Computers Math. Applic. Vol. 20, No. 1, pp. 57–60, 1990
Printed in Great Britain.

0097-4943/90 $3.00 + 0.00

A HYPO-QUADRATIC CONVERGENCE METHOD FOR SYSTEMS OF NONLINEAR EQUATIONS

T. ALTMAN
Department of Computer Science, University of Kentucky, Lexington, KY 40506, U.S.A.

(*Received* 17 *July* 1989)

Abstract—In this paper we investigate a class of inexact Newton methods which are hypo-quadratically convergent. The methods are designed for solving nonlinear systems of algebraic equations. An interesting feature is the flexibility of these algorithms: there is a trade-off between the rate of convergence and the accuracy required for solving the linearized equations at each iteration.

1. INTRODUCTION

The classical one-dimensional Newton method was extended to Banach spaces by Kantorovitch in 1948 (for references, see Ref. [1]). It was a new truly new development in the area of solving nonlinear operator equations in general and, in particular, provided a powerful tool for solving nonlinear systems of algebraic equations.

One of the advantages of the Newton–Kantorovitch method is its rapid convergence, that is, the convergence is quadratic if the initial guess, $\mathbf{x}_0$, is sufficiently good. However, one of the drawbacks of this method is that is (theoretically) requires an exact solution of the linearized equation at each iteration step. It is well-known that, in practice, solving the system of linear equations exactly is usually not possible. This means that the method which is in practical use, and is supposed to be the Newton–Kantorovitch method, is in fact a different one.

Krasnosel'skii and Rutickii [2] have proved the convergence of the Newton–Kantorovitch method with approximate solutions of the linearized equations for nonlinear operators with Hölder continuous Fréchet derivatives. Since then, a number of mathematicians have investigated the Newton–Kantorovitch method with various procedures for the approximate solutions of the linearized equations. A more detailed description of these *inexact Newton methods* may be found in Ref. [3].

In this paper we investigate a class of inexact Newton methods which are hypo-quadratically convergent. These methods are designed for the solution of nonlinear systems of algebraic equations. In additional feature is the flexibility of the proposed algorithms from the practical point of view. For instance, in large scale problems the linearized equations at each iteration are usually solved by an iterative method, where the residuum can be easily monitored. In the proposed method, the "smallness" of the residuum can be adjusted so as to obtain the desired rate of convergence. This is due to the fact that the rate of convergence depends on a parameter $1 < t < 2$. Thus, there is a trade-off between the rate of convergence and the accuracy required for solving the linearized equations.

2. DESCRIPTION OF THE METHOD

Consider the system of nonlinear equations

$$F(x) = 0 \tag{1}$$

where

$$F(x) = (f_1(x), f_2(x), \ldots, f_m(x))^{\mathrm{T}}, \quad x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}.$$

Denote by $F'(x)$ the Jacobian of F at x:

$$F'(x) = \frac{\mathrm{d}(f_1(x), \ldots, f_m(x))}{\mathrm{d}(x_1, x_2, \ldots, x_n)} = \left(\frac{\partial f_j}{\partial x_k}\right)_{\substack{j=1,\ldots,m,\\ k=1,\ldots,n.}} \tag{2}$$

The following is the iterative method under consideration:

$$x_{i+1} := x_i + h_i, \tag{3}$$

where h_i is a solution of the equation

$$F'(x_i)h_i + F(x_i) = r_i, \tag{4}$$

x_0 being the initial approximate solution. If $r_i = 0$ for $i = 0, 1, \ldots$, then the iterative method (3) is the *exact* Newton method.

Denote by $B(x_0, R)$ the ball with center x_0 and radius R for some given $R > 0$:

$$B(x_0, R) = \{x \mid \|x - x_0\| \leqslant R\}.$$

We assume that F is defined on a domain containing the ball $B(x_0, R)$.

Definition 2.1

Consider the numerical series

$$\sum_{i=0}^{\infty} a_i, \quad a_i > 0. \tag{5}$$

Assume that series (5) is convergent and set

$$R_i = \sum_{k=i}^{\infty} a_k.$$

Then series (5) is:

(1) *linearly convergent* if

$$R_{i+1} \leqslant q R_i, \quad \text{for some } 0 < q < 1,$$

(2) *superlinearly convergent* if

$$R_{i+1} \leqslant c_i R_i,$$

where c_i is such that $c_i \to 0$ as $i \to \infty$,

(3) *hypo-quadratically convergent* if

$$R_{i+1} \leqslant a^{t^i} R_i,$$

for some $0 < a < 1$ and $1 < t < 2$.

Suppose that the iterative method (3) converges to a solution x of the equation (1). Then, we have

$$x_i = x_0 + \sum_{k=0}^{i-1} h_k \quad \text{and} \quad x = x_0 + \sum_{i=0}^{\infty} h_i. \tag{6}$$

Now, assume that the series in equation (6) is *dominated* by the series in condition (5), that is,

$$\|h_i\| \leqslant D\, a_i \ (i = 0, 1, \ldots), \tag{7}$$

for some constant $D > 0$.

Then the series

$$\sum_{i=0}^{\infty} a_i$$

is called the *iterative majorant* for method (3), and we have

$$\|x - x_i\| \leqslant R_i. \tag{8}$$

We may assume that relation (8) holds for almost all i.

Definition 2.2

The iterative method (3) is convergent: linearly, superlinearly or hypo-quadratically, if there exists an iterative majorant series with the same character of convergence, respectively.

We shall now show how to make the inexact Newton method (3) hypo-quadratically convergent. Consider the sequence

$$a_i = a_0^{t^i}, \tag{9}$$

where $1 < t < 2$ and $a_0 = \| F(x_0) \|$ is sufficiently small, and set

$$R = \sum_{i=0}^{\infty} a_0^{t^i}. \tag{10}$$

Assume that the following conditions are satisfied:

(A_1) There exists a constant $C > 0$ such that

$$\| F(x+h) - F(x) - F'(x)h \| \leqslant C \| h \|^2 \tag{11}$$

for all x and $x + h$ belonging to $B(x_0, R)$.

(A_2) For every $x \in B(x_0, R)$ and $g \in \mathbf{R}^m$, if h is a solution of the equation

$$F'(x)h + g = 0, \tag{12}$$

then

$$\| h \| \leqslant K \| g \|, \tag{13}$$

for some independent constant $K > 0$.

Moreover, assume that equations (4) can be solved with arbitrary accuracy, that is, $\| r_i \|$ can be made arbitrarily small. Thus, we assume that

$$\| r_i \| \leqslant a_i^p, \tag{14}$$

for $i = 0, 1, \ldots,$ and $p > 2$.

Lemma 2.1

The following estimate holds

$$\| F(x_i) \| \leqslant a_i, \tag{15}$$

where $\{x_i\}$ is determined by equation (3) and $\{a_i\}$ is given by equation (9), provided that $a_0 = \| F(x_0) \|$ is sufficiently small.

Proof. We prove condition (15) by induction. Suppose that series (5) holds for all $k \leqslant i$. We get

$$\begin{aligned} \| F(x_{i+1}) \| = \| F(x_i + h_i) \| &= \| F(x_i + h_i) - F(x_i) - F'(x_i)h_i + r_i \| \\ &\leqslant \| F(x_i + h_i) - F(x_i) - F'(x_i)h_i \| + \| r_i \| \leqslant C \| h_i \|^2 + a_i^p. \end{aligned}$$

But

$$\| h_i \| \leqslant K(a_i + a_i^p) \leqslant 2K\, a_i.$$

Hence,

$$\| F(x_{i+1}) \| \leqslant (4C\, K^2 + 1) a_i^2 \leqslant a_{i+1} = a_i^t, \tag{16}$$

if $1 < t < 2$ and $a_0 = \| F(x_0) \| < 1$ is so small that

$$C_1 a_0^2 \leqslant a_0^t, \tag{17}$$

since then we have

$$a_i^{2-t} \leqslant a_0^{2-t} \leqslant C_1^{-1}, \tag{18}$$

where $C_1 = 4C\, K^2 + 1$. Thus, if a_0 satisfies condition (17), then estimate (15) holds. □

Lemma 2.2

Relation (7) holds with $\{a_i\}$ determined by equation (9) and $D = 2K$.

Proof. The proof follows immediately condition (13) with $h = h_i$ and $g = F(x_i) + r_i$, and conditions (15), (14). □

We can now prove the following theorem.

Theorem 2.1

In addition to assumptions (A_1), (A_2) and condition (14), suppose that $a_0 = \| F(x_0) \|$ is so small that condition (17) holds, then the sequence of approximate solutions $\{x_i\}$ determined by equation (3) converges to a solution x of equation (1) and $x_i, x \in B(x_0, R)$ with R given by equation (10). The convergence is hypo-quadratic.

Proof. The convergence of $\{x_i\}$ follows immediately from Lemmas 2.1 and 2.2.

It can be easily seen that

$$R_{i+1} < a_0^{t^i} R_i, \tag{19}$$

where the remainder

$$R_i = \sum_{k=i}^{\infty} a_k = \sum_{k=0}^{\infty} a_0^{t^k}$$

by equation (9). Inequality (19) together with condition (7) show the hypo-quadratic convergence of $\{x_i\}$. This completes the proof of the theorem. □

Condition (11) is satisfied if $F'(x)$ is Lipschitz continuous. Now, suppose that $F'(x)$ is Hölder continuous. Then, condition (11) is replaced by the following.

$$\| F(x+h) - F(x) - F'(x)h \| \leqslant C \| h \|^{1+\alpha}, \tag{20}$$

for all x and $x + h$ belonging to $B(x_0, R)$ and some $0 < \alpha < 1$. Then the condition imposed on the parameter t is

$$1 < t < 1 + \alpha. \tag{21}$$

The other assumptions concerning the method under consideration remain unchanged, except for p in condition (14), which should be

$$p > 1 + \alpha. \tag{22}$$

After these changes, all that was said remains valid. It is easy to see that condition (11) may yield a slower convergence of the method in question. However, conditions (14) and (22) indicate that less accuracy is then required for the solutions of the linearized equations (4), that is, larger $\| r_i \|$ is admissible if $p > 1 + \alpha$, than in the case where $p > 2$.

REFERENCES

1. L. V. Kantorovitch and G. P. Akilov, *Functional Analysis in Normed Spaces*. Pergamon Press, New York (1964).
2. M. A. Krasnosel'skii and Y. A. Rutickii, Some approximate methods of solving non-linear opertor equations based on linearization. *Soviet Math. Dok.* **2**(6) (1961).
3. J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall. Englewood Cliffs, N.J. (1983).