INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 20, Number 5, Pages 667–692

doi: 10.4208/ijnam2023-1029

NEWTON-ANDERSON AT SINGULAR POINTS

MATT DALLAS AND SARA POLLOCK

Abstract. In this paper we develop convergence and acceleration theory for Anderson acceleration applied to Newton's method for nonlinear systems in which the Jacobian is singular at a solution. For these problems, the standard Newton algorithm converges linearly in a region about the solution; and, it has been previously observed that Anderson acceleration can substantially improve convergence without additional a priori knowledge, and with little additional computation cost. We present an analysis of the Newton-Anderson algorithm in this context, and introduce a novel and theoretically supported safeguarding strategy. The convergence results are demonstrated with the Chandrasekhar H-equation and a variety of benchmark examples.

Key words. Anderson acceleration, Newton's method, safeguarding, singular problems.

1. Introduction

Given a nonlinear function $f : \mathbb{R}^n \to \mathbb{R}^n$ and root x^* for which $f(x^*) = 0$, it is well known that if the derivative of f at the root is nonsingular, then Newton's method exhibits quadratic convergence in a sufficiently small ball centered at the root. On the other hand, if the derivative is singular at x^*, e.g., at a bifurcation point [35], Newton's method converges linearly in a star-like region containing the root [18]. This singular setting has been studied in great detail [8, 9, 18, 32, 33], and a number of acceleration schemes have been proposed and analyzed [4, 5, 8, 10, 19, 20, 22, 34]. For example, the Levenberg-Marquardt method featured in [4, 5, 20] is known to be effective for solving nonlinear systems with singular Jacobians under the local-error bound condition. The focus of this paper is the analysis and demonstration of an extrapolation scheme called Anderson acceleration, sometimes called Anderson mixing, e.g.,[36], applied to Newton's method, for singular problems. We will show that with the proposed safeguarding strategy, the method is both theoretically sound and can be beneficial in practice for singular problems.

Anderson acceleration was first proposed in [2], in the context of integral equations, to improve the convergence of fixed-point iterations. Anderson acceleration is an attractive method to improve the convergence of linearly converging fixed-point iterations due to its low computational cost, ease of implementation, and track record of success over a wide range of problems. The method recombines the m most recent update steps and iterates to form an accelerated iterate at each stage of a given fixed-point method, where the particular combination is generally given as the solution to a least-squares problem. Here m may be referred to as the algorithmic depth, which is often chosen small, say less than five [37], but may sometimes benefit from being taken substantially larger [31, 38]. The method has been found beneficial in diverse applications, such as the computation of canonical tensor decompositions [39], the study of block copolymer systems [36], geometry optimization and simulation [27], flow problems [24, 29], and electronic-structure computations [1, 14], to name a few. Substantial advances in understanding the

Received by the editors on February 3, 2023 and, accepted on July 4, 2023.
2000 *Mathematics Subject Classification.* 65J15.

method in relation to generalized Broyden methods and (nonlinear) GMRES are developed in [13, 14, 38]. Recently, significant effort has been devoted to analyzing Anderson acceleration applied to contractive and noncontractive operators with certain nondegeneracy assumptions [12, 29, 31, 37].

Here, we will focus on the analysis of Anderson acceleration applied to Newton's method for a problem of the form $f(x) = 0$, when the derivative f' is singular at a root x^*. Rapid convergence of the accelerated scheme in comparison with standard Newton has been demonstrated numerically in this singular case [30], where it is also observed that it is generally both sufficient and preferable to set the algorithmic depth to $m = 1$. It was also found in [12] in a nondegenerate setting that Anderson accelerated Newton iterations with algorithmic depth $m = 1$ could converge where Newton iterations failed, but that increasing m only slowed convergence. In the remainder, we will consider Anderson acceleration with depth $m = 1$ applied to Newton iterations, which we will refer to simply as Newton-Anderson. In comparison to the accelerated Newton methods of [8, 10, 22], Newton-Anderson may be seen as advantageous as it does not require explicit knowledge of the order of the root (defined in section 7), or construction of an approximate projection mapping onto the null space of $f'(x)$. In further contrast to these predictor-corrector methods, Newton-Anderson requires a single linear solve per iteration. An analysis of Newton-Anderson in the one-dimensional singular case is presented in [28]; however, to our knowledge no previous convergence theory has been developed for dimension $n > 1$. The goal of this paper is to provide such a theory.

The remainder of the paper is organized as follows. The underlying foundation of the analysis relies on a technique for approximating the inverse of the derivative near a given point as developed in [8]. We discuss this technique in section 2, and in section 3 apply it to a Newton-Anderson step to obtain an expansion of the error at step k. We then analyze this expansion in sections 4 and 5 with a one-step analysis of the error based on previous consecutive error-pairs, revealing the mechanism behind the changes in convergence rate demonstrated by the method. The main challenge of proving convergence for any Newton-like method in the singular case is that the geometry of the region of invertibility is more complex. To handle this, in section 6 we introduce a novel safeguarding scheme, which we call γ-safeguarding. This technique leads to the main results of this paper: when the null space of the derivative at the root is one-dimensional, then under the same conditions implying local convergence of the standard Newton method e.g., [9, Theorem 1.2], Newton-Anderson with γ-safeguarding exhibits local convergence, and in general the rate of convergence is improved. We extend these results to high order roots in section 7. These results are then demonstrated numerically in section 8 with several standard benchmark problems, both singular and nonsingular, including the Chandrasekhar H-equation [6, 21]. The introduced γ-Newton-Anderson is further shown to perform favorably in comparison to existing methods developed for the problem class both in terms of robustness and efficiency.

2. Preliminaries

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a C^3 function such that $f(x^*) = 0$ with $x^* \in \mathbb{R}^n$. This regularity assumption is standard for the problem class; see, for example [8, 9, 10, 16, 32, 33]. Suppose $N = \text{null}\left(f'(x^*)\right)$ is nontrivial, let $R = \text{range}\left(f'(x^*)\right)$, and let $\mathbb{R}^n = N \oplus R$. Throughout this paper, $B_r(x)$ denotes a ball of radius $r > 0$ centered at x, P_N and P_R denote the orthogonal projections onto N and R respectively. Denote the error by $e_k = x_k - x^*$, and the Newton update step by