Newton's Method for a Class of Nonsmooth Functions*

STEPHEN M. ROBINSON
Department of Industrial Engineering, University of Wisconsin-Madison, 1513 University Avenue, Madison, WI 53706-1572, U. S. A.

(Received: 1 December 1993)

Abstract. This paper presents and justifies a Newton iterative process for finding zeros of functions admitting a certain type of approximation. This class includes smooth functions as well as nonsmooth reformulations of variational inequalities. We prove for this method an analogue of the fundamental local convergence theorem of Kantorovich including optimal error bounds.


Key words: Newton's method, nonlinear equations, nonsmooth functions, complementarity problems, variational inequalities, generalized equations, point-based approximation, normal map.

1. Introduction

This paper presents and justifies a Newton iterative process for finding zeros of a class of functions admitting a certain type of approximation. In particular, this class includes smooth functions, for which the Newton process is well known; it also includes a formulation of the variational inequality (generalized equation) problem in a Hilbert space. The basic result we establish here is an analogue of the fundamental local convergence theorem of Kantorovich [18, Th. 6(1.XVIII)], including the optimal error bounds established by Gragg and Tapia [7] and discussed in detail in the book of Potra and Pták [25].

The classical Newton method is very well known and widely used for finding zeros of functions having Lipschitz continuous Fréchet derivatives. For an excellent treatment of this method and many references, see the book of Ortega and Rheinboldt [21]. However, when the functions being dealt with do not possess Fréchet derivatives, it is not so clear how a Newton algorithm should be designed. In recent years many investigators have worked on this question, and several methods have

* The research reported here was sponsored by the National Science Foundation under Grants CCR-8801489 and CCR-9109345, by the Air Force Systems Command, USAF, under Grants AFOSR-88-0090 and F49620-93-1-0068, by the U. S. Army Research Office under Grant No. DAAL03-92-G-0408, and by the U. S. Army Space and Strategic Defense Command under Contract No. DASG60-91-C-0144. The U. S. Government has certain rights in this material, and is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.
been presented and justified for special cases of importance in applications. Many papers of this kind are listed in the references at the end of this paper.

In particular, as far back as 1963 Wilson [39] suggested solving nonlinear programming problems by replacing the original problem with a sequence of quadratic programming problems whose data depended on the progress of the solution. In the 1970's, Robinson [34] established a local convergence theorem explaining the quadratic convergence observed in Wilson's method, and Eaves [6] and Robinson [35] each suggested Newton-like linearization methods for solving nonlinear variational inequalities in finite-dimensional spaces. This approach was justified by Josephy [14], who proved implementability and local R-quadratic convergence for a Newton method applied to the nonlinear variational problem in $\mathbb{R}^n$. He also extended his analysis to quasi-Newton methods [15], and applied it to a particular problem in energy modeling [16, 17]. For more recent work along these lines, including computational results, see the survey paper of Harker and Pang [10].

All of these methods used conceptually similar techniques, replacing smooth functions occurring in the original problem by their linearizations, while attempting to retain the form of the problem: thus, for example, a nonlinear complementarity problem is replaced by a sequence of linear complementarity problems. These methods can also be regarded as the simplest versions of a much larger family of solution methods in which the Newton step is only approximated rather than exactly computed. This family includes the extremely popular sequential quadratic programming (SQP) methods for nonlinear programming; for a discussion of these, see Powell [26].

The methods discussed above were developed by taking a specific problem, such as a variational inequality, and attempting to apply to it the central idea (linearization) of the smooth Newton process, while retaining the basic form of the problem. In this paper we take a different approach: we try to identify what properties of an abstract problem will be needed for applying the Newton method, and we then prove that the Newton algorithm will be applicable to any problem having those properties. Thus, although the analysis here is quite different, the viewpoint taken here is close to those of Gwinner [8] and of Kummer [20].

Numerous other authors have recently investigated Newton-type methods for solving various problems with some types of nonsmoothness: see for example [2, 12, 13, 19, 22–24, 29, 30]. Also, methods of damping and other modifications have been proposed for ensuring convergence: see [9, 11, 33].

The rest of this paper consists of three sections: Section 2 discusses the concept of point-based approximation and proves some simple but useful results about solvability and uniqueness when such approximations are available. Section 3 introduces a technical device (nondiscrete induction), then states and proves the main result of the paper, a local convergence theorem of Kantorovich type for Newton's method using point-based approximations. Section 4 exhibits an important class of practical problems (variational inequalities) and shows how these can