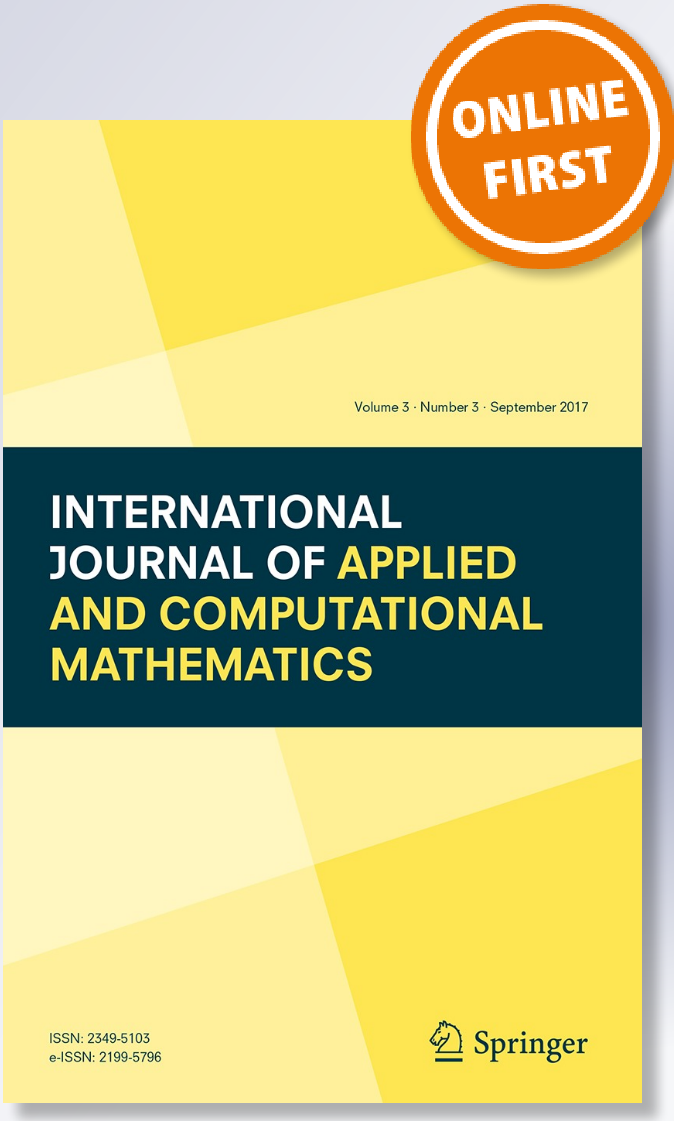
Inexact Newton's Method to Nonlinear Functions with Values in a Cone Using Restricted Convergence Domains

Ioannis K. Argyros, Santhosh George & Shobha M. Erappa

International Journal of Applied and Computational Mathematics

ISSN 2349-5103

Int. J. Appl. Comput. Math
DOI 10.1007/s40819-017-0392-7



Springer

Int. J. Appl. Comput. Math
DOI 10.1007/s40819-017-0392-7

ORIGINAL PAPER

Inexact Newton's Method to Nonlinear Functions with Values in a Cone Using Restricted Convergence Domains

Ioannis K. Argyros[1] · Santhosh George[2] · Shobha M. Erappa[3]

Abstract Using our new idea of restricted convergence domains, a robust convergence theorem for inexact Newton's method is presented to find a solution of nonlinear inclusion problems in Banach space. Using this technique, we obtain tighter majorizing functions. Consequently, we get a larger convergence domain and tighter error bounds on the distances involved. Moreover, we obtain an at least as precise information on the location of the solution than in earlier studies. Furthermore, a numerical example is presented to show that our results apply to solve problems in cases earlier studies cannot.

Keywords Inclusion problems · Inexact Newton's method · Restricted convergence domains · Semi-local convergence

Mathematics Subject Classification 40D25 · 49M15 · 26A16 · 47A05

Introduction

Let X and Y are Banach spaces, X is reflexive, $D \subseteq X$ an open set and $C \subset Y$ a nonempty closed convex cone. We consider the inexact Newton's method considered for solving the nonlinear inclusion problem

✉ Shobha M. Erappa
shobha.me@manipal.edu

Ioannis K. Argyros
iargyros@cameron.edu

Santhosh George
sgeorge@nitk.ac.in

1 Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA

2 Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, Mangalore 575025, India

3 Department of Mathematics, Manipal Institute of Technology, Manipal, Karnataka 576104, India

Published online: 28 July 2017

$$F(x) \in C, \tag{1}$$

where $F : D \to Y$ is a nonlinear continuously differentiable function. Importance of the nonlinear inclusion problems of the form (1) can be found in [1,3,6–8,13–15] and [16]. To solve (1), the following Newton-type iterative method was proposed in [17]:

$$x_{k+1} = x_k + d_k, \quad d_k \in \arg\min_{d\in X}\{\|d\| : F(x_k) + F'(x_k)d \in C\}, \quad k = 0, 1, \ldots. \tag{2}$$

In general, the above algorithm may fail to converge and may even fail to be well defined. Hence Robinson [17], made the following two assumptions to ensure that the method is well defined and converges to a solution of the nonlinear inclusion:

H1. There exists $x_0 \in X$ such that $rgeT_{x_0} = Y$, where $T_{x_0} : X \rightrightarrows Y$ is the convex process given by $T_{x_0}d := F'(x_0)dC, \quad d \in X$, and $rgeT_{x_0} = \{y \in Y : y \in T_{x_0}(x)$ for some $\in X\}$, see [8] for additional details.
H2. F' is Lipschitz continuous with modulo L, i.e., $\|F'(x) - F'(y)\| \le L\|x - y\|$, for all $x, y \in X$.

Under these assumptions, it was proved in [17], that the sequence $\{x_k\}$ generated by (2) is well defined and converges to x_* satisfying $F(x_*) \in C$, provided that the convergent criterion $\|x_1 - x_0\| \le \frac{1}{2L\|T_{x_0}^{-1}\|}$ is satisfied. Further the results in [17] are extended in [12].

The inexact Newton's method for solving nonlinear inclusion is defined by

$$x_{k+1} = x_k + d_k, \quad d_k \in \arg\min_{d\in X}\{\|d\| : F(x_k) + F'(x_k)d + r_k \in C\}, \tag{3}$$

$$\max_{w\in\{-r_k,r_k\}} \|T_{x_0}^{-1}w\| \le \theta\|T_{x_0}^{-1}[-F(x_k)]\| \tag{4}$$

for $k = 0, 1, \ldots, 0 \le \theta < 1$ is a fixed suitable tolerance, and

$$T_{x_0}^{-1}(y) := \{d \in X : F'(x_0)d - y \in C\}, \quad y \in Y,$$

where x_0 is the initial point, $\{r_k\}$ is the residual sequence in C chosen so that conditions (3) and (4) are satisfied. Moreover, $w \in \{-r_k, r_k\}$ means that w is one of these values (see also [14] and the conditions of Theorem 1). Notice that (1) is a particular case of the following generalized equation

$$F(x) + T(x) \ni 0, \tag{5}$$

when $C(x) \equiv C$ and $C : X \rightrightarrows Y$ is a set valued mapping. In [9] (see also [2]), the following Newton-type method was considered for solving (5):

$$(F(x_k) + F'(x_k)(x_{k+1} - x_k) + C(x_{k+1})) \cap R_k(x_k, x_{k+1}) \ne \emptyset, \quad k = 0, 1, \ldots, \tag{6}$$

where $R_k : X \times X \rightrightarrows Y$ is a sequence of set-value mappings with closed graphs. Note that, in the case, when $C(x) \equiv 0, \quad \theta \equiv \eta_k$ and

$$R_k(x_k, x_{k+1}) \equiv B_{\eta_k\|F(x_k)\|}(0).$$

As in the particular case $C(x) \equiv C$, the iteration (6) has (3) and (4) as a minimal norm affine invariant version.

In the present paper using our new idea of restricted convergence domains we extended the applicability of the results in [14]. That is we find a more precise location, where the iterates are located than in [14]. This way the majorant functions are tighter leading to the advantages as already stated in the abstract of this study. These advantages are obtained under the same computational cost, since the new majorant functions are special cases of the majorant functions used in [14].

The organization of the paper is as follows. In "Preliminaries" section, we give the preliminaries and in "Semilocal Convergence" section, we establish the semi-local convergence of inexact Newton's method. In "Numerical Examples" section, the advantages of the new approach are justified with a numerical example.

Preliminaries

Let $U(w, \xi)$, $\overline{U}(w, \xi)$, be the open and closed balls in X, respectively, with center $w \in X$ and of radius $\xi > 0$. A set valued mapping $T : X \rightrightarrows Y$ is called sublinear or convex process when its graph is a convex cone, i.e., $0 \in T(0), \quad T(\lambda x) = \lambda T(x), \quad \lambda > 0, \quad T(x + x') \supseteq T(x) + T(x'), x, x' \in X$, [8,18] and [19]. The domain and range of a sublinear mapping T are defined, respectively, by $domT := \{d \in X : Td \neq \emptyset\}$ and $rgeT := \{y \in Y : y \in T(x)\}$ for some $x \in X$. The norm [8] of a sublinear mapping T is defined by $\|T\| := \sup\{\|Td\| : d \in domT, \|d\| \leq 1\}$ where $\|Td\| := \inf\{\|v\| : v \in Td\}$ for $Td \neq \emptyset$. We use the convention $\|Td\| = +\infty$ for $Td = \emptyset$, $Td + \emptyset = \emptyset$ for all $d \in X$. Let $S, T : X \rightrightarrows Y$ and $U : Y \rightrightarrows Z$ be sublinear mappings. The scalar multiplication, addition and composition of sublinear mappings are sublinear mappings defined, respectively, by $(\alpha S)(x) := \alpha S(x)$, $(S+T)(x) := S(x) + T(x)$, and $UT(x) : \cup\{U(y) : y \in T(x)\}$, for all $x \in X$ and $\alpha > 0$ and the following norm properties there hold $\|\alpha S\| = |\alpha|\|S\|$, $\|S + T\| \leq \|S\| + \|T\|$ and $\|UT\| \leq \|U\|\|T\|$.

The linear map $F(x) : X \rightarrow Y$ denotes the Fréchet derivative of $F : D \rightarrow Y$ at $x \in D$. Let $C \subset Y$ be a nonempty closed convex cone, $z \in D$ and $Tz : X \Rightarrow Y$ a mapping defined as

$$T_z d := F'(z)d - C. \tag{7}$$

It is known that the mappings T_z and T_z^{-1} are sublinear with closed graph, $domT_z = X, \quad \|T_z\| < +\infty$ and, moreover, $rgeT_z = Y$ if and only if $\|T^{-1}\| < +\infty$ (see [8]). Note that $T_z^{-1}y := \{d \in X : F'(z)d - y \in C\}$, $z \in D$, $y \in Y$.

Lemma 1 (c.f., [8]) *There holds* $T_z^{-1}F'(v)T_v^{-1}w \subseteq T_z^{-1}w$, *for all* $v, z \in D$, $w \in Y$. *As a consequence,* $\|T_z^{-1}[F'(y) - F'(x)]\| \leq \|T_z^{-1}F'(v)T_v^{-1}[F'(y) - F'(x)]\|$.

Semilocal Convergence

The semilocal convergence of the inexact Newton's method is based on the hypotheses (H): Let X, Y be Banach spaces, X reflexive, $D \subseteq X$ an open set, $F : D \rightarrow Y$ a continuously Fréchet differentiable function.

(h_0) The function F satisfies the Robinson's Condition at $x_0 \in D$ if $rge \quad Tx_0 = Y$, where $Tx_0 : X \Rightarrow Y$ is a sublinear mapping as defined in (7).

(h_1) Let $R > 0$ be a scalar constant. There exist continuously differentiable majorant functions $f_0, f : [0, R) \rightarrow R$ at a point $x_0 \in D$ for F such that

$$\|T_{x_0}^{-1}[F'(y) - F'(x_0)]\| \leq f_0'(\|y - x_0\|) - f_0'(0) \tag{8}$$

for all $y \in B(x_0, R)$ and for each $x, y \in D_0 = D \cap \cup(x_0, R)$

$$\|T_{x_0}^{-1}[F'(y) - F'(x)]\| \leq f'(\|x - x_0\| + \|y - x\|) - f'(\|x - x_0\|). \tag{9}$$

(h_2) $f_0(0) > 0,\ f_0'(0) = -1,\ f(0) > 0,\ f'(0) = -1,\ f_0(t) \le f(t)$, $f_0'(t) \le f'(t)$ for $t \in (0, R)$.

(h_3) f_0' and f' are convex and strictly increasing.

(h_4) $f(\bar{t}) = 0$ for some $\bar{t} \in (0, R)$.

(h_5) $U(x_0, \bar{t}) \subseteq D$.

We also need the following condition on the majorant condition f which will be considered to hold only when explicitly stated.

(h_6) $f(t) < 0$ for $t \in (0, R)$.

Note that the condition h_6 implies the condition h_4. The sequence $\{z_k\}$ generated by inexact Newton's method for solving the inclusion $F(x) \in C$ with starting point z_0 and residual relative error tolerance θ is defined by:

$$
\begin{aligned}
&z_{k+1} := z_k + d_k,\\
&d_k \in \arg\min_{d\in X}\{\|d\| : F(z_k) + F'(z_k)d + r_k \in C\},\\
&\max_{w\in\{-r_k, r_k\}} \|T_{x_0}^{-1}\| \le \theta \|T_{x_0}^{-1}[-F(z_k)]\|\}, \quad k = 0, 1, \ldots
\end{aligned}
$$

Next we state the main result under the (H) hypotheses.

Theorem 1 *Suppose that (H) hypotheses hold. Let $C \subset Y$ a nonempty closed convex cone and $R > 0$. Suppose that $x_0 \in D$, satisfies $\|T_{x_0}^{-1}[-F(x_0)]\| \le f(0)$. Let $\beta := \sup\{-f(t) : t \in [0, R)\}$. Take $0 \le \rho < \beta/2$ and define the constants $\kappa_\rho := \sup_{\rho<t<R} \frac{-(f(t)+2\rho)}{|f_0'(\rho)|(t-\rho)}$, $\lambda_\rho := \sup\{t \in [\rho, R) : \kappa_\rho + f'(t) < 0\}$, $\tilde{\theta}_\rho := \frac{\kappa_\rho}{2-\kappa_\rho}$. Then for any $\theta \in [0, \tilde{\theta}_\rho]$ and $z_0 \in B(x_0, \rho)$, the sequence $\{z_k\}$, is well defined, for any particular choice of each d_k, $\|T_{z_0}^{-1}[-F(z_k)]\| \le \left(\frac{1+\theta^2}{2}\right)^k [f(0)+2\rho]$, $\{z_k\}$ is contained in $B(z_0, \lambda_\rho)$ and converges to a point $x_* \in B[x_0, \lambda_\rho]$ such that $F(x_*) \in C$. Moreover, if*

(h_7) $\lambda_\rho < \bar{t} - \rho$,

then the sequence $\{z_k\}$ satisfies, for $k = 0, 1, 2, \ldots$,

$$
\|z_k - z_{k+1}\| \le \alpha_k \|z_k - z_{k-1}\| \le \beta_k \|z_k - z_{k-1}\| \tag{10}
$$

where $\beta_k := \alpha_k(f', f')$, and $\alpha_k := \alpha_k(f_0', f') = \frac{1+\theta}{1-\theta}[\frac{1+\theta}{2}\frac{D^- f'(\lambda_\rho+\rho)}{|f_0'(\lambda_\rho+\rho)|}\|z_k - z_{k-1}\| + \theta\frac{2+f_0'(\lambda_\rho+\rho)}{|f_0'(\lambda_\rho+\rho)|}]$. If, additionally, $0 \le \theta < [-2(\kappa_\rho+1)+\sqrt{4(\kappa_\rho+1)^2+\kappa_\rho(4+\kappa_\rho)}]/[4+\kappa_\rho]$ then $\{z_k\}$ converges Q-linearly as follows $\lim_{k\to\infty}\sup \frac{\|x_*-z_{k+1}\|}{\|x_*-z_k\|} \le \frac{1+\theta}{1-\theta}[\frac{1+\theta}{2} + \frac{2\theta}{\kappa_\rho}]$, $k = 0, 1, \ldots$.

Remark 1 (a) The introduction of the center-Lipschitz-type majorant condition (8) (i.e, function f_0) leads to the introduction of restricted Lipschitz-type majorant condition (9). This introduction was not possible before [14], since only the condition

$$
\|T_{x_0}^{-1}[F'(y) - F'(x)]\| \le f_1'(\|x - x_0\| + \|y - x\|) - f_1'(\|x - x_0\|) \tag{11}
$$

for each $x, y \in D$ was used instead of (9). Notice, that

$$
f_0'(t) \le f_1'(t) \tag{12}
$$

and

$$
f'(t) \le f_1'(t) \tag{13}
$$

hold for each $t \in [0, R)$ since $D_0 \subseteq D$. If $f_0'(t) = f'(t) = f_1'(t)$ for each $t \in [0, R)$, then our results reduce to the corresponding ones in [14]. Otherwise (i.e, if strict inequality holds in (12) or (13)) then, we obtain advantages as already stated in the abstract of this study. Indeed notice that by (10), (12) and (13) $\alpha_k \leq \beta_k \leq \gamma_k$ where $\gamma_k = \alpha(f_1', f_1')$. Let $\bar{\bar{t}}$ be the smallest zero of function f_1 on $(0, R)$. We can assume without loss of generality that $f(t) \leq f_1(t)$. Then, we have $f(\bar{\bar{t}}) \leq f_1(\bar{\bar{t}}) = 0$ and $f(0) > 0$ so $\bar{t} \leq \bar{\bar{t}}$. It is also worth noticing that these advantages are obtained under the same computational cost as in [14], since in pratice the computation of the function f_1 requires the computation of functions f_0 and f as special cases (see also the numerical examples).

(b) If $f'(t) \leq f_0'(t)$ and $f(t) \leq f_0(t)$, then the results obtained here hold with f_0 replacing f.

(c) Clearly, our results improve the specializations of Theorem 1, [10,12], if we take $\theta = 0$ or $\theta_k = 0$, respectively.

Next we present some auxillary results needed for the proof of Theorem 1. First, we need a Banach-type perturbation result.

Lemma 2 *Let* $S = \sup\{t \in [0, R) : f'(t) < 0\}$ *and suppose* $x \in \bar{\mho}(x_0, t)$, $t \in (0, S)$. *Then the folowing hold:*

$$dom[T_x^{-1} F'(x_0)] = X, \quad rge T_x = Y,$$

$\|T_x^{-1} F'(x_0)\| \leq -\frac{1}{f_0'(t)} \leq -\frac{1}{f'(t)}$ *and* $\|T_{x_0}^{-1} F'(x)\| \leq 2 + f_0'(t)$.

Proof Use the needed (8) instead of the less precise (11) employed in ([10], Prop.12), (12) and (13). □

Remark 2 If $f_0'(t) = f'(t) = f_1'(t)$, we obtain the corresponding results in [14]. Otherwise, our results are tighter, since $-\frac{1}{f_0'(t)} \leq -\frac{1}{f'(t)} \leq -\frac{1}{f_1'(t)}$.

Proof of Theorem 1 Simply notice that the iterates z_k lie in D_0 which is a more accurate location than D used in [14]. Then, employ the proof of Theorem 2 in [14] but use f or f_0' instead of f_1 and Lemma 2 in places, where the old Lemma was used. □

Numerical Examples

Remark 3 Although the advantages of our approach have already been shown in general, we specialize operator F in such a way that f_0, f, f_1 can be defined and satisfy (h_2), (12) and (13) as strict inequalities, so that the advantages will hold, when $T_{x_0} = F'(x_0)$. Suppose that Lipschitz conditions hold: $\|F'(x_0)^{-1}(F'(x) - F'(x_0))\| \leq L_0\|x - x_0\|,\ x \in D$, $\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq L\|x - y\|,\ x, y \in D_0$, $\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq L_1\|x - y\|,\ x, y \in D$ and define functions

$$f_0(t) = \frac{L_0}{2}t^2 - t + b,$$
$$f(t) = \frac{L}{2}t^2 - t + b,$$
$$f_1(t) = \frac{L_1}{2}t^2 - t + b.$$

Let

$$\delta_1 := (1 - \sqrt{2bL_1})/(1 + \sqrt{2bL_1})$$

and

$$\delta := (1 - \sqrt{2bL})/(1 + \sqrt{2bL}).$$

Then the range for θ given in [14] is defined by $0 \le \theta \le \delta_1$. However, in our case $0 \le \theta \le \delta$ which is a better interval, if $L < L_1$.

Example 1 Let us consider the example on $X = Y = R$, $D = U(x_0, 1 - p)$, $p \in (0, 1/2)$, $x_0 = 1$ and define function F on D by $F(x) = x^3 - p$. Then, we have that $b = \frac{1}{3}(1-p)$, $L_0 = 3-p$, $L_1 = 2(2-p)$ and $L = 2(1+\frac{1}{3-p})$. Notice that $L_0 < L < L_1$, so $f_0(t) < f(t) < f_1(t)$.

(a) The results in [12,14,17] cannot be used since $2bL_1 > 1$ for all $p \in (0, 1/2)$. Notice that majorant function f_1 has no real roots and the range for θ does not exist even if $\theta = 0$. However, our results can apply, since $2bL \le 1$, for all $p \in [0.461, 0.5)$ and δ is well defined, so we can choose $\theta \in [0, \delta]$. It is worth noticing, that for $\theta = 0$, we have Newton's method, but if $\theta \in (0, \delta]$, we have the inexact Newton method. Hence, we have shown that our results can be used where as the ones in [12,14,17] cannot be used in both the Newton's and inexact Newton's case.
(b) Let us assume $p \in (0, 1)$. Choose in particular $p = 0.6$. Then both the results in [14] and our results apply but in [14], $0 \le \theta \le \delta_1 = 0.073$ and in our case $0 \le \theta \le \delta = 0.07$, so the range for θ is improved under our approach. Moreover, $\alpha_k < \beta_k < \gamma_k$. Hence, the error bounds are also improved.

Similar advantages are obtained when the majorant functions specialize to the ones given in Smale's alpha theory [20] or Wang's γ-theory [21].

Remark 4 It is worth noticing that inexact Newton methods do not include only Newton's method as a special case but also many popular single step methods such as Stirling's method or the so called Newton-like methods [1,4,5,8,11,14,19]. Therefore, the applicability of the technique introduced in this paper by far surpasses Newton's method. Clearly the benefits extend also in the case of these methods.

References

1. Argyros, I.K., Hilout, S.: Inexact Newton-type methods. J. Complex. **26**(6), 577–590 (2010)
2. Argyros, I.K., George, S., Alberto Magreñán, A.: Local convergence for multi-point-parametric Chebyshev–Halley-type methods of high convergence order. J. Comput. Appl. Math. **282**, 215–224 (2015)
3. Argyros, I.K., George, S.: Ball comparison between two optimal eight-order methods under weak conditions. SeMA Journal Boletin de la Sociedad Espanola de Matematica Aplicada (2015). doi:10.1007/s40324-015-0035-z
4. Argyros, I.K., Magreñán, A.: Iterative Methods and Their Dynamics with Applications. CRC Press, New York (2017)
5. Cibulka, R., Dontchev, A., Geoffroy, M.H.: Inexact Newton methods and Dennis More theorems for nonsmooth generalized equations. SIAM J. Control Optim. **53**(2), 1003–1019 (2015)
6. Daniel, J.W.: Newtons method for nonlinear inequalities. Numer. Math. **21**, 381–387 (1973)
7. Dedieu, J.P., Priouret, P., Malajovich, G.: Newtons method on Riemannian manifolds: convariant alpha theory. IMA J. Numer. Anal. **23**(3), 395–419 (2003)
8. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. SIAM J. Numer. Anal. **19**(2), 400–408 (1982)

9. Deuflhard, P.: Newton methods for nonlinear problems. Affine invariance and adaptive algorithms. Springer Series in Computational Mathematics, 1st edn, vol. 35. Springer, Berlin (2004)
10. Dontchev, A.L., Rockafellar, R.T.: Implicit functions and solution mappings. A view from variational analysis. Springer Monographs in Mathematics. Springer, Dordrecht (2009)
11. Dontchev, A.L., Rockafellar, R.T.: Convergence of inexact Newton methods for generalized equations. Math. Program. **139**(1–2, Ser. B), 115–137 (2013)
12. Ferreira, O.: A robust semi-local convergence analysis of Newtons method for cone inclusion problems in Banach spaces under affine invariant majorant condition. J. Comput. Appl. Math. **279**, 318–335 (2015)
13. Ferreira, O.P., Gonçalves, M.L.N., Oliveira, P.R.: Convergence of the Gauss–Newton method for convex composite optimization under a majorant condition. SIAM J. Optim. **23**(3), 1757–1783 (2013)
14. Ferreira, O.P., Silva, G.N.: Inexact Newton's method to nonlinear functions with values in a cone. arXiv:1510.01947v2 [math.NA], 21 April 2016
15. He, Y., Sun, J.: Error bounds for degenerate cone inclusion problems. Math. Oper. Res. **30**(3), 701–717 (2005)
16. Krantz, S.G., Parks, H.R.: The Implicit Function Theorem. Modern Birkhauser Classics. Birkhauser/Springer, New York, (2013). History, theory, and applications, reprint of the 2003 edition
17. Li, C., Ng, K.F.: Convergence analysis of the Gauss–Newton method for convex inclusion and convex-composite optimization problems. J. Math. Anal. Appl. **389**(1), 469–485 (2012)
18. Magreñán, A.A.: A new tool to study real dynamics. Appl. Math. Comput. **248**, 215–225 (2014)
19. Magreñán, A.A.: Improved convergence analysis for Newton-like methods. Numer. Algorithms **71**, 1–23 (2015)
20. Smale, S.: Newton's method estimates from data at one point. In: Ewing, C.R., Gross, K., Martin, C. (eds.) The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, pp. 185–196. Springer, New York (1986)
21. Wang, X.: Convergence of Newton's method and inverse function in Banach space. Math. Comput. **68**, 169–186 (1999)