Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Vol. No. 01, 1997, pp. 30–36.
Submitted: March 4, 1997. Published: July 31, 1997.

A LOCAL CONVERGENCE THEOREM FOR THE INEXACT NEWTON METHOD AT SINGULAR POINTS

Ioannis K. Argyros

Abstract. If Newton's method is used to find a solution of an equation on a Banach space, and the Fréchet-derivative is singular at the solution, the convergence of Newton iterates to the solution is linear rather than quadratic. In this study we employ the inexact Newton method to restore quadratic convergence, especially in cases when Newton's or the corrected Newton's method cannot achieve this or cannot even be used.

1991 A.M.S. (MOS) Subject Classification Codes. 65J15, 47H17, 49D15.

Key Words and Phrases. Banach space, Newton method, singular points, quadratic convergence.

I. INTRODUCTION

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

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

where F is an operator defined on a closed convex subset D of a Banach space E_1 with values on a Banach space E_2.

Newton's method

$$y_{n+1} = y_n - F'(y_n)^{-1}F(y_n) \quad (n \geq 0) \tag{2}$$

is the most popular way of generating a sequence $\{y_n\}$ $(n \geq 0)$ converging to x^*. According to the well-known Newton-Kantorovich theorem [3], [5], [7] method (2) converges quadratically to x^* provided $\|x_0 - x^*\|$ is small and the Fréchet derivative $F'(x^*)$ is an invertible operator. The solution x^* is called isolated in this case. A solution x^* at which $F'(x^*)$ is not invertible is called singular. Method (2) converges to x^*, only linearly in this case [3], [5], [7].

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

In [6] and the references there the convergence behavior of method (2) has been studied. Roughly speaking the Banach space E_1 is written as a direct, $N \oplus E$, sum of the null space N of $F'(x^*)$ and an appropriate complementing space E. Then, when the iterates converge, the components of the iterate error in N converges linearly, while the E-component of the error converges with asymptotic order 2. This technique is considered rather complicated, and most importantly it is assumed that the iterates can be computed exactly. However this is not the case most of the time. That is why we propose the study of the inexact Newton method of the form

$$x_{n+1} = x_n - P'(x_n)^{-1}F(x_n) - z_n, \tag{3}$$

where

$$F(x) = P(x) + Q(x) \quad (x \in D), \tag{4}$$

$F(x^*) = 0$, $P'(x^*)$ is invertible and $P, Q\colon D \to E_2$ are Fréchet-differentiable operators on D.

The importance of studying the inexact Newton method (3) comes from the fact that many commonly used variants of Newton's method can be considered procedures of this type. Indeed approximation (3) characterizes any iterative process in which corrections are taken as approximate solutions of Newton equations. If for example an equation on the real line is solved, $F(x_n) \geq 0$ $(n \geq 0)$, and $P'(x_n)^{-1}$ $(n \geq 0)$ overestimates the derivative, y_n is always larger than the corresponding Newton iterate. In such cases, a positive z_n $(n \geq 0)$ correction term is appropriate. Points $z_n \in D$ $(n \geq 0)$ are determined in such a way that iteration $\{x_n\}$ $(n \geq 0)$ converges to a solution x^* of equation (1). Sufficient conditions for the convergence of method (3) to x^* have been given in [1]–[4], [7] under various assumptions. Here as in [6] and the references there we will assume that x^* exists but we will not construct N and E which are very expensive to produce in general. A local convergence theorem is given an special choices of the points z_n $(n \geq 0)$ that restore quadratic convergence in some cases will be given.

Finally a numerical example is also provided to show that quadratic convergence can be restored in the case when Newton's method or the corrected Newton's method cannot even be used.

II. CONVERGENCE ANALYSIS

Let F, Q be Fréchet-differentiable on $D \subseteq E_1$, $x^* \in D$ with $F(x^*) = 0$, $P'(x^*)$ invertible on D, a, b, c, c_n $(n \geq 0)$ nondecreasing functions on the nonnegative axis, $R \geq 0$ such that:

$$\|(F'(x^*) - Q'(x^*))^{-1}[F'(x + t(x^* - x)) - F'(x)]\| \leq a(\|x^* - x\|)\|x^* - x\|t, \tag{5}$$

$$\|(F'(x^*) - Q'(x^*))^{-1}(Q'(x^*) - Q'(x))\| \leq b(\|x - x^*\|), \tag{6}$$

and

$$\|(F'(x^*) - Q'(x^*))^{-1}B_n\| \leq c_n(\|x^* - x_n\|)\|x_n - x^*\| \leq c(\|x^* - x_n\|)\|x_n - x^*\|, \tag{7}$$

where

(8) $B_n = (Q'(x_n) - F'(x_n))z_n - Q'(x_n)(x_n - x^*) \quad (n \geq 0)$ with x_n given by (3),

and for all $x, x_n \in U(x^*, r) = \{x \in E_1 \mid \|x - x^*\| \leq r\} \subseteq U(x^*, R) \subseteq D$, $t \in [0,1]$.

It is convenient to define on the nonnegative axis the functions f_1 and f_2 by

$$f_1(r) = 1 - b(r) - a(r)r, \tag{9}$$

and

$$f_2(r) = 3a(r)r + 2c(r) + 2b(r) - 2. \tag{10}$$

We can now state and prove the following local convergence result concerning the inexact Newton method (3):

Theorem. *Assume F, Q, P be Fréchet-differentiable operators defined on a closed convex subset D of a Banach space E_1 with values in a Banach space E_2, $x^* \in D$ with $F(x^*) = 0$, $P'(x^*)$ invertible on D, a, b, c, c_n $(n \geq 0)$ nondecreasing functions on the nonnegative axis and $R \geq 0$ such that:*

(a) *conditions* (5)–(7) *are satisfied;*

(b) *there exists an interval $[0, d]$, $d < R$ such that*

$$f_1(r) > 0, \tag{11}$$

and

$$f_2(r) < 0, \tag{12}$$

for all $r \in [0, d]$.

(c) *$U(x_0, d) \subseteq D$.*

Then the inexact Newton method $\{x_n\}$ $(n \geq 0)$ generated by (3) *is well defined remains in $U(x^*, d)$ for all $n \geq 0$ and converges to x^* provided that $x_0 \in U(x^*, d)$. Moreover the following error bounds hold for all $n \geq 0$*

$$\|x_{n+1} - x^*\| \leq \frac{1}{1 - b(\|x_n - x^*\|) - a(\|x_n - x^*\|)\|x_n - x^*\|}\bar{\varepsilon}_n \tag{13}$$

$$\leq \frac{1}{1 - b(d) - a(d)d}\varepsilon_n, \tag{14}$$

where

$$\bar{\varepsilon}_n = \left[\frac{1}{2}a(\|x_n - x^*\|) + c_n(\|x_n - x^*\|)\right]\|x_n - x^*\| \tag{15}$$

and

$$\varepsilon_n = \left[\frac{1}{2}a(d) + c(d)\right]\|x_n - x^*\|. \tag{16}$$

Proof: By hypothesis the point x_0 belongs in $U(x^*, d)$. Assume $x_k \in U(x^*, d)$ for $k = 0, 1, 2, \ldots, n$. Then using (6), (7), the Banach Lemma on invertible operators [7], the approximation

$$F'(x_k) - Q'(x_k) = F'(x_k) - Q'(x_k) - (F'(x^*) - Q'(x^*)) + (F'(x^*) - Q'(x^*))$$

and the estimate

$$\begin{aligned}&\|(F'(x^*) - Q'(x^*))^{-1}[(F'(x_k) - F'(x^*)) - (Q'(x_k) - Q'(x^*))]\| \\ &\qquad \leq a(\|x_k - x^*\|)\|x_k - x^*\| + b(\|x_k - x^*\|) \leq a(d)d + b(d) < 1 \quad \text{(by (12))},\end{aligned}$$

it follows that $P'(x_k)^{-1}$ exists and

$$\|P'(x_k)^{-1}(F'(x^*) - Q'(x^*))\| \leq [1 - b(\|x_k - x^*\|) - a(\|x_k - x^*\|)\|x^* - x_k\|]^{-1} \tag{17}$$

$$\leq [1 - b(d) - a(d)d]^{-1}. \tag{18}$$

Using (3) and (8) we obtain the approximation

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - P'(x_k)^{-1}F(x_k) - z_k \\ &= [F'(x_k) - Q'(x_k)]^{-1}[F'(x^*) - Q'(x^*)][F'(x^*) - Q'(x^*)]^{-1} \cdot \\ &\quad \cdot \left\{ \int_0^1 [F'(x_k + t(x^* - x_k)) - F'(x_k)](x^* - x_k)dt + B_k \right\}. \end{aligned} \tag{19}$$

Using (5)–(7), (17) and (18) we obtain (13) and (14). Moreover by (12) and (14)

$$\|x_{k+1} - x^*\| \leq \varepsilon\|x_k - x^*\| \leq \cdots \leq \varepsilon^{k+1}\|x_0 - x^*\| \leq \varepsilon^{k+1}d \tag{20}$$

for

$$\varepsilon = \frac{a(d) + 2c(d)}{2[1 - b(d) - a(d)d]} \in [0, 1). \tag{21}$$

It follows from (20) and (21) that

$$x_k \in U(x^*, d) \quad \text{for all } k \geq 0 \text{ and } \lim_{k\to\infty} x_k = x^*.$$

That completes the proof of the Theorem.

Remark 1. We note that in the proof of the Theorem we have not used (or assumed) that $F'(x^*) = 0$.

Remark 2. Set $z_n = 0$ $(n \geq 0)$, $E_1 = E_2 = \mathbf{R}$ and assume x^* is a zero of multiplicity m for F. Choose $Q(x) = (1 - m)F(x)$ to obtain using (4) that $P(x) = mF(x)$. In this case (3) reduces to the corrected Newton's method which converges quadratically to x^* [5], [7].

Remark 3. Choose $z_n = P'(x_n)^{-1}Q'(x_n)(x^* - x_n)$ $(n \geq 0)$ then by (7) we can choose $c_n = c = 0$ $(n \geq 0)$. Moreover if $a(t) = qt$, $t \in [0, d]$ for some $q \geq 0$ we obtain by (13) that the order of convergence of method (3) is quadratic.

Remark 4. Let $z_n = L_n(x_n - x_{n-1})$ $(n \geq 1)$, for some bounded linear operators L_n, $z_0 = 0$, $x_0, x_1 \in U(x^*, d)$, $\|L_n\| \leq \|x_n - x^*\|$ $(n \geq 0)$ and c_n given by

$$\begin{aligned} c_n(\|x_n - x^*\|) = {} & s_1(\|x_n - x^*\|)\|x_n - x^*\|^{p_1} + s_2(\|x_n - x^*\|)\|x_n - x_{n-1}\|^{p_2} \\ & + s_3(\|x_n - x^*\|)\|x_n - x^*\|^{p_3}\|x_n - x_{n-1}\|^{p_4} + s_4(\|x_n - x^*\|), \end{aligned}$$

for all $n \geq 1$ where s_1, s_2, s_3, s_4 are nondecreasing functions on $[0, d]$ and p_1, p_2, p_3 and p_4 are nonnegative numbers. Moreover assume there exists h such that

$$\|(F'(x^*) - Q'(x^*))^{-1}Q'(z)\| \leq h \quad \text{for } z = x^* \text{ or } x_0.$$

Furthermore assume $F'(x^*) = 0$. Then from the approximation $B_n = (Q'(x_n) - F'(x_n))L_n(x_n - x_{n-1}) - Q'(x_n)(x_n - x^*)$, we obtain

$$\begin{aligned} \|(F'(x^*) - Q'(x^*))^{-1}B_n\| \leq {} & a(\|x_n - x^*\|)\|L_n\|\,\|x_n - x^*\| \cdot \|x_n - x_{n-1}\| \\ & + (b(\|x_n - x^*\|) + h)\|L_n\|\,\|x_n - x_{n-1}\| \\ & + (b(\|x_n - x^*\| + h)\|x_n - x^*\|. \end{aligned}$$

Hence we can take $s_1(r) = 0$, $s_2(r) = b(r) + h$, $c_3(r) = a(r)$, $c_4(r) = b(r) + h$, $r \in [0, d]$, $p_2 = p_3 = p_4 = 1$ any $p_1 > 0$, and

$$c(r) = \max\{\bar{c}(r), h\}, \quad r \in [0, d]$$

where

$$\bar{c}(r) = s_1(r)r^{p_1} + 2^{p_2}s_2(r)r^{p_2} + 2^{p_4}s_3(r)r^{p_3+p_4} + s_4(r), \quad r \in [0, d].$$

Remark 5. Our Theorem can be used to solve autonomous differential equations satisfying

$$P'(x) = T(F(x))$$

for some continuous operator T. Since $P'(x^*) = T(F(x^*)) = T(0)$ the inverse of $P'(x^*)$ can be computed without knowing the solution x^* [3], [7].

Remark 6. Our results can be extended to include a wider class of inexact Newton methods of the form

$$x_{n+1} = x_n - P_1(x_n)^{-1}F(x_n) - z_n \quad (n \geq 0) \tag{22}$$

related to

$$v_{n+1} = v_n - A(v_n)^{-1}F(v_n) \quad (n \geq 0),$$

where

$$A(x) = P_1(x) + Q_1(x) \quad (x \in D)$$

and $A(x)$, $P_1(x)$, $Q_1(x) \in L(E_1, E_2)$ $(x \in D)$, with $A(x)$ being an approximation to $F'(x)$ $(x \in D)$. The conclusions of the theorem will hold for method (22) assuming the following changes: In (5) replace $F'(x^*)$, $Q'(x^*)$ and $F'(x)$ by $A(x^*)$, $Q_1(x^*)$ and $A(x)$ respectively; in (6) replace $F'(x^*)$, $Q'(x^*)$ and $Q'(x)$ by $A(x^*)$, $Q_1(x^*)$ and $Q_1(x)$ respectively; in (7) replace $F'(x^*)$, $Q'(x^*)$ by $A(x^*)$ and $Q_1(x^*)$ respectively and in (8) replace $Q'(x_n)$, $F'(x_n)$ by $Q_1(x_n)$ and $A(x_n)$ $(n \geq 0)$ respectively. Then exactly as in the proof of the theorem using the approximation

$$x_{n+1} - x^* = P_1(x_n)^{-1}[F(x^*) - F(x_n) - A(x_n)(x^* - x_n) + B_n]$$

we can show all the conclusions of the theorem. Comments similar to the ones made in Remarks 1–5 can also be made for method (22).

We will finish this study with an example related with Remark 3.

Example. Let $E_1 = E_2 = \mathbf{R}$, $x^* = 0$, $R = d \in [0, \frac{1}{3})$, $U(x^*, d) = D$, consider the operator F on D given by

$$F(x) = x^2. \tag{23}$$

Choose

$$z_n = (1 - 2x_n)^{-1}(x_n - x^*) \quad (n \geq 0) \quad \text{for } |x_0| \in U(0, d). \tag{24}$$

and

$$Q(x) = x \quad (x \in D). \tag{25}$$

Using (5)–(10), (23), (24) and (25) we can choose $a(r) = 2$, $b(r) = c(r) = 0$ for all $r \in [0, d]$. Moreover we have $B_n = 0$ $(n \geq 0)$, $f_1(r) = 1 - 2r > 0$ and $f_2(r) = 6r - 2 < 0$ for all $r \in [0, d]$. Furthermore $q = 0$ in Remark 3. Hence all hypotheses of the Theorem are satisfied and (13) becomes

$$x_{n+1} - x^* = (2x_n - 1)^{-1}(x_n - x^*)^2 \quad (n \geq 0),$$

which shows quadratic convergence of sequence $\{x_n\}$ $(n \geq 0)$ to x^*. We also note that $x^* = 0$ is a zero of multiplicity 2 for F given by (23).

We observe that as it can easily be seen Newton's method cannot be used since $F'(x_n)^{-1} \neq 0$ $(n \geq 0)$ and $y_{n+1} = y_n - F'(y_n)^{-1}F(y_n)$ cannot be computed in any neighborhood of zero, for any starting point y_0. The corrected Newton's method $w_{n+1} = w_n - 2F'(w_n)^{-1}F(w_n)$ $(n \geq 0)$ converges quadratically to x^* [3], [5], [7], provided that $z_n = 0$ $(n \geq 0)$, which is not true in general.

REFERENCES

[1] Argyros, I.K. A convergence theorem for Newton-like methods under generalized Chen-Yamamoto type assumptions, Appl. Math. Comp. 61, 1, (1994), 25–37.

[2] Argyros, I.K. A unified approach for constructing fast two-step Newton-like methods, Monatshefte fur Mathematik, 119, (1995), 1–22.

[3] Argyros, I.K. and Szidarovszky, F. The Theory and Applications of Iteration Methods, C.R.C. Press, Inc. Boca Raton, Florida, U.S.A., 1993.

[4] Argyros, I. K. and Szidarovszky, F. On the convergence of modified contractions. J. Comput. Appl. Math. 55, 2, (1994), 183–189.

[5] Atkinson, K. An Introduction to Numerical Analysis, Wiley Publ., New York, U.S.A., 1988.

[6] Decker, D.W., Keller, H.B. and Kelley, C.T. Convergence rates for Newton's method at singular points, SIAM J. Numer. Anal. 20, 2, (1983), 296–314.

[7] Kantorovich, L.V. and Akilov, G.P. Functional Analysis, Pergamon Press, New York, U.S.A., 1982.