Nonlinear Functional Analysis and Applications
Vol. 18, No. 3 (2013), pp. 411-420

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm
Copyright © 2013 Kyungnam University Press

AN EXTENSION OF A THEOREM BY B.T. POLYAK ON GRADIENT-TYPE METHODS

Ioannis K. Argyros[1] **and Santhosh George**[2]

[1]Department of Mathematicsal Sciences
Cameron University, Lawton, OK 73505, USA
e-mail: `ioannisa@cameron.edu`

[2]Department of Mathematical and Computational Sciences
National Institute of Technology Karnataka, India
e-mail: `sgeorge@nitk.ac.in`

Abstract. We extend the applicability of semilocal convergence results by B.T. Polyak for solving nonlinear equations in a Banach space setting [16]. Our idea uses the center-Lipschitz instead of the Lipschitz condition in the convergence analysis of the iterative procedure involved. Numerical examples are presented to show that the earlier results cannot apply but the new results can apply to solve equations.

1. Introduction

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

$$F(x) = 0, \tag{1.1}$$

where F is a Fréchet-differentiable operator defined on a nonempty convex subset of a Banach space X with values in a Banach space Y.

Many problems in Applied Sciences can be brought in the form of equation (1.1) using mathematical modelling [2], [6]-[7]. The solutions of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative. We shall study the most general one-step method

[0]Received March 20, 2013. Revised April 29, 2013.
[0]2000 Mathematics Subject Classification: 65J20, 65B05, 65G99, 65J15, 65H10, 47H17, 49M15.
[0]Keywords: Gradient methods, semilocal convergence, Banach space, Lipschitz condition, Center-Lipschitz condition, Fréchet-derivative.

for solving equation (1.1) which is defined by

$$x_{n+1} = x_n - A(x_n) \text{ for each } n = 0, 1, 2, \cdots \tag{1.2}$$

where x_0 is an initial point and $A : X \to Y$ is known. If

$$A_n = A(x_n) = a_n F'(x_n)^{-1} F(x_n) \text{ for each } n = 0, 1, 2, \cdots \tag{1.3}$$

where $\{a_n\}$ is a sequence of real numbers, then we obtain the so-called damped Newton method([2], [6], [15], [16]). If we let $a_n = 1$ in (1.3) we obtain Newton's method([1], [2]-[15], [9]-[12]).

If

$$A(x) = BF(x) \text{ for each } x \in D \tag{1.4}$$

where $B : Y \to X$ is known and linear. For $B = F'(x_0)^{-1}$ we obtain the modified Newton method [3], [6], [15], [16] whereas if $X = Y$ and $B = aI$ where a is the identity operator, we obtain the method of successive substitutions [2],[3], [6]-[16]. Finally, method (1.2) specializes to solve (1.1) in Hilbert spaces. If X, Y are Hilbert spaces and

$$A_n = a_n \Gamma(x_n) F(x_n) \tag{1.5}$$

where $\Gamma(x) = [F'(x)]^*$ ([2], [6], [15], [16]).

These iterative methods are naturally called gradient methods [2]- [16] since they are the same time methods of minimizing some auxiliary function which is used in the proof of their convergence.

The study about convergence analysis of these iterative methods is usually divided into two categories: semi-local and local convergence analysis. The semilocal convergence analysis is based upon the information around an initial point to give criteria ensuring the convergence of the iterative procedure. While the local convergence analysis is based on the information around a solution to find estimates of the radii of convergence balls.

There is a plethora of convergence results for method (1.2) and its specializations mentioned above [1]-[19].

In the present paper we are motivated by the elegant works of B.T.Polyak [16] concerning the semilocal convergence of method (1.2) and its specializations. The semilocal convergence in [16] is based on the Lipschitz condition

$$\|F'(x) - F'(y)\| \le L\|x - y\|$$

for $L > 0$ and each x and $y \in D$.

The Lipschitz condition is not satisfied for many interesting classes of operators (see the numerical examples in Section 3). We expand the applicability as well as the convergence domain of method (1.2) by replacing the Lipschitz condition by the weaker center Lipschitz condition

$$\|F'(x) - F'(x_0)\| \le L_0\|x - x_0\|$$

for $L_0 > 0$ and each $x \in D$. Notice that the Lipschitz implies the center-Lipschitz condition but not vice versa. Moreover,

$$L_0 \leq L \tag{1.6}$$

holds in general and $\frac{L}{L_0}$ can arbitrarily large (see the last example in Section 3).

The paper is organized as follows: The semilocal convergence analysis is presented in Section 2. The numerical examples are given in Section 3.

2. Semilocal Convergence

We present a general semilocal convergence result for method (1.2) followed by its specializations and their comparisons to the earlier results which can be found in [16].

Let $U(x, \rho)$ stand for an open ball centered at $x \in X$ and radius $\rho > 0$. We also denote by $\overline{U}(x, \rho)$ its closure. We present the following main semilocal convergence result for method (1.2).

Theorem 2.1. *Let $x_0 \in D$. Suppose that the following conditions hold*

(i) *$F : D \subseteq X \to Y$ is Fréchet-differentiable;*

(ii) *The Fréchet derivative $F'(x)$ satisfies the center-Lipschitz condition*

$$\|F'(x) - F'(x_0)\| \leq L_0 \|x - x_0\|$$

for each $x \in D$ and $L_0 > 0$;

(iii) *There exists $A : D \to Y$ and $\gamma \in (0, 1)$ such that*

$$\|F(x) - F'(x)A(x)\| \leq \gamma \|F(x)\|$$

for each $x \in D$;

(iv) *$\|A(x)\| \leq \lambda \|F(x)\|$ for $\lambda > 0$;*

(v) *$\|F(x_0)\| \leq \mu_0 = \frac{2(9-\gamma-4\sqrt{5-\gamma})}{\lambda^2 L_0}$; and $\overline{U}(x_0, r) \subseteq D$,*

where

$$r = \frac{2\lambda \|F(x_0)\|}{1 - \gamma - \frac{\lambda^2 L_0}{2}\|F(x_0)\| + \sqrt{(1 - \gamma - \frac{\lambda^2 L_0}{2})^2 - 8\lambda^2 L_0 \|F(x_0)\|}}. \tag{2.1}$$

Then, sequence $\{x_n\}$ generated by method (1.2) is well defined, remains in $\overline{U}(x_0, r)$ for each $n = 0, 1, 2, \cdots$ and converges to a solution $x^* \in \overline{U}(x_0, r)$ of equation $F(x) = 0$. Moreover, the following estimates hold for each $n = 0, 1, 2, \cdots$

$$\|x_{n+1} - x_n\| \leq \lambda \delta^n \|F(x_0)\| \tag{2.2}$$

and

$$\|x_n - x^*\| \leq r\delta^n, \tag{2.3}$$

where

$$\delta = \gamma + L_0\lambda\Big(2r + \frac{\lambda\|F(x_0)\|}{2}\Big). \tag{2.4}$$

Proof. It follows from (v) that r is well defined; $r > 0, \lambda\|F(x_0)\| \leq r$ and $\delta \in (0,1)$. Then, $\|x_1 - x_0\| = \|A(x_0)\| \leq \lambda\|F(x_0)\| \leq r$. Hence, we deduce that $x_1 \in \overline{U}(x_0, r)$. Let us assume that $x_k \in \overline{U}(x_0, r)$ for each $k = 0, 1, 2, \cdots, n$. Using the approximation

$$\begin{aligned} F(x_{n+1}) &= [F(x_n) - F'(x_n)A(x_n)] + \int_0^1 \{[F'(x_0) - F'(x_n + \theta(x_{n+1} - x_n))]d\theta \\ &\quad + [F'(x_n) - F'(x_0)]\}(x_{n+1} - x_n), \end{aligned}$$

conditions (iii), (iv) and (2.1) we obtain in turn

$$\begin{aligned} \|F(x_{n+1})\| &\leq \|F(x_n) - F'(x_n)A(x_n)\| + \|F'(x_n) - F'(x_0)\|\|x_{n+1} - x_n\| \\ &\quad + \int_0^1 \|F'(x_0) - F'(x_n + \theta(x_{n+1} - x_n))d\theta\|\|x_{n+1} - x_n\|^2 \\ &\leq \gamma\|F(x_n)\| + L_0[2\|x_n - x_0\| + \frac{1}{2}\|x_{n+1} - x_n\|]\|x_{n+1} - x_n\| \\ &\leq \gamma\|F(x_n)\| + L_0(2\|x_n - x_0\| + \frac{\lambda}{2}\|F(x_n)\|)\lambda\|F(x_n)\| \\ &\leq \delta_n\|F(x_n)\|, \end{aligned}$$

where $\delta_n = \gamma + \lambda L_0(2r + \frac{\lambda\|F(x_n)\|}{2})$. If $\delta_n \leq 1$, then $\delta_{n+1} \leq \delta_n \leq \cdots \leq \delta_0 = \delta < 1$, since $\|F(x_n)\| \leq \delta_n\|F(x_n)\| \leq \|F(x_n)\|$. It follows that

$$\|F(x_n)\| \leq \|F(x_0)\| \prod_{k=0}^{n-1} \delta_k \leq \delta^n\|F(x_0)\|. \tag{2.5}$$

Consequently,

$$\begin{aligned} \delta_n &\leq \gamma + 2L_0\lambda r + \frac{\lambda^2 L_0\|F(x_0)\|\delta^n}{2} \\ &= (\gamma + 2\lambda L_0 r)(1 + \frac{\lambda^2 L_0\|F(x_0)\|\delta^n}{2(\gamma + 2\lambda L_0 r)}) \\ &\leq \gamma_0 \exp\exp \alpha\delta^n \end{aligned}$$

and

$$\|F(x_n)\| \leq \|F(x_0)\|\gamma^n \exp\sum_{k=0}^{n-1} \alpha\delta^k \leq \beta\gamma^n,$$

where

$$\gamma_0 = \gamma_0(r) = \gamma + 2L_0\lambda r, \quad \alpha = \frac{\lambda^2 L_0\|F(x_0)\|}{2\gamma_0}$$

and

$$\beta = \|F(x_0)\| \exp \frac{\alpha}{1-\delta}.$$

We also have

$$\begin{aligned} \|x_{n+1} - x_n\| &= \|A(x_n)\| \leq \lambda \|F(x_n)\| \\ &\leq \lambda \delta^n \|F(x_0)\|. \end{aligned}$$

Hence, for $m \geq n$

$$\begin{aligned} \|x_m - x_n\| &\leq \sum_{k=n}^{m-1} \|x_{k+1} - x_k\| \leq \lambda \|F(x_0)\| \sum_{k=n}^{m-1} \delta^k \\ &= r(\delta^n - \delta^m). \end{aligned}$$

In particular for $n = 0$, we get that

$$\|x_m - x_0\| \leq r(1 - \delta^m) \leq r.$$

That is, sequence $\{x_n\}$ is complete in a Banach space X and as such it converges to some $x^* \in \overline{U}(x_0, r)$ (since $\overline{U}(x_0, r)$ is a closed set). Finally, by letting $n \to \infty$ in (2.5) we deduce that $F(x^*) = 0$. That completes the proof of the Theorem. □

Remark 2.2. Condition (v) of Theorem 2.1 measures the smallness of $\|F(x_0)\|$. The corresponding condition in [16] using the Lipschitz instead of the center-Lipschitz condition is given by, $\|F(x_0)\| \leq \mu = \frac{2(1-\gamma)}{\lambda^2 L}$.

We have that, $\frac{\mu_0}{\mu} = \frac{9-\gamma-4\sqrt{5-\gamma}}{1-\gamma} \frac{L}{L_0} \to \infty$ as $\frac{L}{L_0} \to \infty$. The preceding estimate shows by how many times we have extended the applicability of Theorem 1 in [16]. Note however that if the Lipschitz condition can be verified and $L_0 = L$ we shall use Theorem 1 instead of Theorem 2.1, since the former guarantees quadratic convergence whereas the latter only linear. Note also that in this case $\mu > \mu_0$.

The results obtained here are usefull even if Lipschitz condition holds when the sufficient convergence condition $\|F(x_0)\| \leq \mu$ is not satisfied but condition $\|F(x_0)\| \leq \mu_0$ is satisfied. Indeed we simply start with method (1.2) (or its specializations) until a certain finite iterate x_N (N a finite positive integer) for which $\|F(x_0)\| \leq \mu$ is satisfied with x_N now being the starting point of these methods. Then, Theorem 1 in [16] guarantees the quadratic convergence of method (1.2) for $\gamma \neq 0$. Such an approach has already been reported in [3] where modified Newton's method is used as a predictor for Newton's method.

In the rest of the paper we suppose that hypotheses (i) and (ii) of Theorem 2.1 are satisfied and $D = \overline{U}(x_0, \rho)$. Next from Theorem 2.1, we present some semilocal convergence results on concrete methods for some special choices of operator A. The proofs are analogous to the corresponding ones in [16] are abreviated.

Corollary 2.3. *Suppose that* $\|[F'(x)]^*y\| \geq \xi\|y\|$, $\xi > 0$ *for all* $x \in D$ *and* $y \in Y^*$. *Then, the following assertions hold*

(i) *If* $\xi\|F(x_0)\| < \rho$, *equation* $F(x) = 0$ *has a solution* $x^* \in \overline{U}(x_0, \xi\|F(x_0)\|)$;

(ii) *Let* $0 < \varepsilon_1 \leq a_n \leq 2 - \varepsilon_2, \varepsilon_2 > 0, a_n \neq 1$. *If* δ *and* $r \leq \rho$ *are defined by* (2.4) *and* (2.1), *respectively for all* $r = \max_{\varepsilon_1 \leq a_n \leq 2-\varepsilon_2} |1 - a_n| = \max\{1 - \varepsilon_1, 1 - \varepsilon_2\}$ *and* $\lambda = (2 - \varepsilon_2)\xi$.

Suppose that z_n *is any solution of equation* $F'(x_n)z_n = F(x_n)$ (*guaranteed to exist by the condition* $\|F'(x)^*y\| \geq \xi\|y\|$ [16]) *such that* $\|z_n\| \leq \xi\|F(x_n)\|$. *Then, the sequence* $\{x_n\}$ *generated by*

$$x_{n+1} = x_n - a_n z_n \tag{2.6}$$

converges to a solution $x^* \in \overline{U}(x_0, \rho)$; *of equation* $F(x) = 0$. *Moreover, estimates* (2.2) *and* (2.3) *hold.*

Proof. Simply verify the conditions of Theorem 2.1 for $A_n(x) = a_n(x_n - x_{n+1})$. If we assume that the operator $F'(x)$ has an inverse, then method (2.6) reduces to the so-called damped Newton method

$$x_{n+1} = x_n - a_n F'(x_n)^{-1}F(x_n) \text{ for each } n = 0, 1, 2, \cdots \tag{2.7}$$

since $z = F'(x_n)^{-1}F(x_n)$ in this case. In particular, if $a_n = 1$ for each $n = 0, 1, 2, \cdots$, then (2.7) reduces to Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \text{ for each } n = 0, 1, 2, \cdots . \tag{2.8}$$

Next, we consider another special case of method (1.2) defined by

$$x_{n+1} = x_n - BF(x_n) \text{ for each } n = 0, 1, 2, \cdots \tag{2.9}$$

where $B : Y \to X$ is a linear operator. Notice that if $B = F'(x_0)^{-1}$ we obtain the modified Newton method, whereas if $X = Y$ and $B = aI$, where I is the identity operator, we obtain the method of successive substitutions. □

Corollary 2.4. *Suppose that for each* $x \in D$

$$\|I - F'(x)B\| \leq \gamma < 1,$$

δ *and* $r \leq \rho$ *are given by* (2.4) *and* (2.1), *respectively for* $\lambda = \|B\|$. *Then, the sequence* $\{x_n\}$ *generated by* (2.9) *converges to a solution* $x^* \in \overline{U}(x_0, r)$ *of equation* $F(x) = 0$.

Corollary 2.5. *Let* X *and* Y *be Hilbert spaces and let* $M = \sup_{x \in D} \|F'(x)\|$. *Suppose that* $\|\Gamma(x)y\| \geq \xi\|y\|$, $\xi > 0$ for each $x \in D$ *and each* $y \in Y$. *Then, the following assertions hold*

(i) *If* $\rho_0 = \frac{M\|F(x_0)\|}{\xi^2} < \rho$, *then there exists a solution* $x^* \in \overline{U}(x_0, \rho_0)$ *of equation* $F(x) = 0$.

Moreover, estimates (2.2) *and* (2.3) *hold.*

Proof. We simply verify the conditions of Theorem 2.1 for $A_n(x) = BF(x)$.

Finally, we specialize method (1.2) in the case when X and Y are Hilbert spaces by defining

$$x_{n+1} = x_n - a_n \Gamma(x_n) F(x_n) \text{ for each } n = 0, 1, 2, \cdots \tag{2.10}$$

where $\Gamma(x) = [F'(x)]^*$.
(i) If $0 < \varepsilon_1 \leq a_n \leq \frac{2}{M^2} - \varepsilon_2, \varepsilon_2 > 0$ and $\delta, r \leq \rho$ are given by (2.4) and (2.1), respectively for $\gamma = \max\{1 - \varepsilon_1 \xi^2, 1 - \varepsilon_2 \xi^2\}$, then there exists a solution $x^* \in \overline{U}(x_0, r)$ of equation $F(x) = 0$. Moreover, estimates (2.2) and (2.3) hold. □

Remark 2.6. The advantages of the results after Remark 2.2 over the corresponding ones in [16] have already been explained in Remark 2.2.

3. Examples

We present three examples in this Section. In the first two we show that center-Lipschitz holds but not Lipschitz condition. Whereas in the third example we show that $\frac{L}{L_0}$ can be arbitrarily large.

Example 3.1. Let $X = Y = \mathbb{R}$, $D = [0, \infty), x_0 = 1$ and define function F on D by

$$F(x) = \frac{x^{1+\frac{1}{i}}}{1 + \frac{1}{i}} + c_1 x + c_2, \tag{3.1}$$

where c_1, c_2 are real parameters and $i > 2$ an integer. Then $F'(x) = x^{1/i} + c_1$ is not Lipschitz on D. However central Lipschitz condition holds for $L_0 = 1$.

Indeed, we have

$$\begin{aligned} \|F'(x) - F'(x_0)\| &= |x^{1/i} - x_0^{1/i}| \\ &= \frac{|x - x_0|}{x_0^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}} \end{aligned}$$

so

$$\|F'(x) - F'(x_0)\| \leq L_0 |x - x_0|.$$

Example 3.2. We consider the integral equations

$$u(s) = f(s) + \lambda \int_a^b G(s,t)u(t)^{1+1/n}dt, \quad n \in \mathbb{N}. \tag{3.2}$$

Here, f is a given continuous function satifying $f(s) > 0, s \in [a,b], \lambda$ is a real number, and the kernel G is continuous and positive in $[a,b] \times [a,b]$.

For example, when $G(s,t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$\begin{aligned} u'' &= \lambda u^{1+1/n} \\ u(a) &= f(a), u(b) = f(b). \end{aligned}$$

These type of problems have been considered in [1], [2], [6], [9]-[12].

Equation of the form (3.2) generalize equations of the form

$$u(s) = \int_a^b G(s,t)u(t)^n dt \tag{3.3}$$

studied in [1], [2], [6], [9]-[12]. Instead of (3.2) we can try to solve the equation $F(u) = 0$ where

$$F : \Omega \subseteq C[a,b] \to C[a,b], \Omega = \{u \in C[a,b] : u(s) \geq 0, s \in [a,b]\},$$

and

$$F(u)(s) = u(s) - f(s) - \lambda \int_a^b G(s,t)u(t)^{1+1/n}dt.$$

The norm we consider is the max-norm.

The derivative F' is given by

$$F'(u)v(s) = v(s) - \lambda(1 + \frac{1}{n}) \int_a^b G(s,t)u(t)^{1/n}v(t)dt, \quad v \in \Omega.$$

First of all, we notice that F' does not satisfy a Lipschitz-type condition in Ω. Let us consider, for instance, $[a,b] = [0,1], G(s,t) = 1$ and $y(t) = 0$. Then $F'(y)v(s) = v(s)$ and

$$\|F'(x) - F'(y)\| = |\lambda|(1 + \frac{1}{n}) \int_a^b x(t)^{1/n}dt.$$

If F' were a Lipschitz function, then

$$\|F'(x) - F'(y)\| \leq L_1\|x - y\|,$$

or, equivalently, the inequality

$$\int_0^1 x(t)^{1/n}dt \leq L_2 \max_{x \in [0,1]} x(s), \tag{3.4}$$

would hold for all $x \in \Omega$ and for a constant L_2. But this is not true. Consider, for example, the functions

$$x_j(t) = \frac{t}{j}, \;\; j \geq 1, \;\; t \in [0,1].$$

If these are substituted into (3.4)

$$\frac{1}{j^{1/n}(1+1/n)} \leq \frac{L_2}{j} \Leftrightarrow j^{1-1/n} \leq L_2(1+1/n), \;\; \forall j \geq 1.$$

This inequality is not true when $j \to \infty$.

Therefore, condition (3.4) is not satisfied in this case. However, center-Lipschitz condition holds. To show this, let $x_0(t) = f(t)$ and $\gamma = \min_{s\in[a,b]} f(s)$, $\alpha > 0$. Then for $v \in \Omega$,

$$\begin{aligned} \|[F'(x) - F'(x_0)]v\| &= |\lambda|(1+\frac{1}{n}) \max_{s\in[a,b]} |\int_a^b G(s,t)(x(t)^{1/n} - f(t)^{1/n})v(t)dt| \\ &\leq |\lambda|(1+\frac{1}{n}) \max_{s\in[a,b]} G_n(s,t) \end{aligned}$$

where $G_n(s,t) = \frac{G(s,t)|x(t)-f(t)|}{x(t)^{(n-1)/n}+x(t)^{(n-2)/n}f(t)^{1/n}+\cdots+f(t)^{(n-1)/n}}\|v\|$. Hence,

$$\begin{aligned} \|[F'(x) - F'(x_0)]v\| &= \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}} \max_{s\in[a,b]} \int_a^b G(s,t)dt\|x - x_0\| \\ &\leq L_0\|x - x_0\|, \end{aligned}$$

where $L_0 = \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}N$ and $N = \max_{s\in[a,b]} \int_a^b G(s,t)dt$. Then center-Lipschitz condition holds for sufficiently small λ.

Example 3.3. Define the scalar function F by $F(x) = d_0x + d_1 + d_2 \sin e^{d_3x}$, $x_0 = 0$, where $d_i, i = 0,1,2,3$ are given parameters. Then, it can easily be seen that for d_3 large and d_2 sufficiently small, $\frac{L_0}{L}$ can be arbitrarily small.

References

[1] S. Amat, C. Bermidez, S. Busquier and S. Pluza, *On a third order Newton-type method free of bilinear operators*, Numerical linear Algebra with Applications, **17(4)** (2010), 639–653.

[2] I.K. Argyros, Convergence and Application of Newton-type Iterations, Springer, Berlin, 2008.

[3] I.K. Argyros, *Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point*, Rev. Anal. Numer. Theor. Approx., **36** (2007), 123–138.

[4] I.K. Argyros, *A Semilocal convergence for directional Newton methods*, Math. Comput.(AMS)., **80** (2011), 327–343.

[5] I.K. Argyros and S. Hilout, *Weaker conditions for the convergence of Newton's method*, J. complexity, **28** (2012), 364–387.
[6] I.K. Argyros, Y.J. Cho and S. Hilout, Numerical methods for equations and its applications, CRC Press, Taylor and Francis, New York, 2012.
[7] A.B. Conn, N.I.M. Govld and Ph.L. Toint, Trust region methods, SIAM Philadelphia, 2000.
[8] P. Deulfhard, Newton method for nonlinear problems: Affine invariant and adaptive algorithms, Springer, Berlin, 2004.
[9] J.A. Ezquerro, D. Gonzalez and M.A. Hernandez, *Majorizing sequences for Newton's method from initial value problems*, J. Comput. Appl. Math., **236** (2012), 2246–2258.
[10] J.A. Ezquerro and M.A. Hernandez, *Generalized differentiability conditions for Newton's method*, IMA J. Numer. Anal., **22(2)** (2002), 187–205.
[11] J.A. Ezquerro, D. Gonzalez and M.A. Hernandez, *On the local convergence of Newton's method under generalized conditions of Kantorovich*, Appl. Math. Letters, **26** (2013), 566–570.
[12] J.A. Gotierrez, *A new semilocal convergence theorem for Newton's method*, J. Comput. Appl. Meth., **70** (1997), 131–145.
[13] L.V. Kantorovich and G.P. Akilov, Functional Analysis in Normed Spaces, Pergamon Press, New York, 1964.
[14] Yu. Nesterov, Introductory lectures in convex programming, Kluwer, Boston, 2004.
[15] J.M. Ortega and W.C. Rheinboldt, Iterative solution of nonlinear equations in general variables, Academic Press, New York and London, 1970.
[16] B.T. Polyak, *Gradient methods for solving equations and inequalities*, USSR Comp. Math. Math. Phys., **4** (1994), 17–32.
[17] B.T. Polyak, *Converging of a nonlinear image of a small ball with applications to optimization*, Set-Valued Analysis, **9** (2001), 159–168.
[18] B.T. Polyak, *The convexity principle and its applications*, Bull. Braz. Math. Soc., **34** (2003), 59–75.
[19] B.T. Polyak, *Newton-Kantorovich method and its global convergence*, J. Math. Sciences, **133(4)** (2006), 1513–1523.