Annales Univ. Sci. Budapest., Sect. Comp. 15 (1995) 37-47

STIRLING'S METHOD IN GENERALIZED BANACH SPACES

I.K. Argyros (Lawton, USA)

Abstract. We provide convergence results and error estimates for Stirling's method in generalized Banach spaces. The idea of a generalized norm is used which is defined to be a map from a linear space into a partially ordered Banach space. Convergence results and error estimates are improved compared with the ones of other methods.

I. Introduction

A fixed point x^* of an operator G defined on a convex subset D of a generalized Banach space E (to be precised later) and taking values into itself satisfies the equation

(1) $$F(x^*) = 0 \quad \text{in} \quad D,$$

where

(2) $$F(x) = x - G(x) \quad \text{for all} \quad x \in D.$$

The fixed point x^* can be approximated using a method of the form

(3) $$x_{n+1} = x_n + y_{n'}(I - F'(F(x_n)))y_n + F(x_n) = 0, \quad n \geq 0$$

(4) $$\iff \quad y_n := L_n(y_n) = F'(F(x_n))y_n - F(x_n), \quad n \geq 0, \quad x_0 \in D.$$

The above method is the so-called Stirling's method. Stirling's method can be viewed as a combination of the method of successive substitutions and Newton's method. It is consequently reasonable to examine the convergence of the method of successive substitutions. In terms of computational effort, Stirling's and Newton's method require essentially the same labour per step, as each requires the evaluation of F, F' and the solution of a linear equation, assuming that F and its derivative are evaluated independently. Convergence results and error estimates for Stirling's method are derived. A generalized norm is used which is defined to be a map from a linear space into a partially ordered Banach space. By this tool the metric properties of the problem can be better analyzed. Our results compare favourably with similar ones obtained for Newton's method in [7], [6] and [8]. Finally relevant work can be found in [1], [9] and [11] for the real norm theory.

II. Preliminaries

To make the paper as self-contained as possible we reproduce some variations of definitions 1-5 which can also be found in [8, p.250].

Definition 1. *A generalized Banach space is defined to be a triple* $(E, Z, /\cdot/)$ *with the following properties:*

(a) *E is a linear space over $\mathbb{R}(C)$;*

(b) *$Z = (Z, K, \|\cdot\|)$ is a partially ordered Banach space.*
That is

(b_1) *$(Z, \|\cdot\|)$ is a real Banach space;*

(b_2) *Z is partially ordered by a closed convex cone K;*

(b_3) *the norm $\|\cdot\|$ is monotone on K;*

(c) *The map $/\cdot/ : E \to K$ satisfies for $x, y \in E$, $\lambda \in \mathbb{R}(C)$:*

$$/x/ = 0 \Longleftrightarrow x = 0, \quad /\lambda x/ = /\lambda/ \cdot /x/, \quad /x+y/ \le /x/ + /y/.$$

(d) *E is a Banach space with respect to the induced norm*

$$\|\cdot\|_i := \|\cdot\| \cdot /\cdot/.$$

The map $/\cdot/$ is called a generalized norm. Because of (c) and (b_3) $\|\cdot\|_i$ is indeed a real norm. In the following all topological terms will be understood with respect to this norm.

Definition 2. *We will denote by $L(E^n, E)$ the space of n-linear symmetric bounded operators from E^n to E, where E^n, E are Banach spaces. If E, E are partially ordered denote by $L_+(E^n, E)$ the subset of monotone operators M such that*

$$0 \le a_i \le b_i \Longrightarrow M(a_1, \ldots a_n) \le M(b_1, \ldots, b_n).$$

Definition 3. *For an operator $L \in L(E,E)$ on a generalized Banach space $(E, Z, / \cdot /)$ the set of bounds is defined to be*

$$B(L) := \{M \in L_+(Z,Z) / |Lx| \leq M|x| \text{ for } x \in E\}.$$

III. Convergence conditions and error estimates

In this section we shall study the iterative procedure (3)-(4) for pairs (F, x_0) belonging to the class $C(r, P, Q_1, Q_2)$ defined as follows: we say that a pair (F, x_0) belongs to the class $C(r, P, Q_1, Q_2)$ if

(a) F is a nonlinear operator defined on a convex subset D of a Banach space E and with values in E.

(b) The operator F is Fréchet-differentiable on the interior D^0 of D and the linear operators $I - F'(F(x_n))$, $n \geq 0$ are such that:

(b_1) There exist an operator $P \in B(F'(F(x_0)))$ and operators $Q_1, Q_2 \in$ $\in L(Z^2, Z)$ such that

$$(5) \qquad |F'(v)y - F'(w)y| \leq Q_1(|v-w|, |y|) \quad \text{for} \quad v, w \in D, \quad y \in E$$

and

$$(6) \qquad |F'(x_n)y - (I - F'(F(x_n)))y| \leq Q_2\left(\sum_{j=1}^{n} |x_j - x_{j-1}|, |y|\right), \quad n \geq 0, \quad x_n \in D, \quad y \in E.$$

(c) There exists a solution $r \in K$ of the inequality

$$(7) \qquad R_0(r) := Pr + (Q_1 + Q_2)r^2 + |F(x_0)| \leq r.$$

(d) The ball

$$(8) \qquad U(x_0, r) = \{x \in E|\ |x - x_0| \leq r\} \leq D$$

and

(e) the estimate

$$(9) \qquad (P + 2(Q_1 + Q_2)r)^k r \to 0 \quad \text{as} \quad k \to \infty$$

is true.

We will need the following

Lemma 1. *Let $(Z, K, \|\cdot\|)$ be a partially ordered Banach space, w be an element of K and $P \in L_+(Z, Z)$, $Q \in L_+(Z^2, Z)$.*
(a) If there exists $r \in K$ such that

$$R_0(r) \leq r \quad \text{and} \quad (P + 2(Q_1 + Q_2))^k r \to 0 \quad \text{as} \quad k \to \infty$$

then

$$b := R_0^\infty(0) := \lim_{n\to\infty} (R_0^n(0))$$

is well defined, solves $b = R_0(b)$ and is smaller than any solution of the inequality $R_0(d) \leq d$.
(b) If there exists $q \in k$ and $\lambda \in (0,1)$ such that $R_0(q) \leq \lambda q$ then there exists $r \leq q$ satisfying (a).

Proof. (a) Let us define the sequence $\{b_n\}$, $n \geq 1$ by $b_n = R_0^n(0)$. We first show that the sequence $\{b_n\}$, $n \geq 1$ is bounded above by r. For $n = 0$, $b_1 = = R_0(0) = w \leq r$. Let us assume $b_n = R_0^n(0) \leq r$. Then

$$b_{n+1} = R_0^{n+1}(0) = R_0(R_0^n(0)) = R_0(b_n) = Pb_n + (Q_1 + Q_2)b_n^2 + w \leq$$

$$\leq Pr + (Q_1 + Q_2)r^2 + w = R_0(r) \leq r.$$

That completes the induction.

We now define the sequence $\{P_n\}$, $n \geq 1$ by $P_n = R_0^{n+1}(0) - R_0^n(0)$, $n \geq 1$. We will show that

$$P_n \leq (P + 2(Q_1 + Q_2)r)^n r, \qquad n \geq 1. \tag{10}$$

For $n = 1$ we get

$$P_1 = R_0^2(0) - R_0(0) = R_0(R_0(0)) - R_0(0) = R_0(w) - R_0(0) = \int_0^1 R_0'(tw)w dt \leq$$

$$\leq R_0'(w)w = (P + 2(Q_1 + Q_2)w)w \leq (P + 2(Q_1 + Q_2)r)r.$$

That is (10) is true for $n = 1$. Assume that (10) is true for $k = 1, 2, \ldots, n$. Then

$$P_{k+1} = R_0^{k+2}(0) - R_0^{k+1}(0) = R_0^{k+1}(R_0(0)) - R_0^{k+1}(0) = R_0^{k+1}(w) - R_0^{k+1}(0) =$$

$$= R_0(R_0^k(w)) - R_0(R_0^k(0)) =$$

$$= \int_0^1 R_0'(R_0^k(0) + t(R_0^k(w) - R_0^k(0)))(R_0^k(w) - R_0^k(0))dt \leq$$

$$\leq R_0'(R_0^k(w))(R_0^k(w) - R_0^k(0)) \leq$$

$$\leq R_0'(r)(R_0^{k+1}(0) - R_0^k(0)) \leq (P + 2(Q_1 + Q_2)r)^{k+1}r.$$

That completes the induction. From (10) it now follows that $\{b_n\}$, $n \geq 1$ is a Cauchy sequence in a Banach space as such it converges to some $b = R_0^\infty(0)$ and $b = R_0^\infty(b)$. Since $b_n \leq r$, then $b \leq r$. That is, b is smaller than any solution d of the inequality $R_0(d) \leq d$.

(b) Let us define the sequences $\{v_n\}$, $\{q_n\}$, $n \geq 1$ by $v_0 = 0$, $v_{n+1} = R_0(v_n)$ and $q_0 = q$, $q_{n+1} = R_0(q_n)$. By the monotonicity of R_0 we get

$$0 \leq v_n \leq v_{n+1} \leq q_{n+1} \leq q_n \leq q.$$

We will show the estimate

$$q_n - v_n \leq \lambda^n(q - v_n), \qquad n \geq 0. \tag{11}$$

Inequality (11) is trivially true for $n = 0$. Let us assume that it is true for $k = 0, 1, 2, \ldots, n$. Then

$$q_{n+1} - v_{n+1} = R_0(q_n) - R_0(v_n) = \int_0^1 R_0'(v_n + t(q_n - v_n))(q_n - v_n)dt \leq$$

$$\leq \lambda^n \int_0^1 R_0'(v_n + t(q - v_n))(q - v_n)dt = \lambda^n(R_0(q) - R_0(v_n)) \leq \tag{12}$$

$$\leq \lambda^{n+1}(q - v^{n+1}).$$

The sequence $\{v_n\}$ is monotonically increasing and bounded above by q. Therefore it converges to some $r \leq q$. Since the sequence $\{q_n - v_n\}$, $n \geq 1$ converges to zero by (12), the sequence $\{q_n\}$ must also be convergent and bounded above by r. This completes the proof of the lemma.

We now state and prove a lemma concerning the solvability of the fixed point problems appearing in (4).

Lemma 2. *Let $(E, (Z, K, \|\cdot\|, /\cdot/))$ be a generalized Banach space and $\overline{L} \in B(L)$ be a bound for $L \in L(E, E)$. If for $y \in E$ there exists $q \in K$ such that*

$$\overline{L}q + /y/ \leq q \quad \text{and} \quad \overline{L}^k q \to 0 \quad \text{as} \quad k \to \infty$$

then $e = T^\infty(0)$, $T(x) = Lx + y$ is well defined and satisfies

$$e = Le + y \quad \text{and} \quad /e/ \leq \overline{L}/e/ + /y/ \leq q.$$

Proof. It is enough to show that $\{T^n(0)\}$, $n \geq 1$ is a Cauchy sequence. Using Definition 3 we obtain the estimate

$$/T^{n+k}(0) - T^k(0)/ =$$

$$= /(L^{n+k-1}y + L^{n+k-2}y + \ldots + L^{k-1}y + y) - (L^{k-1}y + L^{k-2}y + \ldots + y)/ \leq$$

$$\leq \overline{L}^{n+k-1}/y/ + \overline{L}^{n+k-2}/y/ + \ldots \overline{L}^k/y/ \leq$$

$$\leq (\overline{L}^{n+k-1} + \overline{L}^{n+k-2} + \ldots + \overline{L}^k)(q - \overline{L}q) = \overline{L}^k q - \overline{L}^{n+k} q \leq \overline{L}q \to 0$$

as $k \to \infty$. The lemma now follows immediately from the above estimate.

We can now formulate the main result.

Theorem 1. *If $(F, X_0) \in C(r, P, Q_1, Q_2)$ then*

(a) the iterative procedure (3)-(4) is well defined and the sequence $\{x_n\}$, $n \geq 0$ produced by it converges to a unique fixed point x^* of G in $U(x_0, r)$.

(b) An a priori estimate is given by the null sequence $\{r_n\}$, $n \geq 0$ defined by $r_0 - r$ and $r_n = P_0^\infty(0)$, where

$$P_n(q) = Pq + (Q_1 + Q_2)(r - r_{n-1})q + Q_1 r_{n-1}^2 + Q_2(r_0 - r_{n-1})r_{n-1}, \quad n \geq 1.$$

(c) An a posteriori estimate is given by the sequence $\{d_n\}$, $n \geq 1$ with

$$d_n = R_0^\infty(0),$$

$$R_n(q) = Pq + (Q_1 + Q_2)\left(\sum_{j=0}^{n} a_j\right) q + Q_1 r_{n-1}^2 + Q_2 \left(\sum_{j=0}^{n} a_j\right)(a_{n-1}),$$

where $a_{n-1} = /x_n - x_{n-1}/$, $n \geq 0$.

Proof. Using induction on n we will show the statement:

(S_n) $x_n \in E$, $r_n \in K$ are well defined and satisfy

$$r_n + a_{n-1} \leq r_{n-1}, \qquad n \geq 1.$$

By Lemma 1, (7) and (9) there exists $q \leq r$ such that

$$Pq + /F(x_0)/ = q \quad \text{and} \quad P^k q \leq P^k r \to 0 \quad \text{as} \quad k \to \infty.$$

By Lemma 2 x_1 is well defined and $a_0 \leq q$. We now have
(13)

$$\begin{aligned} P_1(r-q) &= P(r-q) + (Q_1+Q_2)(r-r)(r-q) + Q_1 r^2 + Q_2(r-r)r = \\ &= Pr - Pq + Q_1 r^2 = Pr + Q_1 r^2 + /F(x_0)/ - q \leq R_0(r) - q \leq r - q. \end{aligned}$$

Using Lemma 1 and (13) we get that r_1 is well defined and

$$r_1 + a_0 \leq r - q + q = r_0.$$

That is, the statement (S_1) is true. Assume now that (S_1), (S_2),...,(S_n) are true. Let us observe that

$$Pr_n + (Q_1+Q_2)(r-r_n)r_n + Q_1(r_{n-1}-r_n)^2 + Q_2(r-r_{n-1})(r_{n-1}-r_n) \leq$$

$$\leq P_n r_n \leq r_n.$$

Using Lemma 1 we can find $q \leq r_n$ such that

(14) $q = Pq + (Q_1+Q_2)(r-r_{n-1})q + Q_1(r_{n-1}-r_n)^2 + Q_2(r_0-r_{n-1})(r_{n-1}-r_n).$

We also have that

$$t_n = /x_n - x_0/ \leq \sum_{j=0}^{n-1} a_j \leq \sum_{j=0}^{n-1}(r_j - r_{j+1}) = r - r_n \leq r.$$

That is, $x_n \in U(x_0, r) \subset D$. Let us note that for all $n \geq 1$ (5), (6) give

$$\begin{aligned} /F'(F(x_n))y/ &\leq /(I - F'(x_0))y/ + \\ &+ /(F'(x_0) - F'(x_n))y/ + /(F'(x_n) - (I - F'(F(x_n)))y/ \leq \\ &\leq P/y/ + Q_1(t_n, /y/) + Q_2\left(\sum_{j=1}^{n} /x_j - x_{j-1}/, /y/\right) \leq \\ &\leq P/y/ + Q_1(r - r_n, /y/) + Q_2(r - r_n, /y/). \end{aligned}$$

That is, the operator $P + (Q_1 + Q_2)(r - r_n)$ is a bound for $F'(F(x_n))$, $n \geq 1$. Moreover, using (5), (6) and the induction hypothesis we get

(15)
$$\begin{aligned} /F(x_n)/ &\leq /F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1})/ + \\ &+ /(F'(x_{n-1})(I - F'(F(x_{n-1})))(x_n - x_{n-1})/ \leq \\ &\leq Q_1(/x_{n-1} - x_n/^2) + Q_2\left(\sum_{j=1}^{n-1} /x_j - x_{j-1}/\right)(/x_{n-1} - x_n/) \leq \\ &\leq Q_1(r_{n-1} - r_n)^2 + Q_2(r - r_{n-1})(r_{n-1} - r_n). \end{aligned}$$

By (14) and (15) we get

$$PQ + (Q_1 + Q_2)(r - r_n)q + /F(x_n)/ \leq q.$$

Hence, by Lemma 2, x_{n+1} is well defined and $a_n \leq q \leq r_n$.

As in (13) it is simple calculus to show

$$P_{n+1}(r_n - q) \leq P_n(r_n) - q \leq r_n - q.$$

By Lemma 1 r_{n+1} is well defined and

$$r_{n+1} + a_n \leq r_n - q + q = r_n.$$

Hence (S_{n+1}) is true. For $m \geq n$ we get

$$/x_{m+1} - x_n/ \leq \sum_{j=n}^{m} a_j \leq \sum_{j=1}^{m}(r_j - r_{j+1}) = r_n - r_{m+1} \leq r_n. \tag{16}$$

Moreover,

$$\begin{aligned} r_{n+1} &= P_{n+1}(r_{n+1}) \leq \\ \leq P_{n+1}(r_n) \leq (P + 2(Q_1 + Q_2)r)r_n &\leq \ldots \leq (P + 2(Q_1 + Q_2)r)^{n+1}r. \end{aligned}$$

By (9) the sequence $\{r_n\}$ converges to zero. Using (16) we derive that the sequence $\{x_n\}$ is a Cauchy sequence in a Banach space and as such it converges to some $x^* \in E$. Let $m \to \infty$ in (16) to obtain $x^* \in U(x_n, r_n)$. From (15) we get that x^* is a fixed point of G by letting $n \to \infty$. The uniqueness of x^* in $U(x_n, r_n)$ follows easily as in [7, Theorem 4.1]. Thus we have shown (a) and (b).

To show (c) note that the sequence $\{d_n\}$ is well defined by Lemma 1, since

$$R_n(r_n) \leq P_n(r_n) \leq r_n.$$

Hence, $d_n \leq r_n$ in general. We can easily show that using (16) that

$$(F, x_0) \in C\left(d_n,\ P + (Q_1 + Q_2)\left(\sum_{j=1}^{n} a_j\right),\ Q_1,\ Q_2\right).$$

Hence by (a) $x^* \in U(x_n, d_n)$ which shows (c). That completes the proof of the theorem.

For practical purposes the a posteriori estimate is of interest mostly. Condition (9) can be avoided then. In particular, following the techniques used in Lemmas 1, 2 and Theorem 1 we can easily prove the following theorems.

Theorem 2. *Assume*

(a) the conditions (5) and (6) are satisfied;

(b) there exist $d \in K$, $\lambda \in (0,1)$ such that

$$R_0(d) = Pd + (Q_1 + Q_2)d^2 + /F(x_0)/ \leq \lambda d;$$

(c) the ball $U(x_0, d) \subset D$.

Then there exists $r \leq d$ such that $(F, x_0) \in C(r, P, Q_1, Q_2)$. The fixed point x^* of G is unique in $U(x_0, d)$.

Theorem 3. *Let the conditions of Theorem 1 be satisfied. If $d \in K$ solves $R_n(d) \leq d$ then $q = d - a_n \in K$ and solves $R_{n+1}(q) \leq q$.*

Remarks. (a) Note that this solution might be improved by $R^k_{n+1}(q) \leq q$ for any $k \in N$.

(b) In case of a real-normed space (i.e. $Z = R$) by (5), (6) an operator norm $P \in R \geq 0$ and Lipschitz-constants $Q_1, Q_2 \in R \geq 0$ are defined. The condition (6) becomes a real quadratic inequality then and the smallest solution is used [1]-[5], [9]-[12].

(c) The motivation for (6) comes from conditions of the form

$$/P(x_n)y - A_n y/ \leq Q_2\left(\alpha \sum_{j=1}^{n} |x_j - x_{j-1}| + v_n - \beta,\ |y|\right),$$

where $\alpha \in R^+$, $v_n, \beta \in K$ for all $n \geq 0$, which were introduced for the real norm theory (see e.g. [4, p.431] and the references there.

(d) Let c_n denote upper bounds for the distances $|z_{n+1} - z_n|$, $n \geq 0$, where $\{z_n\}$, $n \geq 0$ is the corresponding Newton-sequence which was produced to approximate a fixed point x^* of (1) in [8, p.251]. Then it can be easily seen by induction on n (see also relevant work in [1], [9]) that if r_0 is sufficiently smaller than c_0, then $r_n \leq c_n$ for all $n \geq 0$. That is Stirlings's method converges faster to the same fixed point x^* of (1) than Newton's method (assuming that we start from the same initial guess x_0).

References

[1] **Argyros I.K.**, On the solution of equations with nondifferentiable operators and Ptak error estimates, *BIT,* **30** (1990), 752-754.

[2] **Argyros I.K.**, On the Secant method and fixed points of nonlinear equations, *Monatshefte für Mathematik,* **106** (1988), 85-94.

[3] **Balazs M. and Goldner G.**, On the method of the cord and on a modification of it for the solution of nonlinear operator equations, *Stud. Cerc. Mat.,* **20** (1968), 981-990.

[4] **Dennis J.E.**, Toward a unified convergence theory for Newton-like methods, *Nonlinear Functional Analysis and Applications,* ed. L.B. Rall, Academic Press, New York, 1971, 425-472.

[5] **Kantorovich L.V. and Akilov G.P.**, *Functional analysis in normed spaces,* Academic Press, New York, 1978.

[6] **Meyer P.W.**, *Die Anwendung verallgemeinerter Normen zur Fehlerabschätzung bei Iterationsverfahren,* Dissertation, Düsseldorf, 1980.

[7] **Meyer P.W.**, Die modifizierten Newton-Verfahren in verallgemeinerten Banach-Räumen, *Numer. Math.,* **43** (1984), 91-104.

[8] **Meyer P.W.**, Newton's method in generalized Banach spaces, *Numer. Funct. Anal. and Optimiz.,* **9** (3-4) (1987), 249-259.

[9] **Rall L.B.**, Convergence of Stirling's method in Banach spaces, *Aequationes Math.,* **12** (1975), 12-20.

[10] **Schröder J.**, Über das Newtonsche Näherungsverfahren, *Arch. Rat. Mech. Anal.,* **1** (1957), 154-180.

[11] **Stirling J.**, *Methodus differentialis: sire tractatus de summatione et interpolatione serierum infinitarum,* W.Boyer Pub., London, 1730.

[12] **Yamamoto T.**, Componentwise error estimates for approximate solutions of nonlinear equations, *J.Inf.Proc.*, **2** (1979), 121-126.

(Received October 30, 1991)

I.K. Argyros
Department of Mathematics
Cameron University
Lawton, OK 73505
USA