AN ITERATIVE SOLVING OF NONLINEAR EQUATIONS

by

BÉLA SZABÓ

Central Mining Development Institute, 1037 Budapest, Mikoviny u. 2–4.

(Received March 21, 1983 – revised January 18, 1984)

Introduction. There are different methods for solving nonlinear operator equations $P(x) = 0$ [1]. In many cases a functional equation $F(x) = 0$ equivalent to an operator equation $P(x) = 0$ can be constructed. Thus it seems to be convenient to formulate the methods of solution of nonlinear operator equations for functionals as well. With these methods functionals need less calculations than do operators.

The generalized tangent hyperbola method for nonlinear operator equations is:

$$x_{n+1} = x_n - \left[I - \frac{1}{2}\Gamma_n P''(x_n)\Gamma_n P(x_n)\right]^{-1} \Gamma_n P(x_n) \quad (n = 0, 1, 2, \dots)$$

where Γ_n denotes the inverse of $P'(x_n)$.

It can be seen that two inverse operators should be calculated at every iteration step.

The present paper discusses the generalized tangent hyperbola method for functionals.

Let X be a complete, linear normed space.

Let us consider a

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

nonlinear functional equation, where $F(x): D \to \mathbf{R}$, $D \subset X$ convex and $F(x)$ three times differentiable in D, in Fréchet's sense. The generalized tangent hyperbola method for functionals was constructed as follows: [1]

$$x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)y_n - \frac{1}{2} \cdot \frac{F''(x_n)y_n^2}{F'(x_n)y_n} F(x_n)} y_n, \quad n = 0, 1, 2, \dots \tag{2}$$

where $F'(x_n)$ and $F''(x_n)$ represent first and second order derivatives, respectively, in Fréchet's sense, for $x_n \in D$; $\{y_n\} \subset X$ are chosen so that the following condition can be satisfied:

$$F'(x_n)\cdot y_n = \|F'(x_n)\|$$

where $\|y_n\| = 1$, $n = 0, 1, 2, \ldots$ [1], [2].

Theorem 1. *Let us assume that the following conditions are satisfied:* (x_0 *is a starting approximation*)

1.) $$\frac{1}{\|F'(x_0)\|} \le B_0 \quad < +\infty;$$

2.) $$\frac{\|F(x_0)\|}{\|F'(x_0)\|} \le \eta_0 \quad < +\infty;$$

3.) $$\|F''(x)\| \le M \text{ and } \|F'''(x)\| \le N \quad x \in D(x_0, \varrho) \text{ where}$$

$$\varrho := \frac{16}{9}\eta_0, \quad D(x_0, \varrho) := \{x : \|x - x_0\| \le \varrho\} \text{ and}$$

$F''(x)$, $F'''(x)$ are 2., 3. *order derivatives*;

4.) $$h_0 := B_0 M \eta_0 \le \frac{1}{2};$$

5.) $$\sigma_0 := \frac{\dfrac{1}{2} + \dfrac{N}{3B_0M^2\left(1 - \dfrac{1}{2}h_0\right)}}{\left(1 - \dfrac{1}{2}h_0\right)\left(1 - \dfrac{3}{2}h_0\right)} \le 2;$$

then there exist $x^* \in D$, $F(x^*) = 0$ *and*

$$\|x^* - x_n\| \le \frac{2^2(2h_0)^{3^n - 1}\eta_0}{2^{2n}\left(1 - \dfrac{1}{2}h_0\right)}. \quad \square$$

Proof. First of all we prove that conditions 1.), – 5.), are satisfied for x_1, too, where the first approximation x_1 is given by iteration (2), for $n = 0$.

a.) $$\|F'(x_1)\| \ge \|F'(x_0)\|\left[1 - \frac{\|F'(x_0) - F'(x_1)\|}{\|F'(x_0)\|}\right].$$

Using the Lagrange's generalized formula and conditions 1.), 3.),: we have

$$\|F'(x_1)\| \ge \|F'(x_0)\| \cdot [1 - B_0 M \|x_1 - x_0\|],$$

and using (2) and the conditions 1), – 4), we can establish the estimations:

$$\|x_1 - x_0\| \le \frac{|F(x_0)|}{\|F'(x_0)\|} \cdot \frac{1}{1 - \dfrac{1}{2}\dfrac{\|F''(x_0)\|}{\|F'(x_0)\|} \cdot \dfrac{|F(x_0)|}{\|F'(x_0)\|}} \le \frac{\eta_0}{1 - \dfrac{1}{2}B_0 M \eta_0} \tag{3}$$

so

$$\|F'(x_1)\| \geq \|F'(x_0)\| \left[1 - \frac{h_0}{1 - \frac{1}{2}h_0} \right]$$

and

(4) $$\frac{1}{\|F'(x_1)\|} \leq \frac{2-h_0}{2-3h_0} =: B_1 \quad (> B_0)$$

Thus the condition 1.), is valid for x_1 too and we have

$$B_0 < B_1 < 4B_0.$$

b) Applying the generalized Taylor's formula for $F(x)$:

$$\left| F(x_1) - F(x_0) - F'(x_0)(x_1-x_0) - \frac{1}{2} F''(x_0)(x_1-x_0)^2 \right| \leq \frac{1}{6} \|F'''(\bar{x})\| \cdot \|x_1 - x_0\|^3$$

where $\bar{x} := x_0 + \vartheta(x_1 - x_0)$, $0 \leq \vartheta \leq 1$,
we get

$$|F(x_1)| \leq \left| F(x_0) + F'(x_0)(x_1 - x_0) + \frac{1}{2} F''(x_0)(x_1-x_0)^2 \right| + \\ + \frac{1}{6} \|F'''(\bar{x})\| \cdot \|x_1 - x_0\|^3.$$

Using the identity

$$F(x_0) + F'(x_0)(x_1-x_0) + \frac{1}{2} F''(x_0)(x_1-x_0)^2 =$$

$$= \frac{F^3(x_0)[F''(x_0)y_0^2]^2}{4 \cdot [F'(x_0)y_0]^2 \cdot \left[F'(x_0)y_0 - \frac{1}{2} \frac{F''(x_0)y_0^2}{F'(x_0)y_0} F(x_0) \right]}$$

and the conditions 1.)–4.) we obtain

$$\frac{F^3(x_0) \cdot [F''(x_0)y_0^2]^2}{4[F'(x_0)y_0]^2 \cdot \left[F'(x_0)y_0 - \frac{1}{2} \frac{F''(x_0)y_0^2}{F'(x_0)y_0} F(x_0) \right]} \leq \frac{1}{4} \frac{B_0 M^2 \eta_0^3}{\left(1 - \frac{1}{2}h_0\right)^2}$$

so

(5) $$|F(x_1)| \leq \frac{1}{4} \frac{B_0 M^2}{\left(1 - \frac{1}{2}h_0\right)^2} \eta_0^3 + \frac{1}{6} \frac{N}{\left(1 - \frac{1}{2}h_0\right)^3} \eta_0^3.$$

Using (4) and the condition 5.),:

$$\frac{\|F(x_1)\|}{|F'(x_1)|} \leq \frac{h_0^2 \eta_0}{2\left(1 - \frac{1}{2}h_0\right) \cdot \left(1 - \frac{3}{2}h_0\right)} \cdot \left[\frac{1}{2} + \frac{N}{3B_0 M^2 \left(1 - \frac{1}{2}h_0\right)} \right] \leq \frac{\eta_0}{4}$$

so condition 2.), is valid for x_1 too and

(6) $$\eta_1 \le h_0^2 \eta_0 \le \frac{\eta_0}{4}.$$

c) Later we shall prove $\{x_1, x_2, \dots, x_n, \dots\} \in D$ where $\{x_n\}$ is given by the method (2), so the condition 3.) is valid for x_1 too.

d) Because of estimation (4) and relations (6)

$$h_1 := B_1 M \eta_1 < 4B_0 M \frac{\eta_0}{4} = h_0$$

$$h_0 \le \frac{1}{2} \quad \text{so} \quad h_1 \le 4\, h_0^3.$$

e) The condition 5.) is valid for x_1 too because of

$$h_1 < h_0, \quad B_1 > 0 \quad \text{so} \quad \sigma_1 < \sigma_0 \quad (\le 2).$$

So the conditions 1.)—5.) are valid for x_1, too.

The following inequalitites can be proved by induction:

$$B_n \le 4B_{n-1}$$

where

$$B_n := B_{n-1} \frac{2 - h_{n-1}}{2 - 3h_{n-1}},$$

$$\eta_n \le h_{n-1}^2 \eta_{n-1},$$

where

$$\eta_n := \frac{|F(x_n)|}{\|F'(x_n)\|}, \quad h_n := B_n M \eta_n;$$

$$\sigma_n \le \sigma_{n-1} \le 2,$$

where

$$\sigma_n := \frac{\dfrac{1}{2} + \dfrac{N}{3B_n M\left(1 - \dfrac{1}{2} h_n\right)}}{\left(1 - \dfrac{1}{2} h_n\right)\left(1 - \dfrac{3}{2} h_n\right)}.$$

From that

$$\eta_n \le h_{n-1}^2 h_{n-2}^2 \dots h_0^2 \eta_0,$$

but

$$h_n \le 4h_{n-1}^3 \le \frac{1}{2}(2h_{n-1})^3,$$

$$h_{n-1} \le \frac{1}{2}(2h_{n-2})^3,$$

that is why

$$h_n \le \frac{1}{2}(2h_0)^{3^n}$$

so

$$\eta_n \le \frac{(2h_0)^{3^n-1}}{2^{2n}}\eta_0.$$

Because of the induction and the (2)

$$\|x_{n+1}-x_n\| \le \frac{\eta_n}{1-\frac{1}{2}h_n},$$

$$\|x_{n+k}-x_n\| \le \|x_{n+k}-x_{n+k-1}\| + \ldots + \|x_{n+1}-x_n\| \le$$

$$\le \frac{\eta_0}{1-\frac{1}{2}h_0} \cdot \frac{(2h_0)^{3^n-1}}{2^{2n}}\left(1-\frac{1}{2^k}\right)2^2.$$

Thus the sequence $\{x_n\}$ represents a Cauchy-sequence and the space X being complete, there exists the limit $x^* := \lim\limits_{n\to\infty} x_n$ and the estimation

$$\|x^*-x_n\| \le \frac{2^2\eta_0}{2^{2n}\left(1-\frac{1}{2}h_0\right)} \cdot (2h_0)^{3^n-1}.$$

Now we prove that $\{x_n\} \subset D \quad n = 0, 1, 2, \ldots$

$$\|x_0-x_n\| \le \|x_0-x_1\| + \|x_1-x_2\| + \ldots + \|x_{n-1}-x_n\| \le$$

$$\le \frac{1}{1-\frac{1}{2}h_0}(\eta_0+\eta_1+\ldots+\eta_n) \le$$

$$\le \frac{1}{1-\frac{1}{2}h_0}\left(\eta_0+\frac{\eta_0}{4}+\frac{\eta_0}{4^2}+\ldots+\frac{\eta_0}{4^n}\right) \le \frac{16}{9}\eta_0 =: \varrho$$

and $\qquad \|x_0-\dot{x}\| \le \varrho \quad$ too.

The limit x^* satisfies the $F(x) = 0$, because using the inequality (5)

$$|F(x_{n+1})| \le \frac{1}{4}\frac{B_n M^2 \eta_n^3}{\left(1-\frac{1}{2}h_n\right)^2} + \frac{1}{6}\frac{N\eta_n^3}{\left(1-\frac{1}{2}h_n\right)^3} \le$$

$$\le \left(\frac{4}{3}\right)^3 \frac{h_n^2\eta_n}{B_n}\left[1+\frac{N}{6B_nM^2}\right] \le C\cdot\frac{\eta_0}{4^n}$$

where C = constant. If $n\to\infty$ then $|F(x_{n+1})| \to 0$.

$F(x)$ being continuous so

$$\lim_{n\to\infty} |F(x_n)| = |F(x^*)| = 0 \quad \text{so} \quad F(x^*) = 0. \qquad \square$$

A similar theorem was established in [1] for the cases of operator equations.

The following theorem can be proved similarly, using $\dfrac{F(x)}{\|F'(x)\|}$ instead of $F(x)$.

Theorem 2. *Let us assume that the following conditions are satisfied:*

1.) $$F'(x_0)y_0 \neq 0, \quad \|y_0\| = 1$$

2.) $$\|x_1 - x_0\| \leq \eta_0;$$

3.) $$\frac{\|F''(x)\|}{\|F'(x_0)\|} \leq M \quad \text{and} \quad \frac{\|F'''(x)\|}{\|F'(x_0)\|} \leq N \quad \text{if} \quad x \in D$$

where $D := \{x : \|x - x_0\| \leq 2\eta_0\}$;

4.) $$h_0 := M\eta_0 \leq \frac{1}{2};$$

5.) $$\sigma_0 = \frac{5}{4 - 2h_0} + \frac{5N}{6M^2} \leq 4;$$

(*where* x_0 *is a starting approximation*).

Then there exists $x^* \in D$, $\lim_{n\to\infty} x_n = x^*$, $F(x^*) = 0$ *and*

$$\|x_n - x^*\| \leq 2^{1-n}(2h_0)^{3^n - 1}\eta_0 \qquad \square$$

Application. Consider the non-linear operator equation $P(x) = 0$ $x \in X$ and the equivalent functional equation $F(x) = 0$ choosing the $F(x)$ in the following way

$$F(x) := \|P(x)\|^2.$$

Putting $Q(x) := \bar{P}'(x)P(x)$ where $\bar{P}'(x)$ is the adjoint of $P'(x)$ Fréchet differential, we can be given y_n in the form [1], [2]

$$y_n := \frac{Q(x_n)}{\|Q(x_n)\|} \qquad n = 0, 1, 2, \dots$$

$$F'(x_n)y_n = 2 \cdot \|Q(x_n)\| = \|F'(x_n)\|,$$

$$F''(x_n)y_n^2 = \frac{2}{\|Q(x_n)\|^2}\langle Q'(x_n)\, Q(x_n),\, Q(x_n)\rangle$$

where

$\langle , \rangle$ is a scalar product.

So (2) is in the following form:

$$(7)\qquad x_{n+1} = x_n - \frac{\|P(x_n)\|^2}{2\|Q(x_n)\|^2 - \frac{1}{2}\frac{\|P(x_n)\|^2}{\|Q(x_n)\|^2}\langle Q'(x_n)Q(x_n), Q(x_n)\rangle} Q(x_n)$$

$(n = 0, 1, 2, \ldots)$.

Illustrative example. We apply (7) for Altman's example [3]. Consider the system of equations

$$f :\equiv u^2+u+v+1{,}5 = 0$$
$$g :\equiv u^2+u-v-1 \quad = 0$$

Put $\quad x := (u, v),\ x_0 := (-0{,}5, -1),\ P(x) := (f(u, v), g(u, v))$

$$Q(x) := \overline{P'}(x)\,P(x) = \begin{pmatrix} 2u-1 & 2u+1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} u^2+u+v+1{,}5 \\ u^2+u-v-1 \end{pmatrix} =$$

$$= \begin{pmatrix} 4u^3+6u^2+3u+0{,}5 \\ 2v+2{,}5 \end{pmatrix}.$$

Apply theorem 2.:

$$\|Q(x_0)\| = 0{,}5$$

$$\|x_1 - x_0\| = \frac{1}{8},\quad r = \frac{1}{4},\quad \eta_0 = \frac{1}{8}$$

$$Q'(x) = \begin{pmatrix} 12u^2+12u+3 & 0 \\ 0 & 2 \end{pmatrix},\quad \|Q'(x)\| \le 2 \quad x \in D$$

$$\|F''(x)\| \le 4 \quad x \in D \quad M = 4,$$

$$Q''(x) = \begin{bmatrix} \begin{pmatrix} 24u+12 & 0 \\ 0 & 0 \end{pmatrix} \\ \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{bmatrix}$$

$$\|Q''(x)\| \le 6,\quad N = 12,\quad h_0 = 0{,}5,\quad \sigma_0 = \frac{55}{24} < 4.$$

Theorem 2 can be applied. In our example $h_0 = 0{,}5$ and method (2) produces the solution with 8 digits after 14 iteration steps, while Altman's method [3] needs 22 steps. Although because of the second term in the denominator method (2) uses some more operations than the Altman method, in computation time method (2) is shorter by more than 25%. If the functional $F(x)$ is chosen that $h_0 < \frac{1}{2}$ then the stated order of convergence is 3.

If we know nothing about the roots of the operator equation $P(x) = 0$ how shall we find an initial approximation x_0?

The set of numbers used by any computer can be given by

$$T := [-10^t, -10^{-t}] \cup \{0\} \cup [10^{-t}, 10^t]$$

where t is a characterizing integer.

As T is not connected at zero, we look for a suitable x_0 only in $[10^{-t}, 10^t]^k$ (k is the number of components of x_0) and using simple mirror transformations we can also find x_0 if it is somewhere else in T^k. Our method in $[10^{-t}, 10^t]^k$ was as follows: let L be the set $\{10^{-t}, 10^{-t+1}, \ldots, 10^0, 10^1, \ldots, 10^t\}$. First we find the minimum of $F(x)$ over L^k. If it is not good for x_0 we continue approaching the minimum of $F(x)$ over $[10^{-t}, 10^t]^k$ by using gradient method until we find an x_0 satifying the conditions (2) of the fast method.

Acknowledgement The author would like to thank Dr. Béla Jankó for his help in the elaboration of the presented theorems.

REFERENCES

[1] *Jankó B.*, The Solving of Nonlinear Operator Equations in Banach Spaces (in Roumanian). Publishing House of the Roumanian Academy, Bucharest 1969.

[2] *Altman M.*, On the Approximate Solution of Non-Linear Functional Equations in Banach Spaces. Bull. Acad. Polon. **5** (1958).

[3] *Altman M.*, An Iterative Method of Solving Functional Equations, Bull. Acad. Polon, **9** (2) (1961).