Commun. Optim. Theory 2018 (2018), Article ID 3 https://doi.org/10.23952/cot.2018.3

Communications in Optimization Theory
Available online at http://cot.mathres.org

EXTENDED OPTIMALITY OF THE SECANT METHOD ON BANACH SPACES

IOANNIS K. ARGYROS[1], SANTHOSH GEORGE[2]

[1]Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA
[2]Department of Mathematical and Computational Sciences, NIT Karnataka, 575 025, India

Abstract. The aim of this paper is to extend the applicability of the secant-type method investigated in Galperin (2014) based on restricted convergence domain. Optimality of the secant methods are also presented in this study.

Keywords. Secant method; Banach space; Local convergence.

2010 Mathematics Subject Classification. 65H10, 65J15.

1. INTRODUCTION

Let $\mathbb{B}_1$ and $\mathbb{B}_2$ be two Banach spaces. Let Ω be an open and convex subset of $\mathbb{B}_1$. Secant-type methods are studied for solving the operator equation

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

where $F : \Omega \longrightarrow \mathbb{B}_2$ is nonlinear operator. Due to its wide applications, finding solutions to (1.1) is an important problem in applied mathematics. Most of the solution methods for (1.1) are iterative. So the optimality of iterative methods for solving nonlinear equations is an issue. In [2], Maistrovsky introduced the concept of the optimality for iterative methods. Using this concept, he has proved that no method using the same information at each iteration as the Newton's method is convergent faster than the Newton's method. In [3], Galperin introduced another concept of the optimality of an iterative method using the concept of the entropy of solution's position within a set of its guaranteed the existence and uniqueness. In this concept, the merit of the method is measured by how much this entropy is reduced via one iteration. Using this concept, Galperin designed a new iterative method for nonlinear equations with regularly smooth functions in [3]. He further extended the ideas in [2] to study the secant-type method defined by

$$x_{n+1} = x_n - [x_n, x_{n-1}; F]^{-1} F(x_n), \; n = 0, 1, 2, \ldots, \tag{1.2}$$

*Corresponding author.
E-mail addresses: iargyros@cameron.edu (I.K. Argyros), sgeorge@nitk.ac.in (S. George).
Received June 19, 2017; Accepted December 20, 2017.

where x_0, x_{-1} are some initial guess and $[.,.;F]$ is the divided difference of order one defined by

$$F(x)-F(y)=[x,y;F](x-y).$$

Galperin, in [3], proved an existence and uniqueness theorem for method (1.2), when the divided difference of F are Lipschitz continuous on Ω. Using the idea of restricted convergence domain, we extend the applicability of method (1.2) considered in [3].

The paper is structured as follows. In Section 2, we present the local convergence analysis. In Section 3, the optimality of the secant method is considered.

2. Local Convergence

Suppose that the divided difference $[x,y;F]^{-1} \in \mathscr{L}(\mathbb{B}_2,\mathbb{B}_1)$, the space of bounded linear operators from $\mathbb{B}_2$ to $\mathbb{B}_1$. Set $F_0=[x_),x_{-1};F]^{-1}F$ to normalize F. Without loss of generality, we assume that F is already normalized, i.e., $F:\Omega \longrightarrow \mathbb{B}_1, [x_0,x_{-1};F]=I$. Let $\gamma_{-1},\gamma_0,\gamma,\delta_{-1},\delta_0$ and δ be positive parameters. It is helpful for the analysis that follows to define functions $\alpha,\beta,\varphi,\psi,\psi_1$ and h on the interval $[0,+\infty)$ by

$$\begin{aligned}
\alpha(t) &:= \|F(x_0)\|t,\ \beta(t):=\|x_0-x_{-1}\|t, &(2.1)\\
\varphi(t) &:= \frac{2\alpha(t)}{\sqrt{(1+\beta(t))^2+4\alpha(t)}+1+\beta(t)}, &(2.2)\\
\psi(t) &:= \frac{2\alpha(t)}{1-\beta(t)+\sqrt{(1-\beta(t))^2-4\alpha(t)}}, &(2.3)\\
\psi_1(t): &= \frac{2\alpha(t)}{1-\beta(t)-\sqrt{(1-\beta(t))^2-4\alpha(t)}},\\
h(t) &:= \beta(t)+2\sqrt{\alpha(t)}, &(2.4)
\end{aligned}$$

parameter r_0 by

$$r_0:=\frac{1-\gamma\|x_0-x_{-1}\|}{\gamma_0+\gamma}, \tag{2.5}$$

and the open ball $U(x_0,r)$ by $U(x_0,r):=\{x\in\Omega:\|x-x_0\|<r\}$. Denote by $\bar{U}(x_0,r)$ the closure of $U(x_0,r)$. Suppose that

$$\gamma\|x_0-x_{-1}\|<1. \tag{2.6}$$

Set

$$D_0=D\cap U(x_0,r_0) \tag{2.7}$$

and

$$\delta=\max\{\delta_{-1},\delta_0\}. \tag{2.8}$$

The set D_0 exists by (2.5) and (2.6). We can show the existence and uniqueness results using the preceding notation.

Theorem 2.1. *Suppose that there exist $x^*\in D,\gamma_{-1}\geq 0$ such that $F(x^*)=0$ and*

$$\|[x_0,x^*;F]-[x_0,x_{-1};F]\|\leq\gamma_{-1}\|x_{-1}-x^*\|. \tag{2.9}$$

Then, the following assertions hold

(a)
$$\|x_0 - x^*\| \geq \varphi(\gamma_{-1}). \tag{2.10}$$

(b) If $h(\gamma_{-1}) \leq 1$, then
$$\varphi(\gamma_{-1}) \leq \|x_0 - x^*\| \leq \psi(\gamma_{-1}) \tag{2.11}$$
and x^* is the unique solution of equation $F(x) = 0$ in $U(x_0, \psi_1(\gamma_{-1}))$.

(c) The bounds in (2.11) are sharp.

Proof. We can write by hypothesis $F(x^*) = 0$ and the definition of the divided difference that
$$F(x_0) = F(x_0) - F(x^*) = [x_0, x^*; F](x_0 - x^*). \tag{2.12}$$
Using (2.9), (2.12) and the triangle inequality, we have in turn that
$$\begin{aligned} |\|F(x_0)\| - \|x_0 - x^*\|| &\leq \|F(x_0) - (x_0 - x^*)\| \\ &= \|([x_0, x^*; F] - I)(x_0 - x^*)\| \\ &\leq \|([x_0, x^*; F] - [x_0, x_{-1}; F])\| \|x_0 - x^*\| \\ &\leq \gamma_{-1}\|x^* - x_{-1}\| \|x_0 - x^*\| \\ &\leq \gamma_{-1}(\|x_{-1} - x_0\| + \|x_0 - x^*\|)\|x_0 - x^*\|. \end{aligned} \tag{2.13}$$
Set $\lambda := \gamma_{-1}\|x_0 - x^*\|$. Then (2.13) can be written as
$$|\alpha(\gamma_{-1}) - \lambda| \leq \lambda(\lambda + \beta(\gamma_{-1})) \Leftrightarrow \lambda^2 + (1 + \beta(\gamma_{-1}))\lambda - \alpha(\gamma_{-1}) \geq 0 \tag{2.14}$$
and
$$\lambda^2 - (1 - \beta(\gamma_{-1}))\lambda + \alpha(\gamma_{-1}) \geq 0.$$
The first inequality in the right hand side of equivalence (2.14) shows the left hand side inequality in (2.11). The rest of the proof is obtained from the proof of Theorem 2.3 that follows [3], if we simply replace $a, b, c^{-1}v, c^{-1}w$, by $\alpha(\gamma_{-1}), \beta(\gamma_{-1}), \varphi(\gamma_{-1}), \psi(\gamma_{-1})$, respectively. □

Theorem 2.2. *Suppose that*
$$\|[x, y; F] - [x_0, x_{-1}; F]\| \leq \gamma_0\|x - x_0\| + \gamma\|y - x_{-1}\| \text{ for each } x, y \in D, \tag{2.15}$$
$$\|[x, y; F] - [u, z; F]\| \leq \delta_{-1}\|x - u\| + \delta_0\|y - z\| \text{ for each } x, y, u, z \in D_0, \tag{2.16}$$
$$h(\delta) \leq 1 \tag{2.17}$$
and (2.6) holds. Then, equation (1.1) has a solution in the set
$$A_0 = A_0(x_0, x_{-1}) := \{x \in D_0 / \varphi(\gamma_{-1}) \leq \|x - x_0\| \leq \psi(\gamma_{-1})\}. \tag{2.18}$$

Proof. It follows from (2.5), (2.6) and (2.15) that, for each $x, y \in D(r) := D \cap U(x_0, r_0), r \leq r_0$,
$$\begin{aligned} \|[x, y; F] - [x_0, x_{-1}; F]\| &\leq \gamma_0\|x - x_0\| + \gamma\|y - x_{-1}\| \\ &\leq \gamma_0\|x - x_0\| + \gamma\|y - x_0\| + \gamma\|x_0 - x_{-1}\| \\ &\leq (\gamma_0 + \gamma)r + \gamma\|x_0 - x_{-1}\| < 1. \end{aligned} \tag{2.19}$$
If follows from (2.19) and the Banach perturbation Lemma [5, 6] that $[x, y; F]$ is invertible and
$$\|[x, y; F]^{-1}\| \leq \frac{1}{1 - \gamma_0\|x - x_0\| - \gamma(\|y - x_0\| + \|x_0 - x_{-1}\|)}. \tag{2.20}$$

Therefore, the secant iterates $\{x_n\}$ lie in D_0 which is a more precise domain than D used in the literature, [1]-[12] since $D_0 \subseteq D$. The rest of the proof for the convergence of the secant method to some x_∞ follows e.g. as the proof in [3] by using the weaker (2.16) instead of (2.21) (see Theorem 2.1 below). Clearly, x_∞ lie in A_0 by Theorem 2.1 (b). □

We now state the following result from [3], so that we can compare it with ours.

Theorem 2.3. *Suppose:*

$$\|[x,y;F]-[u,z;F]\| \leq c(\|x-u\|+\|y-z\|) \text{ for each } x,y,u \in D. \tag{2.21}$$

Define

$$a := c\|F(x_0)\|,\, b := c\|x_0-x_{-1}\|,$$

$$v := \sqrt{(1+b)^2+4a}-1-b,$$

$$w := \frac{1-b-\sqrt{(1-b)^2-4a}}{2}.$$

(a) If the equation (1.1) has a solution $x^* \in D$, then $\|x_0-x^*\| \geq c^{-1}v$.

(b) If the equation (1.1) has a solution $x^* \in D$ and

$$b+2\sqrt{a} \leq 1, \tag{2.22}$$

then

$$c^{-1}v \leq \|x_0-x^*\| \leq c^{-1}w \tag{2.23}$$

and x^* is the only solution in $U(x_0, \frac{1-b-\sqrt{(1-b)^2-4a}}{2c})$.

(c) If (2.22) holds, then the equation (1.1) has a solution in the set

$$A = A(x_0,x_{-1}) := \{x \in D | / c^{-1}v \leq \|x-x_0\| \leq c^{-1}w\}. \tag{2.24}$$

(d) The bounds (2.23) are sharp.

Remark 2.4. If $\gamma_{-1} = \gamma_0 = \gamma = \delta_{-1} = \delta_0 = \delta = c$ and $U(x_0,r_0) \subseteq D$, then Theorem 2.1 and Theorem 2.2 reduce to Theorem 2.3. However, in general $\delta \leq c, \gamma_{-1} \leq \gamma \leq c, \gamma_0 \leq \gamma_{-1}$ and $\gamma \leq \delta$. Therefore, we have that

$$h(c) \leq 1 \Longrightarrow h(\delta) \leq 1, \tag{2.25}$$

$$h(c) \leq 1 \Longrightarrow h(\gamma_{-1}) \leq 1, \tag{2.26}$$

$$c^{-1}v \leq \varphi(\gamma_{-1}) \leq \|x_0-x^*\| \leq \psi(\gamma_{-1}) \leq c^{-1}w, \tag{2.27}$$

and

$$A_0 \subseteq A, \tag{2.28}$$

since $D_0 \subseteq D$, function φ is decreasing and function ψ is increasing. Hence, the applicability of the secant method is expanded. Notice also that the advantages (2.25)-(2.28) are obtained under the weaker hypotheses and less computational cost, since in practice the computation of c requires the computation of the rest of the Lipschitz constants as special cases (see also the numerical examples).

3. Optimal Secant methods

In this section, we improve the results of Section 3 in [3] along the lines of Section 2. Suppose that

$$[x_0, x_{-1}; F] := I \text{ and } h(\delta) \leq 1, \tag{3.1}$$

where $\alpha_1(t) := \|F(x_1)\|t, \beta_2(t) := \|x_1 - x_0\|t$. Using the replacements of $c^{-1}v, c^{-1}w, a, b, h(c)$ given in Section 2, we obtain the following results corresponding to Lemma 3.1 and Proposition 3.3 in [3], respectively.

Lemma 3.1. *If $\alpha(\delta) > 0, \beta(\delta) > 0$ and $h(\delta) \leq 1$, then*

(a) $a < \frac{c(\varphi(\delta)+\psi(\delta))}{2}$,

(b) $\mu(\delta) := (3 - 2\alpha(\delta))^2 - 5(1 + 4\alpha(\delta)) > 0 \Longrightarrow c\psi(\delta) < \frac{3-2\beta(\delta)-\sqrt{\mu(\delta)}}{5}$,

(c) the inequalities $s \geq 0$ and $c\varphi(\delta) \leq t \leq c\psi(\delta)$ and

$$\lambda_1(\delta) := 4(|\alpha(\delta) - \delta| - \delta(\delta + \beta(\delta)) \leq 4s \leq \lambda_2(\delta) := 4(\alpha(\delta) + \delta + \delta(\delta + \beta(\delta))$$

are consistent.

Proposition 3.2. *The optimal choice of x_1 satisfies:*

$$\delta\|x_1 - x_0\| = \begin{cases} \delta\psi(\delta), & if\ \lambda_3(\delta) \leq \delta\varphi(\delta) \vee \delta\varphi(\delta) \leq \lambda_3(\delta) \leq \delta\psi(\delta), \\ \text{and} & \lambda_4(\delta) \geq \sqrt{2 + 2(\delta\psi(\delta))^2}, \\ \delta\varphi(\delta), & if\ \lambda_3(\delta) \geq \delta\psi(\delta) \vee \delta\varphi(\delta) \leq \lambda_3(\delta) \leq \delta\psi(\delta), \\ \text{and} & \lambda_4(\delta\varphi(\delta)) \geq \sqrt{2 + 2(\delta\varphi(\delta))^2}, \end{cases}$$

where $\lambda_3(\delta) := \frac{\sqrt{(3+2\beta(\delta))^2+3(1-4\alpha(\delta))}-3-2\beta(\delta)}{3}$ *and*

$$\lambda_3(t) := \sqrt{1 + t^2 + (4\alpha(\delta) + 2(3 + 2\beta(\delta))t + 4t^2)} + \sqrt{1 + t^2 - (4\alpha(\delta) + 2(3 + 2\beta(\delta))t + 4t^2)}.$$

Remark 3.3. If $\delta = c$, the preceding results reduce to the corresponding ones in [3]. Otherwise, i.e., $\delta < c$, then our results improve the corresponding ones in [3]. Moreover, the results in this paper reduce to results for the Newton's method, if $x_{-1} = x_0$ and $\delta_{-1} = \delta_0 = \delta$, which in turn improve the results in [3]. Numerical examples, where the first inequalities in Remark 2.4 are strict can be found in [5, 6, 9] and the references therein.

References

[1] A. S. Sergeev, On the method of chords (Russian), Sibirsk. Math. J. 11 (1961), 471-484.

[2] G.D. Maistrovsky, The optimality of Newton method, Sovit Math. Dobl. 13 (1972), 838-840.

[3] A. Galperin, Optimal secant-type methods for operator equations, Numer. Funct. Anal. Optim. 32 (2014), 177-188.

[4] J. W. Schmidt, Eine Ubertragung der Regula Falsi auf Gleichungen in Banach Raumen I, II Z. Angew. Math. Mech. 43 (1963), 97-110

[5] I.K. Argyros, Computational theory of iterative methods, Series: Studies in Computational Mathematics 15, Editors, C.K. Chui and L. Wuytack, Elservier Publ. Co. New York, USA, 2007.

[6] I. K. Argyros, Convergence and applications of Newton-type iterations, Springer-Verlag, New York, 2008.

[7] S. Amat, S. Busquier, J.M. Gutiérrez, Geometric constructions of iterative functions to solve nonlinear equations, J. Comput. Appl. Math. 157 (2003), 197-205.

[8] I. K. Argyros, A.A. Magreñán, Improved local convergence analysis of the Gauss-Newton method under a majorant condition, Comput. Optim. Appl. 60 (2015), 423-439.

[9] I. K. Argyros, A.A. Magreñán, Iterative methods and their dynamics with applications, A contemporary study, CRC Press, Boca Raton Florida, 2017.
[10] J. M. Gutiérrez, M. A. Hernández, An acceleration of Newton's method: Super-Halley method, Appl. Math. Comput. 117, (2001), 223-239.
[11] A. A. Magrenãn, Different anomalies in a Jarratt family of iterative root-finding methods, Appl. Math. Comput. 233 (2014), 29-38.
[12] A.A. Magrenãn, A new tool to study real dynamics: The convergence plane, Appl. Math. Comput. 248 (2014), 215-224.