreasonable, however, since there are many other factors that influence a numerical algorithm that are not taken into account here. More significantly, our analysis treats the sequence of steps $\{s_k\}$ as generated independently of the sequence $\{H_k\}$; whereas in optimization algorithms the determination of s_{k+1} is directly influenced by H_k. This being said, it is hoped that some insight into the appropriateness or limitations of a particular update in a certain situation may be gained from this type of analysis. For example, the remarks above suggest that the updating methods that preserve positive definiteness, long favored for unconstrained optimization algorithms, cannot guarantee superlinear convergence when applied in SQP algorithms for constrained optimization problems.

One intriguing question that has inspired many research efforts (e.g., [6]) is to explain the experimentally observed superiority of the BFGS updating scheme over the DFP scheme in quasi-Newton algorithms for unconstrained optimization problems. In the context of system (2.2) the BFGS update corresponds to the C update and the DFP to the L update (where H_k represents a direct approximation of the Hessian of the objective function). If one interprets the system (2.2) as a local approximation to the system (1.5) then one might hope to observe better convergence results for the C update than for the L update. Theorem 7.4 suggests, however, that their theoretical convergence

properties are identical for positive definite initial matrices.

Another way to compare the local convergence properties of the Broyden class of updates is in terms of their relation to the SR1 update. As was noted in Section 2 the SR1 updating scheme used in (2.2) yields finite convergence in N steps (if $Es \neq 0$ at an earlier step) if the sequence $\{s_k\}_{k=0}^{N-1}$ is linearly independent, while the other updates require that the s_k be orthogonal to obtain the same result. This property can be expressed in terms of the system (2.5) by observing that if s_{k-1}, s_k, and E_k are given with $E_k s_{k-1} = 0$, $s_k = \alpha s_{k-1} + v$, and $v^t s_{k-1} = 0$ then, in the notation of (6.1),

$$E_{k+1} = G(s_k; E_k) - \phi_k H(s_k; E_k) = G(s_k; E_k) - \phi_S H(s_k; E_k) + (\phi_S - \phi_k) H(s_k; E_k)$$

(8.4)

$$= G(v; E_k) - \phi_S H(v; E_k) + (\phi_S - \phi_k) H(s_k; E_k)$$

where ϕ_S is the SR1 update parameter. Since the last term is zero if $\phi_k = \phi_S$ it is seen that the SR1 parameter is the only one that causes the update formula to be a function of only the component, v, of s_k perpendicular to the preceding step, s_{k-1}. Thus we can think of the last term involving $(\phi_S - \phi_k)$ as the error in the update that is preventing finite convergence. Clearly, in terms of speeding up convergence for system (2.5) (and hence (2.2)) it is best to choose $\phi_k = \phi_S$. If, however, it is desired that positive definiteness of the H_k matrices be preserved, then the value ϕ_S cannot always be chosen. A reasonable strategy in this case might be to choose the value of ϕ_k as close to ϕ_S as possible consistent with preserving positive definiteness, thus minimizing the error term in (8.4). From Propostion 2.1 it is seen that positive definiteness is preserved if and only if $\phi \leq \phi_C$ where ϕ_C is a constant greater than one whose exact value depends on the current s and H. For the system (2.2)

$$\phi_S = \frac{s^t H s}{s^t H s - 1}$$

so that if $s^t H s < 1$ then $\phi_S < 0$ and hence positive definiteness is preserved, while if $s^t H s > 1$ then $\phi_S > 1$ and positive definiteness may or may not be preserved depending on the value of ϕ_C. These observations lead to the consideration of four strategies for choosing the update parameter to maintain positive definiteness for the system (2.2):

Strategy I: (pure DFP) Choose $\phi = 0$;

Strategy II: (pure BFGS) Choose $\phi = 1$;

Strategy III: (mixed BFGS and DFP) If $s^t H s < 1$ choose $\phi = 0$, otherwise choose $\phi = 1$;

Strategy IV: (mixed BFGS and SR1) If $s^t H s < 1$ choose $\phi = \phi_S$, otherwise choose $\phi = 1$.

Based on the arguments given above, one would expect Strategy IV on average to lead to faster

convergence of the matrices than Strategy III and the latter to be better than either of the strategies I or II. Very limited numerical experimentation on randomly generated problems has supported these conjectures. This testing also gave an edge to Strategy II over Strategy I. A reasonable explanation for this is that the eigenvalues of the randomly generated initial matrices were preponderantly greater than one and, hence, so were those of the succeeding matrices. Therefore, on average, $s^t H s$ was more often than not greater than one and thus the value $\phi = 1$ was closer to ϕ_S more often than was the value $\phi = 0$. This situation would also be expected to occur for almost any realistic distribution of positive definite matrices; thus one should expect that the strategies above are listed in increasing order of general effectiveness. In particular, the BFGS strategy should generally outperform the DFP strategy when there are numerous large eigenvalues (cf., [6]). Presumably, the best strategy would be to choose ϕ as close to the upper bound ϕ_C as possible whenever ϕ_S is greater than ϕ_C and $\phi = \phi_S$ otherwise. Whether or not this is computationally attractive is not clear.

It should be emphasized that the above remarks apply only to the system (2.2) and not to the system (1.5). Further experimentation is necessary to determine if these conjectures can be validated for the latter system and applied in a useful way to quasi-Newton algorithms for unconstrained optimization.

A generalization of the results of this paper that also might yield interesting insights would be the relaxation of the secant condition (1.2). It seems clear that the analysis carried out here could also work if another term were added to (2.5), say $\hat{U}(s_k, E_k)$, with $\hat{U} s_k$ converging to, but not equal to, zero. Such a class of updates satisfying a type of asymptotic secant condition would yield the same convergence results and, by the remarks above, preserve superlinear convergence. Such updates could potentially broaden the choices of updating schemes available for use in special classes of problems.

A final question that is unanswered by this analysis is if the global convergence properties enjoyed by the L and C updates when $s_k = y_k$ can be extended to other updates. In particular, it would be reasonable to expect that the restricted Broyden class of updates would have this property.

References

[1] Dennis, J. and Schnabel, R., Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983.

[2] Fletcher, R., Practical Methods of Optimization, Vol. 1: Unconstrained Optimization, John Wiley and Sons, Chichester, England, 1980.

3] Schuller, G., "On the order of convergence of certain quasi-Newton methods", *Numerische Mathematik*, vol. 23, pp. 181-192, 1974.

[4] Ge, R. and Powell, M., "The convergence of variable metric matrices in unconstrained optimization", *Mathematical Programming*, vol. 27, pp.123-143, 1983.

[5] Stoer, J., "The convergence of matrices generated by rank-2 methods from the restricted β-class of Broyden", *Numerische Mathematik*, vol. 44, pp. 37-52, 1984.

[6] Byrd, R., Nocedal, J., and Yuan, Y., "Global convergence of a class of quasi-Newton methods on convex problems", *SIAM Journal of Numerical Analysis*, vol. 24, pp. 1171-1190, 1987.

[7] Conn, A., Gould, N., and Toint, Ph., "Convergence of quasi-Newton matrices generated by the symmetric rank one update", Technical Report 87/12, Department of Computer Sciences, Waterloo, University, Waterloo, Ontario, 1987.

[8] Wilkinson, J., The Algebraic Eigenvalue Problem, Oxford University Press, London, 1965.

[9] Ortega, J. and Rheinboldt, W., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[10] Boggs, P., and Tolle, J., "Convergence of quasi-Newton methods", Technical Report 90-17, Department of Operations Research, University of North Carolina, Chapel Hill, NC, 1990.

[11] Boggs, P., Tolle, J., and Wang, P., "On the local convergence of quasi-Newton methods for constrained optimization", *SIAM Journal on Control and Optimization*, vol. 20, pp. 161-171, 1982

[12] Stoer, J., and Tapia, R., "On the characterization of Q-superlinear convergence of quasi-Newton methods for constrained optimization", *Mathematics of Computation*, vol. 49, pp. 581-584, 1987.

[13] Boggs, P., and Tolle, J., "A strategy for global convergence in a sequential quadratic programming algorithm," *SIAM Journal of Numerical Analysis*, vol. 26, pp. 1-23, 1989.

[14] Han, S., "Superlinearly convergent variable metric algorithms for general nonlinear programming problems", *Mathematical Programming*, vol. 11, pp. 263-382, 1976.

Appendix

Here we address the proofs of the theorems in Section 3. We will not prove the theorems stated there in their full generality. Rather we will sketch the proof of Theorem 3.1 and prove one theorem from which the proof of Theorem 3.2 can be derived by generalizing the arguments.

For Theorem 3.1 we have two uncoupled systems

$$z_{k+1} = A_k z_k + Z_k(t_k)$$

and

$$w_{k+1} = w_k + W_k(t_k)$$

where, because of the R-linear convergence of the $\{t_k\}$,

$$|t_k| \leq |t_0| \cdot \tau^k \tag{A.1}$$

for some τ, $0 < \tau < 1$, and some t_0. For the first equation it can be shown by induction, using the fact that $||A_i|| \leq 1$ for all i,

$$|z_{k+r}| \leq (\,||\prod_{j=0}^{r-1} (A_{k+j})\,||)\,|z_k| + \kappa_1 \cdot \sum_{j=0}^{r-1} |t_{k+j}|.$$

$$\leq \beta_{k,r}\,|z_k| + \frac{\kappa_1 \cdot |t_0|}{1-\tau}\,\tau^k \tag{A.2}$$

Assuming, without loss of generality, that $k_{j+1} \geq k_j + m$, condition (ii) of Theorem (3.1) implies

$$|z_{k_{j+1}}| \leq \beta |z_{k_j}| + K \cdot \tau^j .$$

where $K = \frac{\kappa_1 \cdot |t_0|}{1-\tau}$. Let $\{\xi_j\}$ satisfy

(A.3) $$\xi_{j+1} = \beta \xi_j + K \cdot \tau^j$$

with $\xi_0 = |z_{k_0}|$. Then for each j, $|z_{k_j}| \leq \xi_j$. The solution to (A.3) satisfies

$$\xi_j = \beta^j \xi_0 + K \cdot \sum_{i=1}^{j} \beta^{j-i} \tau^i \leq \rho^j [\xi_0 + K \cdot j]$$

where $\rho = \max[\beta, \tau]$. So from (A.2), for $k_j < k \leq k_{j+1}$,

$$|z_k| \leq |z_{k_j}| + K \cdot \tau^{k_j} \leq \rho^j [|z_{k_0}| + K \cdot (j+1)].$$

Thus $\{z_k\} \to 0$. If condition (ii)′ of Theorem 3.1 holds, then without loss of generality, we can assume that it holds for all k. Then for $k = j \cdot m + r$, $0 < r \leq 1$, we have from (A.2)

$$|z_k| = |z_{r+j\cdot m}| \leq \beta |z_{r+(j-1)\cdot m}| + K \cdot \tau^{r+(j-1)\cdot m}.$$

It can now be established by induction on j that for each r, $0 < r \leq m$,

$$|z_{r+j\cdot m}| \leq K_r \rho^{r+j\cdot m}$$

where $\rho = \max[\beta, \tau]$. The R-linear convergence follows. If $Z_k = 0$, then for each k

$$|z_{k+m}| \leq \beta |z_k|$$

which gives the m-step convergence.

For the second system, to show that the $\{w_k\}$ converge to some vector R-linearly we note that for any positive integers k and i,

$$|w_{k+i} - w_k| \leq \kappa_2 \cdot |t_0| \cdot \tau^k \cdot \sum_{j=0}^{i-1} \tau^j \leq K \tau^k,$$

and hence $\{w_k\}$ is a Cauchy sequence converging to some w^*. Since

$$|w^* - w_k| = \lim_{i\to\infty} |w_{k+i} - w_k| \leq K \tau^k$$

it follows that the convergence is R-linear.

To show how Theorem 3.2 can be derived we consider the system of coupled difference equations in two variables:

$$\begin{aligned} z_{k+1} &= \alpha z_k + \beta_1 (z_k)^2 + \gamma_2 \cdot (z_k + w_k) \cdot t_k \\ w_{k+1} &= w_k + \beta_2 (z_k)^2 + \gamma_2 \cdot (z_k + w_k) \cdot t_k \end{aligned} \tag{A.4}$$

where $0 < \alpha < 1$, β_1, β_2, γ_1, and γ_2 are nonnegative constants, and $\{t_k\}$ converges to zero R-linearly, i.e., (A.1) holds.

Theorem A.1: There exists positive constants ξ_0 and τ_0 such that if $|z_0| < \xi_0$ and $|t_0| < \tau_0$ then $|z_k| \leq |z_0|$ and $|w_k| \leq 3 \cdot |w_0|$ for all k. Moreover, there exists a w^* such that

$$\{z_k\} \to 0$$

and

$$\{w_k\} \to w^*$$

R-linearly .

Proof: First we show, by induction, that the hypotheses imply that $|w_k| \leq 3 \cdot |w_0|$ and that there is a ρ, $0 < \rho < 1$, such that

$$|z_k| \leq |z_0| \rho^k.$$

We let $\beta = \max\{\beta_1, \beta_2\}$ and $\gamma = \max\{\gamma_1, \gamma_2\}$ and set

(i) $\hat{\alpha} = (1 + \alpha)/2,$

(ii) $\rho = \max \{ \hat{\alpha}^{1/2}, \tau^{1/2} \},$

(iii) $\xi_0 = \min \{ \frac{1-\alpha}{2\beta}, [\frac{|w_0| \cdot (1-\rho^2)}{\beta}]^{1/2}, |w_0| \},$

(iv) $\tau_0 = \min \{ \frac{(1-\rho)|z_0|}{8|w_0|\gamma}, \frac{(1-\rho^2)}{4\gamma} \}.$

Clearly the inequalities are true for $k = 0$. Assume them to be true for $j = 0, 1, \ldots, k$. Then using (i), (iii), and the induction hypotheses

$$|z_{k+1}| \leq \alpha |z_k| + \beta |z_k| \cdot |z_k| + \gamma (|z_k| + |w_k|) |t_0| \tau^k$$

$$\leq \hat{\alpha} |z_0| \rho^k + 4\gamma |w_0| \cdot \tau_0 \cdot \tau^k.$$

Using (ii) and (iv) we have (since $k \geq 1$)

$$(A.5) \qquad |z_{k+1}| \leq \rho^{k+1} \{ \rho |z_0| + \frac{\rho(1-\rho)|z_0|}{2} \} \leq |z_0| \rho^{k+1}.$$

Then for each j, $j = 0, \ldots, k$, we have

$$|w_{j+1}| \leq |w_j| + \beta |z_0|^2 \rho^{2j} + \gamma (|z_0| \rho^j + 3|w_0|) |t_0| \rho^{2j}$$

$$(A.6) \qquad \leq |w_j| + \{ |w_0| (1-\rho^2) + 4\gamma |w_0| \tau_0 \} \rho^{2j}$$

$$\leq |w_j| + 2|w_0| (1-\rho^2) \rho^{2j}.$$

Clearly, the sequence $\{ |w_j| \}$ is majorized by the sequence $\{ \zeta_j \}$ satisfying

$$\zeta_{j+1} = \zeta_j + 2\zeta_0 (1-\rho^2) \rho^{2j}, \qquad \zeta_0 = |w_0|$$

for $j = 0, \ldots, k$. But then

$$\zeta_{k+1} = \zeta_0 + \sum_{j=0}^{k} [\zeta_{j+1} - \zeta_j] = \zeta_0 + 2\zeta_0 (1-\rho^2) \sum_{j=0}^{k} \rho^{2j} \leq 3\zeta_0.$$

Thus

$$|w_{k+1}| \leq 3 \cdot |w_0|$$

which completes the induction step. Now it follows from (A.5) that $\{z_k\} \to 0$ at an R-linear rate. It remains to show that the $\{w_k\}$ converge to some vector R-linearly. From the difference equation for the $\{w_k\}$ and the estimates in (A.6) it is seen that for any positive integers k and i,

$$|w_{k+i} - w_k| \leq 2|w_0|(1-\rho^2) \sum_{j=0}^{i-1} \rho^{2(k+j)} \leq 2|w_0| \rho^{2k}$$

and hence $\{w_k\}$ is a Cauchy sequence converging to some w^*. Since

$$|w^* - w_k| = \lim_{i \to \infty} |w_{k+i} - w_k|$$

it follows from the above inequality that the convergence is R-linear. •

Theorem 3.2 can now be proved by generalizing the above theorem to the case where z_k and w_k are vectors and the condition $0 < \alpha < 1$ is replaced by conditions (i) and (ii)′ of Theorem 3.1.

NIST-114A
(REV. 3-90)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

BIBLIOGRAPHIC DATA SHEET

1. PUBLICATION OR REPORT NUMBER
NISTIR 4554

2. PERFORMING ORGANIZATION REPORT NUMBER

3. PUBLICATION DATE
APRIL 1991

4. TITLE AND SUBTITLE

Convergence Properties of a Class of
Rank-Two Updates

5. AUTHOR(S)

Paul T. Boggs and Jon W. Tolle

6. PERFORMING ORGANIZATION (IF JOINT OR OTHER THAN NIST, SEE INSTRUCTIONS)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
GAITHERSBURG, MD 20899

7. CONTRACT/GRANT NUMBER

8. TYPE OF REPORT AND PERIOD COVERED

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (STREET, CITY, STATE, ZIP)

10. SUPPLEMENTARY NOTES

11. ABSTRACT (A 200-WORD OR LESS FACTUAL SUMMARY OF MOST SIGNIFICANT INFORMATION. IF DOCUMENT INCLUDES A SIGNIFICANT BIBLIOGRAPHY OR LITERATURE SURVEY, MENTION IT HERE.)

Many optimization algorithms generate, at each iteration, a pair (x_k, H_k) consisting of an approximation to the solution, x_k, and a Hessian matrix approximation, H_k, which contains local second order information about the problem. Much is known about the convergence of the x_k to the solution of the problem but relatively little about the behavior of the sequence of matrix approximations. We analyze the sequence $\{H_k\}$ generated by the extended Broyden class of updating schemes independently of the optimization setting in which they are used, deriving various conditions under which convergence is assured and delineating the structure of the limits. Rates of convergence are also obtained. Our results extend and clarify those already in the literature.

12. KEY WORDS (6 TO 12 ENTRIES; ALPHABETICAL ORDER; CAPITALIZE ONLY PROPER NAMES; AND SEPARATE KEY WORDS BY SEMICOLONS)

convergence analysis; matrix updates; optimization;
Quasi-Newton algorithms

13. AVAILABILITY

X	UNLIMITED
	FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NATIONAL TECHNICAL INFORMATION SERVICE (NTIS).
	ORDER FROM SUPERINTENDENT OF DOCUMENTS, U.S. GOVERNMENT PRINTING OFFICE, WASHINGTON, DC 20402.
X	ORDER FROM NATIONAL TECHNICAL INFORMATION SERVICE (NTIS), SPRINGFIELD, VA 22161.

14. NUMBER OF PRINTED PAGES
48

15. PRICE
A03

ELECTRONIC FORM