ORDER INTERVAL SECANT METHOD FOR NONLINEAR SYSTEMS*

LI QING-YANG (李庆扬)

(Qinghua University, Beijing, China)

Abstract

An order interval secant method is given. Its rate of convergence is faster than that of order interval Newton method in [1]. The existence and uniqueness of a solution to nonlinear systems and convergence of the interval iterative sequence are also proved.

Suppose $f:D\subset R^n\to R^n$, $X=[\underline{x},\overline{x}]=\{x\in R^n|\underline{x}\leqslant x\leqslant \overline{x}\}\subset D$. A simple interval Newton method for testing the existence and uniqueness of solutions to the nonlinear equations

$$f(x) = 0 \tag{1}$$

and an order interval Newton method for solving nonlinear equations (1), were given in [1]. This paper is to give an order interval secant method without the derivative calculation of f, which converges faster than the order interval Newton method. The convergence of this iterative method, as well as the existence and uniqueness of solutions to (1), are proved.

Notations used in this paper are the same as those in [1].

First, the definition of order convexity in [2] has to be generalized.

Definition 1. If there exist a nonsingular matrix $P \in L(\mathbb{R}^n)$ and $\lambda \in (0, 1)$ so that

$$Pf(\lambda x + (1-\lambda)y) \leq \lambda Pf(x) + (1-\lambda)Pf(y)$$
 (2)

for $f: D \subset \mathbb{R}^n \to \mathbb{R}^n$ in $X = [x, \overline{x}]$ and for any comparable $x, y \in X (x \leq y \text{ or } x \geq y)$, then f is called P-order convex. Moreover, if

$$Pf(x) \leqslant Pf(y) \tag{3}$$

for any $x \le x \le y \le \overline{x}$, then f is called P-isotone convex in X. If (2) is replaced by

$$Pf(\lambda x + (1-\lambda)y) \geqslant \lambda Pf(x) + (1-\lambda)Pf(y), \tag{2'}$$

then f is called P-order concave (or P-order upper convex) in X, Moreover, if (3) is valid then f is called P-isotone concave in X.

Remark 1. If P=I (the identity matrix), it is the order convexity defined by [2]. That f is P-order convex in X implies that F=Pf is order convex in X.

Definition 2. Suppose P is nonsingular, and $x \le x < y \le \overline{x}$ for any two points $x = (x_1, \dots, x_n)^T$ and $y = (y_1, \dots, y_n)^T$ on the interval $X = [\underline{x}, \overline{x}]$. Let F = Pf, $A(x, y) = (a_{ij})_{n \times n}$ is called an nth-order difference matrix, where

^{*} Received March 21, 1984.

$$a_{ij} = \frac{F_i \left(y - \sum_{k=0}^{j-1} (y_k - x_k) e_k \right) - F_i \left(y - \sum_{k=1}^{j} (y_k - x_k) e_k \right)}{y_j - x_j},$$

$$a_{i1} = \frac{F_i (y) - F_i (y - (y_1 - x_1) e_1)}{y_1 - x_1}, \quad i = 1, \dots, n; j = 2, \dots, n$$
(4)

If $A_i(x, y)$ is the i-th row of matrix A(x, y), by (4),

$$A_{i}(x, y)(y-x) = (a_{i1}, \dots, a_{in}) \begin{pmatrix} y_{1}-x_{1} \\ \vdots \\ y_{n}-x_{n} \end{pmatrix} = F_{i}(y) - F_{i}(x), \quad i=1, \dots, n,$$

then

$$A(x, y)(y-x) = F(y) - F(x)$$
 (5)

Lemma 1. Suppose f is P-order convex in $X = [\underline{x}, \overline{x}]$ and F = Pf; then, for any $x \le x < y \le \overline{x}$,

$$A(x, z')(y-x) \leq F(y) - F(x) \leq A(z, y)(y-x),$$
 (6)

where z=y-t(y-x)< y, z'=x+t(y-x), $t\in(0,1)$, A(x,z') and A(z,y) are nth-order difference matriced of F defined as (4).

If f is P-order concave, then for any $x \le x < y \le \overline{x}$,

.
$$A(x, z')(y-x) \ge F(y) - F(x) \ge A(z, y)(y-x)$$
. (6')

Proof. Because f is P-order convex on $X = [x, \bar{x}]$, by (2),

$$F(z) \leqslant tF(x) + (1-t)F(y),$$

or

$$F(z)-F(y) \leqslant t(F(x)-F(y))$$

for z=y-t(y-x) and $t\in(0, 1)$. Then

$$F(y) - F(x) \le \frac{1}{t} (F(y) - F(z)) = \frac{1}{t} (F(y) - F(y - t(y - x))).$$

By (4) and (5), we have

$$F(y) - F(x) \le \frac{1}{t} (F(y) - F(z)) = \frac{1}{t} A(z, y) (y - z) = A(z, y) (y - x).$$

Therefore, the right inequality in (6) holds. From same reason, if we set z'=x+t(y-x), by (2) we have

$$F(y) - F(x) \geqslant \frac{1}{t} (F(z') - F(x)) = \frac{1}{t} A(x, z') (z' - x) = A(x, z') (y - x).$$

Hence (6) holds.

(6') can be proved similarly.

Lemma 2. Suppose f is P-order convex in $X = [x, \bar{x}]$ and G-differentiable, and set z = y - t(y - x), z' = x + t(y - x). Then, for any $x \le x < y \le \bar{x}$ and $t \in (0, 1)$,

(i)
$$F'(x)(y-x) \leq F(y) - F(x) \leq F'(y)(y-x)$$
 (7)

and

$$\lim_{t\to 0} A(x, z') = F'(x), \quad \lim_{t\to 0} A(z, y) = F'(y),$$

(ii)
$$F'(x) \leq A(x, z') \leq A(x, y) \leq A(z, y) \leq F'(y)$$
. (8)

Moreover, suppose F'(x) and A(x, y) are nonsingular for any $x \le x < y \le \overline{x}$, $F'(x)^{-1} \ge 0$ and $A(x, y)^{-1} \ge 0$. Then

$$F'(x)^{-1} \geqslant A(x, z')^{-1} \geqslant A(x, y)^{-1} \geqslant A(z, y)^{-1} \geqslant F'(y)^{-1} \geqslant 0$$
 (9)

and

$$A(x, z')A(x, y)^{-1} \leq I$$
, $A(x, y)A(x, y)^{-1} \leq I$. (10)

Proof. Because F = Pf is G-differentiable in X and by Lemma 1. Letting $t \to 0$ in (6) we have (7). By (7) and (5),

$$F(y) - F(x) = A(x, y)(y-x) \leqslant F'(y)(y-x)$$
.

Since y-x>0, hence

$$A(x, y) \leqslant F'(y)$$
.

Similarly, if we set z=y-t(y-x),

$$F(y) - F(z) = A(z, y)(y-z) \le F'(y)(y-z)$$
.

Since y-z>0, hence

$$A(z, y) \leqslant F'(y)$$
.

By (5) and (6), we get

$$A(x, y) \leqslant A(z, y)$$
.

Therefore, the right inequality in (8) holds. The left inequality in (8) can be proved similarly.

Left multiplicating the right inequality in (8) with $F'(y)^{-1}$, we have

$$F'(y)^{-1}A(z, y) \leqslant I$$
.

Right multiplicating the above inequality with $A(x, y)^{-1}$, we have

$$A(z, y)^{-1} \geqslant F'(y)^{-1} \geqslant 0.$$

Similarly for the other parts of (9) and (10).

Remark 2. If f is P-order concave in X, results similar to Lemma 2 can be obtained, i.e.

$$F'(y)(y-x) \leqslant F(y) - F(x) \leqslant F'(x)(y-x),$$
 (7')

$$F'(y) \leq A(z, y) \leq A(x, y) \leq A(x, z') = F'(x),$$
 (8')

$$F'(y)^{-1} \geqslant A(z, y)^{-1} \geqslant A(x, y)^{-1} \geqslant A(x, z')^{-1} \geqslant F'(x)^{-1},$$
 (9')

$$A(z, y)A(x, y)^{-1} \leqslant I, \quad A(x, y)A(x, z')^{-1} \leqslant I.$$
 (10')

Order Interval Secant Method. When f is P-order convex in $X = [x, \bar{x}]$ and satisfy

$$Pf(x) \leqslant 0 \leqslant Pf(\bar{x}), \tag{11}$$

the order interval secant method is defined as follows.

Algorithm 1. Let $X^0 = [\underline{x}^0, \overline{x}^0] = [\underline{x}, \overline{x}] = X$. Define

$$X^{k+1} := [S_k x^k, \ \overline{S}_k \overline{x}^k] = [x^{k+1}, \ \overline{x}^{k+1}]$$
 (12)

for $k=0, 1, \dots$, where

and the auxiliary points $z^k = \bar{x}^k - t_k(\bar{x}^k - x^k)$, $0 < t_k < 1$, are chosen such that $F(z^k) \ge 0$. Because F is order convex in X, F is continuous in X. Therefore, when $F(\bar{x}^k) \ge 0$ there always exists a point z^k such that $F(z^k) \ge 0$.

Theorem 1. Let $f: X \subset \mathbb{R}^n \to \mathbb{R}^n$ be P-isotone convex in $X = [\underline{x}, \overline{x}]$ and satisfy

condition (11). If F=Pf is G-differentiable in X, F' and the nth-order difference matrix A(x, y) of F are nonsingular, $F'(y)^{-1} \geqslant 0$ and $A(x, y)^{-1} \geqslant 0$ for any $x \leqslant x < y \leqslant x$, then the interval sequence $\{X^k, k=0, 1, \cdots\}$ defined by Algorithm 1 converges to the unique solution x^* of (1) in X, and

$$W(X^{k+1}) = W(\overline{N}_k X^k), \quad k = 0, 1, \dots,$$
 (13)

where \overline{N}_k is the Newton operator $\overline{N}_k x = x - f'(\overline{x}^k)^{-1} f(x)$, for $\forall x \in [\underline{x}^k, \overline{x}^k]$. Moreover, if

$$||F'(y) - F'(x)|| \leq L||y - x|| \tag{14}$$

for $\forall x, y \in X$, then

$$||W(X^{k+1})|| \leq ||W(N_k X^k)|| \leq C||W(X^k)||^2, \tag{15}$$

where C is a positive constant.

Proof. First, we prove by induction that the interval sequence of (12) satisfies

$$X^{k+1} = [S_k x, \overline{S}^k \overline{x}^k] \subset \overline{N}_k X^k \subset X^k, \quad k = 0, 1, \dots,$$
 (16)

where $\overline{N}_k X^k = [\overline{N}_k x^k, \overline{N}_k \overline{x}^k]$, and

$$Pf(x^k) \leqslant 0 \leqslant Pf(\overline{x}^k), \quad k=0, 1, \cdots$$
 (17)

(16) has an equivalent form

$$x^{k} \leqslant \overline{N}_{k} x^{k} = S_{k} x^{k} = x^{k+1} \leqslant \overline{x}^{k+1} = \overline{S}_{k} x^{k} \leqslant \overline{N}_{k} x^{k} \leqslant \overline{x}^{k}, \quad k = 0, 1, \cdots.$$
 (16')

When k=0, (17) holds by assumption. By (8)—(11), we have

$$\begin{split} & \underline{x}^1 = S_0 \underline{x}^0 = \underline{x}^0 - [A(\underline{x}^0, \ \overline{x}^0)]^{-1} F(\underline{x}^0) \geqslant \underline{x}^0 - [Pf'(\overline{x}^0)]^{-1} Pf(\underline{x}^0) \geqslant \overline{N}_0 \underline{x}^0 \geqslant \underline{x}^0, \\ & \overline{x}^1 = \overline{S}_0 \overline{x}^0 = \overline{x}^0 - A(z^0, \ \overline{x}^0)^{-1} F(\overline{x}^0) \leqslant \overline{x}^0 - [Pf'(\overline{x}^0)]^{-1} Pf(\overline{x}^0) = \overline{N}_0 \overline{x}^0 \leqslant \overline{x}^0, \\ & \overline{x}^1 - \underline{x}^1 = \overline{x}^0 - \underline{x}^0 + A(\underline{x}^0, \ \overline{x}^0)^{-1} F(\underline{x}^0) - A(z^0, \ \overline{x}^0)^{-1} F(\overline{x}^0) \\ & = \overline{x}^0 - \underline{x}^0 - A(\underline{x}^0, \ \overline{x}^0)^{-1} [F(\overline{x}^0) - F(\underline{x}^0)] + [A(\underline{x}^0, \ \overline{x}^0)^{-1} - A(z^0, \ \overline{x}^0)^{-1}] F(\overline{x}^0) \\ & = A(\underline{x}^0, \ \overline{x}^0)^{-1} [I - A(\underline{x}^0, \ \overline{x}^0) A(z^0, \ \overline{x}^0)^{-1}] F(\overline{x}^0) \geqslant 0, \end{split}$$

i.e. (16) holds when k=0.

Now we suppose (17) and (16) hold for k, and we will prove them for k+1. By induction hypothesis and (8), (10), (11), we have

$$\begin{split} F(\underline{x}^{k+1}) &= F(\underline{x}^k) + A(\underline{x}^k, \ \underline{x}^{k+1}) \, (\underline{x}^{k+1} - \underline{x}^k) \\ &= F(\underline{x}^k) - A(\underline{x}^k, \ \underline{x}^{k+1}) \, A(\underline{x}^k, \ \overline{x}^k)^{-1} F(\underline{x}^k) \\ &= [I - A(\underline{x}^k, \ \underline{x}^{k+1}) \, A(\underline{x}^k, \ \overline{x}^k)^{-1}] \, F(\underline{x}^k) \leqslant 0, \\ F(\overline{x}^{k+1}) &= F(\overline{x}^k) + A(\overline{x}^{k+1}, \ \overline{x}^k) \, (\overline{x}^{k+1} - \overline{x}^k) \\ &= F(\overline{x}^k) - A(\overline{x}^{k+1}, \ \overline{x}^k) \, A(z^k, \ \overline{x}^k)^{-1} F(\overline{x}^k) \\ &= [I - A(\overline{x}^{k+1}, \ \overline{x}^k) \, A(z^k, \ \overline{x}^k)^{-1}] \, F(\overline{x}^k) \geqslant 0, \end{split}$$

i.e. (17) holds for k+1. By (9) and (10), we have

$$\begin{split} \underline{x}^{k+2} &= \underline{x}^{k+1} - A(\underline{x}^{k+1}, \ \overline{x}^{k+1})^{-1} F(\underline{x}^{k+1}) \geqslant \underline{x}^{k+1} - F'(\overline{x}^{k+1})^{-1} F(\underline{x}^{k+1}) = \overline{N}_{k+1} \underline{x}^{k+1} \geqslant \underline{x}^{k+1}, \\ \overline{x}^{k+2} &= \overline{x}^{k+1} - A(z^{k+1}, \ \overline{x}^{k+1})^{-1} F(\overline{x}^{k+1}) \leqslant \overline{x}^{k+1} - F'(\overline{x}^{k+1})^{-1} F(\overline{x}^{k+1}) = \overline{N}_{k+1} \overline{x}^{k+1} \leqslant \overline{x}^{k+1}, \\ \overline{x}^{k+2} - \underline{x}^{k+2} &= \overline{x}^{k+1} - \underline{x}^{k+1} - A(\underline{x}^{k+1}, \ \overline{x}^{k+1})^{-1} [F(\overline{x}^{k+1}) - F(\underline{x}^{k+1})] \\ &+ A(\underline{x}^{k+1}, \ \overline{x}^{k+1})^{-1} [I - A(\underline{x}^{k+1}, \ \overline{x}^{k+1}) A(z^{k+1}, \ \overline{x}^{k+1})^{-1}] F(\overline{x}^{k+1}) \geqslant 0, \end{split}$$

i.e. (16) holds for k+1. Therefore, (16) and (17) hold for any k. Thus

$$\overline{N}_{k}\underline{x}^{k} \leq \underline{x}^{k+1} \leq \overline{x}^{k+1} \leq \overline{N}_{k}\overline{x}^{k}$$
.

Therefore,

$$W(X^{k+1}) = \overline{x}^{k+1} - \underline{x}^{k+1} \leqslant \overline{N}_k \overline{x}^k - \overline{N}_k x^k = W(\overline{N}_k X^k).$$

By Theorem 3.2 in [1], we have

$$\lim_{k\to\infty}\, \overline{N}_k\underline{x}^k = \lim_{k\to\infty}\, \overline{N}_k\overline{x}^k = x^*.$$

Thus

$$\lim_{k\to\infty} x^{k+1} = \lim_{k\to\infty} \bar{x}^{k+1} = x^*,$$

i.e. $\{X^k\} \rightarrow x^*$ when $k \rightarrow \infty$.

If (14) holds, let $||F'(\underline{x}^0)^{-1}|| \leq \delta$. By the mean value theorem,

$$\begin{split} \| \overline{N}_{k} \overline{x}^{k} - \overline{N}_{k} \underline{x}^{k} \| &= \| F'(\overline{x}^{k})^{-1} (F'(\overline{x}^{k}) (\overline{x}^{k} - \underline{x}^{k}) - [F(\overline{x}^{k}) - F(\underline{x}^{k})]) \| \\ &= \delta \| F(\underline{x}^{k}) - F(\overline{x}^{k}) - F'(\overline{x}^{k}) (\underline{x}^{k} - \overline{x}^{k}) \| = \frac{1}{2} \| \delta L \| \overline{x}^{k} - \underline{x}^{k} \|^{2}. \end{split}$$

By (13), we get (15), where $C = \frac{1}{2} \delta L$ is a constant.

(15) indicates that the order interval secant method (12) converges faster than the order interval Newton method. In computation the auxiliary points z^k must be so chosen that $F(z^k) \ge 0$. Ordinarily $t_k \ll 1$, i.e. z^k is very close to \bar{x}^k .

It must be noticed in Theorem 1 that condition (11) is necessary. However, if f is P-isotope concave, the order interval secant method (12) has to be modified as follows.

Algorithm 2. Let
$$X^0 = [\underline{x}^0, \overline{x}^0] = [\underline{x}, \overline{x}] = X$$
. Define
$$X^{k+1} := [\underline{S}_k \underline{x}^k, S_k \overline{x}^k] = [\underline{x}^{k+1}, \overline{x}^{k+1}]$$
(18)

for $k=0, 1, \dots,$ where

The auxiliary point $z^k = \underline{x}^k + t_k(\overline{x}^k - \underline{x}^k)$, $0 < t_k \ll 1$, are so chosen that $F(z^k) \leq 0$.

We have, as similar to Theorem 1, the following convergence theorem about Algorithm 2.

Theorem 2. Let $f: X \subset \mathbb{R}^n \to \mathbb{R}^n$ is P-isotone concave on $X = [x, \overline{x}]$ and satisfy condition (11). If F = Pf is G-differentiable in X, F'(x) and the nth-order difference matrix A(x, y) of F are non-singular, $F'(x)^{-1} \geqslant 0$ and $A(x, y)^{-1} \geqslant 0$ for any $x \leqslant x \leqslant y \leqslant \overline{x}$, then the interval sequence $\{X^k, k=0, 1, \cdots\}$ define by Algorithm 2 converges to the unique solution x^* of (1) in X, and

$$W(X^{k+1}) = W(\underline{N}_k X^k),$$

where

$$N_k x = x - f'(x^k)^{-1} f(x), \quad k = 0, 1, \dots$$

The proof of this theorem is similar to that of Theorem 1. The two theorems turn out the sufficient condition for the existence and uniqueness of the solution of equation (1) in X. In actual computation it is required that the Jacobian matrix F'(x) of F=Pf be a monotone matrix (or $A(x, y)^{-1} \geqslant 0$ for any $x, y \in X$). However, condition (11) is in fact equivalent to $NX^0 \subset X^0$, where $Nx=x-A(x, x)^{-1}Pf(x)$ is a simple Newton operator. By Theorem 2.1 in [1], only if $NX^k \subset X^k$ for a certain k can the solution of (1) in X exist, i.e. condition (11) can be replaced by the following condition

$$Pf(x^k) \leq 0 \leq Pf(\bar{x}^k)$$
.

Now we compute an example, which is Example in [1], with the order interval secant method, and compare the computational results with those in [1],

Example.

$$f(x) = \begin{pmatrix} -x_1^3 + 5x_1^2 - x_1 + 2x_2 - 3 \\ x_2^3 + x_2^2 - 14x_2 - x_1 - 19 \end{pmatrix} = 0, \quad X^0 = \left[\begin{pmatrix} 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \end{pmatrix} \right].$$

Let $P = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, then $F = Pf = \begin{pmatrix} -f_1 \\ f_2 \end{pmatrix}$. It is easy to verify that $F'(x)^{-1} \ge 0$

and F is order convex in X^0 , and

$$F(\underline{x}^0) = \begin{pmatrix} -18 \\ -28 \end{pmatrix} \leqslant 0 \leqslant F(\overline{x}^0) = \begin{pmatrix} 35 \\ 55 \end{pmatrix}.$$

Therefore, the conditions of Theorem 1 are satisfied. We solve the equations with Algorithm 1. The computational results are given in Table 1, where the results of the interval Newton method are taken from [1].

Table 1 Comparison between the interval secant method and the interval Newton method

k	interval secant method $X^k = [\underline{x}^k, \overline{x}^k]$	interval Newton method $X^k = [\underline{x}^k, \overline{x}^k]$
0	$\left[\left(\begin{array}{c}3\\3\end{array}\right),\left(\begin{array}{c}6\\5\end{array}\right)\right]$	$\left[\left(\begin{smallmatrix}3\\3\end{smallmatrix}\right),\left(\begin{smallmatrix}6\\5\end{smallmatrix}\right)\right]$
1	$\begin{bmatrix} (3.99999999), & (5.1891260) \\ (3.49999992), & (4.1773150) \end{bmatrix}$	$\begin{bmatrix} \begin{pmatrix} 3.38358 \\ 3.39978 \end{pmatrix}, \begin{pmatrix} 5.25384 \\ 4.21482 \end{pmatrix} \end{bmatrix}$
2	$\begin{bmatrix} \begin{pmatrix} 4.81643659 \\ 3.96558983 \end{pmatrix}, \begin{pmatrix} 5.00515699 \\ 4.00574929 \end{pmatrix} \end{bmatrix}$	$\begin{bmatrix} \begin{pmatrix} 4.01412 \\ 3.81375 \end{pmatrix}, \begin{pmatrix} 5.02251 \\ 4.01346 \end{pmatrix} \end{bmatrix}$
3	$\begin{bmatrix} \binom{4.999610719}{3.999928998}, & \binom{5.00000741}{4.00000967} \end{bmatrix}$	$\left[\begin{pmatrix} 4.65040 \\ 3.97972 \end{pmatrix}, \begin{pmatrix} 5.00020 \\ 4.00006 \end{pmatrix}\right]$
4	$\begin{bmatrix} \begin{pmatrix} 4.99999999995 \\ 8.99999999999 \end{pmatrix}, \begin{pmatrix} 5.00000000001 \\ 4.0000000001 \end{pmatrix} \end{bmatrix}$	$\left[\binom{4.954496}{3.998790}, \binom{5.0000004}{4.0000047}\right]$

References

- [1] Li Qing-yang, Order interval test and iterative methods for nonlinear systems, J. of Computational Math., 2 (1984), 50-55.
- [2] J. M. Ortega, W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic press, New York, 1970, 448—454.