MATHEMATICS OF COMPUTATION, VOLUME 32, NUMBER 142
APRIL 1978, PAGES 447–451

On the Convergence of a Quasi-Newton Method for Sparse Nonlinear Systems

By Binh Lam

Abstract. It is shown that an algorithm for solving a system of nonlinear equations where the Jacobian is known to be sparse, converges locally and Q-superlinearly.

1. **Introduction.** Consider the problem of finding the solution of a system of nonlinear equations $F(x) = 0$, where F and x are n-dimensional vectors. Broyden [1] derived a quasi-Newton method using an iteration of the form

$$x_{k+1} = x_k + t_k p_k, \tag{1.1}$$

where t_k is a scalar and p_k is given by

$$B_k p_k = -F(x_k), \tag{1.2}$$

B_k being an approximation to the Jacobian. To avoid solving the system of linear equations (1.2), an approximation to the inverse Jacobian which is updated at every iteration by a single rank correction is used. However, this method has a drawback when applied to a system where the Jacobian is known to be sparse since the inverse of a sparse matrix is generally not sparse. Schubert [7] modified this method by updating B_k so that the sparsity is retained. It has been proved by Broyden [3] that the modified algorithm is locally convergent when the Jacobian satisfies a Lipschitz condition. He also reported that numerical results suggested that the convergence is superlinear in most cases. In this note, we show that the modified algorithm in fact preserves the convergence properties of the original method. It has a Q-superlinear rate of convergence when applied to linear systems. Furthermore, under certain conditions, the convergence is also Q-superlinear for nonlinear cases.

2. **Main Results.** Let S_j be a diagonal matrix whose (l, l) element is zero if the (j, l) element of the Jacobian is zero, and unity otherwise. To simplify the notation, we let B and B_1 denote the approximation to the Jacobian at kth and $(k + 1)$st step, respectively. Let B have the same sparseness characteristic as the Jacobian and B_1 be given by

$$B_1 = B - \sum_{j=1}^{n} u_j u_j^T (Bp_j - t^{-1} y) \frac{p_j^T}{p^T p_j}, \tag{2.1}$$

where $y = F(x_1) - F(x)$, $p_j = S_j p$ and u_j is the jth unit vector. We note that the sparsity is preserved in B_1 since $u_j^T B_1 = u_j^T B_1 S_j$.

Received January 25, 1977.
AMS (MOS) subject classifications (1970). Primary 65H10, 65K05.

We now prove that the convergence of the algorithm defined by (1.1), (1.2) and (2.1) is Q-superlinear for a linear system $F(x) = Ax + b$, where A is an $n \times n$ matrix.

Let

$$E = B - A, \qquad e = x - x^*,$$

and

$$\phi = \|E\|_F, \quad \epsilon = \|e\|,$$

where x^* is the solution of $F(x) = 0$. We use $\|\cdot\|$ to denote the Euclidean norm and $\|\cdot\|_F$ the Frobenius norm.

THEOREM 2.1. *If* $t_k = 1$ *for all* k *and* $\alpha\phi_0 < 1$, *where* $\alpha = \|A^{-1}\|_F$ *then*

$$\epsilon_k \leqslant (K/k^{1/2})^k \epsilon_0,$$

where $K = \alpha\phi_0/(1 - \alpha\phi_0)$, *when the algorithm is applied to the linear system* $F(x) = Ax + b$.

Proof. Since $t = 1$ and $y = Ap$, from (2.1) we have

$$E_1 = E - \sum_{j=1}^{n} u_j u_j^T E \frac{p_j p_j^T}{p_j^T p_j}. \tag{2.2}$$

Thus,

$$\|u_j^T E_1\|^2 = \left(u_j^T E - u_j^T E \frac{p_j p_j^T}{p_j^T p_j}\right)\left(u_j^T E - u_j^T E \frac{p_j p_j^T}{p_j^T p_j}\right)^T = \|u_j^T E\|^2 - \frac{|u_j^T E p_j|^2}{\|p_j\|^2},$$

on expanding the terms on the right-hand side. As

$$u_j^T E p_j = u_j^T E S_j p = u_j^T E p \quad \text{and} \quad \|p_j\|^2 \leqslant \|p\|^2,$$

$$\|u_j^T E_1\|^2 \leqslant \|u_j^T E\|^2 - \frac{|u_j^T E p|^2}{\|p\|^2}.$$

Summing over j, we obtain

$$\phi_1^2 \leqslant \phi^2 - \|Ep\|^2/\|p\|^2. \tag{2.3}$$

Since

$$\|Ep\|^2/\|p\|^2 \leqslant \|E\|_2^2 \leqslant \|E\|_F^2 \equiv \phi^2,$$

we have

$$\|Ep\|^2/\|p\|^2 = \theta\phi^2,$$

for some θ such that $0 \leqslant \theta \leqslant 1$. Hence, $\phi_1^2 \leqslant \phi^2(1 - \theta)$. In general, we have

$$\phi_{k+1}^2 \leqslant \phi_k^2(1 - \theta_k),$$

which implies

$$\phi_k^2 \leqslant \phi_0^2 \prod_{j=1}^{k} (1 - \theta_{k-j}).$$

Since $E = B - A$, $B^{-1} = (I + A^{-1}E)^{-1}$, and

$$p = -B^{-1}Ae = -(I - A^{-1}E)^{-1}e. \tag{2.4}$$

Thus, if $\alpha\phi < 1$, then

$$\|Ep\| \leqslant \|E\|_F \|p\| \leqslant \phi\epsilon/(1 - \alpha\phi). \tag{2.5}$$

From (2.4), we also have $e_{k+1} = -A^{-1}E_k p_k$; from this and (2.5),

$$\epsilon_{k+1}^2 \leqslant \theta_k \alpha^2 \frac{\phi_k^2 \epsilon_k^2}{(1 - \alpha\phi_k)^2} \leqslant \theta_k \alpha^2 \frac{\phi_k^2 \epsilon_k^2}{(1 - \alpha\phi_0)^2}$$

as $\phi_k \leqslant \phi_0$ by (2.3).

The proof now proceeds in the same fashion as that of Theorem 2 in [2]. We note that Frobenius norm is used here but this change of norm has no effect on the proof.

To analyze the convergence for nonlinear systems, we assume that F satisfies the following conditions:

(a) F is differentiable in an open convex set D in R^n, the linear space of n-dimensional vectors.

(b) For some $x^* \in D$ such that $F(x^*) = 0$, $F'(x^*)$ is nonsingular and F' is continuous at x^*.

(c) F' satisfies a Lipschitz condition of order one at x^* so there exists a positive constant L such that

$$\|F'(x) - F'(x^*)\| \leqslant L\|x - x^*\|. \tag{2.6}$$

We need the following result which is a special case of a more general theorem proved by Broyden, Dennis and Moré [4].

THEOREM 2.2. *Suppose F satisfies assumptions* (a), (b), (c), *and for all k,*

$$\|B_{k+1} - F'(x^*)\|_F \leqslant \|B_k - F'(x^*)\|_F + \alpha\sigma_k, \tag{2.7}$$

where α is some constant and $\sigma_k = \max\{\|x_{k+1} - x^*\|, \|x_k - x^*\|\}$. *Then there exist positive constants ϵ and δ such that if* $\|x_0 - x^*\| < \epsilon$ *and* $\|B_0 - F'(x^*)\|_F < \delta$, *the sequence* (1.1) *with $t_k = 1$ is well defined and converges linearly to x^*.*

We have the following result.

THEOREM 2.3. *If F satisfies assumptions* (a), (b) *and* (c), *then the algorithm defined by* (1.1), (1.2), (2.1) *with $t_k = 1$ is locally convergent.*

Proof. We want to prove that the algorithm satisfies (2.7). From (2.1), we have

$$B_1 - F'(x^*) = \sum_{j=1}^{n} u_j u_j^T \left\{ [B - F'(x^*)] \left(I - \frac{pp_j^T}{p_j^T p_j} \right) + [y - F'(x^*)p] \left(\frac{p_j^T}{p_j^T p_j} \right) \right\}. \tag{2.8}$$

Thus, since $u_j^T Bp = u_j^T Bp_j$ and $u_j^T F'(x^*)p = u_j^T F'(x^*)p_j$,

$$u_j^T[B_1 - F'(x^*)] = u_j^T[B - F'(x^*)]\left(I - \frac{p_jp_j^T}{p_j^Tp_j}\right) + u_j^T[y - F'(x^*)p_j]\left(\frac{p_j^T}{p_j^Tp_j}\right). \tag{2.9}$$

Since $u_j^Ty = u_j^TF'(x + \lambda_jp)p_j$, where $0 < \lambda_j < 1$ (see [6, p. 660]) and F' satisfies Lipschitz condition (2.6),

$$\begin{aligned} |u_j^T[y - F'(x^*)p_j]| &\leq L\|x + \lambda_jp - x^*\|\,\|p_j\| \\ &\leq L\|\lambda_j(x_1 - x^*) + (1 - \lambda_j)(x - x^*)\|\,\|p_j\| \\ &\leq L\sigma\|p_j\|. \end{aligned} \tag{2.10}$$

From (2.9), we obtain

$$\|u_j^T[B_1 - F'(x^*)]\|^2 \leq \|u_j^T[B - F'(x^*)]\|^2 + L^2\sigma^2.$$

Summing over j,

$$\|B_1 - F'(x^*)\|_F^2 \leq \|B - F'(x^*)\|_F^2 + nL^2\sigma^2.$$

Hence,

$$\|B_1 - F'(x^*)\|_F \leq \|B - F'(x^*)\|_F + nL\sigma,$$

as $(\alpha^2 + \beta^2)^{1/2} \leq \alpha + \beta$ for $\alpha, \beta \geq 0$. The result then follows from Theorem 2.2.

To obtain the Q-superlinear convergence of the algorithm, we need the characterization given by Dennis and Moré [5].

THEOREM 2.3. *Suppose F satisfies assumptions* (a) *and* (b), *and for some $x_0 \in D$, the sequence* (1.1) *with $t_k = 1$ is such that $x_k \neq x^*$, $x_k \in D$ and $\{x_k\}$ converges to x^*. Then $\{x_k\}$ converges Q-superlinearly to x^* if and only if*

$$\lim_{k\to\infty} \frac{\|[B_k - F'(x^*)](x_{k+1} - x_k)\|}{\|x_{k+1} - x_k\|} = 0. \tag{2.11}$$

We have the following result.

THEOREM 2.4. *Suppose F satisfies assumptions* (a), (b), (c), *then the algorithm defined by* (1.1), (1.2) *and* (2.1) *with $t_k = 1$ generates a sequence that converges Q-superlinearly.*

Proof. We note that since $\{x_k\}$ is linearly convergent,

$$\sum_{k=1}^{\infty} \sigma_k < \infty. \tag{2.12}$$

We need to prove that (2.11) is satisfied. Let

$$C = \sum_{j=1}^{n} u_ju_j^T[B - F'(x^*)]\left(I - \frac{pp_j^T}{p_j^Tp_j}\right);$$

Setting $E = B - F'(x^*)$ in (2.2) and (2.3), we obtain

$$\|C\|_F^2 \leqslant \|B - F'(x^*)\|_F^2 - \frac{\|[B - F'(x^*)]p\|^2}{\|p\|^2}.$$

Since $(\alpha^2 - \beta^2)^{1/2} \leqslant \alpha - \beta^2/2\alpha$,

$$\|C\|_F \leqslant \eta - \frac{1}{2\eta}\psi^2, \tag{2.13}$$

where $\eta = \|B - F'(x^*)\|_F$ and $\psi = \|[B - F'(x^*)]p\|/\|p\|$.

By using (2.10), (2.13) in (2.8) we obtain

$$\eta_1 \leqslant \eta - \frac{1}{2\eta}\psi^2 + L\sigma.$$

Thus, in general, we have

$$\eta_{k+1} \leqslant \eta_k - \frac{1}{2\eta_k}\psi_k^2 + L\sigma_k\eta^{1/2}.$$

In particular,

$$\eta_{k+1} \leqslant \eta_k + L\sigma_k\eta^{1/2},$$

which implies that $\{\eta_k\}$ is bounded due to (2.12). Let M be its upper bound; then

$$\frac{1}{2M}\psi_k^2 \leqslant -\eta_{k+1} + \eta_k + L\sigma_k\eta^{1/2}.$$

Hence, for any $m \geqslant 0$,

$$\frac{1}{2M}\sum_{k=0}^{m} \psi_k^2 \leqslant \eta^{1/2}L\sum_{k=0}^{m} \sigma_k + \eta_0 - \eta_{m+1}.$$

Since $\Sigma_{k=0}^{\infty} \sigma_k < \infty$, $\Sigma_{k=0}^{\infty} \psi_k^2$ is bounded. Furthermore, as $\Sigma_{k=0}^{m} \psi_k^2$ is monotonic increasing, $\lim_{m\to\infty} \Sigma_{k=0}^{m} \psi_k^2$ exists, we therefore must have $\lim_{k\to\infty} \psi_k = 0$. The result now follows from Theorem 2.3.

Computer Centre
Australian National University
P. O. Box 4
Canberra A. C. T. 2601, Australia

1. C. G. BROYDEN, "A class of methods for solving nonlinear simultaneous equations," *Math. Comp.*, v. 19, 1965, pp. 577–593.

2. C. G. BROYDEN, "The convergence of single-rank quasi-Newton methods," *Math. Comp.*, v. 24, 1970, pp. 365–382.

3. C. G. BROYDEN, "The convergence of an algorithm for solving sparse nonlinear systems," *Math. Comp.*, v. 25, 1971, pp. 285–294.

4. C. G. BROYDEN, J. E. DENNIS & J. J. MORÉ, "On the local and superlinear convergence of quasi-Newton methods," *J. Inst. Math. Appl.*, v. 12, 1973, pp. 223–246.

5. J. E. DENNIS & J. J. MORÉ, "A characterization of superlinear convergence and its application to quasi-Newton methods," *Math. Comp.*, v. 28, 1974, pp. 549–560.

6. L. V. KANTOROVIC & G. P. AKILOV, *Functional Analysis in Normed Spaces*, English transl., Pergamon Press, 1964.

7. L. K. SCHUBERT, "Modification of a quasi-Newton method for nonlinear equations with a sparse Jacobian," *Math. Comp.*, v. 24, 1970, pp. 27–30.