

THE “GLOBAL” CONVERGENCE OF BROYDEN-LIKE METHODS WITH A SUITABLE LINE SEARCH

ANDREAS GRIEWANK¹

(Received 24 July 1985; revised 14 October 1985)

Abstract

Iterative methods for solving a square system of nonlinear equations $g(x) = 0$ often require that the sum of squares residual $\gamma(x) \equiv \frac{1}{2}\|g(x)\|^2$ be reduced at each step. Since the gradient of γ depends on the Jacobian ∇g , this stabilization strategy is not easily implemented if only approximations B_k to ∇g are available. Therefore most quasi-Newton algorithms either include special updating steps or reset B_k to a divided difference estimate of ∇g whenever no satisfactory progress is made. Here the need for such back-up devices is avoided by a derivative-free line search in the range of g . Assuming that the B_k are generated from an arbitrary B_0 by fixed scale updates, we establish superlinear convergence from within any compact level set of γ on which g has a differentiable inverse function g^{-1} .

1. Introduction and assumptions

The computational problem of solving a nonlinear system

$$g(x) = 0 \text{ with } g: \mathcal{D} \subset \mathbf{R}^n \rightarrow \mathbf{R}^n$$

is closely related to that of minimising the sum of squares residual

$$\gamma(x) \equiv \frac{1}{2}\|g(x)\|^2: \mathcal{D} \rightarrow \mathbf{R}. \quad (1.1)$$

Here and throughout $\|\cdot\|$ denotes the 2-norm of real vectors and matrices. Due to the nonnegativity of γ , a point $x_* \in \mathcal{D}$ is a root of g if and only if it globally minimizes γ , i.e. $\gamma(x_*) = 0$. Unless g satisfies certain topological boundary

¹ Centre for Mathematical Analysis, Australian National University, G.P.O. Box 4, Canberra A.C.T. 2600, Australia. Present address: Mathematics Department, Southern Methodist University, Dallas, TX 75275.

© Copyright Australian Mathematical Society 1986, Serial-fee code 0334-2700/86

conditions [2], the existence of a root can usually not be guaranteed a priori. Hence we adopt the customary approach of minimizing $\gamma(x)$, which may however lead to the boundary of \mathcal{D} or a local minimizer \hat{x} with $\gamma(\hat{x}) > 0$. Since such points represent least squares solution of $g(x) = 0$ they may still be of some practical interest.

In our convergence analysis we avoid these difficulties by making the following assumptions.

–All points $x \in \mathbf{R}^n$ can be tested for feasibility, i.e. whether they belong to the open domain \mathcal{D} . Hence we may set

$$\gamma(x) \equiv \infty \quad \text{if } x \notin \mathcal{D}. \quad (1.2)$$

–The starting point $x_0 \in \mathbf{R}^n$ belongs to a compact level set

$$\Omega \equiv \{x \in \mathcal{D}: \gamma(x) \leq \omega\}$$

on which g has a right inverse function, i.e.

$$x, z \in \Omega, g(x) = g(z) \Rightarrow x = z \quad (1.3)$$

–The vector function $g(x)$ has a Jacobian $G(x) = \nabla g(x)$ that is nonsingular and Lipschitz continuous at all $x \in \Omega$.

By standard arguments these assumptions imply the following properties of g on Ω .

LEMMA 1. *There exist positive constants $\sigma_0 \leq \sigma_1$ and τ such that for all $x, z \in \Omega$ with $x \neq z$*

$$\|G(x) - G(z)\| \leq \tau \|x - z\| \quad (1.4)$$

$$\sigma_0 \leq \|G(x)^{-1}\|^{-1} \leq \|G(x)\| \leq \sigma_1 \quad (1.5)$$

$$\sigma_0 \leq \|g(x) - g(z)\| / \|x - z\| \leq \sigma_1 \quad (1.6)$$

which implies in particular

$$\sigma_0 \leq \|g(x)\| / \|x - x_*\| \leq \sigma_1 \quad (1.7)$$

where

$$\{x_*\} = \{x \in \mathcal{D}: g(x) = 0\} = \gamma^{-1}(0)$$

is the unique zero of g in \mathcal{D} .

PROOF. Using the compactness of Ω all assertions are easily established by contradiction.

It is well known [11] that under the above assumptions, Newton's method converges provided it is "damped" or stabilized with a suitable line search. Here this result is generalized to certain quasi-Newton methods based on fixed-scale least-change secant updates, as defined in [4]. While allowing x_0 to be any point in Ω , we place no restrictions on the initial approximation $B_0 \in \mathbf{R}^{n \times n}$ at all.

Our assumptions on g are certainly quite strong, but can perhaps be compared to the conditions under which Powell could show the “global” convergence of the *BFGS* method for unconstrained optimization [15]. The same author has devised a rather complicated modification of Broyden’s method [13] that requires the monitoring of previous steps with respect to their independence. A simpler and more popular approach is to estimate the Jacobian by differencing whenever no satisfactory progress is made. In either case one has some control over the gradient of γ , and can therefore enhance the global convergence properties by introducing an element of steepest descent with respect to γ . This allows, in particular, progress through points at which $G(x)$ is singular but γ not stationary. For the method to be proposed, the assumption of nonsingularity is essential, although modifications along the lines mentioned above are certainly possible. However, since the thrust of this paper is mainly theoretical, we prefer to maintain the surprising simplicity of the approach at this stage. Abbreviating $\gamma_k \equiv \gamma(x_k)$, $g_k \equiv g(x_k)$ and $G_k \equiv G(x_k)$, we may specify the basic structure of the scheme as follows.

- (i) Initialise $k = 0$, $x_0 \in \Omega$ and $B_0 \in \mathbb{R}^{n \times n}$.
- (ii) Compute a nonzero search direction s_k such that

$$B_k s_k = 0 \quad \text{or} \quad B_k s_k = -g_k. \tag{1.8}$$

- (iii) Determine a step multiplier $\alpha_k \in \mathbb{R}$ such that

$$x_{k+1} = x_k + \alpha_k s_k \in \mathcal{D}, \quad \gamma_{k+1} \leq \gamma_k$$

and some other conditions are satisfied.

- (iv) Update B_k to B_{k+1} such that

$$B_{k+1} s_k = y_k = \begin{cases} (g_{k+1} - g_k) / \alpha_k & \text{if } \alpha_k \neq 0, \\ G_k s_k & \text{if } \alpha_k = 0. \end{cases} \tag{1.9}$$

- (v) Provided $g_{k+1} \neq 0$ increment $k = k + 1$ and to to (ii).

The only nonstandard features of this algorithm are that s_k can be chosen as a null vector of B_k , and that the step multiplier α_k may be negative or zero. The last case never occurs in the presence of rounding errors, so that $G_k s_k$ is effectively replaced by a divided difference along s_k . The updating techniques discussed in Section 3 are quite conventional, but their analysis in close interplay with the line search introduced in Section 2 appears to be new. Abbreviating

$$\delta_k \equiv \|g_k + G_k s_k\| / (\|g_k\| + \sigma_0 \|s_k\|), \tag{1.10}$$

we may list the main stages of the analysis as follows.

- (a) The line search ensures that always

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^2 < \infty. \tag{1.11}$$

(b) Then the updating procedure yields

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} \delta_k^2 = 0. \quad (1.12)$$

(c) Since the line search avoids unnecessarily small steps it follows that

$$\lim_{k \rightarrow \infty} x_k = x_* \in g^{-1}(0). \quad (1.13)$$

(d) Because the full step $\alpha_k = 1$ is taken when possible we have

$$\gamma_{k+1}/\gamma_k \leq c(\delta_k^2 + \gamma_k). \quad (1.14)$$

(e) This implies with (b) and (c) R -superlinear convergence, i.e.

$$\lim_{k \rightarrow \infty} \|x_k - x_*\|^{1/k} = 0. \quad (1.15)$$

(f) The resulting summability of the steps ensures

$$\lim_{k \rightarrow \infty} \delta_k = 0. \quad (1.16)$$

(g) Finally it follows by standard arguments that

$$\lim_{k \rightarrow \infty} \|x_{k+1} - x_*\| / \|x_k - x_*\| = 0. \quad (1.17)$$

The assertions (a), (c), (d), (e) and (g) are developed in Section 2 and the remaining points (b) and (f) are treated in Section 3. The paper concludes with a summary and discussion in the final Section 4.

2. A derivative-free line search

Given a search direction $s_k \neq 0$ at the current iterate $x_k \in \Omega$, we would like to find a step multiplier $\alpha_k \in \mathbf{R}$ that minimizes the residual

$$\gamma_k(\alpha) \equiv \gamma(x_k + \alpha s_k).$$

Since the iterative solution of this one-dimensional minimization problem could be quite costly, one prefers to accept values for α_k that satisfy considerably weaker line-search conditions. Because most standard step-size rules were designed for unconstrained optimization, they typically involve the directional derivative

$$\left. \frac{d}{d\alpha} \gamma_k(\alpha) \right|_0 = \nabla \gamma(x_k)^T s_k.$$

Moreover, a search direction s_k is usually only accepted if it can be ensured explicitly or implicitly that the angle between s_k and the direction of steepest descent $-\nabla \gamma(x_k)$ is sufficiently acute.

For example, convergence to a stationary point of γ can be established if the line search is “efficient” as defined in [15], and the search directions satisfy the Zoutendijk condition [16], i.e.

$$\sum_{k=0}^{\infty} \left(\frac{\nabla\gamma(x_k)^T s_k}{\|\nabla\gamma(x_k)\| \|s_k\|} \right)^2 = \infty. \tag{2.1}$$

Here it is assumed that $\nabla\gamma(x_k)s_k \leq 0$ for all k , which cannot be guaranteed in our case, as the gradient takes the form

$$\nabla\gamma(x)^T = g(x)^T G(x) \tag{2.2}$$

and is therefore, like the Jacobian $G(x) = \nabla g(x)$, unknown. However, the derivative-free line search developed below achieves much the same effect as, for example, the popular Goldstein test [5].

Following a suggestion of Lindström and Wedin [10], as well as Al-Baali and Fletcher [1], we interpolate the vector function

$$g_k(\alpha) \equiv g(x_k + \alpha s_k): \mathbf{R} \rightarrow \mathbf{R}^n \tag{2.3}$$

rather than just the scalar $\gamma_k(\alpha)$. Based on the values of $g_k(\alpha)$ for $\alpha = 0$ and some $\alpha_c \neq 0$ we can form the linear approximation

$$g_k(\alpha) \approx g_k + [g_k(\alpha_c) - g_k] \alpha / \alpha_c \tag{2.4}$$

where $g_k = g(x_k) = g_k(0)$ as before. It is easily seen that the Euclidean norm of the right-hand side has the unique minimizer $q_k(\alpha_c) \cdot \alpha_c$ where

$$q_k(\alpha) \equiv \frac{g_k^T(g_k - g_k(\alpha))}{\|g_k - g_k(\alpha)\|^2}: \mathbf{R} - \{0\} \rightarrow \mathbf{R} \tag{2.5}$$

and in agreement with the convention (1.2)

$$q_k(\alpha) = 0 \quad \text{if } g_k(\alpha) \notin \mathcal{D}. \tag{2.6}$$

As depicted in Figure 1 for fixed g_k , the vectors $g_k(\alpha)$ yielding a certain value q of $q_k(\alpha)$ form spheres with radius $1/|2q|$ about the centre $(1 - \frac{1}{2}/q)g_k$.

Now our first and main requirement on the step multipliers α_k is that for some constant $\epsilon \in (0, 1/6)$

$$q_k(\alpha) \geq \frac{1}{2} + \epsilon \quad \text{or} \quad \alpha_k = 0. \tag{2.7}$$

Since the reduction of $\|g\|$ can be expressed as

$$\|g_k\|^2 - \|g_k(\alpha_k)\|^2 = [2q_k(\alpha_k) - 1] \cdot \|g_k - g_k(\alpha_k)\|^2 \geq 0 \tag{2.8}$$

it follows from (1.6) and (2.7) for $x_{k+1} = x_k + \alpha_k s_k$ that

$$\sigma_0^2 \|\alpha_k s_k\|^2 \leq \|g_k - g_{k+1}\|^2 \leq (\gamma_k - \gamma_{k+1})/\epsilon.$$

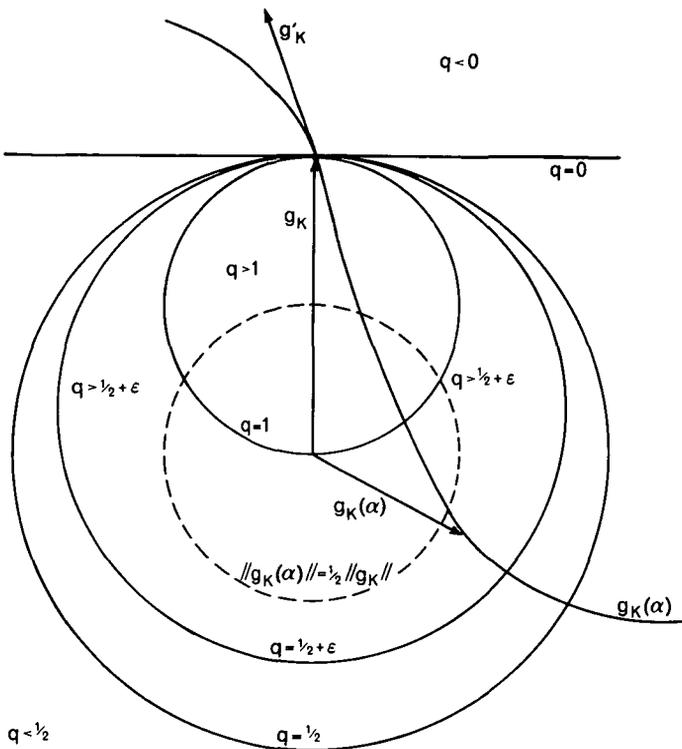


FIGURE 1

Now summing over k yields

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^2 \leq \gamma_0 / (\epsilon \sigma_0^2) < \infty \tag{2.9}$$

as claimed in equation (1.11) of the introduction. As will be shown in Section 3 the square summability of the steps implies that fixed scale secant updates produce infinitely many good, i.e. Newton-like, directions. To utilise these effectively we must ensure that the step multipliers α_k are not chosen unnecessarily small. This second condition is implicitly imposed by the following line search algorithm.

Starting from $\alpha_k^{(0)} = 1$ perform the iteration

$$\alpha_k^{(i+1)} = \alpha_k^{(i)} h(q_k(\alpha_k^{(i)})) \tag{2.10}$$

where the upper semi-continuous function $h(q): \mathbf{R} \rightarrow \mathbf{R}$ is defined by its graph in Figure 2.

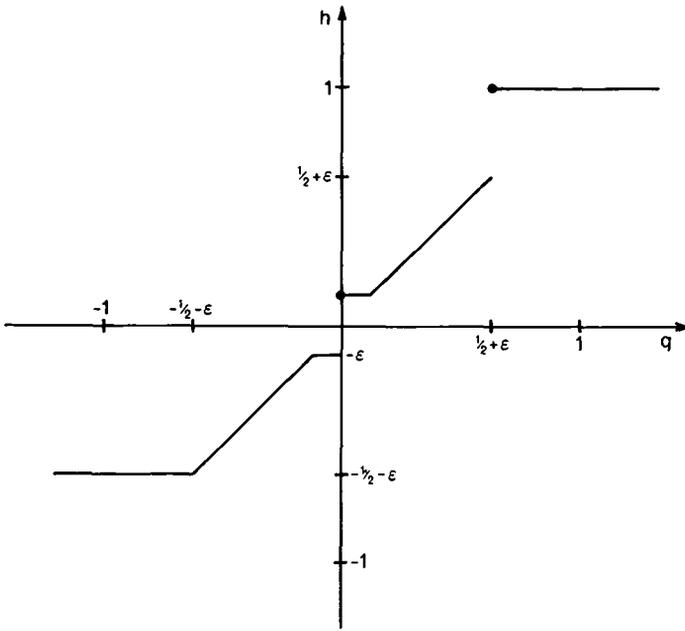


FIGURE 2

Since for all $i \geq 0$

$$\alpha_k^{(i+1)} = \alpha_k^{(i)} \quad \text{or} \quad 0 < \epsilon \leq |\alpha_k^{(i+1)}|/|\alpha_k^{(i)}| \leq \frac{1}{2} + \epsilon \leq \frac{2}{3}, \tag{2.11}$$

there exists obviously a limit

$$\alpha_k = \lim_{i \rightarrow \infty} \alpha_k^{(i)} \in [-1, 1]. \tag{2.12}$$

Normally one would expect that, after only a few readjustments, the sequence $\langle \alpha_k^{(i)} \rangle_{i=1}^\infty$ becomes stationary at some limit $\alpha_k \neq 0$. This follows by inspection of Figure 1 if the tangent

$$g'_k \equiv \left. \frac{d}{d\alpha} g_k(\alpha) \right|_0 = G_k s_k \tag{2.13}$$

is nonorthogonal to g_k , or equivalently

$$\nabla \gamma(x_k)^T s_k = g_k^T g'_k \neq 0 \tag{2.14}$$

Then the smooth curve $g_k(\alpha)$ must intersect the open ball $q > \frac{1}{2} + \epsilon$ for all sufficiently small α with $\text{sign}(\alpha) = -\text{sign}(g_k^T g'_k)$. The $\alpha_k^{(i)}$ cannot approach 0 from the wrong side since a negative value of $q_k(\alpha_k^{(i)})$ will immediately cause a sign change. The only potentially troublesome case arises when $\nabla \gamma(x_k)^T s_k$ is exactly equal to zero. Then the curve $g_k(\alpha)$ may intersect the closure of the ball

$q > \frac{1}{2} + \epsilon$ only at g_k , in which case the $\alpha_k^{(i)}$ form an infinite sequence converging to $\alpha_k = 0$. However, in the presence of rounding errors, the denominator of $q_k(\alpha_k^{(i)})$ as defined in (2.5) declines faster than its numerator, whose sign varies at random. Therefore some small $\alpha_k^{(i)}$ will always be accepted as a suitable choice for α_k . Then the actual step from x_k to x_{k+1} will be negligible but the update from B_k to B_{k+1} should result in a better search direction at the next iteration. Using only the assumption that there is a subsequence of increasingly Newton-like search directions, we obtain the following “global” convergence result.

THEOREM 2. *Stages (a), (c) and (d).*

Let $\langle x_k \rangle_{k=0}^\infty \subset \mathbf{R}^n$ be any sequence generated from some $x_0 \in \Omega$ by the recurrence

$$x_{k+1} = x_k + \alpha_k s_k \text{ with } s_k \neq 0 \text{ arbitrary}$$

and $\alpha_k \in \mathbf{R}$ calculated according to (2.10) and (2.12). Then the steps are square summable, i.e.,

$$\sum_{k=0}^\infty \|x_{k+1} - x_k\|^2 < \infty \tag{2.15}$$

and the property

$$\liminf_{k \rightarrow \infty} \delta_k = 0 \tag{2.16}$$

ensures convergence as such, i.e.,

$$\lim_{k \rightarrow \infty} x_k = x_* \in g^{-1}(0). \tag{2.17}$$

Moreover we have for all k greater than some $k_0 \geq 0$

$$\gamma_{k+1}/\gamma_k \leq (36\delta_k^2 + 4\tau^2\gamma_k/\sigma_0^4) \tag{2.18}$$

where δ_k is defined in (1.10).

PROOF. The first assertion (2.15) has already been established in (2.9). Since the residuals $\gamma_k = \frac{1}{2}\|g_k\|^2$ form a nonincreasing sequence, all x_k must belong to the level set $\Omega = \gamma^{-1}[0, \omega]$. Because of (2.16), there are infinitely many indices k with $\delta_k \leq 1/6$, which implies by elementary arguments

$$\sigma_0/1.4 \leq \|g_k\|/\|s_k\| \leq 1.2(\sigma_0 + \sigma_1). \tag{2.19}$$

An immediate consequence of the uniform Lipschitz continuity property (1.4) is that for all pairs $x, x + s \in \Omega$

$$\|g(x + s) - g(x) - G(x)s\| \leq \frac{1}{2}\tau\|s\|^2. \tag{2.20}$$

Hence we find that for all $\alpha \in [-1, 1]$ with $g_k(\alpha)$ as defined in (2.3)

$$\begin{aligned} \|g_k(\alpha) - (1 - \alpha)g_k\| &= \|g_k(\alpha) - g_k - \alpha G_k s_k + \alpha(g_k + G_k s_k)\| \\ &\leq \frac{1}{2}\tau\alpha^2\|s_k\|^2 + |\alpha|\|g_k + G_k s_k\| \leq \tau\alpha^2\|g_k\|^2/\sigma_0^2 + \delta_k(\|g_k\| + \sigma_0\|s_k\|)\alpha \\ &\leq \frac{1}{2}|\alpha|\|g_k\|(2|\alpha|\tau\omega/\sigma_0^2 + \frac{1}{3}(1 + 1.4)) \leq \frac{1}{2}|\alpha|\|g_k\|, \end{aligned}$$

where the last inequality holds provided

$$|\alpha| \leq \bar{\alpha} \equiv \sigma_0^2/\max\{10\omega\tau, 4\sigma_1^2\}.$$

Consequently we have for the numerator of $q_k(\alpha)$

$$|g_k^T[g_k - g_k(\alpha)] - \alpha g_k^T g_k| \leq \frac{1}{2}|\alpha|\|g_k\|^2 = |\alpha|\gamma_k$$

so that $q_k(\alpha)$ itself satisfies by (2.5)

$$\alpha q_k(\alpha) \geq \alpha^2\gamma_k/\|g_k - g_k(\alpha)\|^2 \geq \gamma_k(\sigma_1\|s_k\|)^{-2} \geq \frac{1}{4}(\sigma_0/\sigma_1)^2,$$

where the last two inequalities follow from (1.6) and (2.19). Thus we have for all $\alpha \neq 0$ with $|\alpha| < \bar{\alpha}$

$$|q_k(\alpha)| \geq 1 \text{ and } \alpha q_k(\alpha) > 0. \tag{2.21}$$

Now suppose that $|\alpha_k| \leq \bar{\alpha}\epsilon^2$. Then the line search procedure must have generated two previous trials points $\alpha_k^{(-1)}$ and $\alpha_k^{(-2)}$ with

$$|\alpha_k^{(-2)}| \leq \bar{\alpha} \text{ and } |\alpha_k^{(-1)}| \leq (2/3)|\alpha_k^{(-2)}|. \tag{2.22}$$

However, if $\alpha_k^{(-2)}$ was positive the relations (2.21) guarantee $q_k(\alpha_k^{(-2)}) \geq 1 > \frac{1}{2} + \epsilon$ so that it should have been accepted as α_k . On the other hand if $\alpha_k^{(-2)} < 0$ then $q_k(\alpha_k^{(-2)}) < -1$ would have resulted in $\alpha_k^{(-1)}$ being positive so that it should have been accepted as α_k . Thus we have a contradiction and $|\alpha_k|$ must be greater than $\bar{\alpha}\epsilon^2$ whenever $\delta_k \leq 1/6$. Since there are infinitely many such steps and $\alpha_k s_k \rightarrow 0$ it follows with (2.19) that

$$\lim_{k \rightarrow \infty} \inf \|g_k\| \leq 1.2(\sigma_0 + \sigma_1) \cdot \liminf_{k \rightarrow \infty} \|s_k\| = 0.$$

Due to the compactness of Ω and the steady reduction of $\|g_k\|$ this implies that the whole sequence converges to the unique root x_* . At all steps with $\delta_k \geq 1/6$ the final assertion is trivial. Otherwise it follows from (2.20) that

$$\begin{aligned} \|g(x_k + s_k)\| &= \|g(x_k + s_k) - g_k - G_k s_k + g_k + G_k s_k\| \\ &\leq \|g_k + G_k s_k\| + \frac{\tau}{2}\|s_k\|^2 \leq \delta_k(\|g_k\| + \sigma_0\|s_k\|) + \tau\|g_k\|^2/\sigma_0^2 \\ &\leq \|g_k\|[2.4\delta_k + \tau\|g_k\|/\sigma_0^2] \leq \|g_k\|[.4 + \tau\|g_k\|/\sigma_0^2]. \end{aligned} \tag{2.23}$$

Because of (2.17) we have $\tau\|g_k\|/\sigma_0^2 \leq .1$ for all sufficiently large k , so that then

$$\|g(x_k + s_k)\| \leq \frac{1}{2}\|g_k\|,$$

which implies by inspection of Figure 1 that $q_k(1) \geq 2/3 \geq \frac{1}{2} + \varepsilon$, since always $\varepsilon \in (0, 1/6)$. Hence the full step $\alpha_k = 1$ must be accepted when $\delta_k \leq 1/6$ and k is sufficiently large. Thus (2.18) follows from (2.23) by the Hölder inequality.

The final inequality (2.18) holds because once the iterates x_k are reasonably close to x_* then the line search accepts the full step $\alpha_k = 1$ whenever the search direction s_k is similar to the Newton correction $-G_k^{-1}g_k$. Provided this occurs sufficiently often one obtains superlinear convergence as shown below.

COROLLARY 3. *Stages (e) and (g).*

Under the assumptions of Theorem 2 the condition

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} \delta_k^2 = 0 \tag{2.24}$$

implies R-superlinear convergence in that

$$\lim_{k \rightarrow \infty} \|x_k - x_*\|^{1/k} = 0 \tag{2.25}$$

which ensures in particular that

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\| < \infty. \tag{2.26}$$

If furthermore

$$\lim_{k \rightarrow \infty} \delta_k \rightarrow 0 \tag{2.27}$$

then the rate of convergence is in fact Q-superlinear in that

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0. \tag{2.28}$$

PROOF. By the inequality of the geometric and arithmetic means we derive from (2.18)

$$\begin{aligned} (\gamma_l/\gamma_0)^{1/l} &= \left(\prod_{k=0}^{l-1} \gamma_{k+1}/\gamma_k \right)^{1/l} \leq \frac{1}{l} \sum_{k=0}^{l-1} \gamma_{k+1}/\gamma_k \\ &\leq 36 \frac{1}{l} \sum_{k=0}^{l-1} \delta_k^2 + \frac{\tau^2}{\sigma_0^4} \frac{4}{l} \sum_{k=0}^{l-1} \gamma_k \end{aligned}$$

where we have assumed without loss of generality that $k_0 = 0$. Since $\gamma_k \rightarrow 0$ the last term goes to zero as l tends to infinity, so that the condition (2.24) implies

$$\lim_{l \rightarrow \infty} (\gamma_l)^{1/l} / \lim_{l \rightarrow \infty} (\gamma_0)^{1/l} = \lim_{l \rightarrow \infty} (\gamma_l)^{1/l} = 0,$$

which proves assertion (2.25) as by (1.7)

$$\|x_k - x_*\|^2 \leq 2\gamma_k/\sigma_0^2.$$

Since then for all k greater than or equal to some $k_0 \geq 0$

$$\|x_k - x_\star\|^{1/k} \leq \frac{1}{2},$$

it follows by the triangle inequality that

$$\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\| \leq 2 \sum_{k=k_0}^{\infty} \|x_k - x_\star\| \leq 2.$$

The last assertion (2.28) is a consequence of (2.18), (2.27) and the fact that again by (1.7)

$$\left(\frac{\sigma_0}{\sigma_1}\right)^2 \leq \left(\frac{\gamma_{k+1}}{\gamma_k}\right) / \frac{\|x_{k+1} - x_\star\|^2}{\|x_k - x_\star\|^2} \leq \left(\frac{\sigma_1}{\sigma_0}\right)^2,$$

which completes the proof.

The condition (2.27) was originally shown by Dennis and Moré [3] to be necessary and sufficient for the Q -superlinear rate of a convergent sequence. Usually their result is applied after at least R -linear convergence has been established by some other technique. Here we prove directly R -superlinear convergence using the condition (2.24) which is sufficient but not necessary. The same condition was employed in [6] to establish local and R -superlinear convergence of Broyden’s method assuming merely continuity of the Jacobian $G(x) = \nabla g(x)$ near x_\star . However, as will be seen in the next section, Lipschitz continuity of $G(x)$ seems required for the global part of our argument. Otherwise even fixed-scale updates may fail to produce any Newton-like search direction at all.

3. Asymptotic properties of fixed scale updates

At first we will consider a slightly more general situation than in the first two sections. Let $g: \mathcal{D} \subset \mathbf{R}^n \rightarrow \mathbf{R}^n$ be some vector-function whose Jacobian $G(x) \equiv \nabla g(x)$ belongs to all $x \in \mathcal{D}$ to some subspace $\mathcal{B} \subset \mathbf{R}^{n \times n}$. For example $G(x)$ could be sparse, symmetric or have a certain nullspace at all $x \in \mathcal{D}$. Moreover we require that $G(x)$ is uniformly Lipschitz continuous on some subset $\Omega \subset \mathcal{D}$ such that for some constant $\tau > 0$

$$\|G(x) - G(z)\| \leq \tau \|x - z\|, \tag{3.1}$$

where x and z range over all elements in Ω .

Now given $x_0 \in \Omega$ let $\langle s_k \rangle_{k=0}^\infty$ and $\langle \alpha_k \rangle_{k=0}^\infty$ be any two sequences of nonzero search directions s_k and real multipliers α_k such that

$$x_{k+1} = x_k + \alpha_k s_k \in \Omega \quad \text{for } k = 0, 1, \dots$$

Starting from some matrix $B_0 \in \mathbb{R}^{n \times n}$ we may then recursively generate a sequence of approximating Jacobians $\langle B_k \rangle_{k=0}^\infty \subset \mathcal{B}$ such that for each k

$$B_{k+1}s_k = y_k \equiv \begin{cases} (g_{k+1} - g_k)/\alpha_k & \text{if } \alpha_k \neq 0, \\ G_k s_k & \text{if } \alpha_k = 0, \end{cases} \tag{3.2}$$

where naturally $g_k = g(x_k)$ and $G_k = G(x_k)$.

By the fundamental theorem of calculus, this so-called secant condition on B_{k+1} is satisfied by the matrix

$$M_{k+1} = \int_0^1 G(x_{k+1}(1 - \mu) + \mu x_k) d\mu \in \mathcal{B}. \tag{3.3}$$

It is customary to remove the remaining freedom in choosing B_{k+1} by minimising the correction matrix

$$C_k = B_{k+1} - B_k \in \mathcal{B} \tag{3.4}$$

with respect to some inner product norm

$$|C| = \langle C, C \rangle^{1/2} \text{ for } C \in \mathcal{B}.$$

Due to the norm equivalence on finite-dimensional spaces we may assume without loss of generality that for some constant ρ

$$1 \leq |C|/\|C\| \leq \rho < \infty, \tag{3.5}$$

where $\|\cdot\|$ denotes the 2-norm as before. This relation holds with $\rho = \sqrt{n}$ in case of the popular Frobenius norm

$$|C|_F^2 = \text{Trace}(CC^T) = \text{Trace}(C^TC).$$

An important consequence of (3.5) is that for any $s \in \mathbb{R}^n - \{0\}$

$$\|Cs\|/\|s\| \leq \|C\| \leq |C|. \tag{3.6}$$

Moreover the affine variety

$$[y_k/s_k] \equiv \{B \in \mathcal{B} : Bs_k = y_k\} \tag{3.7}$$

is closed and nonempty as it contains the matrix M_{k+1} . Hence there exists a unique “least-change update” $B_{k+1} = B_k + C_k$ such that for all $B \in [y_k/s_k]$

$$|C_k| = |B_{k+1} - B_k| \leq |B - B_k|.$$

Since $|\cdot|$ is an inner product norm this condition is equivalent to the requirement that

$$\{C_k\} = [r_k/s_k] \cap [0/s_k]^\perp,$$

where

$$r_k = y_k - B_k s_k = C_k s_k \tag{3.8}$$

and the affine varieties $[r_k/s_k]$ and $[0/s_k]$ are defined in analogy to (3.7). Now since C_k is orthogonal to the discrepancy

$$D_{k+1} \equiv B_{k+1} - M_{k+1} \in [0/s_k]$$

we obtain by Pythagoras' theorem the crucial inequality

$$|D_{k+1}|^2 = |M_{k+1} - B_k|^2 - |C_k|^2. \tag{3.9}$$

Using the shift matrices

$$E_k \equiv M_{k+1} - M_k \in \mathcal{B}$$

we derive from (3.9) by the triangle inequality

$$\begin{aligned} |D_{k+1}| - |D_k| &\leq |E_k| \leq |M_{k+1} - G_k| + |M_k - G_k| \\ &\leq \tau p (\|x_{k+1} - x_k\| + \|x_k - x_{k-1}\|), \end{aligned} \tag{3.10}$$

where the last inequality is an elementary consequence of (3.1) and (3.5).

This form of bounded deterioration [3] does not apply to iteratively rescaled updates [4] where the inner product norm $|\cdot|$ changes slightly from step to step. For example the convex class updates for unconstrained minimisation [3] can be interpreted in this way. These formulae are not covered by the following result, as it is very much dependent on the constancy of $|\cdot|$.

THEOREM 4. *Under the assumptions of this section the condition*

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^p < \infty \quad \text{for } p = 2 \text{ or } p = 1 \tag{3.11}$$

ensures that

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} (\|r_k\|/\|s_k\|)^2 = 0 \quad \text{if } p = 2 \tag{3.12}$$

and

$$\lim_{k \rightarrow \infty} \|r_k\|/\|s_k\| = 0 \quad \text{if } p = 1. \tag{3.13}$$

PROOF. As an immediate consequence of (3.10) we note that

$$|D_k| \leq |D_0| + \sum_{j=0}^{k-1} |E_j|.$$

Substituting this estimate into a rearranged version of (3.9) we find that

$$\begin{aligned} |C_k|^2 &\leq |D_k|^2 + 2|E_k||D_k| + |E_k|^2 - |D_{k+1}|^2 \\ &\leq |D_k|^2 - |D_{k+1}|^2 + |E_k|^2 + 2|E_k| \left(|D_0| + \sum_{j=0}^{k-1} |E_j| \right). \end{aligned} \tag{3.14}$$

Now summing over $k = 0, 1, \dots, l - 1$ and collecting terms one obtains the simple expression

$$\begin{aligned} \sum_{k=0}^{l-1} |C_k|^2 &\leq \left(|D_0| + \sum_{k=0}^{l-1} |E_k| \right)^2 \\ &\leq 2 \left(|D_0| + \sum_{k=0}^{m-1} |E_k| \right)^2 + 2 \left(\sum_{k=m}^{l-1} |E_k| \right)^2, \end{aligned} \tag{3.15}$$

where m is some integer less than l . With m constant and l tending to infinity we derive using Hölder’s inequality that

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} |C_k|^2 \leq \lim_{l \rightarrow \infty} 2 \frac{(l - m)}{l} \sum_{k=m}^l |E_k|^2 \leq 2 \sum_{k=m}^{\infty} |E_k|^2.$$

Since (3.10) and (3.11) imply that the $|E_k|$ are square summable, the rightmost sum is finite and goes to zero as m tends to infinity. Thus we have shown that

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} (\|r_k\| / \|s_k\|)^2 \leq \lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} |C_k|^2 = 0,$$

where the first inequality follows from (3.6) and (3.8). Since the stronger condition (3.11) with $p = 1$ implies by (3.10) the summability of the $|E_k|$ themselves, the second assertion (3.13) is an immediate consequence of (3.15).

In this section we have so far not assumed any particular relationship between the search directions s_k and the approximations B_k or even the iterates x_k . Instead of applying Theorem 4 directly to the vector function g whose zero x_* is sought we may also consider the partitioned case

$$g(x) \equiv \sum_{i=1}^m g^{(i)}(x^{(i)}) \quad \text{with } x^{(i)} \equiv S^{(i)}x. \tag{3.16}$$

Here the element functions $g^{(i)}$ are defined on the range of the linear mappings $S^{(i)}: \mathbf{R}^n \rightarrow \mathbf{R}^n$. Provided the Jacobians $G^{(i)} = \nabla g^{(i)}$ are restricted to subspaces $\mathcal{B}^{(i)} \subset \mathbf{R}^{n \times n}$ one can then form individual approximations $B_k^{(i)} \in \mathcal{B}^{(i)}$ based on the secant conditions

$$B_{k+1}^{(i)} s_k^{(i)} = y_k^{(i)} \equiv \begin{cases} (g_{k+1}^{(i+1)} - g_k^{(i)}) / \alpha_k & \text{if } \alpha_k \neq 0, \\ G_k^{(i)} s_k^{(i)} & \text{if } \alpha_k = 0, \end{cases} \tag{3.17}$$

where naturally $g_k^{(i)} = g^{(i)}(x_k)$. Then the overall approximation

$$B_{k+1} = \sum_{i=1}^m B_{k+1}^{(i)} S^{(i)} \simeq G(x_{k+1}) \tag{3.18}$$

satisfies the usual secant condition

$$B_{k+1} s_k = \sum_{i=1}^m y_k^{(i)} = y_k = M_{k+1} s_k.$$

Similarly we find that the vectors

$$r_k^{(i)} = y_k^{(i)} - B_k^{(i)}s_k^{(i)}$$

add up to the overall approximation error

$$r_k = \sum_{i=1}^m r_k^{(i)} = y_k - B_k s_k.$$

Hence we have by the triangle inequality

$$\|r_k\|/\|s_k\| \leq \sum_{i=1}^m \|r_k^{(i)}\|/\|s_k\| \leq \sum_{i=1}^m \|S^{(i)}\|^{-1} \|r_k^{(i)}\|/\|s_k^{(i)}\|, \tag{3.19}$$

where the last inequality follows from

$$\|s_k^{(i)}\| = \|S^{(i)}s_k\| \leq \|S^{(i)}\| \|s_k\|.$$

Now if the steps $x_{k+1} - x_k$ are p summable with $p = 2$ or $p = 1$ the same is true for the ‘‘projected’’ steps $S^{(i)}(x_{k+1} - x_k)$ and Theorem 4 ensures

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} \|r_k^{(i)}\|/\|s_k^{(i)}\| = 0 \quad \text{if } p = 2 \tag{3.20}$$

or

$$\lim_{k \rightarrow \infty} \|r_k^{(i)}\|/\|s_k^{(i)}\| = 0 \quad \text{if } p = 1 \tag{3.21}$$

provided the $g^{(i)}$ are uniformly Lipschitz continuous. By combining (3.19) with (3.20) or (3.21) we can therefore conclude that Theorem 4 remains valid if the Jacobian of g is approximated in the partitioned form (3.18). Obviously no harm is done if some of the element functions $g^{(i)}$ are differentiated explicitly so that $B_k^{(i)} = \nabla g^{(i)}(x_k^{(i)})$ for all k .

For example the discretization of a boundary-value problem

$$u' = f(t, u), \quad A_0 u(0) + A_1 u(1) = b \in \mathbf{R}^m$$

by the box scheme [9] yields algebraic equations of the form

$$g(x) \equiv Ax + F(Sx) = 0 \quad \text{with } g: \mathbf{R}^{m \times p} \rightarrow \mathbf{R}^{m \times p} \tag{3.22}$$

where

$$(x)_i \approx u(t_i) \text{ at } 0 = t_1 < \dots < t_{p-1} < t_p = 1.$$

The only nonlinear component $F: \mathbf{R}^{m \times (p-1)} \rightarrow \mathbf{R}^{m \times (p-1)}$ has a block diagonal Jacobian ∇F so that the approximations $B_k \approx F'(Sx_k)$ are naturally restricted to the space \mathcal{B} of real matrices with that sparsity pattern. The least-change criterion with respect to the Frobenius norm on \mathcal{B} defines simply the Schubert [14] or sparse Broyden update. However it was found by Hart and Soul [8] that this approach yields much better numerical results than the direct approximation of

the overall Jacobian $\nabla g(x)$ by the same update formula. Some local convergence results for iteratively rescaled partitioned updates were given in [7]. Since those methods are mainly geared towards unconstrained optimization one can easily ensure convergence to stationary points, e.g., by a crude angle test. For the nonlinear equations considered here, the remaining gaps in our line of argument are closed by the following result.

COROLLARY 5. *Stages (b) and (f).*

Suppose g satisfies the assumptions listed in the introduction and the search direction $s_k \neq 0$ are chosen such that for all k

$$B_k s_k = 0 \quad \text{or} \quad B_k s_k = -g_k.$$

Then the condition

$$\sum_{k=0}^{\infty} |x_{k+1} - x_k|^p < \infty \quad \text{for } p = 2 \text{ or } p = 1 \tag{3.23}$$

implies

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=0}^{l-1} \delta_k^2 = 0 \quad \text{if } p = 2 \tag{3.24}$$

and

$$\lim_{k \rightarrow \infty} \delta_k = 0 \quad \text{if } p = 1. \tag{3.25}$$

PROOF. The assertions (3.24) and (3.25) follow directly from (3.12) and (3.13) as we will show that

$$\delta_k^2 \leq (\sigma_1/\sigma_0)^2 \left[2\|r_k\|^2/\|s_k\|^2 + \tau^2\|x_{k+1} - x_k\|^2 \right] / \sigma_0^2. \tag{3.26}$$

Since by (1.5) always

$$\delta_k \leq \|G_k s_k\| / \|\sigma_0 s_k\| \leq (\sigma_1/\sigma_0),$$

the above inequality holds trivially if $\|r_k\| \geq \|s_k\| \cdot \sigma_0/\sqrt{2}$. This applies in particular if $B_k s_k = 0$, since then

$$\|r_k\|/\|s_k\| = \|y_k\|/\|s_k\| \geq \sigma_0.$$

Otherwise we must have

$$r_k = y_k - B_k s_k = (g_{k+1} - g_k)/\alpha_k + g_k,$$

so that by (2.20) and the inverse triangle inequality

$$\|g_k + G_k s_k\| - \|r_k\| \leq \|G_k s_k - (g_{k+1} - g_k)/\alpha_k\| \leq \frac{1}{2}\tau|\alpha_k|\|s_k\|^2.$$

Thus division by $\|g_k\| + \sigma_0\|s_k\|$ yields

$$\delta_k \leq \|r_k\|/\|\sigma_0 s_k\| + \frac{1}{2}\tau\|x_{k+1} - x_k\|/\sigma_0,$$

which implies (3.26) by the Hölder inequality and the fact that $\sigma_1/\sigma_0 \geq 1$.

Numerical experiments on the boundary-value problems (3.22) treated in reference [8] confirmed the validity of our analysis. In particular it was found that the line search in combination with sparse Broyden approximations to ∇F yielded very similar global convergence properties to the same scheme with ∇F evaluated analytically. Naturally either scheme can get stuck near a local minimiser \hat{x} . In this situation a great number of function evaluations is wasted, with very little progress being made as the directions s_k become larger and larger, so that the α_k must tend to zero. Whereas this problem can be detected quite easily if ∇g and consequently $\nabla \gamma$ are explicitly available, there appears to be no natural remedy in the quasi-Newton case. This is probably not surprising since nobody has so far succeeded in devising a quasi-Newton method for the solution of inconsistent nonlinear least-squares problems. If the Jacobian ∇g is nonsingular at all x in Ω , but the restriction of g to Ω is not 1-1, the quasi-Newton scheme like Newton's method appears to converge without hesitation to the unique root in Ω . In other words, the injectivity assumption (1.3) seems to be merely technical. It would be eliminated by the inclusion of a trust-region strategy, that localises the problem whenever nonlinearities invalidate the Newton step.

4. Summary and discussion

It was shown that Broyden's method and similar fixed-scale quasi-Newton schemes can be made to converge from within any compact level set on which the vector function under consideration has a continuously differentiable inverse. This is achieved by a derivative-free line search, whose acceptance test involves not only the norms but also the angle between consecutive residual vectors. By allowing null vectors as search directions, any restriction on the initial Jacobian approximation could be avoided. The key to the convergence analysis was that fixed-scale secant updates produce on average Newton-like search directions whenever the steps are square summable.

It would appear intuitively clear that Broyden's method produces infinitely many Newton-like steps even if the steps tend to zero at a slower rate or the Jacobian is merely Hölder rather than Lipschitz continuous. However this does not follow from our analysis as the inequalities in the proof of Theorem 4 are all quite sharp. In contrast, an extension of the results to Hilbert spaces should provide little difficulty since the corresponding local convergence theory is fairly complete [6]. From a practical point of view, the main challenge lies in an efficient implementation of the line search, probably using quadratic rather than

linear interpolation. In the long run, one might hope for a generalization to genuine nonlinear least squares problems.

Acknowledgements

This work was partially supported by NSF grant DMS-8401023.

References

- [1] M. Al-Baali and R. Fletcher, "An efficient line search for nonlinear least squares", *Numerical Analysis Report NA/78*, Univ. of Dundee, 1984.
- [2] E. Allgower and K. Georg, "Simplicial and continuation methods for approximating fixed points and solutions to systems of equations", *SIAM Rev.* 22 (1980), 28–85.
- [3] J. E. Dennis and J. J. Moré, "Quasi-Newton methods, motivation and theory", *SIAM Rev.* 19 (1977), 46–89.
- [4] J. E. Dennis and J. F. Walker, "Convergence theorems for least change secant update methods", *SIAM J. Numer. Anal.* 18, (1982), 949–987.
- [5] A. A. Goldstein, "On steepest descent", *SIAM J. Control Optim.* 3 (1965), 147–151.
- [6] A. Griewank, "The local convergence of Broyden's method on Lipschitzian problems in Hilbert spaces", to appear in *SIAM J. Numer. Anal.*
- [7] A. Griewank and Ph.L. Toint, "Local convergence analysis for partitioned quasi-Newton updates", *Numer. Math.* 39 (1982), 429–448.
- [8] W. F. Hart and S. O. W. Soul, "Quasi-Newton methods for discretized nonlinear boundary problems", *Journal Institute of Maths. Applications* 11 (1973), 351–359.
- [9] H. B. Keller, "Numerical solution of two point boundary value problems", CBMS-NSF Regional Conference Series in Applied Mathematics, *SIAM*, Philadelphia, 1976.
- [10] P. Lindström and P. A. Wedin, "A new line search algorithm for nonlinear least squares problems", Report UMINF-82-81, University of Umea, 1981.
- [11] J. C. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).
- [12] M. R. Osborne, "An efficient weak line search with guaranteed termination", MRC Technical Summary Report #1870, University of Wisconsin, 1978.
- [13] M. J. D. Powell, "A hybrid method for nonlinear equations" in *Numerical Methods for Nonlinear Algebraic Equations*, (ed. P. Rabinowitz), (Gordon and Breach, London, 1970).
- [14] L. K. Schubert, "Modification of a quasi-Newton method for nonlinear equations with a sparse Jacobian", *Math. Comp.* 24, (1970), 27–30.
- [15] W. Warth and J. Werner, "Effiziente Schrittweitenfunktionen bei unrestringierten Optimierungsaufgaben", *Computing*, 19 (1) (1977), 59–72.
- [16] P. Wolfe, "Convergence conditions for ascent methods", *SIAM Rev.* 11 (2) (April 1969).