Commun Nonlinear Sci Numer Simulat 19 (2014) 2958–2973

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

An adaptive Newton-method based on a dynamical systems approach

Mario Amrein, Thomas P. Wihler *

Mathematics Institute, University of Bern, CH-3012 Bern, Switzerland

ARTICLE INFO

Article history:
Received 19 August 2013
Received in revised form 29 January 2014
Accepted 10 February 2014
Available online 19 February 2014

Keywords:
Newton–Raphson methods
Continuous Newton–Raphson method
Adaptive step size control
Nonlinear differential equations
Chaotic behavior

ABSTRACT

The traditional Newton method for solving nonlinear operator equations in Banach spaces is discussed within the context of the continuous Newton method. This setting makes it possible to interpret the Newton method as a discrete dynamical system and thereby to cast it in the framework of an adaptive step size control procedure. In so doing, our goal is to reduce the chaotic behavior of the original method without losing its quadratic convergence property close to the roots. The performance of the modified scheme is illustrated with various examples from algebraic and differential equations.

1. Introduction

Let $X,\ Y$ be two Banach spaces, with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. Given an open subset $\Omega \subset X$, and a continuous (possibly nonlinear) operator $\mathsf{F}:\Omega \to Y$, we are interested in finding the zeros $x \in \Omega$ of F, i.e., we aim to solve the operator equation

$$x \in \Omega: \quad \mathsf{F}(x) = 0. \tag{1}$$

Supposing that the Fréchet derivative F' of F exists in Ω (or in a suitable subset), the classical Newton–Raphson method for solving (1) starts from an initial guess $x_0 \in \Omega$, and generates the iterates

$$x_{n+1} = x_n + \delta_n, \tag{2}$$

where the update $\delta_n \in X$ is implicitly given by the *linear* equation

$$\mathsf{F}'(x_n)\delta_n = -\mathsf{F}(x_n),$$

for $n \geqslant 0$. Naturally, we need to assume that $\mathsf{F}'(x_n)$ is invertible for all $n \geqslant 0$, and that $\{x_n\}_{n\geqslant 0} \subset \Omega$.

Newton's method features both local as well as global properties. On the one hand, on a *local* level, the scheme is often celebrated for its quadratic convergence regime 'sufficiently close' to a root. From a *global* perspective, on the other hand, the Newton method is well-known to exhibit chaotic behavior. Indeed, the original works of Fatou [4] and Julia [5], for instance, revealed that applying the Newton method to algebraic systems of equations may result in highly complex or even fractal attractor boundaries of the associated roots. This was confirmed in the 1980s when computer graphics were employed to illustrate the theoretical results numerically; see, e.g., [13].

* Corresponding author. Tel.: +41 316318829.
E-mail addresses: mario.amrein@math.unibe.ch (M. Amrein), wihler@math.unibe.ch (T.P. Wihler).

http://dx.doi.org/10.1016/j.cnsns.2014.02.010
1007-5704/

In order to tame the chaotic behavior of Newton's method a number of different ideas have been proposed in the literature. In particular, the use of damping aiming to avoid the appearance of possibly large updates in the iterations, constitutes a popular approach in practical applications. More precisely, (2) is replaced with

$$x_{n+1} = x_n + \alpha\delta_n,$$

for a possibly small damping parameter $0 < \alpha < 1$. More sophisticatedly, variable damping may lead to more efficient results; see, e.g., the extensive overview [1] or [2,3,16] for different variations of the classical Newton scheme. The idea of adaptively adjusting the magnitude of the Newton updates has also been studied in the recent article [14]; there, following, e.g., [9,13,15], the Newton method was identified as the numerical discretization of a specific ordinary differential equation (ODE)–the so-called continuous Newton method—by the explicit Euler scheme, with a fixed step size $h = 1$. Then, in order to tame the chaotic behavior of the Newton iterations, the idea presented in [14] is based on discretizing the continuous Newton ODE by the explicit Euler method with variable step sizes, and to combine it with a simple step size control procedure; in particular, the resulting procedure retains the optimal step size 1 whenever sensible and is able to deal with singularities in the iterations more carefully than the classical Newton scheme. In fact, numerical experiments revealed that the new method is able to generate attractors with almost smooth boundaries where the traditional Newton method produces fractal Julia sets. Moreover, the numerical tests demonstrated an improved convergence rate not matched on average by the classical Newton method.

The goal of the present paper is to continue the work in [14] on simple algebraic systems, and to extend it to the context of general Banach spaces; in particular, nonlinear boundary value problems will be focused on, and an empirical investigation demonstrating the ability of the proposed approach to tame chaos in attractor boundaries will be provided in such situations. Furthermore, in contrast to the adaptive control mechanism in [14], which is based on an intermediate step technique, we develop and test a pure prediction scheme in the present article. This will make it possible to compute the individual iterations much more efficiently. Indeed, this is most relevant in more complex applications such as in the numerical approximation of nonlinear ordinary and partial differential equations.

Finally, let us remark that there is a large application and research area where methods related to the continuous version of the Newton method are considered in the context of nonlinear optimization. Some of these schemes count among the most efficient ones available for special purposes; see, e.g., [12] and the references therein for details.

2. An adaptive Newton method

The aim of this section is to develop an adaptive Newton method based on a simple prediction strategy. To this end, we will first recall the continuous Newton ODE.

2.1. Discrete vs. continuous Newton method

In order to improve the convergence behavior of the (*discrete*) Newton method (2) in the case that the initial guess is far away from a root $x_\infty \in \Omega$, it is classical to consider a damped version of the Newton sequence. More precisely, given a possibly small $t_n > 0$, we consider the iteration

$$x_{n+1} = x_n - t_n \mathsf{F}'(x_n)^{-1}\mathsf{F}(x_n). \tag{3}$$

Rearranging terms results in

$$\frac{x_{n+1} - x_n}{t_n} = -(\mathsf{F}'(x_n))^{-1}\mathsf{F}(x_n),$$

we observe that (3) can be seen as the discretization of the initial value problem

$$\begin{cases} \dot{x}(t) &= \mathsf{N}_\mathsf{F}(x(t)), \quad t \geqslant 0, \\ x(0) &= x_0, \end{cases} \tag{4}$$

by the explicit Euler scheme with step size t_n. Here, $\mathsf{N}_\mathsf{F}(x) = -\mathsf{F}'(x)^{-1}\mathsf{F}(x)$ is the so-called Newton Raphson transform (NRT, for short; see [14]) of F. The system (4) is called *continuous Newton method*. It is noteworthy that, if N_F is of class C^1 on some neighborhood of $x_\infty \in \Omega$, then we have $\mathsf{D}(\mathsf{N}_\mathsf{F})(x_\infty) = -\mathsf{Id}$. In particular, by the Poincaré-Ljapunow Theorem (see, e.g., [17]) we conclude that each regular zero of F is located in an attracting neighborhood contained in Ω when the NRT is applied. Furthermore, hoping that a sufficiently smooth solution of (4) exists, and that $\lim_{t\to\infty} x(t) = x_\infty \in \Omega$ is well-defined with $\mathsf{F}(x_\infty) = 0$, one can readily infer that

$$\mathsf{F}(x(t)) = \mathsf{F}(x_0)e^{-t}. \tag{5}$$

The solvability of (4) within the framework of Banach spaces has been addressed in [10,11]. Note that the trajectory of a solution of (4) either ends at the solution point x_∞ which is located closest to the initial value x_0, or at a some point close to a critical point x_c with non-invertible derivative F', or at some point on the boundary $\partial\Omega$ of the domain of F; see [7,8].