ON THE SOLUTION OF NONLINEAR EQUATIONS
IN INTERVAL ARITHMETIC

KAJ MADSEN

Abstract.

We consider the problem of finding a simple zero of a continuously differentiable
function $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$. There is given an interval vector $X_0$ containing one zero of $f$, and we will construct a contracting sequence of interval vectors enclosing this zero.

This can be done by Newton's method, which gives quadratic convergence, but requires inversion of an interval matrix at each step of the iteration. Alefeld and Herzberger, [1], give a modification of Newton's method, without the necessity of inversion, the convergence being superlinear. We give a slight modification of the latter method, with the property that the sequence of interval widths is dominated by a quadratically convergent sequence.

1.

First, we give some definitions concerning interval matrices and interval vectors. By an interval matrix (interval vector) we mean a matrix (vector) whose components are real-valued intervals. If $A = ([a_{ij}^{(1)}, a_{ij}^{(2)}])$ and $B = ([b_{ij}^{(1)}, b_{ij}^{(2)}])$ are interval matrices, and $A = (a_{ij})$ and $B = (b_{ij})$ are real matrices, we define

$$A \leq B \iff a_{ij} \leq b_{ij}$$

for all components

$$A \in A^I \iff a_{ij} \in [a_{ij}^{(1)}, a_{ij}^{(2)}]$$

$$A^I \subseteq B^I \iff [a_{ij}^{(1)}, a_{ij}^{(2)}] \subseteq [b_{ij}^{(1)}, b_{ij}^{(2)}]$$

If $\{A_k^I\}$ is a sequence of interval matrices, we define

$$A_k^I \rightarrow A^I \iff a_{ij}^{(s)} \rightarrow a_{ij}^{(s)}, \quad s = 1, 2,$$

for all components.

Further

$$A^I \cap B^I = ([a_{ij}^{(1)}, a_{ij}^{(2)}] \cap [b_{ij}^{(1)}, b_{ij}^{(2)}])$$

$$|A^I| = \max_{s=1,2} |a_{ij}^{(s)}|$$

$$d(A^I) = (a_{ij}^{(2)} - a_{ij}^{(1)})$$

Received February 9, 1973.
The sum and product of two interval matrices are
\[
A^I + B^I \equiv \left( [a^{(1)}_{ij}, a^{(2)}_{ij}] + [b^{(1)}_{ij}, b^{(2)}_{ij}] \right)
\]
\[
A^I \times B^I \equiv \left( \sum_k [a^{(1)}_{ik}, a^{(2)}_{ik}] \times [b^{(1)}_{kj}, b^{(2)}_{kj}] \right).
\]
Finally we define
\[
M(A^I) \equiv \left( (a^{(1)}_{ij} + a^{(2)}_{ij})/2 \right)
\]
and
\[
N(A, B^I) \equiv (b'_{ij})
\]
where
\[
|b'_{ij} - a_{ij}| = \min_{b_{ij} \in [b^{(1)}_{ij}, b^{(2)}_{ij}]} |b_{ij} - a_{ij}|
\]
i.e. \(N(A, B^I)\) is the matrix \(B \in B^I\) which is "nearest" to \(A\).

For interval vectors we make definitions similar to those mentioned above.

2.

Now consider the problem
\[
(1) \quad f(x) = 0
\]
where \(f: \mathbb{R}^n \rightarrow \mathbb{R}^n\) is continuously differentiable. If \(x^*\) is a solution to (1) and \(x \in \mathbb{R}^n\), we have by the mean value theorem
\[
(2) \quad 0 = f(x^*) = f(x) + J(x, x^*)(x^* - x)
\]
where
\[
J(x, x^*) = \left( \frac{\partial f_i}{\partial x_j} \mid x + \theta_i(x^* - x) \right), \quad 0 < \theta_i < 1.
\]
By means of (2) we can prove the following:

**Theorem 1.** Let \(f: \mathbb{R}^n \rightarrow \mathbb{R}^n\) be continuously differentiable. Let \(X^I_0\) be an interval vector with the property that \(x^*\) is the only zero of \(f\) in \(X^I_0\). Furthermore, suppose that \(J(x, x^*)\) and \(f'(x)\) are nonsingular for \(x \in X^I_0\), and let \(A^I_0\) be an interval matrix for which
\[
\{J^{-1}(x, x^*) \mid x \in X^I_0\} \subseteq A^I_0.
\]
Let \(\tilde{x}_0 \in X^I_0\), and define for \(k = 0, 1, 2, \ldots\) the sequences
\[
(3) \quad X^I_{k+1} = (\tilde{x}_k - A^I_k f(\tilde{x}_k)) \cap X^I_k
\]
\[
(4) \quad A^I_{k+1} = (A_k + A^I_k (E - H^I_k A_k)) \cap A^I_k
\]
\[
(5) \quad \tilde{x}_{k+1} = N\left( \tilde{x}_k - (f'(\tilde{x}_k))^{-1} f(\tilde{x}_k), X^I_{k+1} \right)
\]