Chin. Ann. of Math.
14B: 2(1993), 213-224.

ON THE CONVERGENCE ANALYSIS OF THE NONLINEAR ABS METHODS

HUANG ZHIJIAN*

Abstract

In order to complete the convergence theory of nonlinear ABS algorithm, through a careful investigation to the algorithm structure, the author converts the nonlinear ABS algorithm into an inexact Newton method. Based on such equivalent variation, the Kantorovich type convergence of the ABS algorithm is established and the convergence conditions of the algorithm that only depend on the initial conditions are obtained, which provides a useful basis for the choices of initial points of the ABS algorithm.

§1. Introduction

The ABS methods for solving systems of linear algebraic equations were presented by Abaffy, Broyden and Spedicato [1] in 1984. In 1987, Abaffy, Galantai and Spedicato [2] applied the methods to systems of nonlinear algebraic equations

$$F(x)=0, \qquad F: D\subset \mathbb{R}^n \to \mathbb{R}^n \tag{1.1}$$

and established the following nonlinear ABS algorithm model:

Algorithm (NABS)

Step 0: Choose an initial point $x_0 \in \mathbb{R}^n$, and let $k := 0$.

Step 1: Let $H_1^{(k)} = I$, $y_1^{(k)} = x_k$, $j := 1$.

Step 2: Choose $z_j^{(k)} \in \mathbb{R}^n$, such that

$$z_j^{(k)^T} H_j^{(k)} \nabla f_j(y_j^{(k)}) \neq 0 \tag{1.2}$$

and let

$$p_j^{(k)} = H_j^{(k)^T} z_j^{(k)}, \tag{1.3}$$

$$y_{j+1}^{(k)} = y_j^{(k)} - f_j(y_j^{(k)}) p_j^{(k)} / p_j^{(k)^T} \nabla f_j(y_j^{(k)}). \tag{1.4}$$

Step 3: If $j < n$, choose $w_j^{(k)} \in \mathbb{R}^n$ such that

$$w_j^{(k)^T} H_j^{(k)} \nabla f_j(y_j^{(k)}) = 1, \tag{1.5}$$

and update

$$H_{j+1}^{(k)} = H_j^{(k)} - H_j^{(k)} \nabla f_j(y_j^{(k)}) w_j^{(k)^T} H_j^{(k)}. \tag{1.6}$$

Let $j := j+1$, go to step 2; otherwise, go to Step 4.

Step 4: Let $x_{k+1} = y_{n+1}^{(k)}$, $k := k+1$, go to Step 1.

Manuscript received October 30, 1990.

*Department of Statistics and Operations Research, Fudan University, Shanghai 200433, China .

The parameter vectors $z_j^{(k)}$, $w_j^{(k)}$ in the algorithm can be arbitrarily chosen except for conditions (1.2) and (1.5). In [2], the authors proved that if the parameter vectors are suitably chosen, some well-known methods, such as Gay-Brown class, can be derived from the ABS methods. Furthermore, the authors studied the local convergence of the ABS methods, and proved that the ABS algorithm quadratically converges to a solution x^* of (1.1) if the initial point x_0 is sufficiently close to x^*, which provided an important basis for the reliability and efficiency analysis of the ABS algorithm.

In this paper, in order to complete the convergence theory of the nonlinear ABS algorithm, through a careful investigation to the algorithm structure, we convert the nonlinear ABS algorithm into an inexact Newton method. Based on such equivalent variation, we establish the Kantorovich type convergence of the nonlinear ABS algorithm and obtain the convergence conditions of the algorithm that only depend on the initial conditions, which provides a useful basis for the choices of initial points of the algorithm.

§2. Preliminaries

In this paper, we take the Frobenius norm as matrix norm, and the Euclidean norm as vector norm. For the sake of brevity, we denote $\nabla f_j(y_j^{(k)})$ by $a_j^{(k)}$.

Denote

$$A_k^T = [a_1^{(k)}, \cdots, a_n^{(k)}],$$

$$D_k = [p_1^{(k)}/p_1^{(k)^T} a_1^{(k)}, \cdots, p_n^{(k)}/p_n^{(k)^T} a_n^{(k)}].$$

Suppose that $F'(x_k)^{-1}$ and A_k^{-1} exist. Then the ABS algorithm can be expressed as

$$\begin{aligned} x_{k+1} = y_{n+1}^{(k)} &= y_1^{(k)} - \sum_{s=1}^{n} f_s(y_s^{(k)}) p_s^{(k)} / p_s^{(k)^T} a_s^{(k)} \\ &= x_k - [p_1^{(k)}/p_1^{(k)^T} a_1^{(k)}, \cdots, p_n^{(k)}/p_n^{(k)^T} a_n^{(k)}] \begin{bmatrix} f_1(y_1^{(k)}) \\ \vdots \\ f_n(y_n^{(k)}) \end{bmatrix} \\ &= x_k - F'(x_k)^{-1}F(x_k) + [F'(x_k)^{-1} - A_k^{-1}]F(x_k) \\ &\quad + A_k^{-1}F(x_k) - D_k \begin{bmatrix} f_1(y_1^{(k)}) \\ \vdots \\ f_n(y_n^{(k)}) \end{bmatrix}. \end{aligned} \tag{2.1}$$

Denote

$$\eta_k = U_k + A_k^{-1} V_k, \tag{2.2}$$

where

$$U_k = [F'(x_k)^{-1} - A_k^{-1}]F(x_k), \tag{2.3}$$

$$V_k = F(x_k) - A_k D_k \begin{bmatrix} f_1(y_1^{(k)}) \\ \vdots \\ f_n(y_n^{(k)}) \end{bmatrix}. \tag{2.4}$$

Now we give the following basic properties of the ABS algorithm. The proof can be found in [1, 2], or can be obtained by simple calculation from the algorithm and the known properties.

Property 2.1. $H_i^{(k)}a_j^{(k)} = 0,\ 1 \le j < i \le n.$

Property 2.2. $H_j^{(k)^2} = H_j^{(k)},\ j = 1, \cdots, n.$

Property 2.3. If $a_1^{(k)}$, $a_2^{(k)}$, $\cdots$, $a_n^{(k)}$ are linearly independent, then $p_1^{(k)}$, $p_2^{(k)}$, $\cdots$, $p_n^{(k)}$ are also linearly independent.

Property 2.4. $a_j^{(k)^T}(y_{j+1}^{(k)} - y_j^{(k)}) = -f_j(y_j^{(k)}),\ 1 \le j \le n.$

Property 2.5.

$$A_kD_k = \begin{bmatrix} 1 & & & \\ & 1 & & \\ l_{ij}^{(k)} & & \ddots & \\ & & & 1 \end{bmatrix},$$

where

$$l_{ij}^{(k)} = a_i^{(k)^T}p_j^{(k)}/p_j^{(k)^T}a_j^{(k)}, \qquad 1 \le j < i \le n.$$

§3. Variation of the Algorithm

First of all, we make the following assumption on $F(x)$:

(A1): $F(x)$ is continuously differentiable on some open convex set $D_0 \subset D$, and there exists a constant $K_0 \ge 0$ such that

$$\|F'(x) - F'(y)\| \le K_0\|x - y\|, \quad \forall x, y \in D_0. \tag{3.1}$$

(A2): There exists $x_0 \in D_0$ such that $F'(x_0)^{-1}$ exists. Let $\|F'(x_0)^{-1}\| \le b_0$.

Under the above assumptions, by the continuity of $F(x)$, there exists a constant $r_0 > 0$ such that

$$\bar{B}(x_0, 2r_0) \subset D_0, \tag{3.2}$$

$$\|F(x)\| \le 2\|F(x_0)\|, \quad \forall x \in B(x_0, r_0). \tag{3.3}$$

Moreover, from (A1), we have

$$\|F(y) - F(x) - F'(x)(y - x)\| \le \frac{1}{2}K_0\|x - y\|, \quad \forall x, y \in D_0. \tag{3.4}$$

We also make the following assumption to the choices of parameter vectors $z_j^{(k)}$ and $w_j^{(k)}$:

(A3): There exists a constant $\Gamma > 0$ such that

$$\|D_k\| \le \Gamma, \qquad k = 0, 1, \cdots. \tag{3.5}$$

Under the above assumptions, we have the following lemmas.

Lemma 3.1. *Suppose that* (A1) *and* (A2) *hold. Then there exist two constants* $r_1 > 0$, $\alpha > 0$ *such that for any* $z_1, \cdots, z_n \in B(x_0, r_1)$, *the matrix* $[\nabla f_1(z_1), \cdots, \nabla f_n(z_n)]$ *is invertible and satisfies*

$$\|[\nabla f_1(z_1), \cdots, \nabla f_n(z_n)]^{-1}\| \le \alpha.$$

The proof of Lemma 3.1 can be obtained by simple calculations.

The following lemma is direct from Lemma 3.1.

Lemma 3.2. *Suppose that* (A1) *and* (A2) *hold. Then, for any* $x \in B(x_0, r_1)$, $F'(x)^{-1}$ *exists and satisfies*

$$\|F'(x)^{-1}\| \le \alpha.$$

Let

$$\begin{aligned} r &= \min\{r_0, r_1\}, \\ K_1 &= \max\{\|F'(x)\| \,|\, x \in \bar{B}(x_0, 2r_0)\}, \\ K_2 &= \min\{r, (\sqrt{n}\alpha K_0)^{-1}\}, \\ K_3 &= (1 + \Gamma K_1)^{n-1}. \end{aligned}$$

Lemma 3.3. *Suppose that* (A1), (A2) *and* (A3) *hold and that*

$$\|F(x_0)\| < \frac{1}{2} K_1 K_2 K_3^{-1}.$$

If $x_k \in B(x_0, r)$, *then for any* j, *there hold*

(i) $|f_j(y_j^{(k)})| \le (1 + \Gamma K_1)^{j-1} \|F(x_k)\|$,

(ii) $\|y_{j+1}^{(k)} - y_j^{(k)}\| \le K_1^{-1}[(1 + \Gamma K_1)^j - 1]\|F(x_k)\|$,

(iii) $y_{j+1}^{(k)} \in B(x_0, 2r)$.

Proof. By induction. For $j = 1$, (i) obviously holds. Since

$$\|y_2^{(k)} - y_1^{(k)}\| = \| - f_1(y_1^{(k)}) p_1^{(k)} / p_1^{(k)T} a_1^{(k)}\| \le \Gamma |f_1(y_1^{(k)})| \le \Gamma \|F(x_k)\|,$$

(ii) also holds for $j = 1$. From (ii) and (3.3), we have

$$\begin{aligned} \|y_2^{(k)} - y_1^{(k)}\| &\le \Gamma 2 \|F(x_0)\| \\ &< 2\Gamma \cdot K_1 K_2 K_3^{-1}/2 \\ &\le K_2 \le r. \end{aligned}$$

It follows that

$$\|y_2^{(k)} - x_0\| \le \|y_2^{(k)} - y_1^{(k)}\| + \|y_1^{(k)} - x_0\| < 2r.$$

It implies that $y_2^{(k)} \in B(x_0, 2r)$. Thus, (i), (ii) and (iii) hold for $j = 1$.

Suppose that (i), (ii) and (iii) hold for $j = s$. Then for $j = s + 1$, by the inductive assumption,

$$\|y_{s+1}^{(k)} - y_1^{(k)}\| \le K_1^{-1}[1 + \Gamma K_1)^s - 1]\|F(x_k)\|.$$

Thus,

$$\begin{aligned} |f_{s+1}(y_{s+1}^{(k)})| &\le |f_{s+1}(y_{s+1}^{(k)}) - f_{s+1}(y_1^{(k)})| + |f_{s+1}(y_1^{(k)})| \\ &\le K_1 \|y_{s+1}^{(k)} - y_1^{(k)}\| + \|F(x_k)\| \\ &\le K_1 K_1^{-1}[(1 + \Gamma K_1)^s - 1]\|F(x_k)\| + \|F(x_k)\| \\ &= (1 + \Gamma K_1)^s \|F(x_k)\|. \end{aligned}$$

For (ii), we have

$$\begin{aligned}\|y_{s+2}^{(k)} - y_1^{(k)}\| &\le \|y_{s+2}^{(k)} - y_{s+1}^{(k)}\| + \|y_{s+1}^{(k)} - y_1^{(k)}\| \\ &\le \| - f_{s+1}(y_{s+1}^{(k)})p_{s+1}^{(k)}/p_{s+1}^{(k)T} a_{s+1}^{(k)}\| + \|y_{s+1}^{(k)} - y_1^{(k)}\| \\ &\le \Gamma|f_{s+1}(y_{s+1}^{(k)})| + \|y_{s+1}^{(k)} - y_1^{(k)}\| \\ &\le \Gamma(1+\Gamma K_1)^s\|F(x_k)\| + K_1^{-1}[(1+\Gamma K_1)^s - 1]\|F(x_k)\| \\ &= K_1^{-1}[(1+\Gamma K_1)^{s+1} - 1]\|F(x_k)\|.\end{aligned}$$

By (ii) and (3.3), we have

$$\begin{aligned}\|y_{s+2}^{(k)} - y_1^{(k)}\| &\le K_1^{-1}[(1+\Gamma K_1)^{s+1} - 1]2\|F(x_0)\| \\ &< 2K_1^{-1}[(1+\Gamma K_1)^{s+1} - 1]\cdot K_1K_2K_3^{-1}/2 \\ &\le K_2 \le r.\end{aligned}$$

It follows that

$$\|y_{s+2}^{(k)} - x_0\| \le \|y_{s+2}^{(k)} - y_1^{(k)}\| + \|y_1^{(k)} - x_0\| < 2r,$$

i.e., $y_{s+2}^{(k)} \in B(x_0, 2r)$. Therefore, the proof is completed.

Lemma 3.4. *suppose that* (A1), (A2) *and* (A3) *hold, and that*

$$\|F(x_0)\| < K_1K_2K_3^{-1}/2.$$

If $x_k \in B(x_0, r)$, *then*

(i) $\|A_k - F'(x_k)\| \le \sqrt{n}K_0K_1^{-1}[(1+\Gamma K_1)^{n-1} - 1]\|F(x_k)\|$,

(ii) A_k *is invertible and satisfies*

$$\|A_k^{-1}\| \le K_4,$$

where K_4 *is a constant.*

Proof. (i) Under the assumptions of Lemma 3.4, by Lemma 3.3, we have

$$y_j^{(k)} \in B(x_0, 2r) \subset D_0, \quad \forall j.$$

Thus, again from Lemma 3.3, for any j we have

$$\begin{aligned}\|a_j^{(k)} - \nabla f_j(x_k)\| &= \|\nabla f_j(y_j^{(k)}) - \nabla f_j(y_j^{(k)})\| \\ &\le K_0\|y_j^{(k)} - y_1^{(k)}\| \\ &\le K_0K_1^{-1}[(1+\Gamma K_1)^{j-1} - 1]\|F(x_k)\| \\ &\le K_0K_1^{-1}[(1+\Gamma K_1)^{n-1} - 1]\|F(x_k)\|.\end{aligned}$$

It follows that

$$\|A_k - F'(x_k)\| \le \sqrt{n}K_oK_1^{-1}[(1+\Gamma K_1)^{n-1} - 1]\|F(x_k)\|.$$

(ii) Let $\beta = 2\sqrt{n}\alpha K_0K_1^{-1}K_3\|F(x_0)\|$. From (i) we have

$$\begin{aligned}\|F'(x_k)^{-1}\|\|A_k - F'(x_k)\| &\le \alpha\sqrt{n}K_oK_1^{-1}[(1+\Gamma K_1)^{n-1} - 1]2\|F(x_0)\| \le \beta \\ &< 2\sqrt{n}\alpha K_0K_1^{-1}K_3\cdot K_1K_2K_3^{-1}/2 \\ &< \sqrt{n}\alpha K_0K_2 \le 1.\end{aligned}$$

Thus, by the Banach Perturbation Lemma, we see that A_k is invertible and satisfies

$$\|A_k^{-1}\| \le \|F'(x_k)^{-1}\|/(1-\beta) \overset{\text{def.}}{=} K_4.$$

Based on the above lemmas, we can establish the following variation theorem.

Theorem 3.1. *Suppose that* (A1), (A2) *and* (A3) *hold, and that*

$$\|F(x_0)\| < K_1K_2K_3^{-1}/2.$$

If $x_k \in B(x_0, r)$, *then the iterative procedure of the ABS algorithm at the k-th iteration can be expressed as*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k) + \eta_k, \tag{3.6}$$

where $\|\eta_k\| \le M\|F(x_k)\|^2$, $M > 0$ *is a constant.*

Proof. From (2.3), we have

$$\begin{aligned}\|U_k\| &= \|F'(x_k)^{-1}[A_k - F'(x_k)]A_k^{-1}F(x_k)\| \\ &\le \|F'(x_k)^{-1}\|\|A_k - F'(x_k)\|\|A_k^{-1}\|\|F(x_k)\| \\ &\le \sqrt{n}\alpha K_0K_1^{-1}[(1+\Gamma K_1)^{n-1} - 1]\|F(x_k)\|K_4\|F(x_k)\| \\ &= \sqrt{n}\alpha K_0K_1^{-1}K_4[(1+\Gamma K_1)^{n-1} - 1]\|F(x_k)\|^2.\end{aligned} \tag{3.7}$$

Denote $V_k^T = [V_1^{(k)}, \cdots, V_n^{(k)}]$. Then, making use of Property 2.5, we have

$$\begin{aligned}V_j^{(k)} &= f_j(x_k) - [l_{j1}^{(k)}, \cdots, l_{j,j-1}^{(k)}, 1, 0, \cdots, 0]\begin{bmatrix} f_1(y_1^{(k)}) \\ \vdots \\ f_n(y_n^{(k)}) \end{bmatrix} \\ &= f_j(x_k) - [\sum_{s=1}^{j-1} l_{js}^{(k)} f_s(y_s^{(k)}) + f_j(y_j^{(k)})] \\ &= f_j(y_1^{(k)}) - f_j(y_j^{(k)}) - a_j^{(k)T}\sum_{s=1}^{j-1} f_s(y_s^{(k)})p_s^{(k)}/p_s^{(k)T}a_s^{(k)} \\ &= f_j(y_1^{(k)}) - f_j(y_j^{(k)}) - a_j^{(k)}(y_1^{(k)} - y_j^{(k)}).\end{aligned}$$

Thus, from (3.4) and Lemma 3.3, we obtain

$$\begin{aligned}\|V_j^{(k)}\| &\le K_0\|y_1^{(k)} - y_j^{(k)}\|^2/2 \\ &\le K_0K_1^{-2}[(1+\Gamma K_1)^{n-1} - 1]^2\|F(x_k)\|^2/2.\end{aligned}$$

It follows that

$$\|V_k\| \le \sqrt{n}K_0K_1^{-2}[(1+\Gamma K_1)^{n-1} - 1]^2\|F(x_k)\|^2/2. \tag{3.8}$$

Combining (2.2), (3.7) and (3.8), we obtain

$$\begin{aligned}\|\eta_k\| &\le \|U_k + A_k^{-1}V_k\| \le \|U_k\| + \|A_k^{-1}\|\|V_k\| \\ &\le \sqrt{n}\alpha K_0K_1^{-1}K_4[(1+\Gamma K_1)^{n-1} - 1]\|F(x_k)\|^2 \\ &\quad + K_4 \cdot \sqrt{n}K_0K_1^{-2}[(1+\Gamma K_1)^{n-1} - 1]^2\|F(x_k)\|^2/2 \\ &= \sqrt{n}K_0K_1^{-1}K_4[(1+\Gamma K_1)^{n-1} - 1][\alpha + K_1^{-1}((1+\Gamma K_1)^{n-1} - 1)/2]\|F(x_k)\|^2.\end{aligned}$$

Let

$$M = \sqrt{n}K_0K_1^{-1}K_4[(1+\Gamma K_1)^{n-1} - 1][\alpha + K_1^{-1}((1+\Gamma K_1)^{n-1} - 1)/2].$$

Therefore, Theorem 3.1 is proved.

Through the above analysis, we know that the equivalent variation (3.6) of the ABS algorithm can be viewed as an inexact Newton method, and that the error term satisfies

$$\|\eta_k\| \le M\|F(x_k)\|^2.$$

Therefore, the further investigation on the convergence of the nonlinear ABS algorithm can be made to the inexact Newton variation (3.6).

§4. The Kantorovich Type Convergence Theorem

In this section, based on the algorithm analysis in the last section, we establish the Kantorovich type convergence theorem of the nonlinear ABS algorithm.

First of all, we prove the following lemmas.

Lemma 4.1. *Suppose that* (A1), (A2) *and* (A3) *hold and that*

$$\|F(x_0)\| < \min\{K_1K_2K_3^{-1}/2, (2MK_1)^{-1}\}.$$

If $x_k \in B(x_0, r)$, *then*

$$\|F(x_k)\| \le \frac{K_1}{1-2MK_1\|F(x_0)\|}\|x_{k+1}-x_k\|. \tag{4.1}$$

Proof. From the iterative procedure (3.7), we have

$$F(x_k) = F'(x_k)(x_k - x_{k+1}) + F'(x_k)\eta_k.$$

Then by (3.3), we have the following inequality

$$\begin{aligned}\|F(x_k)\| &\le \|F'(x_k)\|\|x_k - x_{k+1}\| + \|F'(x_k)\|\|\eta_k\|\\ &\le K_1\|x_{k+1}-x_k\| + K_1M\|F(x_k)\|^2\\ &\le K_1\|x_{k+1}-x_k\| + 2MK_1\|F(x_0)\|\|F(x_k)\|.\end{aligned}$$

It follows that

$$(1-2MK_1\|F(x_0)\|)\|F(x_k)\| \le K_1\|x_{k+1}-x_k\|.$$

Since $1-2MK_1\|F(x_k)\| > 0$, from the above inequality we obtain

$$\|F(x_k)\| \le \frac{K_1}{1-2MK_1\|F(x_0)\|}\|x_{k+1}-x_k\|.$$

Lemma 4.2. *Let the function* $g(t)$ *be of the form*

$$g(t) = \frac{\tau}{2(1-\tau)} + at + \frac{2-\tau}{1-\tau}\cdot\frac{at}{1-at}, \qquad 0 < t < 1/a,$$

where $a > 0$, $0 < \tau < 1/2$. *Then, for any* $0 < t < \bar{t} = (1-2\tau)/4a$, $g(t) < 1-\tau$.

Proof. It is easy to see that $0 < \bar{t} < 1/a$. Solving the inequality

$$g(t) - (1-\tau) < 0, \tag{4.2}$$

we see that for any

$$0 < t < \tilde{t} = \frac{4\tau^2 - 10\tau + 4}{[2\tau^2 - 9\tau + 8 + \sqrt{(2\tau^2-9\tau+8)^2 - 8(1-\tau)(2\tau^2-5\tau+2)}]a}$$

inequality (4.2) holds. Since for $0 < \tau < 1/2$

$$\tilde{t} > \frac{2\tau^2 - 5\tau + 2}{(2\tau^2 - 9\tau + 8)a} \ge (1-2\tau)/4a = \bar{t},$$

we have, for any $0 < t < \bar{t}$, $g(t) < 1 - \tau$.

Now we establish the following Kantorovich type convergence theorem of the nonlinear ABS algorithm.

Consider $x_k \in D_0$. If $F'(x_k)^{-1}$ exists, we assume that $\|F'(x_k)^{-1}\| \le b_k$.

Denote

$$\begin{aligned} \Delta_k &= x_{k+1} - x_k, \\ \theta_k &= b_k K_0 \|\Delta_k\|, \\ K_5 &= \max\{\|F(x)\| \,|\, x \in \bar{B}(x_0, 2r)\}. \end{aligned}$$

Theorem 4.1. *Suppose that* (A1), (A2) *and* (A3) *hold and that*

$$\theta_0 = b_0 K_0 \|\Delta_0\| \le \tau < 1/2, \tag{4.3}$$

$$\|F(x_0)\| < \min\{K_1 K_2 K_3^{-1}/2, (1-2\tau)/8MK_1, \tau r/(b_0 + MK_5)\}. \tag{4.4}$$

Then, starting from x_0, the sequence $\{x_k\}$ generated by the nonlinear ABS algorithm remains in $B(x_0, r)$ and converges to the unique solution x^* of (1.1) *in $\bar{B}(x_0, r_1)$.*

Proof. It is easy to see that under the assumptions of Theorem 4.1 the conditions of Lemma 4.1 hold.

First of all, we prove the following consequence: under the assumptions of the theorem, if x_k, $x_{k+1} \in B(x_0, r)$ and $\theta_k \le \tau$, then

$$\|\Delta_{k+1}\| < (1-\tau)\|\Delta_k\|. \tag{4.5}$$

Since x_k, $x_{k+1} \in B(x_0, r)$, $F'(x_k)$ and $F'(x_{k+1})$ are invertible. Let

$$G_k = F'(x_k)^{-1}[F'(x_k) - F'(x_{k+1})].$$

Since

$$\begin{aligned} \|G_k\| &\le \|F'(x_k)^{-1}\| \|F'(x_k) - F'(x_{k+1})\| \\ &\le b_k K_0 \|x_k - x_{k+1}\| \\ &= b_k K_0 \|\Delta_k\| = \theta_k, \end{aligned}$$

by Neumann lemma we obtain

$$\|(I - G_k)^{-1}\| \le 1/(1-\theta_k).$$

Let $b_{k+1} = b_k/(1-\theta_k)$. From the identity

$$F'(x_{k+1})^{-1} = (I - G_k)^{-1} F'(x_k)^{-1},$$

we have

$$\begin{aligned} \|F'(x_{k+1})^{-1}\| &\le \|(I - G_k)^{-1}\| \|F'(x_k)^{-1}\| \\ &\le b_k/(1-\theta_k) = b_{k+1}. \end{aligned}$$

From (3.6), we have

$$F(x_{k+1}) = F(x_{k+1}) - F(x_k) - F'(x_k)(x_{k+1} - x_k) + F'(x_k)\eta_k.$$

Thus,

$$\begin{aligned}\Delta_{k+1}&=-F'(x_{k+1})^{-1}F(x_{k+1})+\eta_{k+1}\\&=-(I-G_k)^{-1}F'(x_k)^{-1}[F(x_{k+1})-F(x_k)-F'(x_k)(x_{k+1}-x_k)+F'(x_k)\eta_k]+\eta_{k+1}\\&=-(I-G_k)^{-1}\{F'(x_k)^{-1}[F(x_{k+1})-F(x_k)-F'(x_k)(x_{k+1}-x_k)+\eta_k\}+\eta_{k+1}.\end{aligned}\tag{4.6}$$

Since

$$\begin{aligned}\|\eta_{k+1}\|&\le M\|F(x_{k+1})\|^2\\&\le M(\|F(x_{k+1})-F(x_k)\|+\|F(x_k)\|)\|F(x_{k+1})\|\\&\le 2M(K_1\|x_{k+1}-x_k\|+\|F(x_k)\|)\|F(x_0)\|\\&=2MK_1\|F(x_0)\|\|\Delta_k\|+2M\|F(x_0)\|\|F(x_k)\|,\end{aligned}\tag{4.7}$$

by (4.6), (4.8) and Lemma 4.1 we have

$$\begin{aligned}\|\Delta_{k+1}\|&\le\frac{1}{1-\theta_k}[\|F'(x_k)^{-1}\|\|F(x_{k+1})-F(x_k)-F'(x_k)(x_{k+1}-x_k)\|+\|\eta_k\|]+\|\eta_{k+1}\|\\&\le\frac{1}{1-\theta_k}[b_k^{1/2}K_0\|x_{k+1}-x_k\|^2+M\|F(x_k)\|^2]\\&\quad+2MK_1\|F(x_0)\|\|\Delta_k\|+2M\|F(x_0)\|\|F(x_k)\|\\&\le\frac{1}{1-\theta_k}[\theta_k\|\Delta_k\|/2+2M\|F(x_0)\|\|\frac{K_1}{1-2MK_1\|F(x_0)\|}\|\Delta_k\|]\\&\quad+2MK_1\|F(x_0)\|\|\Delta_k\|+2M\|F(x_0)\|\frac{K_1}{1-2MK_1\|F(x_0)\|}\|\Delta_k\|\\&=\frac{\theta_k}{2(1-\theta_k)}\|\Delta_k\|+2MK_1\|F(x_0)\|\|\Delta_k\|+\frac{2-\theta_k}{1-\theta_k}\cdot\frac{2MK_1\|F(x_0)\|}{1-2MK_1\|F(x_0)\|}\|\Delta_k\|\\&\le[\frac{\tau}{2(1-\tau)}+2MK_1\|F(x_0)\|+\frac{2-\tau}{1-\tau}\cdot\frac{2MK_1\|F(x_0)\|}{1-2MK_1\|F(x_0)\|}]\|\Delta_k\|.\end{aligned}$$

By (4.4) and Lemma 4.2, we obtain

$$\|\Delta_{k+1}\|<(1-\tau)\|\Delta_k\|.$$

By making use of the above consequence, we can prove by induction that

(i) $\theta_k\le\tau$,

(ii) $x_{k+1}\in B(x_0,r)$.

For $k=0$, (i) obviously holds. Since

$$\begin{aligned}\|\Delta_0\|&=\|-F'(x_0)^{-1}F(x_0)+\eta_0\|\\&\le[\|F'(x_0)^{-1}\|\|F(x_0)\|+M\|F(x_0)\|^2]\\&\le(b_0+MK_5)\|F(x_0)\|\\&<(b_0+MK_5)\tau r/(b_0+MK_5)\\&=\tau r<r,\end{aligned}\tag{4.8}$$

(ii) holds for $k=0$.

Suppose that (i) and (ii) hold for $k\le m$. Then for $k=m+1$, by the inductive assumptions and (4.5), we have

$$\|\Delta_{s+1}\|<(1-\tau)\|\Delta_s\|,\qquad 1\le s\le m.$$

It follows that

$$\begin{aligned}\theta_{m+1} &= b_{m+1}K_0\|\Delta_{m+1}\| \\ &\le \frac{b_m}{1-\theta_m}K_0(1-\tau)\|\Delta_m\| \\ &\le b_mK_0\|\Delta_m\| = \theta_m \le \tau.\end{aligned} \tag{4.9}$$

Moreover, from (4.9), we have

$$\begin{aligned}\|x_{m+2}-x_0\| &\le \sum_{s=0}^{m+1}\|x_{s+1}-x_s\| = \sum_{s=0}^{m+1}\|\Delta_s\| \\ &< \sum_{s=0}^{m+1}(1-\tau)^s\|\Delta_0\| < \sum_{s=0}^{\infty}(1-\tau)^s\|\Delta_0\| \\ &= \frac{1}{\tau}\|\Delta_0\| \le \frac{1}{\tau}\tau r = r.\end{aligned}$$

It implies that $x_{m+2} \in B(x_0, r)$. Thus, (i) and (ii) hold for any k, which implies that (4.5) hold for any k. Hence, we have

$$\|\Delta_k\| \le (1-\tau)\|\Delta_{k-1}\| \le \cdots \le (1-\tau)^k\|\Delta_0\|.$$

It follows that $\|\Delta_k\| \to 0$ as $k \to \infty$, which implies that $\{x_k\}$ is a Cauchy sequence. Therefore, $\{x_k\}$ is convergent.

Suppose that $x_k \to x^*$. Letting $m \to \infty$ at both sides of (4.9) and making use of (4.8), we have

$$\begin{aligned}\|x^*-x_0\| &\le \sum_{s=0}^{\infty}\|\Delta_s\| < \sum_{s=0}^{\infty}(1-\tau)^s\|\Delta_0\| \\ &= \frac{1}{\tau}\|\Delta_0\| \le r.\end{aligned}$$

It implies that $x^* \in B(x_0, r)$. By Lemma 4.1, we have

$$\begin{aligned}\|F(x_k)\| &\le \frac{K_1}{1-2MK_1\|F(x_0)\|}\|x_{k+1}-x_k\| \\ &= \frac{K_1}{1-2MK_1\|F(x_0)\|}\|\Delta_k\|.\end{aligned}$$

Letting $k \to \infty$, we have

$$\lim_{k\to\infty}\|F(x_k)\| = 0.$$

By the continuity of $F(x)$, we obtain $F(x^*) = 0$, that is, x^* is a solution of (1.1).

Suppose that there exists another solution $\widetilde{x} \in \bar{B}(x_0, r_1)$. By the mean value theorem, we have

$$\begin{aligned}f_j(x_k) &= f_j(x_k) - f_j(\widetilde{x}) \\ &= \nabla f_j[x_k + \theta_j^{(k)}(\widetilde{x}-x_k)]^T(x_k-\widetilde{x}), \qquad \forall j,\end{aligned}$$

where $\theta_j^{(k)} \in (0,1)$. It follows that

$$F(x_k) = \begin{bmatrix}\nabla f_1[x_k + \theta_1^{(k)}(\widetilde{x}-x_k)]^T \\ \vdots \\ \nabla f_n[x_k + \theta_n^{(k)}(\widetilde{x}-x_k)]^T\end{bmatrix}(x_k-\widetilde{x}). \tag{4.10}$$

Since $x_k \in B(x_0, r) \subset B(x_0, r_1)$, we have $x_k + \theta_j^{(k)}(\widetilde{x} - x_k) \in B(x_0, r_1)$, $\forall j$. Thus, by Lemma 3.1,

$$R_k = \begin{bmatrix} \nabla f_1[x_k + \theta_1^{(k)}(\widetilde{x} - x_k)]^T \\ \vdots \\ \nabla f_n[x_k + \theta_n^{(k)}(\widetilde{x} - x_k)]^T \end{bmatrix} = [\nabla f_1(x_k + \theta_1^{(k)}(\widetilde{x} - x_k)), \cdots, \nabla f_n(x_k + \theta_n^{(k)}(\widetilde{x} - x_k))]^T$$

is invertible and satisfies $\|R_k^{-1}\| \le \alpha$. Hence, from (4.10), we have

$$x_k - \widetilde{x} = R_k^{-1} F(x_k).$$

It follows that

$$\|x_k - \widetilde{x}\| \le \|R_k^{-1}\| \|F(x_k)\| \le \alpha \|F(x_k)\|.$$

Letting $k \to \infty$, we obtain

$$\|x^* - \widetilde{x}\| \le \lim_{k\to\infty} \alpha \|F(x_k)\| = \alpha \|F(x^*)\| = 0.$$

It implies that $\widetilde{x} = x^*$. Therefore, x^* is the unique solution of (1.1) in $\bar{B}(x_0, r_1)$.

Based on (3.6), we can also give the convergence rate of the sequence $\{x_k\}$.

Theorem 4.2. *Suppose that the assumption of Theorem 4.1 hold. Then, starting from* x_0*, the sequence* $\{x_k\}$ *generated by the nonlinear ABS algorithm quadratically convergences to the unique solution* x^* *of* (1.1) *in* $\bar{B}(x_0, r_1)$.

Proof. From Theorem 4.1, we know that the sequence $\{x_k\}$ converges to the unique solution x^* of (1.1) in $\bar{B}(x_0, r_1)$.

From the iterative procedure (3.6), we have

$$\begin{aligned} x_{k+1} - x^* =& x_k - x^* - F'(x_k)^{-1} F(x_k) + \eta_k \\ =& - F'(x_k)^{-1}[F(x_k) - F(x^*) - F'(x_k)(x_k - x^*)] + \eta_k \\ =& - F'(x_k)^{-1}[F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)] \\ &+ F'(x_k)^{-1}[F'(x_k) - F'(x^*)](x_k - x^*) + \eta_k. \end{aligned}$$

Thus,

$$\begin{aligned} \|x_{k+1} - x^*\| \le& \|F'(x_k)^{-1}\| \|F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)\| \\ &+ \|F'(x_k)^{-1}\| \|F'(x_k) - F'(x^*)\| \|x_k - x^*\| + \|\eta_k\| \\ \le& \alpha K_0 \|x_k - x^*\|^2/2 + \alpha K_0 \|x_k - x^*\|^2 + M\|F(x_k)\|^2 \\ \le& \frac{3}{2}\alpha K_0 \|x_k - x^*\|^2 + M\|F(x_k) - F(x^*)\|^2 \\ \le& \frac{3}{2}\alpha K_0 \|x_k - x^*\|^2 + MK_1^2 \|x_k - x^*\|^2 \\ \le& (\frac{3}{2}\alpha K_0 + MK_1^2)\|x_k - x^*\|^2. \end{aligned}$$

It implies that $\{x_k\}$ quadratically converges to x^*.

REFERENCES

[1] Abaffy, J., Broyden, C. G. & Spedicato, E., A class of direct methods for linear systems, *Numerische Mathematik*, **45** (1984), 361-376.

[2] Abaffy, J., Galantai, A. & Spedicato, E., The local convergence of ABS methods for nonlinear algebraic equations, *Numerische Mathematik*, **51** (1987), 429-439.

[3] Abaffy, J. & Spedicato, E., A generalization of Huang's method for solving systems of algebraic equations, *BUMI.*, **3-B** (1984), 517-529.

[4] Brent, R. P., Some efficient algorithms for solving systems of nonlinear equations, *SIAM J. Numer. Anal.*, **10** (1973), 327-344.

[5] Brown, K. M., A quadratically convergent Newton-like method upon Gaussian-elimination, *SIAM J. Numer. Anal.*, **6** (1969), 560-569.

[6] Dembo, R. S., Eisenstat, S. C. & Steihaug, T., Inexact Newton methods, *SIAM J. Numer. Anal.*, **19** (1982), 400-408.

[7] Dennis, J. E. & Jr. Schnabel, R. B., Numerical method for unconstrained optimization and nonlinear equations, *Prentice Hall, Inc., Englewood Cliffs, New Jersey*, **1983**.

[8] Gay, D. M., Brown's method and some generalizations, with applications to minimization problems, *Cornell Univ., Comput. Sci. Techn. Rep.*, (1975), 75-225.

[9] Kantorovich, L. V., On Newton's method for functional equations, *Dokl. Akad. SSSR*, **59** (1948), 1237-1240.

[10] Moret, I., On the behavior of approximate Newton methods, *Computing*, **37** (1986), 185-193.

[11] Ortega, J. M. & Rheinboldt, W. C., Iterative solution of nonlinear equations in several variables, *Academic Press*, **1970**.