TAMKANG JOURNAL OF MATHEMATICS
Volume 25, Number 4, Winter 1994

ON THE CONVERGENCE OF SOME PROJECTION METHODS AND INEXACT NEWTON-LIKE ITERATIONS

IOANNIS K. ARGYROS

Abstract. We provide a general theorem for the convergence of some projection methods for inexact Newton-like iterations under Yamamoto-type assumptions. Our results extend and improve several situations already appeared in the literature.

1. Introduction

We consider the inexact Newton-like method

$$x_{n+1} = x_n + y_{n'} \quad PA(x_n)y_n = -(F(x_n) + G(x_n)) + r_{n'} \quad n \geq 0 \tag{1}$$

for some $x_0 \in U(x_0, R), R > 0$, to approximate a solution x^* of the equation

$$F(x) + G(x) = 0, \quad \text{in} \quad \bar{U}(x_0, R). \tag{2}$$

Here $A(x), F, G$ denote operators defined on the closed ball $\bar{U}(x_0, R)$ with center x_0 and radius R, of a Banach space E with values in a Banach space $\hat{E}$, whereas r_n are suitable points in $\hat{E}$. The operator $A(x)(\cdot)$ is linear and approximates the Frechet derivative of F at $x \in U(x_0, R)$. P is a projection operator in E such that $P^2 = P$. We will assume that for any $x,\ y \in \bar{U}(x_0, r) \subseteq \bar{U}(x_0, R)$ with $0 \leq ||x - y|| \leq R - r$,

$$||P(F'(x + t(x - y)) - A(x))|| \leq B_1(r, ||x - x_0|| + t||y - x||), t \in [0, 1] \tag{3}$$

and

$$||P(G(x) - G(y))|| \leq B_2(r, ||x - y||). \tag{4}$$

Received June 14, 1993.
1991 *Mathematics Subject Classification.* 47D15, 47H17, 65R20, 65J10.
Key words and phrases. Banach space, Newton-like method.

The functions $B_1(r, r')$ and $B_2(r, r')$ defined on $[0, R] \times [0, R]$ and $[0, R] \times [0, R - r]$ are respectively nonnegative, continuous and non-decreasing functions of two variables. Moreover B_2 is linear in the second variable.

Note that the Newton method, the modified Newton method and the secant method are special cases of (1) with $A(x_n) = F'(x_n), A(x_n) = F'(x_0)$ and $A(x_n) = S(x_n, x_{n-1})$ respectively (when $P = I$ the identity operator on E, or not).

If we take

$$w(r') + c \tag{5}$$

and

$$e(r'), \tag{6}$$

where w, e are nonnegative, nondecreasing functions on $[0, R - r]$, to be the right hand sides of (3) and (4) respectively, then we obtain conditions similar but not identical to the Zabrejko-Nguen-type assumptions considered by Chen and Yamamoto [2]. They provided sufficient conditions for the convergence of the sequence $\{x_n\}, n \geq 0$ generated by (1) to solution x^* of equation (2), when $r_n = 0, n \geq 0$ and $P = I$.

Moret [5] also studied (1), when $G = 0$ and condition (5) is satisfied. Further work on this subject but for even more special cases than the ones considered by the above authors can be found in [1], [3], [4], [5], [6], [7], [8], [9], [10].

In this paper we will derive a criterion for controlling the residuals r_n in such a way that the convergence of the sequence $\{x_n\}$, $n \geq 0$ to a solution x^* of equation (2) is ensured.

We believe that conditions of the form (3)-(4) are useful not only because we can treat a wider range of problems than before, but it turns out that under natural assumptions we can find better error bounds on the distances $||x_n - x^*||, n \geq 0$.

The iterates $\{x_n\}$ generated by (1) when $P = I$ can rarely be computed in infinite dimensional spaces. It may be difficult or impossible to compute the inverses $A(x_n)$ at each step. It is easy to see however that the solution of equation (1) reduces to solving certain operator equations in the space E_p. If, moreover, E_p is a finite dimensional space of dimension N, we obtain a system of linear algebraic equations of at most order N.

2. Convergence Theorems.

Throughout the paper the notation $|| \cdot ||$ will stand both for norms in E (or in $\hat{E}$) and also for the induced operator norms $L(E, \hat{E})$, where $L(E, \hat{E})$ denotes the space of bounded linear operators from E to $\hat{E}$.

We will need the following proposition.

Proposition. *Let $a \geq 1, \sigma > 0, 0 \leq \mu < 1, 0 \leq \rho < R, s > 0$ be real constants such that the equation*

$$\varphi(t) := a\sigma[\int_0^t B_1(R, \rho + \theta)d\theta + B_2(R, t)] - t(1 - \mu) + s = 0 \tag{7}$$

has solutions in the interval $[0, R)$ *and let us denote by* x^* *the least of them.*
Let $v > 0, \mu^1 \geq 0$ *such that*

$$v(1-\mu) - (1-\mu^1) \leq 0. \tag{8}$$

Then, for every s^1 *satisfying*

$$0 < s^1 \leq v[\sigma(\int_0^s B_1(R, \rho+\theta)d\theta + B_2(R, s)) + s\mu] \tag{9}$$

and for every ρ^1 *such that*

$$0 \leq \rho^1 \leq \rho + s, \tag{10}$$

the equation

$$\varphi^1(t) := av\sigma[\int_0^t B_1(R, \rho^1+\theta)d\theta + B_2(R, t)] - t(1-\mu^1) + s^1 = 0 \tag{11}$$

has nonnegative solutions and at least one of them, denoted by t^{**}*, lies in the interval* $[s^1, t^* - s]$.

Proof. We first observe that since $\varphi(t^*) = 0$ and $0 \leq \mu < 1$, we obtain from (7) that $s \leq t^*$. We will show that

$$\varphi^1(t^* - s) \leq 0. \tag{12}$$

Using (7)-(11), we obtain

$$\begin{aligned}
\varphi^1(t^*-s) =& av\sigma[\int_0^{t^*-s} B_1(R, \rho^1+\theta)d\theta + B_2(R, t^*-s)] - (t^*-s)(1-\mu^1) + s^1 \\
\leq& v[a\sigma(\int_s^{t^*} B_1(R, \rho+\theta)d\theta + B_2(R, t^*) - B_2(R, s)) \\
&+ \sigma(\int_0^s B_1(R, \rho+\theta)d\theta + B_2(R, s)) + s\mu - \frac{(t^*-s)}{v}(1-\mu^1)] \\
\leq& v[a\sigma(\int_0^{t^*} B_1(R, \rho+\theta)d\theta + B_2(R, t^*)) - t^*(1-\mu) + s \\
&+ t^*(1-\mu) - s + s\mu - \frac{(t^*-s)}{v}(1-\mu^1)] \\
\leq& v(t^*-s)[(1-\mu) - \frac{(1-\mu^1)}{v}] \leq 0, \qquad \text{by}(8).
\end{aligned}$$

Hence, $\varphi^1(t)$ has nonnegative real roots and for the least of them t^{**}, it is

$$s^1 \leq t^{**} \leq t^* - s.$$

Moreover, from (11) we get $\mu^1 < 1$.

That completes the proof of the proposition.

We can now prove the following result.

Theorem 1. *Let* $\{s_n\}, \{\mu_n\}, \{\sigma_n\}, n \geq 0$ *be real sequences, with* $s_n > 0, \mu_n \geq 0, \sigma_n > 0$. *Let* $\{\rho_n\}$ *be a sequence on* $[0, R)$, *with* $\rho_0 = 0$ *and*

$$\rho_{n+1} \leq \sum_{j=0,1,2,\dots,n} s_j,\ n \geq 0. \tag{13}$$

Suppose that $1 - \mu_0 > 0$ *and that, for a given constant* $a \geq 1$, *the function*

$$\varphi_0(t) := a\sigma_0[\int_0^t B_1(R, \rho_0 + \theta)d\theta + B_2(R, t)] - t(1 - \mu_0) + s_0 \tag{14}$$

has roots on $[0, R)$.

Assume that for every $n \geq 0$ *the following conditions are satisfied*

$$s_{n+1} \leq v_n[\sigma_n(\int_0^{s_n} B_1(R, \rho_n + \theta)d\theta + B_2(R, s_n)) + s_n\mu_n], \tag{15}$$

$$v_n(1 - \mu_n) - (1 - \mu_n) \leq 0, \tag{16}$$

where $v_n = \frac{\sigma_{n+1}}{\sigma_n}$.

Then,

(a) *for every* $n \geq 0$, *the equation*

$$\varphi_n(t) := av_n\sigma_n[\int_0^t B_1(R, \rho_n + \theta)d\theta + B_2(R, t)] - t(1 - \mu_n) + s_n \tag{17}$$

has solutions in $[0, R)$ *and, denoting by* t_n^* *the least of them, we have*

$$\sum_{j=n,\dots,\infty} s_j \leq t_n^*. \tag{18}$$

(b) *Let* $\{x_n\}, n \geq 0$ *be a sequence in a Banach space such that* $||x_{n+1} - x_n|| \leq s_n$. *Then, it converges and denoting its limit by* x^*, *the error bounds*

$$||x^* - x_n|| \leq t_n^* \tag{19}$$

and

$$||x^* - x_{n+1}|| \leq t_n^* - s_n \tag{20}$$

are true for all $n \geq 0$.

(c) *If there exists* $h_0 \in [0, R)$ *such that*

$$\varphi_0(h_0) \leq 0, \tag{21}$$

then $\varphi_0(t)$ *has roots on* $[0, R)$.

Proof.

(a) We use induction on n. Let us assume that for some $n \geq 0$, $1-\mu_n > 0$, $\varphi_n(t)$ has roots on $[0, R)$ and t_n^* is the least of them. This is true for $n = 0$. Then, by (13), (15), (16) and the proposition, by setting $s = s_n, s^1 = s_{n+1}, \mu = \mu_n, \mu^1 = \mu_{n+1}$ and $v = v_n$, it follows that t_{n+1}^* exists, with

$$s_{n+1} \leq t_{n+1}^* \leq t_n^* - s_n$$

and $1 - \mu_{n+1} > 0$.

That completes the induction and proves (a).

(b) This part follows easily from part (a).

(c) Using (21), we deduce immediately that $\varphi_0(t)$ has roots on $[0, R)$.

That completes the proof of theorem.

We can now prove the main result.

Theorem 2. *Consider the method* (1). *Assume that for* $s_0 > 0, \sigma_0 > 0, 0 \leq \mu_0 < 1$ *and* $a \geq 1$, (21) *is true. Then, the function* $\varphi_0(t)$ *defined by* (14) *has roots on* $[0, R)$. *Denote by* t_0^* *the least of them and suppose that*

$$t_0^* < R_0 \leq R. \tag{22}$$

Let $s_n > 0, \mu_n \geq 0, \sigma_n > 0, n \geq 0$ *be such that* $\liminf \sigma_n > 0$ *as* $n \to \infty$ *and condition* (15) *is true for all* $n \geq 0$.

Assume that, for all $n \geq 0$,

$$||y_n|| \leq s_n \leq \sigma_n ||P(F(x_n) + G(x_n))|| \tag{23}$$

and

$$||Pr_n|| \leq \frac{\mu_n s_n}{\sigma_n}. \tag{24}$$

Then the sequence $\{x_n\}, n \geq 0$ *generated by* (1) *remains in* $U(x_0, t_0^*)$ *and converges to a solution* x^* *of equation* (2). *Moreover, the error bounds* (19) *and* (20) *are true for all* $n \geq 0$, *where* t_n^* *is the least root in* $[0, R)$ *of the function* $\varphi_n(t)$ *defined by* (17), *with* $\rho_n = ||x_n - x_0||, n \geq 0$.

Proof. The existence of t_0^* is guaranteed by (21). Let us assume that $x_n, x_{n+1} \in U(x_0, t_0^*)$. We will show that for every $n \geq 0$, condition (15) is true. Since $||y_0|| \leq s_0$, this is true for $n = 0$.

Using the identity

$$P(F(x_{n+1}) + G(x_{n+1})) = \int_0^1 P[F'(x_n + t(x_{n+1} - x_n)) - A(x_n)](x_{n+1} - x_n)dt$$
$$+P(G(x_{n+1} - G(x_n)) + Pr_n,$$

(3), (4), (23), (24), setting $\rho_n = ||x_n - x_0||$ and by taking norms in the above identity we get

$$s_{n+1} \leq \sigma_{n+1}||P(F(x_{n+1})+G(x_{n+1}))|| \leq v_n[\sigma_n(\int_0^{s_n} B_1(R, \rho_n+\theta)d\theta+B_2(R, s_n))+s_n\mu_n]$$

which shows (15) for all $n \geq 0$.

The hypothesis (b) of Theorem 1 can now easily be verified by induction and thus, by(18) and (23), the sequence $\{x_n\}$, $n \geq 0$ remains in $U(x_0, t_0^*)$, converges to x^* and (19) and (20) hold. Finally, from the inequality

$$||P(F(x_n) + G(x_n))|| \leq ||P(A(x_n) - F'(x_0))|| \, ||y_n|| + ||PF'(x_0)|| \, ||y_n|| + ||Pr_n||,$$

(3), (24) and the continuity of F and G, as $\liminf \sigma_n > 0$ and $s_n \to 0$, as $n \to \infty$ it follows that $P(F(x^*) + G(x^*)) = 0$. Hence $F(x^*) + G(x^*) = 0$.

That completes the proof of the theorem.

Remark.

(a) In the special case when B_1 and B_2 are given (5) and (6) respectively, then our results can be reduced to the ones obtained by Moret [5, p.359] (when $G = 0$ and $P = I$).

(b) Let $G = 0, P = I, A(x) = F'(x)$ and define the functions $\bar{\varphi}_0(t), \bar{\varphi}_n(t)$ by

$$\bar{\varphi}_0(t) = a\sigma_0 \int_0^t (t-\theta)k(\theta)d\theta - t(1-\mu_0) + s_0,$$

$$\bar{\varphi}_n(t) = av_n\sigma_n \int_0^t (t-\theta)k(\rho_n+\theta)d\theta - t(1-\mu_n) + s_n,$$

where k is a nondecreasing function on $[0, R]$ such that

$$||F'(x) - F'(y)|| \leq k(r)||x-y||, \quad x, y \in \bar{U}(x_0, r) \qquad (r < R_0).$$

Assume that B_1 can be chosen in such a way that

$$\varphi_n(t) \leq \bar{\varphi}_n(t), \; n \geq 0. \tag{25}$$

The under the hypotheses of Theorem 2 above and Proposition 1 in [5, p. 359], using (25) we can show

$$||x^* - x_n|| \leq t_n^* \leq m_n^*, \; n \geq 0$$

and

$$||x^* - x_{n+1}|| \leq t_n^* - s_n \leq m_n^* - s_n, \; n \geq 0$$

where by m_n^*, we denote the least solutions of the equations

$$\bar{\varphi}_n(t) = 0, \; n \geq 0 \quad \text{in} \quad [0, R).$$

References

[1] P.T. Brown, "A local convergence theory for inexact-Newton finite-difference projection methods," *SIAM J. Numer. Anal.*, 24, (1987), 407-437.

[2] X. Chen, T. Yamamoto, "Convergence domains of certain iterative methods for solving nonlinear equations," *Numer. Funct. Anal. and Optimiz.*, 10, (1 and 2), (1989), 37-48.

[3] R.S. Dembo, S.C. Eisenhart, and T. Steihaug, "Inexact Newton methods," *SIAM J. Numer. Anal.*, 19, (1982), 400-408.

[4] M.A. Krasnosel'skii, and Y.B. Rutickii, "Some approximate methods of solving nonlinear equations based on linearization," *Soviet Math. Dokl.*, 2, (1961), 1542-1546.

[5] I. Moret, "A Kantorovich-type theorem for inexact Newton methods," *Numer. Funct. Anal. and Optimiz.*, 10, (3 and 4), (1989), 351-365.

[6] F.A. Potra, and V. Ptàk, *Nondiscrete induction and iterative processes.*, Pitman Advanced Publishing Program, London, 1984.

[7] A.H. Sherman, "On Newton-iterative methods for the solution of systems of equations," *SIAM J. Numer. Anal.*, 15, (1978), 755-771.

[8] P.P. Zabrejko, and D.F. Nguen, "The majorant method in the theory of Newton-Kantorovich approximations and the Ptàk error estimates," *Numer. Funct. Anal. and Optimiz.*, 9, (1987), 671-684.

[9] T. Yamamoto, "A note on a posterior error bound of Zabrejko and Nguen for Zincenko's iteration," *Numer. Funct. Anal. and Optimiz.*, 9, (1987), 987-994.

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

Cameron University, Department of Mathematics, Lawton, OK 73505-6377, U. S. A.