APPLICATIONES MATHEMATICAE
26,4 (1999), pp. 457–465

I. K. ARGYROS (Lawton, OK)

LOCAL CONVERGENCE OF INEXACT NEWTON METHODS UNDER AFFINE INVARIANT CONDITIONS AND HYPOTHESES ON THE SECOND FRÉCHET DERIVATIVE

Abstract. We use inexact Newton iterates to approximate a solution of a nonlinear equation in a Banach space. Solving a nonlinear equation using Newton iterates at each stage is very expensive in general. That is why we consider inexact Newton methods, where the Newton equations are solved only approximately, and in some unspecified manner. In earlier works [2], [3], natural assumptions under which the forcing sequences are uniformly less than one were given based on the second Fréchet derivative of the operator involved. This approach showed that the upper error bounds on the distances involved are smaller compared with the corresponding ones using hypotheses on the first Fréchet derivative. However, the conditions on the forcing sequences were not given in affine invariant form. The advantages of using conditions given in affine invariant form were explained in [3], [10]. Here we reproduce all the results obtained in [3] but using affine invariant conditions.

1. Introduction. In this study we are concerned with approximating a solution x^* of the equation

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

where F is a nonlinear operator defined on a Banach space E_1 with values in a Banach space E_2 with the properties: F belongs to the class of operators $P_\lambda(r)$ defined for any $\lambda \in [0,1]$ and $r > 0$ by $P_\lambda(r) = \{F \mid F : D \subseteq E_1 \to E_2$, where D is open and convex; there exists $x^* \in D$ such that $F(x^*) = 0$; $U(x^*, r) \subseteq D$, where $U(x^*, r) = \{x \in E_1 \mid \|x - x^*\| < r\}$;

1991 *Mathematics Subject Classification*: 65B05, 47H17, 49D15.

Key words and phrases: inexact Newton method, Banach space, forcing sequence, Fréchet derivative, superlinear, strong, weak convergence.

F is twice Fréchet-differentiable on $U(x^*, r)$; F'' is continuous on $U(x^*, r)$; $F'(x^*)^{-1} \in L(E_2, E_1)$, the space of bounded linear operators from E_2 into E_1; there exists $a_\lambda > 0$ such that for all $x \in U(x^*, r)$,

$$\|F'(x^*)^{-1}[F''(x) - F''(x^*)]\| \leq a_\lambda \|x - x^*\|^\lambda\}. \tag{2}$$

Here $F''(x) \in L(E_1, L(E_1, E_2))$ $(x \in D)$ denotes the second Fréchet derivative of F evaluated at $x \in D$ [3], [8].

An *inexact Newton method* is any procedure which, given an initial guess x_0, generates a sequence $\{x_n\}$ $(n \geq 0)$ of approximations to x^* as follows:

FOR $n = 0$ STEP 1 UNTIL Convergence DO.

Find some step s_n which satisfies

$$F'(x_n)s_n = -F(x_n) + r_n \quad (n \geq 0), \tag{3}$$

where

$$\frac{\|F'(x_n)^{-1} r_n\|}{\|F'(x_n)^{-1} F(x_n)\|} \leq c_n \leq c \quad (n \geq 0). \tag{4}$$

Set

$$x_{n+1} = x_n + s_n \quad (n \geq 0). \tag{5}$$

The numbers c_n depend on x_n $(n \geq 0)$. In particular for $c_n = 0$ $(n \geq 0)$ we obtain Newton's method [1]–[3], [7]—[9].

In [5], [6] the local behavior of such inexact Newton methods is analysed in the special case when $E_1 = E_2 = \mathbb{R}^i$ $(i \in \mathbb{N})$. However, instead of condition (2) above they use

$$\|F'(x^*)^{-1}[F'(x) - F'(y)]\| \leq a'_\lambda \|x - y\|^\lambda, \tag{6}$$

which is in some sense stronger than (2). The condition

$$\frac{c_n}{\|F(x_n)\|} \leq \eta_n \leq \eta \quad (n \geq 0) \tag{7}$$

was used in [5], [6], but assumption (4) was employed in [10]. The advantages of using conditions in affine invariant form over the ones that do not have been explained in some detail in [3], [4], [10]. Using (2) and (7) we showed that all results on convergence developed in [5], [6] also hold in our setting [3]. Moreover, we showed that our upper error bounds on the distances involved are smaller. Here we further improve upon these results by using (4) instead of (7). We conclude that all results obtained in [3] also hold in the new setting.

2. Convergence analysis. If $F \in P_\lambda(r)$, then we define

$$m_\lambda(x^*) \equiv \sup \left\{ \frac{\|F'(x^*)^{-1}[F''(x) - F''(x^*)]\|}{\|x - x^*\|^\lambda} \;\middle|\; x \neq x^*,\ x \in U(x^*, r) \right\} \tag{8}$$

and

$$b(x^*) \equiv \|F'(x^*)^{-1}F''(x^*)\|. \tag{9}$$

We need the lemmas:

LEMMA 1. *Let $F \in P_\lambda(r)$. Then there exists $\overline{r}_1 \le r$ such that $F \in P_\lambda(\overline{r}_1)$, $F'(x)$ is nonsingular for all $x \in U(x^*, \overline{r}_1)$, and for all $x, y \in U(x^*, \overline{r}_1)$,*

$$\|F'(y)^{-1}[F''(x) - F''(x^*)]\| \le \frac{m_\lambda(x^*)}{1 - b(x^*)\|y - x^*\| - \frac{m_\lambda(x^*)}{\lambda+1}\|y - x^*\|^{\lambda+1}}\|x - x^*\|^\lambda, \tag{10}$$

$$m_\lambda(x) \le \frac{m_\lambda(x^*)}{1 - b(x^*)\|x - x^*\| - \frac{m_\lambda(x^*)}{\lambda+1}\|x - x^*\|^{\lambda+1}}, \tag{11}$$

$$b(x) \le \frac{b(x^*)}{1 - b(x^*)\|x - x^*\| - \frac{m_\lambda(x^*)}{\lambda+1}\|x - x^*\|^{\lambda+1}}, \tag{12}$$

where

$$m_\lambda(x) \equiv \sup\left\{\frac{\|F'(x)^{-1}[F''(x) - F''(x^*)]\|}{\|x - x^*\|^\lambda}\ \middle|\ x \ne x^*,\ x \in U(x^*, \overline{r}_1)\right\} \tag{13}$$

and

$$b(x) = \|F'(x)^{-1}F''(x^*)\|. \tag{14}$$

Proof. Define the function

$$h(t) = \frac{m_\lambda(x^*)}{\lambda + 1}t^{\lambda+1} + b(x^*)t - 1 \tag{15}$$

for each fixed $\lambda \in [0, 1]$. Since h is continuous, $h(0) = -1$ and $h(t) > 0$ for sufficiently large t, by the intermediate value theorem there exists a minimum positive number $\overline{r}_0$ such that $h(\overline{r}_0) = 0$. Choose $\overline{r}_1 = \min\{r, \overline{r}_0\}$. Then

$$h(t) < 0 \quad \text{for all } t \in [0, \overline{r}_1). \tag{16}$$

Using (8), (9), (15), (16) and the identity

$$\begin{aligned} F'(x^*)^{-1}[F'(x^*) - F'(x)] = & -F'(x^*)^{-1}[F'(x) - F'(x^*) \\ & - F''(x^*)(x - x^*) + F''(x^*)(x - x^*)] \\ = & -\int_0^1 F'(x^*)^{-1}\{F''[x^* + t(x - x^*)] \\ & - F''(x^*)\}(x - x^*)\,dt \\ & - F'(x^*)^{-1}F''(x^*)(x - x^*), \end{aligned}$$

we get

$$\begin{aligned}\|F'(x^*)^{-1}[F'(x^*)-F'(x)]\| &\\ &\le m_\lambda(x^*)\int_0^1 \|t(x-x^*)\|^\lambda\|x-x^*\|\,dt + b(x^*)\|x-x^*\|\\ &\le \frac{m_\lambda(x^*)}{\lambda+1}\|x-x^*\|^{\lambda+1} + b(x^*)\|x-x^*\|\\ &< \frac{m_\lambda(x^*)}{\lambda+1}\overline{r}_1^{\lambda+1} + b(x^*)\overline{r}_1 \le 1,\end{aligned}$$

and

$$\|F'(x)^{-1}F'(x^*)\| \le \left[1 - b(x^*)\|x-x^*\| - \frac{m_\lambda(x^*)}{\lambda+1}\|x-x^*\|^{\lambda+1}\right]^{-1}. \tag{17}$$

It follows by the Banach Lemma on invertible operators [4], [8] that $F'(y)^{-1}$ exists for all $y \in U(x^*, \overline{r}_1)$ so that (10) holds. By (10), (13) and the estimate

$$\begin{aligned}\|F'(x)^{-1}[F''(z)-F''(x^*)]\| &= \|[F'(x)^{-1}F'(x^*)][F'(x^*)^{-1}(F''(z)-F''(x^*))]\|\\ &\le \|F'(x)^{-1}F'(x^*)\|\cdot\|F'(x^*)^{-1}(F''(z)-F''(x^*))\|,\end{aligned} \tag{18}$$

for all $x, z \in U(x^*, \overline{r}_1)$, we obtain (11). Moreover, by (9), (14) and the estimates

$$\begin{aligned}\|F'(x)^{-1}F''(x^*)\| &= \|[F'(x)^{-1}F'(x^*)][F'(x^*)^{-1}F''(x^*)]\|\\ &\le \|F'(x)^{-1}F'(x^*)\|\,\|F'(x^*)^{-1}F''(x^*)\|\end{aligned} \tag{19}$$

for all $x \in U(x^*, \overline{r}_1)$, we obtain (12). ∎

Lemma 2. *Let $F \in P_\lambda(\overline{r}_1)$. Then, for any $x \in U(x^*, \overline{r}_1)$,*

$$\|G(x)-x^*\| \le \frac{1}{\lambda+2}m_\lambda(x)\|x-x^*\|^{\lambda+2} + \frac{1}{2}b(x)\|x-x^*\|^2 \tag{20}$$

and

$$\|G(x)-x^*\| \le q(x)\|x-x^*\|, \tag{21}$$

where

$$q(x) = \frac{\frac{1}{\lambda+2}m_\lambda(x^*)\|x-x^*\|^{\lambda+1} + \frac{1}{2}b(x^*)\|x-x^*\|}{1 - b(x^*)\|x-x^*\| - \frac{m_\lambda(x^*)}{\lambda+1}\|x-x^*\|^{\lambda+1}} \tag{22}$$

and

$$G(x) = x - F'(x)^{-1}F(x) \quad (x \in D). \tag{23}$$

P r o o f. By (13), (14) and (23) we can write

$$G(x)-x^* = F'(x)^{-1}[F(x^*)-F(x)-F'(x)(x^*-x)]$$
$$= F'(x)^{-1}\int_0^1 [F''(x^*+t(x-x^*))-F''(x^*)]t\,dt\,(x-x^*)^2$$
$$+\frac{1}{2}F'(x)^{-1}F''(x^*)(x-x^*)^2.$$

By taking norms above we get

$$\|G(x)-x^*\| \le \frac{1}{\lambda+2}m_\lambda(x)\|x-x^*\|^{\lambda+2}+\frac{1}{2}b(x)\|x-x^*\|^2,$$

which is (20). Estimate (21) follows from (11), (12) and (20). ■

We can prove the following main local convergence theorem for the inexact Newton method $\{x_n\}$ $(n\ge 0)$ generated by (5).

THEOREM 1. *Assume condition* (4) *holds for* $F\in P_\lambda(\bar r_1)$. *Then the inexact Newton method* $\{x_n\}$ $(n\ge 0)$ *generated by* (5) *with* $x_n\in U(x^*,\bar r_1)$ *satisfies*

(24) $$\|x_{n+1}-x^*\|\le d_n\|x_n-x^*\|,$$

where

(25) $$d_n \equiv c_n+(1+c_n)q(x)$$

$(n\ge 0)$, *where* q *is defined in* (22). *Moreover*, *if* $c_n\le c<1$ $(n\ge 0)$, *define the function* g *by*

(26) $$g(t)=\alpha_1 t^{\lambda+1}+\alpha_2 t+\alpha_3,$$

where

(27) $$\alpha_1=\frac{m_\lambda(x^*)[2\lambda+3-c]}{(\lambda+1)(\lambda+2)},\qquad \alpha_2=\frac{b(x^*)}{2}(3-c),\qquad \alpha_3=c-1.$$

Then

(a) *there exists a minimum positive number* $\bar r_2$ *such that* $g(\bar r_2)=0$ *and*

(28) $$g(t)<0,\ h(t)<0 \quad \textit{for all } t\in[0,r^*),\ r^*=\min\{\bar r_1,\bar r_2\},$$

where the function h *is given in* (15);

(b) *for* $x_0\in U(x^*,r^*)$,

(29) $$d_n\le d = c+\frac{(1+c)\big[\frac{1}{s+2}m_\lambda(x^*)\|x_0-x^*\|^{\lambda+1}+\frac{1}{2}b(x^*)\|x_0-x^*\|\big]}{1-b(x^*)\|x_0-x^*\|-\frac{m_\lambda(x^*)}{\lambda+1}\|x_0-x^*\|^{\lambda+1}}\in(0,1)$$

$(n\ge 0)$, *and*

(30) $$\lim_{n\to\infty}x_n=x^*.$$

P r o o f. We use induction on $n \ge 0$ to show that estimate (24) holds and the nth step of the inexact Newton method is defined so that there exist s_n satisfying (3)–(5) for all $n \ge 0$. Assume $\|x_n - x^*\| \le \|x_0 - x^*\|$ for some $n \ge 0$. It follows that $x_n \in U(x^*, \overline{r}_1)$, so $F'(x_n)^{-1}$ exists and $m_\lambda(x_n)$ is defined. Hence, the nth step of the inexact Newton method is defined so that there exists s_n satisfying (3)–(5). Since $s_n = F'(x_n)^{-1}(-F(x_n) + r_n)$, we get

$$x_{n+1} - x^* = F'(x_n)^{-1}[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n) + r_n]. \tag{31}$$

By (4) we also have

$$\|F'(x_n)^{-1} r_n\| \le c_n \|F'(x_n)^{-1} F(x_n)\|$$

and

$$\begin{aligned} \|F'(x_n)^{-1} F(x_n)\| \le \|F'(x_n)^{-1}[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)]\| \\ + \|x_n - x^*\|, \end{aligned} \tag{32}$$

and by (21) and (25),

$$\begin{aligned} \|x_{n+1} - x^*\| &\le c_n \|x_n - x^*\| \\ &\quad + (1 + c_n)\|F'(x_n)^{-1}[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)]\| \\ &\le d_n \|x_n - x^*\|, \end{aligned}$$

which shows (24) for all $n \ge 0$.

As with the function h in Lemma 1, we can find a minimum positive number $\overline{r}_2$ such that $g(\overline{r}_2) = 0$. This shows (28). Since $c_n \le c < 1$, it can easily be seen that $d_n \in [0, 1)$ $(n \ge 0)$ if

$$g(\|x_0 - x^*\|) < 0 \quad \text{and} \quad h(\|x_0 - x^*\|) < 0, \tag{33}$$

which is true by (28) and the choice of r^*.

The induction is now complete.

Moreover, by (24) and (29) we get

$$\|x_{n+1} - x^*\| \le d^{n+1} \|x_0 - x^*\| \le d^{n+1} r^* \to 0 \quad \text{as } n \to \infty,$$

which shows (30). ∎

Defining rates of convergence in the same way as in [3], [6], [10] we can extend the results obtained in [3].

THEOREM 2. *Let $F \in P_\lambda(\overline{r}_1)$. Assume the inexact Newton method $\{x_n\}$ $(n \ge 0)$ generated by* (5) *converges to x^*. Then*

(a) $\{x_n\}$ $(n \ge 0)$ *converges superlinearly if and only if*

$$\limsup_{n \to \infty} \frac{\|F'(x_n)^{-1} r_n\|}{\|F'(x_n)^{-1} F(x_n)\|} = 0,$$

or if and only if $\limsup_{n \to \infty} c_n = 0$;

(b) $\{x_n\}$ $(n \geq 0)$ *converges with order* $1+\lambda$ *if and only if*

$$\limsup_{n\to\infty} \frac{\|F'(x_n)^{-1} r_n\|}{\|F'(x_n)^{-1} F(x_n)\|^{1+\lambda}} < \infty,$$

or if and only if

$$\limsup_{n\to\infty} \frac{c_n}{\|F'(x_n)^{-1} F(x_n)\|^{\lambda}} < \infty;$$

(c) $\{x_n\}$ $(n \geq 0)$ *converges with weak order at least* $1+\lambda$ *if and only if*

$$\lim_{n\to\infty} \|F'(x_n)^{-1} r_n\|^{(1+\lambda)^{-n}} < 1,$$

or if

$$\limsup_{n\to\infty} c_n^{(1+\lambda)^{-1}} < 1.$$

Proof. The results follow directly using the techniques of [3]. Alternatively they can also be regarded as corollaries of the results in [3], by noting that if β is a bound on the condition number (see (4)) of $F'(x)$ in a neighborhood of x^*, then by Lemma 1 in [3] or Lemma 3.1 in [6],

$$\frac{1}{\beta} \cdot \frac{\|r_n\|}{\|F(x_n)\|} \leq \frac{\|F'(x_n)^{-1} r_n\|}{\|F'(x_n)^{-1} F(x_n)\|} \leq \beta \frac{\|r_n\|}{\|F(x_n)\|} \qquad (n \geq 0). \blacksquare$$

3. Applications

Remark 1. As noted in [3]–[6], [10] the results obtained here can be used for projection methods such as Arnoldi's method, the generalized minimum residual method (GMRES), the generalized conjugate residual method (GCR), and for combined inexact-Newton/finite-difference projection methods.

Remark 2. The results obtained here can also be used to solve equations of the form $F(x) = 0$, where F' satisfies the autonomous differential equation

$$F'(x) = T(F(x)), \tag{34}$$

with $T : E_2 \to E_1$ being a known continuously Fréchet-differentiable operator at x^*. Since $F'(x^*) = T(F(x^*)) = T(0)$, $F''(x^*) = F'(x^*)T'(F(x^*)) = T(0)T'(0)$, we can apply the results obtained here without actually knowing the solution x^* of equation (1).

Below, we provide such an example.

Example. Let $E_1 = E_2 = \mathbb{R}$, $D = U(0,1)$, and define the function F on D by

$$F(x) = e^x - 1. \tag{35}$$

Then it can be easily seen that we can take $T(x) = x + 1$ in (34). That is, F' satisfies the autonomous differential equation (34).

For Newton's method, set $c = 0$, and take $\lambda = 1$. Then using (8), (9), (15) and (26), we can easily obtain the following:

$$m_1(x^*) = e, \quad b(x^*) = 1, \quad \bar{r}_2 = .411254048, \quad \bar{r}_1 = .5654448, \quad r^* = \bar{r}_2.$$

Hence, the conclusions of Theorem 1 hold if

$$\|x_0 - x^*\| < r^* = .411254048. \tag{36}$$

To compare our results with the corresponding ones obtained in [5], [6], [10] we first define as in [10]

$$\mu_\lambda(x^*) \equiv \sup\left\{ \frac{\|F'(x^*)^{-1}[F'(y) - F'(z)]\|}{\|y - z\|^\lambda} \,\middle|\, y \neq z,\ y, z \in U(x^*, r) \right\}. \tag{37}$$

Then, by Theorem 3.1 in [10, p. 585] we must have

$$\|x_0 - x^*\| \leq \tfrac{2}{3}\mu_\lambda(x^*)^{-1} \equiv r_1^*. \tag{38}$$

As above, using (35), (37) and (38) we get $\mu_1(x^*) = e$, and $r_1^* = .245253 < r^*$. Hence, our Theorem 1 provides a wider choice for x_0 than the corresponding Theorem 3.1 in [10, p. 585]. This observation is important and finds applications in steplength selection in predictor-corrector continuation procedures [4], [5], [6], [10].

References

[1] I. K. Argyros, *On the convergence of some projection methods with perturbation*, J. Comput. Appl. Math. 36 (1991), 255–258.

[2] —, *Comparing the radii of some balls appearing in connection to three local convergence theorems for Newton's method*, Southwest J. Pure Appl. Math. 1 (1998), 32–43.

[3] —, *Relations between forcing sequences and inexact Newton iterates in Banach space*, Computing 62 (1999), 71–82.

[4] I. K. Argyros and F. Szidarovszky, *The Theory and Application of Iteration Methods*, CRC Press, Boca Raton, FL, 1993.

[5] P. N. Brown, *A local convergence theory for combined inexact-Newton/finite-difference projection methods*, SIAM J. Numer. Anal. 24 (1987), 407–434.

[6] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. 19, (1982), 400–408.

[7] J. M. Gutierrez, *A new semilocal convergence theorem for Newton's method*, J. Comput. Appl. Math. 79 (1997), 131–145.

[8] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.

[9] F. A. Potra, *On Q-order and R-order of convergence*, SIAM J. Optim. Theory Appl. 63 (1989), 415–431.

[10] T. J. Ypma, *Local convergence of inexact Newton methods*, SIAM J. Numer. Anal. 21 (1984), 583–590.

Ioannis K. Argyros
Department of Mathematics
Cameron University
Lawton, OK 73505, U.S.A.
E-mail: ioannisa@cameron.edu

Received on 25.2.1999;
revised version on 30.8.1999 and 8.11.1999