Journal of Computational Mathematics, Vol.25, No.6, 2007, 705–718.

CONVERGENCE OF NEWTON'S METHOD FOR SYSTEMS OF EQUATIONS WITH CONSTANT RANK DERIVATIVES*

Xiubin Xu

(*Department of Mathematics, Zhejiang Normal University, Jinhua 321004, China*
Email: xxu@zjnu.cn)

Chong Li

(*Department of Mathematics, Zhejiang University, Hangzhou 310027, China*
Email: cli@zju.edu.cn)

Abstract

The convergence properties of Newton's method for systems of equations with constant rank derivatives are studied under the hypothesis that the derivatives satisfy some weak Lipschitz conditions. The unified convergence results, which include Kantorovich type theorems and Smale's point estimate theorems as special cases, are obtained.

Mathematics subject classification: 49M15, 65F20, 65H10.
Key words: Newton's method, Overdetermined system, Lipschitz condition with L average, Convergence, Rank.

1. Introduction

Let $\mathbb{X}$ and $\mathbb{Y}$ be Euclidean spaces or more generally Banach spaces, and let

$$f : \mathbb{X} \to \mathbb{Y}$$

be a Fréchet differentiable function. Consider the system of nonlinear equations

$$f(x) = 0. \tag{1.1}$$

Such a system is widely used in both theoretical and applied areas of mathematics. Newton's method is the most efficient method for solving such systems. In the case when $f'(x)$ is an isomorphism, there are two points of view to analyze the convergence properties for Newton's method: the Kantorovich type theorem and Smale's point estimate theory. The Kantorovich theorem gives the convergence criteria based on the boundary of f'' in a neighborhood of the initial point x_0, see Ortega and Rheinboldt [10] or Ostrowski [11]; while Smale's point estimate theory gives that based on the invariant

$$\gamma(f, x_0) = \sup_{k \geq 2} \left\| f'(x_0)^{-1} \frac{f^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}, \tag{1.2}$$

see, e.g., Kim [8], Smale [13], Shub and Smale [12]. The convergence criteria based on the radii around a solution of (1.1) were given independently by Traub and Wozniakowski in [14] and Wang in [16]. Since then, there have been many extensions of the above results, see, e.g., [5, 7, 17–20]. In particular, a great progress was made by Wang [17, 18], where the notions of

* Received June 28, 2005; final revised June 29, 2006; accepted June 29, 2006.

Lipschitz conditions with L average were introduced and Kantorovich's and Smale's convergence criteria were unified, see also [21].

Recent attentions have been focused on the study of convergence properties of Newton's method for the case when $f'(x)$ is not an isomorphism. For example, Dedieu and Shub in [3] (resp. [4]) developed the convergence properties for underdetermined (resp. overdetermined) systems with surjective (resp. injective) derivatives under the hypothesis that f is analytic; Li et al. [9] achieved the convergence for overdetermined systems with injective derivatives under the hypothesis that f' satisfies the Lipschitz conditions with L average.

Dedieu and Kim [2] studied the convergence properties of Newton's iteration for analytic systems of equations with constant rank derivatives. They considered an analytic function $f : \mathbb{X} \to \mathbb{Y}$ between two Euclidean spaces and obtained the convergence theorems for solutions and the least square solutions of $f = 0$, respectively. This case generalizes both the surjective-underdetermined case ($\mathrm{rank} f'(x) = \dim\mathbb{Y}$) and the injective-overdetermined case ($\mathrm{rank} f'(x) = \dim\mathbb{X}$).

In this paper, we will investigate the convergence properties of Newton's method for systems of equations with constant rank derivatives under the hypothesis that the derivatives satisfy Lipschitz conditions with L average. The unified convergence properties are obtained. Our results extend and improve those in [2]. We end this section by briefly describing the organization of this paper. The notion of Lipschitz condition with L average and several preliminary results are given in Section 2. The main convergence theorem is stated and proved in Section 3. In Section 4, we discuss the convergence for two special cases. These discussions result in the Kantorovich type results and Smale's point estimate results, respectively. The latter improves the results in [2].

2. Notions and Preliminary Results

In this section, we give some properties related to the Moore-Penrose inverse, which will be used in the next section. Let $A : \mathbb{X} \to \mathbb{Y}$ be a linear operator (or an $m \times n$ matrix). Recall that an operator (or an $n \times m$ matrix) $A^\dagger : \mathbb{Y} \to \mathbb{X}$ is the Moore-Penrose inverse of A if it satisfies the following four equations,

$$AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (AA^\dagger)^* = AA^\dagger, \quad (A^\dagger A)^* = A^\dagger A,$$

where A^* denotes the adjoint of A. Let $\ker A$ and $\mathrm{im} A$ denote the kernel and image of A, respectively. For a subspace E of $\mathbb{X}$, we use Π_E to denote the projection onto E. Then it is clear that

$$A^\dagger A = \Pi_{\ker A^\perp} \text{ and } AA^\dagger = \Pi_{\mathrm{im} A}. \tag{2.1}$$

The following two lemmas give some perturbation bounds for the Moore-Penrose inverse. The first one is stated in [15, Corollary 7.1.1 (2)] and [15, Corollary 7.1.2], while the second one is a direct consequence of [15, Corollary 7.1.1 (2)] and [15, Corollary 7.1.4].

Lemma 2.1. *Let A and B be $m \times n$ matrices and let $r \le \min\{m, n\}$. Suppose that* $\mathrm{rank} A = r$, $\mathrm{rank}(A + B) \le r$ *and* $\|A^\dagger\|\|B\| < 1$. *Then*

$$\mathrm{rank}(A + B) = r \quad \text{and} \quad \|(A + B)^\dagger\| \le \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|B\|}.$$