Extending the Convexity of Nonlinear Image of a Ball Appearing in Optimization

Ioannis K. Argyros[1*], Yeol Je Cho[2], Santhosh George[3]

[1]Department of Mathematicsal Sciences, Cameron University, Lawton, OK 73505, USA
[2]School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, P. R. China, and Department of Mathematics Education, Gyeongsang National University Jinju 52828, Korea
[3]Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka 757025, India
Email: iargyros@cameron.edu

Abstract: Let X, Y be Hilbert spaces and $F : X \rightarrow Y$ be Fréchet differentiable. Suppose that F' is center-Lipschitz on $U(w, r)$ and $F'(w)$ be a surjection. Then, $S_1 = F(U(w, \varepsilon_1))$ is convex where $\varepsilon_1 \leq r$. The set S_1 contains the corresponding set given in [18] under the Lipschitz condition. Numerical examples where the old conditions are not satisfied but the new conditions are satisfied are provided in this paper.

Keywords: convexity, Newton's method, optimization, control theory, image of a ball

1. Introduction

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

$$F(x) = y_0, \tag{1}$$

where F is a Fréchet-differentiable operator defined on a Hilbert space X with values in a Hilbert space Y and $y_0 \in Y$. We denote by F' the Fréchet derivative of operator F.

Numerical problems from Applied Sciences can be brought in the form of equation (1) using Mathematical Modelling[2,6-9,15-21]. The solutions of these equations can be found in closed form only in special cases. That is why most solution methods for these equations are iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method[1-21].

Let $U(w, r)$ stand for a closed ball centered at $w \in X$ and of radius $r > 0$.

In the present paper we are interested in expanding the applicability of an important theorem by B. T. Polyak[19] with numerous applications in optimization and control theory[6,15-21].

Theorem 1.1 [19]Suppose that the following hold:
there exists a constant $L > 0$ such that F' is Lipschitz on the ball $U(w, r)$, thus

$$\|F'(x) - F'(z)\| \leq L\|x - z\| \text{ for all } x, z \in U(w, r); \tag{2}$$

there exists $\alpha > 0$ such that

$$\|F'(w)^*y\| \geq \alpha \|y\| \text{ for all } y \in Y; \tag{3}$$

$$\varepsilon \leq R = \min\{r, \frac{\alpha}{2L}\}. \tag{4}$$

Then, the image of the ball $U(w, \varepsilon)$ under the map F is convex. That is $S = \{F(x) : x \in U(w, \varepsilon)\}$ is a convex set in Y.

There are many operators F' for which (2) does not hold (see the examples at the end of this study). Therefore Theorem 1.1 cannot apply. Moreover the ball $U(w, \varepsilon)$ is small. Our goal is twofold: On the one hand, we provide an analogous to Theorem 1.1 result to cover the case when (2) does not hold. And on the other hand we enlarge the ball $U(w, \varepsilon)$.

Notice that for x_0 and w fixed in $U(w, r)$ there exist constants L_0 and L_1 such that

$$\|F'(x) - F'(x_0)\| \le L_0 \|x - x_0\| \text{ for all } x \in U(x_0, \varepsilon_0) \tag{5}$$

with $\varepsilon_0 \le r$ and

$$\|F'(x) - F'(w)\| \le L_1 \|x - w\| \text{ for all } x \in U(w, \varepsilon_1) \tag{6}$$

with $\varepsilon_1 \le r$. Clearly,

$$L_0 \le L, \tag{7}$$

$$L_1 \le L \tag{8}$$

hold in general and $\frac{L}{L_0}, \frac{L}{L_1}$ can be arbitrarily large (see also the examples at the end of this study).

The paper is organized as follows: In Section 2 we present a result analogous to Theorem 1.1 but using (5) and (6) instead of (2). The numerical examples where (5) (or (6)) hold but not (2) are given in Section 3.

2. Main result

We need two auxiliary results.

Lemma 2.1 [19] A ball in a Hilbert space is strongly convex: If $x_1, x_2 \in U(w, \varepsilon_1)$, $x_0 = \frac{x_1 + x_2}{2}$, then, $U(x_0, \rho) \subset U(w, \varepsilon_1)$ for $\rho = \frac{\|x_1 - x_2\|^2}{8\varepsilon_1}$.

Lemma 2.2 [7] Suppose there exist $L_0, \rho, \mu > 0$ such that $\|F'(x) - F'(x_0)\| \le L_0 \|x - x_0\|$ for all $x \in U(x_0, \rho)$, $\|F'(x)^* y\| \ge \mu \|y\|$ for all $y \in Y$ and all $x \in U(x_0, \rho)$, and $\|F'(x_0) - y_0\| \le \rho \mu$. Then, the equation $F(x) = y_0$ has a solution $x^* \in U(x_0, \rho)$ and

$$\|x^* - x_0\| \le \frac{\|F(x_0) - y_0\|}{\mu}.$$

Remake 2.3 (a) Notice that Lemma 2.2 extended Corollary 1 of [18] in the case when (2) is not satisfied but (5) is satisfied. Moreover, these results were given in the case when X and Y are Banach spaces.

(b) It is worth noticing that Lemma 2.2 follows from the Grave's theorem [13] without assuming the Lipschitz continuity. It is sufficient F' be continuous at x_0 which is equivalent to strict differentiability of F at x_0. In particular, the injectivity of F' is equivalent to its surjectivity. Hence, Grave's theorem applies.

Next we present the main result.

Theorem 2.4 Suppose that (3), (5) (with $\varepsilon_0 = \rho$) and (6) hold and $\varepsilon_1 \le \min\{r, \frac{\alpha}{L_0 + L_1}\}$. Then, the image of the ball $U(w, \varepsilon_1)$ under the map F is convex. That is $S_1 = \{F(x) : x \in U(w, \varepsilon_1)\}$ is a convex set in Y.

Proof. Let x_1, x_2 be arbitrary points in $U(w, \varepsilon_1) \subset U(w, r)$, $y_i = F(x_i) \in S_1$, $i = 1, 2$. We shall show the hypotheses of Lemma 2.2 for $\rho = \frac{\|x_1 - x_2\|^2}{8\varepsilon_1}$ and $\mu = \alpha - L_1\varepsilon_1$. As in [Theorem 2.1[19]], set $x_0 = \frac{x_1 + x_2}{2}$ and $y_0 = \frac{y_1 + y_2}{2}$. We must find $x^* \in U(w, \varepsilon_1)$ such that $F(x^*) = y_0$. That will show the convexity of S_1. We have that

$$y_i = F(x_0) + F'(x_0)(x_i - x_0) + b_i, \, i = 1, 2. \tag{9}$$

Then, using (5) and (9) we get in turn that

$$\|b_i\| = \|F(x_i) - F(x_0) - F'(x_0)(x_i - x_0)\|$$

$$\le \int_0^1 \|[F'(x_0 + \theta(x_i - x_0)) - F'(x_0)](x_i - x_0)\| \, d\theta$$

$$\le \frac{L_0}{2}\|x_i - x_0\|^2.$$

That is $y_0 = F(x_0) + b_0, b_0 = \frac{b_1 + b_2}{2}$, and $\| b_0 \| \leq \frac{L_0 \| x_1 - x_2 \|^2}{8}$. Using Lemma 2.1, we have that $U(x_0, \rho) \subset U(w, \varepsilon_1)$. Moreover, by the choice of ε_1, ρ and μ

$$\| F(x_0) - y_0 \| = \| b_0 \| \leq \frac{L_0 \| x_1 - x_2 \|^2}{8} = L_0 \rho \varepsilon \leq \rho(\alpha - L_1 \varepsilon) = \rho \mu.$$

Furthermore, we have in turn by (3) and (6) that for all $x \in U(x_0, \rho)$

$$\| F'(x)^* y \| \geq \| F'(w)^* y \| - \| F'(x)^* - F'(w)^*) y \|$$

$$\geq \alpha \| y \| - L_1 \| x - w \| \, \| y \|$$

$$\geq (\alpha - L_1 \varepsilon_1) \| y \| = \mu \| y \|.$$

Hence, all conditions of Theorem 1.1 are satisfied. That completes the proof of the theorem.

Remark 2.5 (a) It follows from the definition of R and R_1 that

$$R \leq R_1. \tag{10}$$

If $L_0 = L_1 = L$, then $R = R_1$. Moreover, if strict inequality holds in (7) or (8), so does in (10). Furthermore, if $R = \frac{\alpha}{2L}$ and $R_1 = \frac{\alpha}{L_0 + L_1}$, then,

$$\frac{R}{R_1} = \frac{1}{2}\left(\frac{L_0}{L} + \frac{L_1}{L}\right) \to 0 \text{ as } \frac{L_0}{L} \text{ and } \frac{L_1}{L_0} \to 0. \tag{11}$$

Estimate (11) shows by how many times (at most) the ball is enlarged under our approach.

(b) As already noted in Remark 2.3 (b) the Lipschitz continuity is not needed in Theorem 2.4. But we are also interested in providing a computable ball based on Lipschitz constants. That is why we use Lemma 2.2 and the Lipschitz conditions and not Grove's theorem (that does not provide a computable ball).

In the next section, we present some examples in a Banach space setting where (2) is not satisfied but (5) (or (6)) is satisfied. Other concrete applications can be also found in [6, 15-21] and the references therin.

3. Examples

We present three examples in this Section. In the first two we show that the center-Lipschitz holds but not the Lipschitz condition. Whereas in the third example we show that $\frac{L}{L_0}$ can be arbitrarily large. Our examples are presented in a Banach space setting which can certainly be specialized in a Hilbert space one.

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

$$F(x) = \frac{x^{1+\frac{1}{i}}}{1 + \frac{1}{i}} + c_1 x + c_2, \tag{12}$$

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

Indeed, we have

$$\left\| F'(x) - F'(x_0) \right\| = \left| x^{1/i} - x_0^{1/i} \right|$$

$$= \frac{|x - x_0|}{x_0^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}}$$

so

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

Example 3.2 We consider the integral equation

$$u(s) = f(s) + \lambda \int_a^b G(s,t) u(t)^{1+1/n} \, dt, \; n \in \mathbb{N}. \tag{13}$$

Here, f is a given continuous function satisfying $f(s) > 0$; $s \in [a, b]$, λ 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's kernel, the corresponding integral equation is equivalent to the boundary value problem

$$u'' = \lambda u^{1+1/n}$$

$$u(a) = f(a), \; u(b) = f(b).$$

These type of problems have been considered in [1-2, 6, 9-15].
Equation of the form (13) generalizes equations like

$$u(s) = \int_a^b G(s,t) u(t)^n \, dt \tag{14}$$

studied in [1-2, 6, 9-14]. Instead of (13) we can try to solve the equation $F(u) = 0$ where

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

and

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

We consider is the max-norm.
The derivative F' is given by

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

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, we have $F'(y)v(s) = v(s)$ and

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

If F' were a Lipschitz function, then

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

or, equivalently, the inequality

$$\int_0^1 x(t)^{1/n}\,dt \le L_2 \max_{x\in[0,1]} x(s), \tag{15}$$

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

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

If these are substituted into (15)

$$\frac{1}{j^{1/n}(1+1/n)} \le \frac{L_2}{j} \Leftrightarrow j^{1-1/n} \le L_2(1+1/n),\ \forall j \ge 1.$$

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

Therefore, condition (15) is not satisfied in this case. However, the 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 $v \in \Omega$, we get

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

$$\le |\lambda|\left(1+\frac{1}{n}\right)\max_{s\in[a,b]} G_n(s,t),$$

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, we obtain

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

$$\le L_0 \|x - x_0\|,$$

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, the center-Lipschitz condition holds for suffciently small λ.

Example 3.3 Define the scalar function F by $F(x) = d_0 x + d_1 + d_2 \sin e^{d_3 x}$, $x_0 = 0$, 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 suffciently small, $\frac{L_0}{L}$ can be arbitrarily small.

References

[1] S. Amat, C. Bermudez, S. Busquier, S. Plaza. On a third order Newton-type method free of bilinear operators. *Numerical linear Algebra with Applications*. 2010; 17(4): 639-653.
[2] I. K. Argyros. *Convergence and Application of Newton-type Iterations*. Berlin: Springer; 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*. 2007; 36: 123-138.
[4] I. K. Argyros. A Semilocal convergence for directional Newton methods. *Math. Comput.(AMS)*. 2011; 80: 327-343.
[5] I. K. Argyros, S. Hilout Weaker conditions for the convergence of Newton's method. *J. Complexity*. 2012; 28: 364-387.
[6] I. K. Argyros, Y. J. Cho, S. *Hilout Numerical methods for equations and its applications*. Taylor and Francis, New York: CRC Publishing; 2012.
[7] I. K. Argyros, S. George. An extension of a theorem by B. T. Polyak on gradient-type methods.
[8] A. B. Conn, N. I. M. Gould, Ph. L. Toint. *Trust region methods*. Philadelphia: SIAM; 2000.
[9] P. Deulfhard. *Newton method for nonlinear problems: Affine invariant and adaptive algorithms*. Berlin: Springer;

2004.

[10] J. A. Ezquerro, D. Gonzalez, M. A. Hernandez. Majorizing sequences for Newton's method from initial value problems. *J. Comput. Appl. Math*. 2012; 236: 2246-2258.

[11] J. A. Ezquerro, M. A. Hernandez. Generalized differentiability conditions for Newton's method. *IMA J. Numer. Anal*. 2002; 22(2): 187-205.

[12] J. A. Ezquerro, D. Gonzalez, M. A. Hernandez. On the local convergence of Newton's method under generalized conditions of Kantorovich. *Appl. Math. Letters*. 2013; 26: 566-570.

[13] L. M. Graves. Some mapping theorems. *Duke Mathematical Journal*. 1950; 17: 111-114.

[14] J. A. Gutierrez. A new semilocal convergence theorem for Newton's method. *J. Comput. Appl. Meth*. 1997; 70: 131-145.

[15] L. V. Kantorovich, G. P. Akilov. *Functional Analysis in Normed Spaces*. New York: Pergamon Publishing; 1964.

[16] Yu. Nesterov. *Introductory lectures in convex programming*. Boston: Kluwer; 2004.

[17] J. M. Ortega, W. C. Rheinboldt. *Iterative solution of nonlinear equations in general variables*. New York and London: Academic Publishing; 1970.

[18] B. T. Polyak. Gradient methods for solving equations and inequalities. *USSR Comp. Math. Math. Phys*. 1994; 4: 17-32.

[19] B. T. Polyak. Converging of a nonlinear image of a small ball with applications to optimization. *Set-Valued Analysis*. 2001; 9: 159-168.

[20] B. T. Polyak. The convexity principle and its applications. *Bull. Braz. Math. Soc*. 2003; 34: 59-75.

[21] B. T. Polyak. Newton-Kantorovich method and its global convergence. *J. Math. Sciences*. 2006; 133(4): 1513-1523.