68. *The Aitken-Steffensen Formula for Systems of Nonlinear Equations. IV*

By Tatsuo NODA
Faculty of Engineering, Toyama Prefectural University

(Communicated by Shokichi IYANAGA, M. J. A., Oct. 12, 1990)

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 $f(x)$ be a real-valued nonlinear function defined on D. We denote by $R^{n\times n}$ the set of all $n\times n$ real matrices. Define an n-dimensional vector $\nabla f(x)$ and an $n\times n$ matrix $H(x)$ by

$$\nabla f(x)=(\partial f(x)/\partial x_i) \qquad (1\leqq i\leqq n)$$

and

$$H(x)=(\partial^2 f(x)/\partial x_j \partial x_k) \qquad (1\leqq j, k\leqq n).$$

For any $x\in R^n$, we shall use the norms $\|x\|$ and $\|x\|_2$ defined by

$$\|x\|=\max_{1\leqq i\leqq n}|x_i| \quad \text{and} \quad \|x\|_2=\left(\sum_{i=1}^n x_i^2\right)^{1/2},$$

respectively. The corresponding matrix norms, denoted by $\|A\|$ and $\|A\|_s$, are defined as

$$\|A\|=\max_{1\leqq i\leqq n}\sum_{j=1}^n |a_{ij}| \quad \text{and} \quad \|A\|_s=\lambda^{1/2},$$

respectively, where $A=(a_{ij})\in R^{n\times n}$, and λ is the maximum eigenvalue of A^*A, A^* being the transposed matrix of A. We also define the matrix norm $\|A\|_E$ by

$$\|A\|_E=\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2\right)^{1/2}.$$

In this section, we shall assume the same conditions (A.1)–(A.4) as in [5] except for (A.1).

(A.1) $f(x)$ is three times continuously differentiable on D.

(A.2) There exists a point $\bar{x}\in D$ satisfying $\nabla f(x)=0$.

(A.3) The $n\times n$ symmetric matrix $H(\bar{x})$ is positive definite.

(A.4) β is a constant satisfying $0<\beta<2$.

We see that $f(x)$ has a local minimum at $\bar{x}$ by conditions (A.1)–(A.3). For computational purpose, we have proposed in [5, (2.1)] an iteration method

$$x^{(k+1)}=x^{(k)}-\frac{\beta}{\|H(x^{(k)})\|_E}\nabla f(x^{(k)}) \tag{1.1}$$

for finding $\bar{x}$ under conditions (A.1)–(A.4).

As mentioned in [2], [3] and [4], Henrici [1, p. 116] has considered a formula, which is called the Aitken-Steffensen formula. Now, we have studied the above Aitken-Steffensen formula for systems of nonlinear equations in [2], [3] and [4], and shown [2, Theorem 2], [3, Theorem 2] and [4, Theorem 1].

The purpose of this paper is to construct a formula by use of (1.1), which we shall also call an Aitken-Steffensen formula, and to show Theorem 1 by using [2, Theorem 2].

2. Statement of results. Define an n-dimensional vector $g(x)=(g_i(x))$ by

$$g(x)=x-\frac{\beta}{\|H(x)\|_E}\nabla f(x). \tag{2.1}$$

Given $x^{(0)}\in R^n$, define $x^{(i)}\in R^n$ $(i=1,2,\cdots)$ by

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

Put $d^{(i)}=x^{(i)}-\bar{x}$ for $i=0,1,2,\cdots$, and then define an $n\times n$ matrix D_k by

$$D_k=(d^{(k)},\ d^{(k+1)},\cdots,d^{(k+n-1)}).$$

In addition to conditions (A.1)–(A.4), we suppose the following two conditions (A.5) and (A.6) which are based on [2, Theorem 2].

(A.5) The vectors $d^{(k)}$, $d^{(k+1)},\cdots,d^{(k+n-1)}$, $k=0,1,2,\cdots$, are linearly independent.

(A.6) $\inf\{|\det D_k|/\|d^{(k)}\|_2^n\}>0$.

As suggested by [2, (1.5)], we can construct an Aitken-Steffensen formula

$$y^{(k)}=x^{(k)}-\Delta X^{(k)}(\Delta^2X^{(k)})^{-1}\Delta x^{(k)} \qquad (k=0,1,2,\cdots), \tag{2.2}$$

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

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

and

$$\Delta^2X^{(k)}=\Delta X^{(k+1)}-\Delta X^{(k)}.$$

In this paper, we shall show the following

Theorem 1. *Under conditions* (A.1)–(A.6), *for* $x^{(k)}\in U(\bar{x};\delta)$, *there exists a constant* M_2 *such that the following property*

$$\|y^{(k)}-\bar{x}\|_2\leqq M_2\|x^{(k)}-\bar{x}\|_2^2$$

holds for sufficiently large k.

3. Proof of Theorem 1. We shall prove Theorem 1. By (A.3),

$$0<(\rho,H(\bar{x})\rho)\leqq\|H(\bar{x})\|_E$$

for any $\rho\in R^n$ with $\|\rho\|_2=1$. Since, by (A.1), $\|H(x)\|_E$ is continuous at every point $x\in D$, there exists a neighbourhood

$$U(\bar{x};\delta_1)=\{x;\|x-\bar{x}\|_2<\delta_1\}\subset D$$

such that $x\in U(\bar{x};\delta_1)$ implies $\|H(x)\|_E>0$. Then, we observe that, by (A.1),

(3.1) $g_i(x)$ $(1\leqq i\leqq n)$ are two times continuously differentiable on $U(\bar{x};\delta_1)$, and, from (2.1), by (A.2),

$$\bar{x}=g(\bar{x}), \tag{3.2}$$

while we have shown in [5] that the following inequality

$$\|G(\bar{x})\|_s<1 \tag{3.3}$$

holds from (A.3) and (A.4), where $G(x)=(\partial g_i(x)/\partial x_j)$ $(1\leqq i,\ j\leqq n)$. Choosing a constant M so as to satisfy $\|G(\bar{x})\|_s<M<1$, we see, by (A.1), that there exists a constant $\delta\leqq\delta_1$ such that $U(\bar{x};\delta)\subset U(\bar{x};\delta_1)$ and $\|G(x)\|_s<M$ for $x\in U(\bar{x};\delta)$. By (1.1), (2.1) and (3.2),

$$x^{(k+1)}-\bar{x}=g(x^{(k)})-g(\bar{x})$$
$$=\int_0^1 G(\bar{x}+t(x^{(k)}-\bar{x}))(x^{(k)}-\bar{x})dt.$$

We note that $\bar{x}+t(x^{(k)}-\bar{x})\in U(\bar{x}\,;\delta)$ $(0\leqq t\leqq 1)$, provided $x^{(k)}\in U(\bar{x}\,;\delta)$. Then, by $\|G(x)\|_s<M$ for $x\in U(\bar{x}\,;\delta)$ shown above,

$$\int_0^1 \|G(\bar{x}+t(x^{(k)}-\bar{x}))\|_s dt\leqq M$$

holds, so that we have

$$\|x^{(k+1)}-\bar{x}\|_2\leqq M\|x^{(k)}-\bar{x}\|_2 \tag{3.4}$$

for $x^{(k)}\in U(\bar{x}\,;\delta)$.

For the proof of Theorem 1, we need the following well-known relations.

(3.5) $n^{-1/2}\|x\|_2\leqq\|x\|\leqq\|x\|_2$ for all $x\in R^n$,

(3.6) $\|I\|=\|I\|_s=1$ for the identity matrix $I\in R^{n\times n}$,

(3.7) $\|A\|_s\leqq\|A\|_E$ for all $A\in R^{n\times n}$

and

(3.8) $n^{-1/2}\|A\|_s\leqq\|A\|\leqq n^{1/2}\|A\|_s$ for all $A\in R^{n\times n}$.

Now, we recall that conditions (A.1)–(A.4) imply (3.1), (3.2) and (3.3) as shown above. Then applying the argument in the proof of [2, Theorem 2] to the norms $\|x\|_2$ and $\|A\|_s$ instead of the norms $\|x\|$ and $\|A\|$, respectively, and using (3.4), (3.5), (3.6), (3.7) and (3.8), we deduce that, for $x^{(k)}\in U(\bar{x}\,;\delta)$, there exists a constant M_2 such that

$$\|y^{(k)}-\bar{x}\|_2\leqq M_2\|x^{(k)}-\bar{x}\|_2^2$$

holds for sufficiently large k. In this way, we have proved Theorem 1, as desired.

4. Numerical example. We deal with a function

$$y(x\,;a,b,c,d)=e^{ax}(c\cos bx+d\sin bx)\qquad(a<0),$$

which is the same as in [5]. In order to show the efficiency of the Aitken-Steffensen formula (2.2), we consider a system of nonlinear equations, Example 4.1. The solution of Example 4.1 using the Aitken-Steffensen formula (2.2) is presented in Table 4.1 below, together with the solution by the iteration method [5, (2.1)].

Example 4.1: $$\begin{cases} y(0.0\,;a,b,c,d)=1.50,\\ y(0.8\,;a,b,c,d)=-0.05,\\ y(1.6\,;a,b,c,d)=-0.12,\\ y(2.4\,;a,b,c,d)=0.04.\end{cases}$$

The solution is $(a,b,c,d)=(-1.50,-2.50,1.50,-0.50)$.

Table 4.1. Computation results for Example 4.1

Methods	Solutions
Iteration method [5, (2.1)] (β=0.99)	(−1.506458, −2.501487, 1.499880, −0.5009617)
Aitken-Steffensen formula (2.2)	(−1.502620, −2.505557, 1.499941, −0.5007080)

$(a^{(0)},b^{(0)},c^{(0)},d^{(0)})=(-1.0,\ -1.0,\ -1.0,\ -1.0)$

The author would like to express his hearty thanks to H. Mine, Professor Emeritus 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] ——: The Aitken-Steffensen formula for systems of nonlinear equations. II. ibid., **38**, 83–85 (1986) (in Japanese).

[4] ——: The Aitken-Steffensen formula for systems of nonlinear equations. III. Proc. Japan Acad., **62A**, 174–177 (1986).

[5] ——: A modification of the gradient method and function extremization. ibid., **65A**, 39–42 (1989).