<table>
<thead>
<tr>
<th>Title</th>
<th>Numerical Validation of Solutions of Nonlinear Complementarity Problems (Self-validating numerical methods and related topics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Alefeld, G.E.; Chen, Xiaojun; Potra, F.A.</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 2000, 1147: 67-77</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2000-04</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/63977">http://hdl.handle.net/2433/63977</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>
Numerical Validation of Solutions of Nonlinear Complementarity Problems

G. E. Alefeld*  X. Chen†  F. A. Potra‡

Abstract

This paper proposes a validation method for solutions of nonlinear complementarity problems.

1 Introduction

Let $f: R^n \rightarrow R^n$ be a continuous function. The nonlinear complementarity problem (NCP) is to find a vector $x \in R^n$ such that

$$x \geq 0, \ f(x) \geq 0, \ x^T f(x) = 0.$$  

The NCP models many important problems in engineering and economy. Moreover, the NCP is a fundamental problem for optimization theory since the first order necessary condition for an optimal point can be reformulated as an NCP.

In section 2, we present a slope for the numerical validation of the solution of NCP. In section 3 we give an interval arithmetic evaluation of the slope.

In this paper we denote an interval by $[x] = \{x \in R^n, \underline{x} \leq x \leq \overline{x}\}$. The nonnegative orthant of $R^n$ is denoted by $R^n_+$.  

2 The slope for NCP

It is well known and easy to verify that the NCP is equivalent to the following system of nonlinear equations

$$F(x) := \min(f(x), x) = 0,$$  

where the "min" operator denotes the componentwise minimum of two vectors.

*Institut für Angewandte Mathematik, Universität Karlsruhe, Kaiserstraße 12, D - 76128 Karlsruhe, Germany.
†Department of Mathematics and Computer Science, Shimane University, Matsue 690-8504, Japan.
‡Department of Mathematics, University of Maryland, Baltimore, U.S.
where the "min" operator denotes the componentwise minimum of two vectors.

The function $F$ is not differentiable at $x$ only if $f(x) = x$. Hence $F$ is differentiable on an interval $[x]$ if $f(x) > x$ for all $x \in [x]$ or $f(x) < x$ for all $x \in [x]$. In other words, if $F$ is differentiable on $[x]$, then

$$F(x) = f(x) \quad \text{if } f(x) < x \text{ for all } x \in [x]$$

or

$$F(x) = x \quad \text{if } f(x) > x \text{ for all } x \in [x].$$

Many existing algorithms for validation of solutions of the system of nonlinear equations assume that the involved function is continuously differentiable. Such algorithms are based on the mean value theorem for differentiable functions and an interval extension of the derivative. For instance, we suppose $F$ is differentiable on $[x]$. Then

$$F(x) - F(y) \in F'(x)(x - y), \quad \text{for all } x, y \in [x]. \quad (2.2)$$

The Krawczyk operator is defined by

$$K(x, A, [x]) = x - A^{-1}F(x) + (I - A^{-1}F'(x))[x] - x,$$

where $A$ is an $n \times n$ nonsingular matrix.

However, in general $F$ is nondifferentiable in an arbitrary interval. Recently, some methods have been proposed for general nondifferentiable equations [5, 20]. In this paper we give a sharp and computable interval operator for the special nondifferentiable system (2.1). Using this interval operator, we can verify the existence of solutions of the NCP numerically.

The first step is to define a slope $\delta F(x, y)$ for $F$ such that for a fixed $x \in [x]$

$$F(x) - F(y) = \delta F(x, y)(x - y), \quad \text{for all } y \in [x]. \quad (2.3)$$

We assume $f$ has a slope $\delta f$ on $[x]$ such that for a fixed $x \in [x]$

$$f(x) - f(y) = \delta f(x, y)(x - y), \quad \text{for all } y \in [x]. \quad (2.4)$$

Let us use the following notations

$$S_{i}^{+} = \{x \in [x] \mid f_i(x) > x_i\}$$

$$S_{i}^{-} = \{x \in [x] \mid f_i(x) < x_i\}$$

$$S_{i}^{0} = \{x \in [x] \mid f_i(x) = x_i\}$$

$$N = \{1, 2, \ldots, n\}.$$

For a vector $x \in [x]$ and an $i \in N$, $x$ is in one of the three sets. Hence for any two vectors $x, y \in [x]$ and an $i \in N$ nine cases can happen. We summarize the nine cases in Table 1.
Table 1: $\alpha_i = \frac{(y - f(y))_i}{(f(x) - f(y) - x + y)_i}$, $\beta_i = \frac{(f(x) - x)_i}{(f(x) - x + y - f(y))_i}$.

Lemma 2.1 Let the $i$th row of $\delta F(x, y)$ be defined by Table 1. Then for every two vectors $x, y \in [x],
F(x) - F(y) = \delta F(x, y)(x - y).

Proof: Let $i \in N$ be fixed. Since $f$ has a slope, it holds $f_i(x) - f_i(y) = \delta f_i(x, y)(x - y).$

Suppose $x \in S_i^+ \cup S_i^0$ and $y \in S_i^+ \cup S_i^0$. Then $F_i(x) - F_i(y) = x_i - y_i = e_i^T(x - y)$.

Suppose $x \in S_i^-$ and $y \in S_i^0 \cup S_i^-$ or $y \in S_i^-$ and $x \in S_i^0$. Then $F_i(x) - F_i(y) = f_i(x) - f_i(y) = \delta f_i(x, y)(x - y)$.

Suppose $x \in S_i^-$ and $y \in S_i^+$. Then $F_i(x) - F_i(y) = x_i - f_i(y)$.

Suppose $x \in S_i^+$ and $y \in S_i^-$. Then $F_i(x) - F_i(y) = x_i - f_i(y) + e_i^T(x - y)$.

\[
\begin{array}{|c|c|c|c|}
\hline
\delta F_i(x, y) & y \\
\hline
S_i^+ & e_i^T & \alpha_i(\delta f_i(x, y) - e_i^T) + e_i^T & e_i^T \\
\hline
S_i^- & \beta_i(\delta f_i(x, y) - e_i^T) + e_i^T & \delta f_i(x, y) & \delta f_i(x, y) \\
\hline
S_i^0 & e_i^T & \delta f_i(x, y) & e_i^T \\
\hline
\end{array}
\]
\[
\frac{(y_i - f_i(y))(\delta f_i(x, y) - e_i^T)(x - y)}{(\delta f_i(x, y) - e_i^T)(x - y)} + e_i^T(x - y)
\]

\[
= \left( \frac{y_i - f_i(y)}{(f(x) - f(y) - X_i + y_i)i} (\delta f_i(x, y) - e_i^T) + e_i^T \right)(x - y)
\]

\[
= (\alpha_i (\delta f_i(x, y) - e_i^T) + e_i^T)(x - y).
\]

### Lemma 2.2
In Table 1, we have

\[
\alpha_i \in (0, 1) \quad \text{and} \quad \beta_i \in (0, 1).
\]

**Proof:** Notice that \( \alpha_i \) is used when \( x \in S_i^+ \) and \( y_i \in S_i^- \). Then from \( y_i - f_i(y) > 0 \) and \( f_i(x) - x_i > 0 \), we have

\[
\alpha_i = \frac{y_i - f_i(y)}{(y_i - f_i(y)) + (f_i(x) - x_i)} \in (0, 1).
\]

Since \( \beta_i \) is used when \( x \in S_i^- \) and \( y_i \in S_i^+ \), \( f_i(x) - x_i < 0 \) and \( y_i - f_i(y) < 0 \). Hence we have

\[
\beta_i = \frac{f_i(x) - x_i}{(f_i(x) - x_i) + (y_i - f_i(y))} \in (0, 1).
\]

Now we study the nonsingularity of \( \delta F(x, y) \). The nonsingularity is dependent on the properties of \( \delta f(x, y) \).

An \( n \times n \) matrix \( A \) is called a \( P_0 \) matrix if all principal minors of \( A \) are nonnegative. A matrix \( A \) is called a \( P \) matrix if its all principal minors are positive [8]. By using two theorems on the \( P_0 \) matrix and the \( P \) matrix given by Gabriel and Moré [10], we have the following proposition.

### Proposition 2.1
1. If \( \delta f(x, y) \) is a \( P \) matrix, then \( \delta F(x, y) \) is nonsingular.

2. If \( \delta f(x, y) \) is a \( P_0 \) matrix and \( S_{i}^+ \) contains \( x \) or \( y \) for every \( i \in N \), then \( \delta F(x, y) \) is nonsingular.

**Proof:** 1. By Lemma 2.1 and Lemma 2.2, \( \delta F(x, y) \) can be written as

\[
\delta F(x, y) = I + D(\delta f(x, y) - I),
\]

where \( D = \text{diag}(d_i) \) is a diagonal matrix with \( 0 \leq d_i \leq 1 \). Hence by Theorem 4.4 in [10], \( \delta F(x, y) \) is nonsingular.

2. If \( S_{i}^+ \) contains \( x \) or \( y \) for every \( i \in N \), then

\[
\delta F(x, y) = I + D(\delta f(x, y) - I),
\]

where \( D = \text{diag}(d_i) \) is a diagonal matrix with \( 0 \leq d_i < 1 \). Hence by Theorem 4.3 in [10], \( \delta F(x, y) \) is nonsingular.
If \( f \) is an affine function, say \( f(x) = Ax + c \), then \( A = \delta f(x, y) \). For a \( P_0 \) matrix \( A \), by Proposition 2.1, if we choose \( x \in S_t^+ \), then \( \delta F(x, y) \) is nonsingular for all \( y \in \mathbb{R}^n \).

For a nonlinear function \( f \), we need the following definitions.

**Definition 1** A mapping \( f \) from an interval \([x]\) in \( \mathbb{R}^n \) into \( \mathbb{R}^n \) is said to be

1. a \( P_0 \) function on \([x]\) if for all \( x, y \in [x] \) with \( x \neq y \), there is an index \( i \) such that
   \[ x_i \neq y_i \quad \text{and} \quad (f_i(x) - f_i(y))(x_i - y_i) \geq 0; \]

2. a \( P \) function on \([x]\) if for all \( x, y \in [x] \) with \( x \neq y \), there is an index \( i \) such that
   \[ x_i \neq y_i \quad \text{and} \quad (f_i(x) - f_i(y))(x_i - y_i) > 0; \]

3. a uniform \( P \) function on \([x]\) if for some \( \gamma > 0 \)
   \[ \max_{i \in N}(f_i(x) - f_i(y))(x_i - y_i) \geq \gamma \|x - y\| \quad \text{for all} \ x, y \in [x]; \]

4. a monotone function on \([x]\) if for all \( x, y \in [x] \),
   \[ (f(x) - f(y))^T(x - y) \geq 0; \]

5. a strictly monotone function on \([x]\) if for all \( x, y \in [x] \),
   \[ (f(x) - f(y))^T(x - y) > 0; \]

6. a strongly monotone function if for some \( \gamma > 0 \)
   \[ (f(x) - f(y))^T(x - y) \geq \gamma \|x - y\| \quad \text{for all} \ x, y \in [x]. \]

It is easy to verify that every monotone function is a \( P_0 \) function, every strictly monotone function is a \( P \) function and every strongly monotone function is a uniform \( P \) function.

For a Fréchet differentiable function \( f \), the following results are known [11, 17].

1. If \( f'(x) \) is a \( P \) matrix for all \( x \in [x] \), then \( f \) is a \( P \) function on \([x]\).
2. If \( f \) is a uniform \( P \) function on \([x]\), then \( f'(x) \) is a \( P \) matrix for all \( x \in [x] \).
3. \( f \) is a \( P_0 \) function on \([x]\) if and only if \( f'(x) \) is a \( P_0 \) matrix for all \( x \in [x] \).

If \( f \) is Fréchet differentiable on \([x]\), by the mean value theorem, for any \( x, y \in [x] \), there is a diagonal matrix \( \Lambda = \text{diag}(\lambda_i) \) with \( \lambda_i \in [0, 1] \) such that

\[ f(x) - f(y) = f'(x + \Lambda(y - x))(x - y). \]

The following proposition is a direct corollary of Proposition 2.1.
Proposition 2.2 Suppose that \( f \) is Fréchet differentiable. Let \( \delta f(x, y) = f'(x + \Lambda(y - x)) \), where \( \Lambda = \text{diag}(\lambda_i) \) is a diagonal matrix with \( 0 \leq \lambda_i \leq 1 \).

1. If \( f \) is a uniform \( P \) function, then \( \delta F(x, y) \) is nonsingular.

2. If \( f \) is a \( P_0 \) function and \( S_i^+ \) contains \( x \) or \( y \) for every \( i \in N \), then \( \delta F(x, y) \) is nonsingular.

For a locally Lipschitzian function \( f \), Song, Gowda and Ravindran gave the following results [22].

Suppose \( f \) is semismooth. \( f \) is a \( P_0 \) function on \([x]\) if and only if the Bouligand subdifferential \( \partial_B f(x) \) consists of \( P_0 \) matrices at all \( x \in [x] \).

Notice that the mean value theorem does not hold for \( \partial_B f \). Moreover, for a \( P_0 \) function, the Clarke generalized Jacobian \( \partial f(x) = \text{co} \partial_B f(x) \) may consists of a matrix which is not \( P_0 \). Hence we generalize Proposition 2.2 to nondifferentiable monotone functions as follows.

Proposition 2.3 Suppose that \( f \) is a locally Lipschitzian function.

1. If \( f \) is a strongly monotone function, then there is a \( \delta f(x, y) \in \partial f(\overline{xy}) \) such that \( \delta F(x, y) \) is nonsingular.

2. If \( f \) is a monotone function and \( S_i^+ \) contains \( x \) or \( y \) for every \( i \in N \), then there is a \( \delta f(x, y) \in \partial f(\overline{xy}) \) such that \( \delta F(x, y) \) is nonsingular.

Here \( \text{co} \partial f(\overline{xy}) \) denotes the convex hull of all points \( Z \in \partial F(u) \) for \( u \in \overline{xy} \), and \( \overline{xy} \) denotes the line segment between \( x \) and \( y \).

Proof: 1. Since \( f \) is a locally Lipschitzian function, \( f \) is differentiable almost every where. Moreover at a point \( z \in [x] \) where \( f \) is differentiable, \( f'(z) \) is a strongly monotone matrix. By definition, the Clarke generalized Jacobian at \( y \) is defined by

\[
\partial f(y) = \text{co} \{ \lim_{k \to \infty} f'(z^k) : z^k \to y, \ f \text{ is differentiable at } z^k \}.
\]

Since \( f'(z^k) \) is a strongly monotone matrix, the limit \( \lim_{z^k \to y} f'(z^k) \) is a strongly monotone matrix. Moreover, the convex combination of strongly monotone matrices is still a strongly monotone matrix.

By Proposition 2.6.5 in [6], there is a matrix \( \delta f(x, y) \in \text{co} \partial f(\overline{xy}) \) such that

\[
f(x) - f(y) = \delta f(x, y)(x - y).
\]

Since a strongly monotone matrix is a \( P \) matrix, by Proposition 2.1, \( \delta F(x, y) \) is nonsingular.

The proof for Part 2 is similar.
3 Interval extension

We assume that $f$ has an interval arithmetic evaluation of the slope $\delta f(x, [x])$ for fixed $x \in [x]$ and all $y \in [x]$.

To define an interval arithmetic evaluation for $F$, we fix $x \in [x]$ and consider the following nonlinear programming problems

$$\min \quad y_i - f_i(y)$$
$$\text{s.t.} \quad y \in [x]$$

and

$$\max \quad y_i - f_i(y)$$
$$\text{s.t.} \quad y \in [x].$$

Let $y_i^{i,1}$ and $y_i^{i,2}$ be solutions of the nonlinear programming problems (3.1) and (3.2), respectively. Let

$$\alpha_i = \frac{(y_i^{i,2} - f(y_i^{i,2}))_i}{(f(x) - x + y_i^{i,2} - f(y_i^{i,2}))_i} \quad \text{if} \quad (f(x) - x + y_i^{i,2} - f(y_i^{i,2}))_i \neq 0$$

and

$$\beta_i = \frac{(f(x) - x)_i}{(f(x) - x + y_i^{i,1} - f(y_i^{i,1}))_i} \quad \text{if} \quad (f(x) - x + y_i^{i,1} - f(y_i^{i,1}))_i \neq 0.$$

Then we can define the interval arithmetic evaluation by

$$\delta F_i(x, [x]) = \{ e_i^T, y_i^{i,2} \in S_i^+ \cup S_i^0 \}$$
$$\delta f_i(x, [x]), \quad [0, \alpha_i](\delta f_i(x, [x]) - e_i^T) + e_i^T, \quad x \in S_i^+ \cup S_i^0, y_i^{i,2} \in S_i^-$$
$$[\beta_i, 1](\delta f_i(x, [x]) - e_i^T) + e_i^T, \quad x \in S_i^-, y_i^{i,1} \in S_i^+.$$

Theorem 3.1 For a fixed $x \in [x]$, we have

$$F(x) - F(y) \in \delta F(x, [x])(x - y), \quad \text{for all} \quad y \in [x].$$

Proof: Suppose $y_i^{i,2} \in S_i^+ \cup S_i^{-}$. Then for all $y \in [x],$

$$y_i - f_i(y) \leq y_i^{i,2} - f_i(y_i^{i,2}) \leq 0.$$

That is, $y \in S_i^+ \cup S_i^0$ for all $y \in [x]$, so $x$. Hence

$$F_i(x) - F_i(y) = x_i - y_i = e_i^T(x - y) = \delta F_i(x, [x])(x - y).$$

Suppose $y_i^{i,1} \in S_i^{-} \cup S_i^0$. Then for all $y \in [x],$

$$y_i - f_i(y) \geq y_i^{i,1} - f_i(y_i^{i,1}) \geq 0.$$
That is, $y \in S_i^- \cup S_i^0$ for all $y \in [x]$, so $x$. Hence

$$F_i(x) - F_i(y) = f_i(x) - f_i(y)$$
$$= \delta f_i(x, y)(x - y)$$
$$\in \delta f(x, [x])(x - y)$$
$$= \delta F_i(x, [x])(x - y).$$

Suppose $y^{i,2} \in S_i^-$ and $x \in S_i^+ \cup S_i^0$. Let $y \in [x]$. If $y \in S_i^+ \cup S_i^0$, then

$$F_i(x) - F_i(y) = e_i^T(x - y)$$
$$\in ([0, \alpha_i] \delta f_i(x, [x]) - e_i^T + e_i^T)(x - y)$$
$$= \delta F_i(x, [x])(x - y),$$

where we use $0 < \alpha_i \leq 1$. If $y \in S_i^-$, then by Lemma 2.1, we have

$$F_i(x) - F_i(y) = \left( \frac{y_i - f_i(y)}{(f(x) - f(y) - x + y)_i} (\delta f_i(x, y) - e_i^T + e_i^T)(x - y) \right).$$

Since $y^{i,2}$ is an optimal solution of (3.2), we have

$$0 < \frac{y_i - f_i(y)}{(f(x) - f(y) - x + y)_i} \leq \frac{y^{i,2} - f_i(y^{i,2})}{(f(x) - f(y^{i,2}) - x + y^{i,2})_i} = \alpha_i \leq 1.$$

Therefore, we have

$$F_i(x) - F_i(y) \in ([0, \alpha_i] \delta f_i(x, [x]) - e_i^T + e_i^T)(x - y) = \delta F_i(x, [x])(x - y).$$

Similarly, we can prove this theorem for the case $x \in S_i^-$, $y^{i,1} \in S_i^+$. \hfill \square

**Remark 3.1** In some case, the nonlinear programming problems with box constraints (3.1) and (3.2) are easy to compute. For example, if $f$ is an affine function. However, if one does want to spend time on computing problems (3.1) and (3.2) for a sharp interval arithmetic evaluation, the following interval arithmetic evaluation can be considered as a simple but overestimated interval arithmetic evaluation:

$$\hat{G}(x, [x]) = [0, 1] \delta f(x, [x]) - I + I.$$

Following the discussion above, we can show

$$\delta F(x, [x]) \subseteq \hat{G}(x, [x])$$

and

$$F(x) - F(y) \in \hat{G}(x, [x])(x - y) \text{ for all } x, y \in [x].$$

**Proposition 3.1** 1. If $\delta f(x, [x])$ consists of $P$ matrices at all $y \in [x]$, then every element in $\hat{G}(x, [x])$ is nonsingular.
2. If $\delta f(x, [x])$ consists of $P_0$ matrices at all $y \in [x]$ and $x \in S_i^+$ for all $i \in N$, then every element in $\delta F(x, [x])$ is nonsingular.

Proof: The proof for the part 1 is similar as the proof for part 1 of Proposition 2.1. For part 2, by Theorem 4.3 in [11], we only need to show $\delta f_i(x, [x])$ is not in $\delta F_i(x, [x])$ for every $i \in N$.

Since $x \in S_i^+$ and $y^{i,1} \in S_i^+$. Hence

$$\delta F_i(x, [x]) \neq \delta f_i(x, [x])$$

and

$$\delta F_i(x, [x]) \neq [\beta_i, 1](\delta f_i(x, [x]) - e_i^T) + e_i^T$$

This implies that

$$\delta F_i(x, [x]) = [0, \alpha_i](\delta f_i(x, [x]) - e_i^T) + e_i^T,$$

where $\alpha_i$ is a number between 0 and 1. Now we show $\alpha \neq 1$.

If $y^{i,2} \in S_i^+ \cup S_i^0$, then $\alpha_i = 0$.

If $y^{i,2} \in S_i^-$, then from $x \in S_i^+$, we have

$$f_i(x) - x_i > 0$$

and

$$0 < \alpha_i = \frac{y^{i,2} - f_i(y^{i,2})}{(f(x) - f(y^{i,2}) - x + y^{i,2})_i} < 1.$$  

We complete the proof.

4 Algorithm

Based on the results in [3, 5], we propose the following verification method.

Algorithm 1 Let $r > 0$ be a given tolerance and let $x$ be an approximate solution of the system (2.1). Calculate

$$[x] = (x + r[-e, e]) \cap R_+ \quad (4.1)$$

where $e = [1, \ldots, 1]^T$ and choose a nonsingular matrix $A$. Compute

$$L(x, A, [x]) = x - A^{-1}F(x) + (I - A^{-1}\delta F(x, [x]))([x] - x). \quad (4.2)$$

- If

$$L(x, A, [x]) \subseteq [x], \quad (4.3)$$

then there is a solution $x^* \in [x]$ of (2.1).

- If

$$L(x, A, [x]) \cap [x] = \emptyset, \quad (4.4)$$

then the interval $[x]$ contains no solution of (2.1).

We will give examples and report numerical results in a new version of this paper.
References


