Nonlinear Analysis 75 (2012) 1324–1340

Contents lists available at SciVerse ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

Generalized Newton's method based on graphical derivatives

T. Hoheisel [a], C. Kanzow [a], B.S. Mordukhovich [b,*], H. Phan [b]

[a] *Institute of Applied Mathematics and Statistics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany*

[b] *Department of Mathematics, Wayne State University, Detroit, MI 48202, USA*

ARTICLE INFO

Communicated by Ravi Agarwal

MSC:
49J53
65K15
90C30

Keywords:
Nonsmooth equations
Optimization and variational analysis
Newton's method
Graphical derivatives and coderivatives
Local and global convergence

ABSTRACT

This paper concerns developing a numerical method of the Newton type to solve systems of nonlinear equations described by nonsmooth continuous functions. We propose and justify a new generalized Newton algorithm based on graphical derivatives, which have never been used to derive a Newton-type method for solving nonsmooth equations. Based on advanced techniques of variational analysis and generalized differentiation, we establish the well-posedness of the algorithm, its local superlinear convergence, and its global convergence of the Kantorovich type. Our convergence results hold with no semismoothness and Lipschitzian assumptions, which is illustrated by examples. The algorithm and main results obtained in the paper are compared with well-recognized semismooth and B-differentiable versions of Newton's method for nonsmooth Lipschitzian equations.

1. Introduction

Newton's method is one of the most powerful and useful methods in optimization and in the related area of solving systems of nonlinear equations

$$H(x) = 0 \tag{1.1}$$

defined by continuous vector-valued mappings $H: \mathbb{R}^n \to \mathbb{R}^n$. In the classical setting when H is a continuously differentiable (smooth, C^1) mapping, Newton's method builds the following iteration procedure

$$x^{k+1} := x^k + d^k \quad \text{for all } k = 0, 1, 2, \ldots, \tag{1.2}$$

where $x^0 \in \mathbb{R}^n$ is a given starting point, and where $d^k \in \mathbb{R}^n$ is a solution to the linear system of equations (often called "Newton equation")

$$H'(x^k)d = -H(x^k). \tag{1.3}$$

A detailed analysis and numerous applications of the classical Newton's method (1.2), (1.3) and its modifications can be found, e.g., in the books [1–3] and the references therein.

However, in the vast majority of applications – including those to optimization, variational inequalities, complementarity and equilibrium problems – the underlying mapping H in (1.1) is nonsmooth. Indeed, the aforementioned optimization-related models and their extensions can be written via Robinson's formalism of "generalized equations", which in turn can

* Corresponding author. Fax: +1 313 577 7596.
E-mail addresses: hoheisel@mathematik.uni-wuerzburg.de (T. Hoheisel), kanzow@mathematik.uni-wuerzburg.de (C. Kanzow), boris@math.wayne.edu (B.S. Mordukhovich), pmhung@wayne.edu (H. Phan).

0362-546X/$ – see front matter
doi:10.1016/j.na.2011.06.039

be reduced to standard equations of the form above (using, e.g., the projection operator) but with *intrinsically nonsmooth* mappings H; see [4–7] for more details, discussions, and references.

Robinson originally proposed (see [8] and also [9] based on his earlier preprint) a point-based approximation approach to solve nonsmooth equations (1.1), which then was developed by his student Josephy [10] to extend Newton's method for solving variational inequalities and complementarity problems. Other approaches replace the classical derivative $H'(x_k)$ in the Newton equation (1.3) by some generalized derivatives. In particular, the B-differentiable Newton method developed by Pang [11,12] uses the iteration scheme (1.2) with d^k being a solution to the subproblem

$$H'(x^k; d) = -H(x^k), \tag{1.4}$$

where $H'(x^k; d)$ denotes the classical directional derivative of H at x^k in the direction d. Besides the existence of the classical directional derivative in (1.4), a number of strong assumptions are imposed in [11,12] to establish appropriate convergence results; see Section 5 below for more discussions and comparisons.

In another approach developed by Kummer [13] and Qi and Sun [14], the direction d^k in (1.2) is taken as a solution to the linear system of equations

$$A_k d = -H(x^k), \tag{1.5}$$

where A_k is an element of Clarke's generalized Jacobian $\partial_C H(x_k)$ of a Lipschitz continuous mapping H. In [15], Qi suggested to replace $A_k \in \partial_C H(x^k)$ in (1.5) by the choice of A_k from the so-called B-subdifferential $\partial_B H(x^k)$ of H at x^k, which is a proper subset of $\partial_C H(x^k)$; see Section 4 for more details. We also refer the reader to [4,16,9] and bibliographies therein for wide overviews, historical remarks, and other developments on Newton's method for nonsmooth Lipschitz equations as in (1.1) and to [17] for some recent applications.

It is proved in [14] and [15] that the Newton-type method based on implementing the generalized Jacobian and B-subdifferential in (1.5), respectively, superlinearly converges to a solution of (1.1) for a class of *semismooth* mappings H; see Section 4 for the definition and discussions. This subclass of Lipschitz continuous and directionally differentiable mappings is rather broad and useful in applications to optimization-related problems. However, not every mapping arising in applications (from both theoretical and practical viewpoints) is either directionally differentiable or Lipschitz continuous. The reader can find valuable classes of functions and mappings of this type in [18,19] and overwhelmingly in spectral function analysis, eigenvalue optimization, studying of roots of polynomials, stability of control systems, etc.; see, e.g., [20] and the references therein.

The main goal and achievements of this paper are as follows. We propose a new Newton-type algorithm to solve nonsmooth equations (1.1) described by general continuous mappings H that is based on *graphical derivatives*. It reduces to the classical Newton method (1.3) when H is smooth, being different from previously known versions of Newton's method in the case of Lipschitz continuous mappings H. Based on advanced tools of variational analysis involving *metric regularity* and *coderivatives*, we justify well-posedness of the new algorithm and its *superlinear local* and *global* (of the Kantorovich type) convergence under verifiable assumptions that hold for semismooth mappings but are *not* restricted to them. Detailed comparisons of our algorithm and results with the semismooth and B-differentiable Newton methods are given and certain improvements of these methods are justified.

Note metric regularity and related concepts of variational analysis has been employed in the analysis and justification of numerical algorithms starting with Robinson's seminal contribution; see, e.g., [21–23] and their references. However, we are not familiar with efficient applications of graphical derivatives and coderivatives for these purposes although the contingent derivative can be included in the general scheme of [16, Chapter 10] developed under a different set of assumptions with no involving metric regularity.

The rest of the paper is organized as follows. In Section 2 we present basic definitions and preliminaries from variational analysis and generalized differentiation widely used for formulations and proofs of the main results.

Section 3 is devoted to the description of the new generalized Newton algorithm with justifying its well-posedness/solvability and establishing its superlinear local and global convergence under appropriate assumptions on the underlying mapping H.

In Section 4 we compare our algorithm with the scheme of (1.5). We also discuss in detail the major assumptions made in Section 3 deriving sufficient conditions for their fulfillment and comparing them with those in the semismooth Newton methods.

Section 5 contains applications of our algorithm to the B-differentiable Newton method (1.4) with largely relaxed assumptions in comparison with known ones. In Section 6 we give some concluding remarks and discussions on further research.

Our notation is basically standard in variational analysis and numerical optimization; cf. [4,18,19]. Recall that, given a set-valued mapping $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, the expression

$$\begin{aligned} &\operatorname{Lim\,sup}_{x\to\bar{x}} F(x) := \{y \in \mathbb{R}^m \,|\, \exists\, x_k \to \bar{x} \quad \text{and } y_k \to y \text{ as } k \to \infty \text{ with} \\ &\qquad y_k \in F(x_k) \text{ for all } k \in \mathbb{N} := \{1, 2, \ldots\}\} \end{aligned} \tag{1.6}$$

defines the *Painlevé–Kuratowski upper/outer limit* of F as $x \to \bar{x}$. Let us also mention that the symbols cone Ω and co Ω stand, respectively, for the conic hull and convex hull of the set in question, that $\operatorname{dist}(x; \Omega)$ denotes the Euclidean distance