6. The Steffensen Iteration Method for Systems of Nonlinear Equations

By Tatsuo NODA
Department of Applied Mathematics, Toyama Prefectural College of Technology

(Communicated by Kôsaku YOSIDA, M. J. A., Jan. 12, 1984)

1. Introduction. Let $x=(x_1, x_2, \cdots, x_n)$ be a vector in R^n and D a region contained in R^n. Let $f_i(x)(1\leq i\leq n)$ be real-valued nonlinear functions defined on D and $f(x)=(f_1(x), f_2(x), \cdots, f_n(x))$ an n-dimensional vector-valued function. Then we shall consider a system of nonlinear equations

$$x=f(x), \tag{1.1}$$

whose solution is $\bar{x}$. Denote by $\|x\|$ and $\|A\|$ the l_∞-norm and the corresponding matrix norm, respectively. That is,

$$\|x\|=\max_{1\leq i\leq n}|x_i| \quad \text{and} \quad \|A\|=\max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|,$$

where $A=(a_{ij})$ is an $n\times n$ matrix.

In generalizing the Aitken δ^2-process in one dimension to the case of n-dimensions, Henrici [1, p. 116] has considered the following formula, which is called the Aitken-Steffensen formula:

$$y^{(k)}=x^{(k)}-\Delta X^{(k)}(\Delta^2 X^{(k)})^{-1}\Delta x^{(k)}. \tag{1.2}$$

Furthermore, he has conjectured the following: We may hope that $y^{(k)}$ defined by (1.2) is closer to $\bar{x}$ than $x^{(k)}$, provided that the matrices $\Delta X^{(k)}$ and $\Delta^2 X^{(k)}$ are invertible. But he has not given mathematical certification to such a conjecture.

In [2], we have studied the above Aitken-Steffensen formula and shown [2, Theorem 2].

The purpose of this paper is to show Theorem 1 by considering a method of iteration, often called the Steffensen iteration method. Theorem 1 is an improvement on the result of [2, Theorem 2].

2. Statement of results. Define $f^{(i)}(x)\in R^n$ $(i=0, 1, 2, \cdots)$ by

$$f^{(0)}(x)=x,$$
$$f^{(i)}(x)=f(f^{(i-1)}(x)) \qquad (i=1, 2, \cdots).$$

Put

$$d^{(0,k)}=x^{(k)}-\bar{x},$$
$$d^{(i,k)}=f^{(i)}(x^{(k)})-\bar{x} \qquad \text{for } i=1, 2, \cdots.$$

Then an $n\times n$ matrix $D(x^{(k)})$ is defined as

$$D(x^{(k)})=(d^{(0,k)}, d^{(1,k)}, \cdots, d^{(n-1,k)}).$$

Throughout this paper, we shall assume the following five conditions (A.1)–(A.5) which are analogous to those of [2].

(A.1) $f_i(x)$ $(1 \leq i \leq n)$ are two times continuously differentiable on D.

(A.2) There exists a point $\bar{x} \in D$ satisfying (1.1).

(A.3) $\|J(\bar{x})\| < 1$, where $J(x) = (\partial f_i(x)/\partial x_j)$ $(1 \leq i, j \leq n)$.

(A.4) The vectors $d^{(0,k)}, d^{(1,k)}, \cdots, d^{(n-1,k)}$, $k = 0, 1, 2, \cdots$, are linearly independent.

(A.5) $\inf \{|\det D(x^{(k)})| / \|d^{(0,k)}\|^n\} > 0$.

Now, we consider Steffensen's iteration method

$$x^{(k+1)} = x^{(k)} - \Delta X(x^{(k)})(\Delta^2 X(x^{(k)}))^{-1} \Delta x(x^{(k)}), \tag{2.1}$$

where an n-dimensional vector $\Delta x(x)$, and $n \times n$ matrices $\Delta X(x)$ and $\Delta^2 X(x)$ are given by

$$\Delta x(x) = f^{(1)}(x) - x,$$
$$\Delta X(x) = (f^{(1)}(x) - x, \cdots, f^{(n)}(x) - f^{(n-1)}(x))$$

and

$$\Delta^2 X(x) = (f^{(2)}(x) - 2f^{(1)}(x) + x, \cdots, f^{(n+1)}(x) - 2f^{(n)}(x) + f^{(n-1)}(x)).$$

In this paper, we show the following

Theorem 1. *Under the conditions* (A.1)–(A.5), *there exists a constant M such that an estimate of the form*

$$\|x^{(k+1)} - \bar{x}\| \leq M \|x^{(k)} - \bar{x}\|^2$$

holds, provided that the $x^{(k)}$ generated by (2.1) *are sufficiently close to the solution $\bar{x}$ of* (1.1).

For the proof of Theorem 1, we need the following four lemmas:

Lemma 1 ([2, Lemma 1]). *Let A and C be $n \times n$ matrices and assume that A is invertible, with $\|A^{-1}\| \leq K_1$. If $\|A - C\| \leq K_2$ and $K_1 K_2 < 1$, then C is also invertible, and $\|C^{-1}\| \leq K_1/(1 - K_1 K_2)$.*

Lemma 2. *Under the conditions* (A.1)–(A.5), *there exists a constant L_1 such that the inequality*

$$\|(D(x^{(k)}))^{-1}\| \leq L_1 \|d^{(0,k)}\|^{-1} \tag{2.2}$$

holds for $x^{(k)}$ sufficiently close to $\bar{x}$.

Lemma 3. *Under the conditions* (A.1)–(A.5), *$n \times n$ matrices $\Delta X(x^{(k)})$ and $\Delta^2 X(x^{(k)})$ are invertible, and there exist constants L_2 and L_5 such that the inequalities*

$$\|(\Delta X(x^{(k)}))^{-1}\| \leq L_2 \|d^{(0,k)}\|^{-1}, \tag{2.3}$$

$$\|(\Delta^2 X(x^{(k)}))^{-1}\| \leq L_5 \|d^{(0,k)}\|^{-1} \tag{2.4}$$

hold for $x^{(k)}$ sufficiently close to $\bar{x}$.

Lemma 4 ([2, Lemma 5]). *Let an $n \times n$ matrix A be invertible. Let U and V be $n \times m$ matrices such as $m \leq n$. Then $A + UV^*$ is invertible if and only if $I + V^* A^{-1} U$ is invertible, and then*

$$(A + UV^*)^{-1} = A^{-1} - A^{-1} U (I + V^* A^{-1} U)^{-1} V^* A^{-1},$$

where V^* is the transposed matrix of V.

Lemmas 1 and 2 are used in proving Lemma 3. Since the proofs of the inequalities (2.2)–(2.4) are similar to those of Lemmas 2–4 in [2], respectively, they will not be given here. Lemma 4 may be used for determining $(\Delta^2 X(x^{(k)}))^{-1}$, and is called the Sherman-Morrison-Woodbury formula [3, p. 50].

Remark 1. By the definition, we have

$$\Delta^2 X(x^{(k)}) = (J(\bar{x}) - I)\Delta X(x^{(k)}) + Y(x^{(k)}), \tag{2.5}$$

where $Y(x^{(k)})$ is an $n\times n$ matrix. By (A.1)–(A.3), we may choose a constant L_3 such that, for $x^{(k)}$ sufficiently close to $\bar{x}$,

$$\|Y(x^{(k)})\| \leq L_3 \|d^{(0,k)}\|^2. \tag{2.6}$$

Here we note that the inequality (2.4) holds with $L_5 = L_2/L_4$ by choosing a constant L_4 so as to satisfy

$$1 - \|J(\bar{x})\| - L_2 L_3 \|d^{(0,k)}\| \geq L_4 > 0. \tag{2.7}$$

3. The proof of Theorem 1. We shall prove Theorem 1. As may be seen by Remark 1 in § 2, we also have

$$\Delta x(x^{(k)}) = (J(\bar{x}) - I)d^{(0,k)} + \xi(x^{(k)}), \tag{3.1}$$

where $\xi(x^{(k)})$ is an n-dimensional vector and

$$\|\xi(x^{(k)})\| \leq L_6 \|d^{(0,k)}\|^2, \tag{3.2}$$

a constant L_6 being suitably chosen.

We observe that, from (2.5), by Lemma 3 and (A.3), $\Delta X(x^{(k)}) + (J(\bar{x}) - I)^{-1} Y(x^{(k)})$ is invertible, while we have shown in Lemma 3 that $\Delta X(x^{(k)})$ is also invertible. Then, we may apply Lemma 4 for $m=n$ to $\Delta X(x^{(k)}) + (J(\bar{x}) - I)^{-1} Y(x^{(k)})$ and obtain

$$\begin{aligned} (\Delta^2 X(x^{(k)}))^{-1} = \{ & (\Delta X(x^{(k)}))^{-1} - (\Delta X(x^{(k)}))^{-1} (J(\bar{x}) - I)^{-1} \\ & \cdot [I + Y(x^{(k)})(\Delta X(x^{(k)}))^{-1} (J(\bar{x}) - I)^{-1}]^{-1} \\ & \cdot Y(x^{(k)})(\Delta X(x^{(k)}))^{-1}\} (J(\bar{x}) - I)^{-1}. \end{aligned} \tag{3.3}$$

Substituting (3.1) and (3.3) into (2.1), it yields

$$x^{(k+1)} - \bar{x} = p(x^{(k)}) + q(x^{(k)}), \tag{3.4}$$

where

$$\begin{aligned} p(x^{(k)}) = (J(\bar{x}) - I)^{-1} [& I + Y(x^{(k)})(\Delta X(x^{(k)}))^{-1} \\ & \cdot (J(\bar{x}) - I)^{-1}]^{-1} Y(x^{(k)})(\Delta X(x^{(k)}))^{-1} d^{(0,k)}, \end{aligned} \tag{3.5}$$

$$q(x^{(k)}) = -\Delta X(x^{(k)})(\Delta^2 X(x^{(k)}))^{-1} \xi(x^{(k)}). \tag{3.6}$$

Now, as for $p(x^{(k)})$, we first obtain an estimate

$$\|p(x^{(k)})\| \leq L_3 L_5 \|d^{(0,k)}\|^2, \tag{3.7}$$

from (3.5), by (2.3), (2.6) and (2.7). Since $\|D(x^{(k)})\| \leq \sum_{i=0}^{n-1} \|d^{(i,k)}\|$, we have $\|D(x^{(k)})\| \leq (\sum_{i=0}^{n-1} M^i) \|d^{(0,k)}\|$, by using the fact that $\|d^{(i+1,k)}\| \leq M\|d^{(i,k)}\|$ $(0<M<1)$ for $i=0, 1, 2, \cdots$, so that

$$\|\Delta X(x^{(k)})\| \leq L_7 \|d^{(0,k)}\| \tag{3.8}$$

holds for a constant L_7 chosen suitably. Hence, as for $q(x^{(k)})$, we next obtain an estimate

$$\|q(x^{(k)})\| \leq L_5 L_6 L_7 \|d^{(0,k)}\|^2, \tag{3.9}$$

from (3.6), by (2.4), (3.2) and (3.8). Consequently, (3.4), together with (3.7) and (3.9), shows that Theorem 1 holds with $M = L_5(L_3 + L_6 L_7)$, as desired.

The author would like to express his hearty thanks to Prof. H. Mine of Kyoto University for many valuable suggestions.

References

[1] P. Henrici: Elements of Numerical Analysis. John Wiley, New York (1964).
[2] T. Noda: The Aitken-Steffensen formula for systems of nonlinear equations. Sûgaku, **33**, 369–372 (1981) (in Japanese).
[3] J. M. Ortega and W. C. Rheinboldt: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York (1970).