J. Nonlinear Funct. Anal. 2017 (2017), Article ID 21 https://doi.org/10.23952/jnfa.2017.21

MATHRES

Journal of Nonlinear Functional Analysis

Available online at http://jnfa.mathres.org

ON THE CONVERGENCE OF BROYDEN'S METHOD WITH REGULARITY CONTINUOUS DIVIDED DIFFERENCES AND RESTRICTED CONVERGENCE DOMAINS

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, 575025, India

Abstract. We present a semilocal convergence analysis for Broyden's method with regularly continuous divided differences in a Banach/Hilbert space setting. By using: center-Lipschitz-type conditions defining restricted convergence domains, at least as weak hypotheses and the same computational cost as in [6] we provide a new convergence analysis for Broyden's method with the following advantages: larger convergence domain; finer error bounds on the distances involved, and at least as precise information on the location of the solution.

Keywords. Banach space; Broyden's method; Majorizing sequence; Convergence domain; Semilocal convergence.

2010 Mathematics Subject Classification. 47J25, 65J15.

1. Introduction

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

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

where F is a continuously Fréchet-differentiable operator defined on a convex subset D of a Banach/Hilbert space X with values in a Hilbert space H.

*Corresponding author.

E-mail addresses: iargyros@cameron.edu (I.K. Argyros), sgeorge@nitk.ac.in (S. George).

Received December 28, 2016; Accepted February 21, 2017.

Numerous problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations [1, 2, 3, 4, 5, 6, 7]. The solution of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative.

Methods are usually studied based on: semi–local and local convergence. The semi–local convergence matter is, based on the information around an initial point, to give conditions ensuring the convergence of the iterative procedure; while the local one is, based on the information around a solution, to find a estimates of the radii of convergence balls.

Newton's method

$$x_+ := x - F'(x)^{-1}F(x), \tag{1.2}$$

is undoubtedly the most popular iterative method for generating a sequence approximating x^*. The computation of the inverse $F'(x)^{-1}$ at every step may be very expensive or impossible. That is why Broyden in [5] (for $X = H = \mathbb{R}^m$) replace the inverse Jacobian $F'(x)^{-1}$ by an $m \times m$ matrix A satisfying the equation

$$A(F(x) - F(x_-)) = x - x_-, \tag{1.3}$$

where x_- denotes the iteration preceding the current one x. This way the quasi-Newton methods were introduced [5].

We study the semilocal convergence of Broyden's method defined by

$$x_+ = x - AF(x), \quad A_+ = A - \frac{AF(x_+)\langle A^*AF(x), \cdot\rangle}{\langle A^*AF(x), F(x_+) - F(x)\rangle}, \tag{1.4}$$

where A^* is the adjoint of A and $\langle \cdot, \cdot \rangle$ is the inner product in H.

Semilocal and local convergence results for Broyden's method (1.4) and more general Broyden–like methods have already been given in the literature under Lipschitz–type conditions and for smooth operators F. Recently, in the elegant study by A. Galperin [6] the semilocal convergence of Broyden's method (1.4) was given for nonsmooth operators using the notion regularly continuous divided differences (RCDD) [1, 3, 7] (to be precised in Definition 2.1). The convergence domain found in [6] is small in general. Hence, it is important to expand this domain without adding hypotheses. This has already be done by us in [1, 2, 3, 4] for Newton's method and the Secant method using the notion of the center regularly continuous divided difference (CRCDD) which is always implied by the (RGDD) but not necessarily vice versa.

The convergence domain for such methods is small in general. In the present study, we extend the convergence domain for Newton's method and the secant method. To achieve this goal, we first introduce the center-Lipschitz condition which determines a subset of the original domain for the operator containing the iterates. The scalar functions are then related to the subset instead of the original domain. This way, the scalar functions are more precise than if they were depending on the original studies. The new technique leads to : weaker sufficient convergence conditions tighter error bounds on the distances involved and an at least as precise information on the location of the solution. These advantages are obtained under the same computational cost as in earlier studies, since in practice the new functions are special cases of the old functions. This idea can be used to study other iterative methods requiring inverses of linear mappings.

2. Semilocal convergence analysis of Broyden's method

In the rest of the study, we use the notation already established in [6].

Let $\underline{h}([x,y|F])$ denote the quantity $\inf_{x,y}\{[x,y|F] : (x,y) \in D^2\}$, and let N be the class of continuous non-decreasing concave functions $v : [0,+\infty) \to [0,+\infty)$, such that $v(0) = 0$.

We need the definition of RCDD.

Definition 2.1. [6] The $dd[x,y|F]$ is said to be ω-regularly continuous on D (ω-RCDD) if there exist an $\omega \in N$ (call it regularity modulus), and a constant $\underline{h} \in [0, \underline{h}([x,y|F])]$ such that for all $x,y,u,v \in D$

$$\begin{aligned}&\omega_1^{-1}\left(\min\{\| [x,y|F] \|, \| [u,v|F] \|\} - \underline{h} + \| [x,y|F] - [u,v|F] \|\right)\\ &- \omega_1^{-1}\left(\min\{\| [x,y|F] \|, \| [u,v|F] \|\} - \underline{h}\right) \leq \| x-u \| + \| y-v \|,\end{aligned} \tag{2.1}$$

where $\underline{h}([x,y|F]) = \inf\{\|[x,y|F]\| : (x,y) \in D^2\}$.

We also say that $dd[x,y|F]$ is regularly continuous on D, if it has a regularity modulus there.

A detailed discussion on the properties of a $dd[x,y|F]$ which is ω-regularly continuous on D is given in [6]. In the same reference a semilocal convergence analysis is provided using only condition (2.1). However, in view of condition (2.1), for $\bar{x}$ and $\bar{y}$ fixed and all $u,v \in D$, there exist $\omega_0 \in \mathbb{N}$ such that condition (2.1) holds with ω_0 replacing function ω_1. In particular, we have the following.

Definition 2.2. Let $\bar{x}, \bar{y} \in D$ be fixed. We say that $\omega_0 \in \mathbb{N}$ is a ω_0-center-regularly continuous on $D(\omega_0 - CRCDD)$, if for all $u, v \in D$

$$\begin{aligned} &\omega_0^{-1}\Big(\min\{\| [\bar{x},\bar{y}|F] \|, \| [u,v|F] \|\} - \underline{h} + \| [\bar{x},\bar{y}|F] - [u,v|F] \| \Big) \\ &- \omega_0^{-1}\Big(\min\{\| [\bar{x},\bar{y}|F] \|, \| [u,v|F] \|\} - \underline{h} \Big) \le \| \bar{x} - u \| + \| \bar{y} - v \| . \end{aligned} \tag{2.2}$$

Clearly,

$$\omega_0(s) \le \omega_1(s) \qquad \text{for all} \quad s \in [0, +\infty), \tag{2.3}$$

holds in general and $\dfrac{\omega_1(s)}{\omega_0(s)}$ can be arbitrarily large [1]-[4]. Notice also that in practice the computation of ω_1 requires the computation of ω_0 as a special case. That is (2.2) is not an additional hypothesis to (2.1).

On the other hand, because of the convexity of ω_1^{-1}, each ω_1-regularly continuous dd is also ω_1-continuous in the sense that

$$\| [x,y|F] - [u,v|F] \| \le \omega_1(\| x-u \| + \| y-v \|) \quad \text{for all} \quad x,y,u,v \in D. \tag{2.4}$$

Let $U(z,\rho), \bar{U}(z,\rho)$ stand, respectively for the open and closed balls in X with $z \in X$ and of radius $\rho > 0$. Define parameter r_0 and the set D_0, by

$$r_0 = \sup\{t \ge 0 : w_0(t) < 1\} \tag{2.5}$$

and

$$D_0 = D \cap U(x_0, r_0). \tag{2.6}$$

Definition 2.3. [6] The $dd[x,y|F]$ is said to be ω-regularly continuous on D_0 (ω-RCDD) if there exist an $\omega \in N$ and a constant $\underline{h}_0 \in [0, \underline{h}([x,y|F])]$ such that for all $x,y,u,v \in D_0$

$$\begin{aligned} &\omega^{-1}\Big(\min\{\| [x,y|F] \|, \| [u,v|F] \|\} - \underline{h}_0 + \| [x,y|F] - [u,v|F] \| \Big) \\ &- \omega^{-1}\Big(\min\{\| [x,y|F] \|, \| [u,v|F] \|\} - \underline{h}_0 \Big) \le \| x-u \| + \| y-v \|, \end{aligned} \tag{2.7}$$

where $\underline{h}_0([x,y|F]) = \inf\{\|[x,y|F]\| : (x,y) \in D_0^2\}$.

We have that

$$\omega(t) \leq \omega_1(t) \text{ for all } t \in [0, r_0), \tag{2.8}$$

since $D_0 \subseteq D$. The construction of function ω depends on function ω_0. The creation of function ω was not possible before in the studies using only function ω_1 [2, 6]. Clearly, in these studies ω can replace ω_1, since the iterates lie in D_0 related to w, which is a more precise location than D used in [6] related to ω_1 leading to the advantages as stated previously, when strict inequality holds in (2.3). From now on we also assume that

$$\omega_0(t) \leq \omega(t) \text{ for all } t \in [0, r_0). \tag{2.9}$$

Similar comments can be made for the dd $[x, y|F]$ in connection with functions ω_0 or ω.

Assume that A_0 is invertible, so that A and F in (1.4) can be replaced by their normalizations AA_0^{-1} and A_0F without affecting method (1.4). As in [6] we suppose that A and F have already been normalized:

$$A_0 = [x_0, x_{-1}|F] = I.$$

Then, the current approximation (x, A) induces the triple of reals, where

$$\bar{t} := \| x - x_0 \|, \ \bar{\gamma} := \| x - x_{-1} \| \quad \text{and} \quad \overline{\delta} := \| x_+ - x \| .$$

From now on the superscript $+$ denotes the non-negative part of real number. That is:

$$r_+ = \max\{r, 0\}.$$

We can have [6]:

$$\bar{t}_+ := \| x_+ - x_0 \| \leq \bar{t} + \overline{\delta},$$

$$\overline{\gamma}_+ := \overline{\delta},$$

and

$$\begin{aligned} \overline{\alpha}_+ &:= \omega^{-1}(\| [x_+, x|F] \| - \underline{h}) \\ &\geq \left(\omega^{-1}(1 - \underline{h}) - \| x_+ - x_0 \| - \| x - x_{-1} \| \right)^+ \\ &\geq \left(\omega^{-1}(1 - \underline{h}) - \bar{t}_+ - \bar{t} - \| x_0 - x_{-1} \| \right)^+ . \end{aligned} \tag{2.10}$$

It is also convenient for us to introduce notations:

$$\alpha := \omega_0^{-1}(1 - \underline{h}_0), \quad \overline{\gamma}_0 := \| x_0 - x_{-1} \| \quad \text{and} \quad a := \alpha - \overline{\gamma}_0. \tag{2.11}$$

We need the following result relating $\overline{\delta}_{++} = \|x_{++} - x_+\| = \|A_+ F(x_+)\|$ with $(\bar{t}, \overline{\gamma}, \overline{\delta})$.

Lemma 2.4. *Suppose that $dd[x_1, x_2|F]$ of F is ω-regularly continuous on D.Then, the $dd[x_1, x_2|F]$ of F is ω_0-regularly continuous on D at a given fixed pair (x_0, x_{-1}). If $\bar{t}_+ + \bar{t} < a$, then*

$$\overline{\delta}_+ \leq \overline{\delta} \left(\frac{\omega(a - \bar{t}_+ - \bar{t} + \overline{\delta} + \overline{\gamma} - \omega(a - \bar{t}_+ - \bar{t})}{\omega_0(a - \bar{t}_+ - \bar{t})} \right). \tag{2.12}$$

Proof. $\overline{\delta}_+ \leq \|A_+\| \|F(x_+)\|$. Using the Banach lemma on invertible operators [1, 2, 3, 4, 7] we get

$$\|A_+\|^{-1} \geq \|A_0\|^{-1} - \|A_+^{-1} - A_0^{-1}\| \geq 1 - \underline{h} - \|[x_+, x|F] - [x_0, x_{-1}|F]\|, \tag{2.13}$$

so, by (2.2), we have that

$$\begin{aligned} \|[x_+, x|F] - [x_0, x_{-1}|F]\| \quad \leq \quad & \omega_0 \left(\min \left\{ \omega_0^{-1}(\|[x_+, x|F]\| - \overline{h}_0), \omega_0^{-1}(\|[x_0, x_{-1}|F]\| - \overline{h}_)) \right\} \right) \\ & + \|x_+ - x_0\| + \|x - x_{-1}\|) \\ & - \omega_0 \left(\min\{ \omega_0^{-1}(\|[x_+, x|F]\| - \underline{h}_0), \omega_0^{-1}(\|[x_0, x_{-1}|F]\| - \underline{h}_0)\} \right). \end{aligned}$$

In view of (2.6) (for $\omega_0 = \omega$), we have

$$\begin{aligned} \omega_0(\|[x_+, x|F]\| - \underline{h}) &\geq (\omega_0^{-1}(1 - \underline{h}_0) - \|x_+ - x_0\| - \|x - x_{-1}\|)^+ \\ &\geq (\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+. \end{aligned} \tag{2.14}$$

By the concavity and monotonicity of ω_0, we have

$$\begin{aligned} \|[x_+, x|F] - [x_0, x_{-1}|F]\| &\leq \omega_0 \left(\min\{(\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+, \alpha\} + \bar{t}_+ + \bar{t} + \overline{\gamma}_0 \right) \\ &\quad - \omega_0 \left(\min\{(\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+, \alpha\} \right) \\ &= \omega_0 \left((\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+ + \bar{t}_+ + \bar{t} + \overline{\gamma}_0 \right) \\ &\quad - \omega_0 \left((\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+ \right). \end{aligned} \tag{2.15}$$

If this difference $< 1 - \underline{h}$, then it follows from (2.13) that

$$\|A_+\| \leq \frac{1}{1 - \underline{h} - \omega_0 \left((\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+ + \bar{t}_+ + \bar{t} + \overline{\gamma}_0 \right) + \omega_0 \left((\alpha - \bar{t}_+ - \bar{t} - \overline{\gamma}_0)^+ \right)}.$$

Notice that the difference (2.15) $< 1 - \underline{h} = \omega_0(\alpha)$ if and only if $\bar{t}_+ + \bar{t} < a$. Hence, this assumptions implies

$$\|A_+\| \leq \frac{1}{1 - \underline{h} - \omega_0(\alpha) + \omega_0(a - \bar{t}_+ - \bar{t})} = \frac{1}{\omega_0(a - \bar{t}_+ - \bar{t})}.$$

Then, as in [6, pages 48 and 49], we obtain

$$\|F(x_+)\| \leq \overline{\delta}\left(\omega(a-\bar{t}_+-\bar{t}_+\overline{\delta}+\overline{\gamma})-\omega(a-\bar{t}_+-\bar{t})\right), \tag{2.16}$$

which together with (2.16) show (2.12). This completes the proof.

Lemma 2.4 motivates us to introduce the following majorant generator $g(t,\gamma,\delta)=(t_+,\gamma_+,\delta_+)$:

$$t_+ := t+\delta, \quad \gamma_+ := \delta,$$

$$\delta_+ := \delta\left(\frac{\omega(a-t_+-t+\delta+\gamma)-\omega(a-t_+-t)}{\omega_0(a-t_+-t)}\right) = \delta\left(\frac{\omega(a-2t+\gamma)-\omega(a-2t-\delta)}{\omega_0(a-2t-\delta)}\right) \tag{2.17}$$

We say that the triple $q'=(t',\gamma',\delta')$ majorizes $q=(t,\gamma,\delta)$ (briefly $q \prec q'$) if

$$t \leq t' \quad \& \quad \gamma \leq \gamma' \quad \& \quad \delta \leq \delta'.$$

Lemma 2.4 states that $\overline{q}_+ \prec g(\overline{q})$. Begin fed with the initial triple q_0, the generator iterates producing a majorant sequence as long as the denominator (2.17) remains defined:

$$\underset{n}{\&} 2t_n+\delta_n < a. \tag{2.18}$$

Remark 2.5. If $\omega_0=\omega$, then the generator g reduces to the generator $\overline{g}$ given in [6] defined by $\overline{g}(u,\gamma,\delta)=(u_+,\gamma_+,\delta_+)$, $u_+=u+\theta$, $\overline{\delta}_+=\theta$,

$$\theta_+ = \theta\left(\frac{\omega(\overline{a}-u_+-u+\theta+\gamma)}{\omega(\overline{a}-u_+-u)}-1\right) = \theta\left(\frac{\omega(\overline{a}-2\theta+\gamma)}{\omega(\overline{a}-2u-\theta)}-1\right), \tag{2.19}$$

where $\overline{a}=\omega^{-1}(1-\underline{h})-\overline{\gamma}_0$. However, if strict inequality holds in (2.3), then (2.17) generates a more precise majorizing sequence than (2.19). That is

$$t<u, \tag{2.20}$$

$$\delta_+ < \theta_+ \tag{2.21}$$

and

$$t_\infty = \lim_{n\to+\infty} t_n \leq \lim_{n\to+\infty} u_n = u_\infty. \tag{2.22}$$

Here, the new semilocal convergence are weaker the error bounds are tighter and the information on the location of the solution at least as precise, if we use the generator g instead of the old generator $\overline{g}$ used in [6].

Under condition (2.18), we can ensure convergence of the sequence (x_n, A_n) generated by the method (1.4) from the starter (x_0, A_0) to a solution of the system

$$F(x) = 0 \quad \& \quad A[x,x|F] = I. \tag{2.23}$$

We present the following semilocal convergence result for method (1.4).

Theorem 2.6. *If q_0 is such that $\overline{q}_0 \prec q_0 \quad \& \quad \underset{n}{\&} 2t_n + \delta_n < a$, then sequence $\{x_n\}$ generated by method (1.4) is well defined and converges to a solution x_∞ which is the only solution of equation $F(x) = 0$ in $U(x_0, a - t_\infty)$. Moreover the following estimates hold*

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n \tag{2.24}$$

and

$$\|x_n - x_\infty\| \leq t_\infty - t_n. \tag{2.25}$$

Furthermore, the sequence $\{A_n\}$ converges to A_∞ so that (x_∞, A_∞) solve the system (2.23).

Proof. Simply replace the old generator $\overline{g}$ used in [6, see Lemma 3.2] by the new generator g defined by (2.17).

Remark 2.7.

(a) The rest of the results in [6] can be adjusted by switching the generators so we can obtain the advantages as stated in the abstract or Remark 2.5 of this study. However, we leave the details to the motivated reader. Instead, we return to Remark 2.5 and assume that ω_0, ω are linear functions defined by $\omega_0(t) = c_0 t$ and $\omega(t) = ct$ with $c_0 \neq 0$ and $c \neq 0$. Then, the generators g and $\overline{g}$ provide, respectively the scalar iterations $\{t_n\}$ and $\{u_n\}$ defined by

$$t_{-1} = \gamma_0, \quad t_0 = \delta_0, \quad t_1 = \delta_0 + \|A_0 F(x_0)\|, \quad a = c_0^{-1} - \gamma_0,$$
$$t_{n+2} = t_{n+1} + \frac{(t_{n+1} - t_n)(t_{n+1} - t_{n-1})}{a - (t_{n+1} + t_n)} \tag{2.26}$$

and

$$u_{-1} = \gamma_0, \quad u_0 = \delta_0, \quad u_1 = \delta_0 + \|A_0 F(x_0)\|, \quad \overline{a} = c^{-1} - \gamma_0,$$
$$u_{n+2} = u_{n+1} + \frac{(u_{n+1} - u_n)(u_{n+1} - u_{n-1})}{\overline{a} - (u_{n+1} + u_n)}\cdot \tag{2.27}$$

Then, we have by (2.26) and (2.27) that $t_{-1} = u_{-1}$, $t_0 = u_0$, $t_1 = u_1$ and if $c_0 = c$, then $a = \overline{a}$ and $t_n = u_n$. However, if $c_0 < c$, then a simple inductive argument shows that

$$\overline{a} < a, \tag{2.28}$$

$$t_n < u_n \quad \text{for each} \quad n = 0, 1, 2 \ldots, \tag{2.29}$$

$$t_{n+1} - t_n < u_{n+1} - u_n \quad \text{for each} \quad n = 1, 2, \ldots \tag{2.30}$$

and

$$t_\infty \leq u_\infty. \tag{2.31}$$

It was shown in [6] that the sufficient convergence condition for sequence $\{u_n\}$ is given by

$$4c^{-1}\delta_0 \leq (c^{-1} - \gamma_0)^2. \tag{2.32}$$

Therefore, according to (2.28)-(2.30), conditions (2.32) is also the sufficient convergence conditions for sequence $\{t_n\}$. Notice however that under our new approach the error (2.24) and (2.25) are tighter and by (2.31) the information on the location of the solution x_∞ is also more precise, since $t_\infty - a \leq u_\infty - a$. Moreover, a direct study of sequence $\{t_n\}$ can lead to even weaker sufficient convergence conditions [1, 2, 3, 4]. Hence, concluding the error bounds and the information on the location of the solution x_∞ are improved under weaker convergence conditions (if strict inequality holds in (2.3) since the convergence condition in [6] is given by

$$2u_n + \theta_n < \overline{a} \tag{2.33}$$

and in this case we have that

$$(2.33) \Rightarrow (2.18) \tag{2.34}$$

but not necesarirly vice versa (unless if $\omega_0 = \omega$).

(b) If $\omega(t) \leq \omega_0(t)$ for all $t \in [0, r_0)$ holds instead of (2.8), then clearly function ω_0 (still at least as small as function ω_1) can replace ω in the previous results.

(c) The results can be extended even further, if we relace D_0 by $D_0^* = U(x_1, r_0 - \|A_0F(x_0)\|)$, since x_1 is still depending on the initial data , since $D_0^* \subseteq D_0$.

Examples, where strict inequality holds in (2.3) can be found in [1, 2, 3, 4].

REFERENCES

[1] I. K. Argyros, Computational theory of iterative methods, Studies in Computational Mathematics, 15, Elseiver New York, 2007.

[2] I. K. Argyros, On the semiocal convergence of the secant method with regularly continuous divided differences, Commun. Appl. Nonlinear Anal. 19 (2012), 55-69.

[3] I. K. Argyros, S. Hilout, Computational methods in nonlinear analysis, World Scientific Publ. Comp. New Jersey, 2013.

[4] I. K. Argyros, S. Hilout, Majorizing sequences for iterative methods, J. Computat. Appl. Math. 236 (2012), 1947-1960.

[5] C. G. Broyden, A class of methods for solving nonlinear simultaneous equations, Math. Comput. 19 (1965), 577-593.

[6] A. M. Galperin, Broyden's method for operators with regularly continuous divided differences, J. Korea Math. Soc. 52 (2015), 43-65.

[7] L. V. Kantorovich, G. P. Akilov, Functional Analysis, Pergamon Press, 1982.