Applications of Mathematics

Günter Mayer; Lars Pieper

A necessary and sufficient criterion to guarantee feasibility of the interval Gaussian algorithm for a class of matrices

Applications of Mathematics, Vol. 38 (1993), No. 3, 205–220

Persistent URL: http://dml.cz/dmlcz/104547

A NECESSARY AND SUFFICIENT CRITERION TO GUARANTEE FEASIBILITY OF THE INTERVAL GAUSSIAN ALGORITHM FOR A CLASS OF MATRICES

GÜNTER MAYER, LARS PIEPER, Karlsruhe

Dedicated to Professor Dr. Erich Martensen on the occasion of his 65^{th} birthday

(Received June 26, 1992)

Summary. A Necessary and Sufficient Criterion to Guarantee Feasibility of the Interval Gaussian Algorithm for a Class of Matrices. We apply the interval Gaussian algorithm to an $n \times n$ interval matrix $[A]$ the comparison matrix $\langle [A] \rangle$ of which is irreducible and diagonally dominant. We derive a new necessary and sufficient criterion for the feasibility of this method extending a recently given sufficient criterion.

Keywords: linear interval equations, Gaussian algorithm, interval Gaussian algorithm, linear systems of equations, criteria of feasibility

AMS classification: 65F05, 65G10

1. INTRODUCTION

Starting with an $n \times n$ interval matrix $[A]$ and a corresponding interval vector $[b]$, the interval Gaussian algorithm produces an interval vector $[x]^G = \text{IGA}([A], [b])$ which contains the solution set

$$\mathbf{S} := \{\tilde{x} \in \mathbf{R}^n \mid \exists \tilde{A} \in [A], \tilde{b} \in [b] \colon \tilde{A}\tilde{x} = \tilde{b}\}.$$

The formulae to describe $[x]^G$ are given analogously to those of the well-known Gaussian elimination process in non-interval analysis, combined with a forward and backward substitution (cf. Section 3). Unlike the non-interval case, $[x]^G$ does not necessarily exist if $[A]$ is nonsingular[1]—even if pivoting is taken into account (cf. [9], [11]). Therefore one looks for classes of interval matrices for which the feasibility of the method is guaranteed. An overview of such classes can be found in [7]. In [5], a sufficient criterion for the existence of $[x]^G$ has been derived if $[A]$ has an irreducible and diagonally dominant comparison matrix $\langle [A] \rangle$ defined in Section 2. Generally such interval matrices are no more H-matrices (cf. Section 2 again) as can be seen by the example

$$[A] = \begin{pmatrix} [1,1] & [1,1] \\ [1,1] & [1,1] \end{pmatrix}.$$

This example also shows that $[x]^G$ may no longer exist.

In the present paper, we modify the criterion from [5] this time ending up with a necessary and sufficient one.

2. Preliminaries

By $\mathbf{R}^n$, $\mathbf{R}^{n\times n}$, $\mathbf{IR}$, $\mathbf{IR}^n$, $\mathbf{IR}^{n\times n}$ we denote the set of real vectors with n components, the set of real $n \times n$ matrices, the set of intervals, the set of interval vectors with n components and the set of $n \times n$ interval matrices, respectively. By 'interval' we always mean a real compact interval. We write interval quantities in brackets with the exception of point quantities (i.e., degenerate interval quantities) which we identify with the element they contain. Examples are the null matrix O and the identity matrix I. We use the notation $[A] = [\underline{A}, \overline{A}] = ([a]_{ij}) = ([\underline{a}_{ij}, \overline{a}_{ij}]) \in \mathbf{IR}^{n\times n}$ simultaneously without further reference, and we proceed similarly for the elements of $\mathbf{R}^n, \mathbf{R}^{n\times n}, \mathbf{IR}$ and $\mathbf{IR}^n$.

By $A \geqslant 0$ we denote a nonnegative $n \times n$ matrix, i.e. $a_{ij} \geqslant 0$ for $i, j = 1, \ldots, n$. We call $x \in \mathbf{R}^n$ positive writing $x > 0$ if $x_i > 0$, $i = 1, \ldots, n$.

We also mention the standard notation from interval analysis ([2], [10])

$$\begin{array}{rll} \check{a} := \mathrm{mid}([a]) := \frac{\underline{a}+\overline{a}}{2} & & \text{(midpoint)} \\ \mathrm{rad}([a]) := \frac{\overline{a}-\underline{a}}{2} & & \text{(radius)} \\ |[a]| := \max\{|\tilde{a}| \mid \tilde{a} \in [a]\} = \max\{|\underline{a}|, |\overline{a}|\} & & \text{(absolute value)} \\ \langle [a] \rangle := \min\{|\tilde{a}| \mid \tilde{a} \in [a]\} = \begin{cases} \min\{|\underline{a}|, |\overline{a}|\} & \text{if } 0 \notin [a] \\ 0 & \text{otherwise} \end{cases} & & \text{(minimal absolute value)} \\ \mathrm{int}([a]) := (\underline{a}, \overline{a}) & & \text{(interior)} \end{array}$$

[1] An $n \times n$ interval matrix $[A]$ is termed *nonsingular* if all real $n \times n$ matrices $\tilde{A} \in [A]$ are nonsingular.

for intervals $[a]$. For $[A] \in \mathbf{IR}^{n\times n}$ we obtain $\check{A}$, $|[A]| \in \mathbf{R}^{n\times n}$ by applying $\check{}$ and $|\cdot|$ entrywise and we define the comparison matrix $\langle [A]\rangle = (c_{ij}) \in \mathbf{IR}^{n\times n}$ by setting

$$c_{ij} := \begin{cases} -|[a]_{ij}| & \text{if } i \neq j \\ \langle [a_{ii}]\rangle & \text{if } i = j \end{cases} .$$

Since real quantities can be viewed as degenerate interval ones, $|\cdot|$ and $\langle\cdot\rangle$ can also be used for them. We call $[a] \in \mathbf{IR}^n$ *symmetric* (with respect to zero) if $[a] = -[a]$, i.e. if $[a] = [-|[a]|, |[a]|]$. Note that $\check{a} = 0$ if and only if $[a]$ is symmetric.

We term $[A] \in \mathbf{IR}^{n\times n}$ an *H-matrix* if $\langle [A]\rangle = (c_{ij})$ is an M-matrix, i.e. if $\langle [A]\rangle$ is nonsingular with $\langle [A]\rangle^{-1} \geqslant 0$.

We define $A \in \mathbf{R}^{n\times n}$ to be ***diagonally dominant*** if for $i = 1, \ldots, n$

(1)
$$|a_{ii}| \geqslant \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|,$$

i.e. $\langle A\rangle e \geqslant 0$; here and in the sequel the vector e is defined by $e := (1, \ldots, 1)^T \in \mathbf{R}^n$. According to [12], A is called ***irreducibly diagonally dominant*** if it is irreducible and diagonally dominant with strict inequality in (1) for at least one index i.

We equip $\mathbf{IR}$, $\mathbf{IR}^n$, $\mathbf{IR}^{n\times n}$ with the usual real interval arithmetic as described e.g. in [2], [3], [8], [10]. Thus, we define the addition, subtraction, multiplication and division of two intervals $[a]$, $[b]$ by

$$[a] + [b] := [\underline{a} + \underline{b}, \overline{a} + \overline{b}],\ \ [a]\cdot[b] := \{\min P, \max P\} \text{ where } P := \{\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}\},$$
$$[a] - [b] := [\underline{a} - \overline{b}, \overline{a} - \underline{b}],\ \ [a]/[b] := [a]\cdot[\tfrac{1}{\overline{b}}, \tfrac{1}{\underline{b}}], \text{ provided } 0 \notin [b] .$$

We set $[x] + [y] := ([x]_i + [y]_i)$ and $[x]^T \cdot [y] := \sum_{i=1}^{n} [x]_i [y]_i$ for the addition and for the scalar product of two vectors $[x], [y] \in \mathbf{IR}^n$, and we proceed similarly for matrices.

We assume that the reader is familiar with the elementary rules of the interval arithmetic.

We now mention some formulae given in [2], [10] for intervals $[a]$, $[b]$:

$$
(2)\qquad \left.\begin{array}{l}
|[a]| = |\check{a}| + \mathrm{rad}([a]) \\
\langle [a]\rangle = |\check{a}| - \mathrm{rad}([a]), \quad \text{if } 0 \notin \mathrm{int}([a]) \\
|[a] \pm [b]| \leqslant |[a]| + |[b]| \\
|[a]\cdot[b]| = |[a]|\cdot|[b]| \\
\langle [a] \pm [b]\rangle \geqslant \langle [a]\rangle - |[b]| \\
\left|\frac{[a]}{[b]}\right| = \frac{|[a]|}{\langle [b]\rangle}, \quad \text{if } 0 \notin [b]. \\
\mathrm{rad}([a] \pm [b]) = \mathrm{rad}([a]) + \mathrm{rad}([b]) \\
\mathrm{mid}([a] \pm [b]) = \check{a} \pm \check{b} \\
\mathrm{mid}([a][b]) = \check{a}\check{b} + \mathrm{sign}(\check{a}\check{b})\times \\
\quad \times \min\{\mathrm{rad}([a])|\check{b}|,\ |\check{a}|\,\mathrm{rad}([b]), \mathrm{rad}([a])\,\mathrm{rad}([b])\ \}, \\
\text{hence } \mathrm{sign}(\mathrm{mid}([a][b])) = \mathrm{sign}(\check{a}\check{b}) = \mathrm{sign}(\check{a})\cdot\mathrm{sign}(\check{b}).
\end{array}\right\}
$$

A simple calculation yields

$$
\mathrm{mid}\left(\frac{1}{[b]}\right) = \frac{\check{b}}{|[b]|\langle [b]\rangle}, \qquad \text{if } 0 \notin [b],
$$

which implies

$$
(3)\qquad \mathrm{sign}\left(\frac{[a]}{[b]}\right) = \mathrm{sign}(\check{a}\check{b}) = \mathrm{sign}(\check{a})\cdot\mathrm{sign}(\check{b}).
$$

With these tools we can prove the following lemma which was stated in a modified form already in [5] and which was proved there in an abbreviated way.

Lemma 2.1. *Let* $[a],[b],[c],[d] \in \mathbf{IR}$ *with* $0 \notin [d]$. *Define* $[a]' \in \mathbf{IR}$ *by* $[a]' := [a] - \frac{[b]\cdot[c]}{[d]}$ *and* Δ *by* $\Delta := \mathrm{sign}(\check{a})\,\mathrm{sign}(\check{b})\,\mathrm{sign}(\check{c})\,\mathrm{sign}(\check{d})$. *Then the following assertions hold:*

(a) *The equation*

$$
(4)\qquad |[a]'| = |[a]| + \frac{|[b]|\cdot|[c]|}{\langle [d]\rangle}
$$

is valid if and only if

$$
(5)\qquad \Delta \neq 1, \quad \text{i.e. } \Delta \leqslant 0.
$$

In this case

$$
(6)\qquad \mathrm{sign}(\check{a}') = \begin{cases} \mathrm{sign}(\check{a}), & \text{if } \check{a} \neq 0 \\ -\mathrm{sign}(\check{b})\,\mathrm{sign}(\check{c})\,\mathrm{sign}(\check{d}), & \text{if } \check{a} = 0. \end{cases}
$$

(b) *If we have*

$$\langle[a]\rangle \geqslant \frac{|[b]||[c]|}{\langle[d]\rangle}, \tag{7}$$

then

$$\langle[a]'\rangle = \langle[a]\rangle - \frac{|[b]||[c]|}{\langle[d]\rangle} \tag{8}$$

is true if and only if

$$\Delta \neq -1, \quad \text{i.e. } \Delta \geqslant 0. \tag{9}$$

In this case

$$\operatorname{sign}(\breve{a}') = \operatorname{sign}(\breve{a}), \quad \textit{provided } [a]' \neq 0. \tag{10}$$

Proof. Let $[t] := \frac{[b][c]}{[d]}$. Then $|[t]| = \frac{|[b]||[c]|}{\langle[d]\rangle}$ by (2).

(a) By (2) we get

$$\begin{aligned} |[a]'| &= |\breve{a}'| + \operatorname{rad}([a]') \\ &= |\breve{a} - \breve{t}| + \operatorname{rad}([a]) + \operatorname{rad}([t]) \\ &= |\breve{a} - \breve{t}| - (|\breve{a}| + |\breve{t}|) + |[a]| + |[t]|. \end{aligned}$$

Therefore (4) holds if and only if

$$|\breve{a} - \breve{t}| = |\breve{a}| + |\breve{t}|,$$

i.e.,

$$\operatorname{sign}(\breve{a}) = 0 \text{ or } \operatorname{sign}(\breve{t}) = 0 \quad \text{or} \quad \operatorname{sign}(\breve{a}) = -\operatorname{sign}(\breve{t}) \neq 0. \tag{11}$$

By (2) and (3) we obtain $\Delta = \operatorname{sign}(\breve{a})\operatorname{sign}(\breve{t})$, hence (11) implies $\Delta \neq 1$, and conversely.

Again by (2) we get

$$\operatorname{sign}(\breve{a}') = \operatorname{sign}(\breve{a} - \breve{t});$$

thus (11) proves (6).

(b) $[t] = 0$ implies $[a]' = [a]$, and (8)–(10) hold. Therefore let $[t] \neq 0$. Together with (7) this implies

$$\langle[a]\rangle \geqslant |[t]| > 0, \tag{12}$$

hence $0 \notin [a]$ and ($\underline{a} \geqslant |[t]|$ or $-\overline{a} \geqslant |[t]|$); thus $0 \notin \mathrm{int}([a] - [t]) = \mathrm{int}([a]')$. By (2) we get

$$\begin{aligned}\langle [a]' \rangle &= |\check{a}'| - \mathrm{rad}([a]') = |\check{a} - \check{t}| - (\mathrm{rad}([a]) + \mathrm{rad}([t]))\\ &= |\check{a} - \check{t}| - (|\check{a}| - |\check{t}|) + \langle [a] \rangle - |[t]|.\end{aligned}$$

Therefore (4) holds if and only if

$$|\check{a} - \check{t}| = |\check{a}| - |\check{t}|,$$

i.e. $\mathrm{sign}(\check{a}) = \mathrm{sign}(\check{t})$, since $|\check{a}| \geqslant |\check{t}| > 0$ by (12). This implies (9), and conversely. Inequality (12) and $[a]' \neq 0$ yield $|\check{a}| > |\check{t}|$, hence

$$\mathrm{sign}(\check{a}') = \mathrm{sign}(\check{a} - \check{t}) = \mathrm{sign}(\check{a}).$$

□

It is obvious that $\mathrm{sign}(\check{a})$ determines whether $|[a]| = \underline{a}$ or $|[a]| = \overline{a}$.

3. Results

Mainly for reasons of notation, we start this section by recalling the formulae describing the interval Gaussian algorithm.

Let $[A]^{(k)} \in \mathbf{IR}^{n\times n}$, $[b]^{(k)} \in \mathbf{IR}^n$, $k = 1, \ldots, n$, and let $[x]^G \in \mathbf{IR}^n$ be defined by

$$[A]^{(1)} := [A], \quad [b]^{(1)} := [b]$$

for $k = 1, \ldots, n-1$,

$$[a]_{ij}^{(k+1)} := \begin{cases} [a]_{ij}^{(k)} & i = 1, \ldots, k; j = 1, \ldots, n,\\ [a]_{ij}^{(k)} - \frac{[a]_{ik}^{(k)} \cdot [a]_{kj}^{(k)}}{[a]_{kk}^{(k)}} & i = k+1, \ldots, n;\ j = k+1, \ldots, n,\\ 0 & \text{otherwise}\end{cases}$$

$$[b]_i^{(k+1)} := \begin{cases} [b]_i^{(k)} & i = 1, \ldots, k,\\ [b]_i^{(k)} - \frac{[a]_{ik}^{(k)}}{[a]_{kk}^{(k)}} \cdot [b]_k^{(k)} & i = k+1, \ldots, n,\end{cases}$$

$$[x]_i^G := \Big([b]_i^{(n)} - \sum_{j=i+1}^{n} [a]_{ij}^{(n)} [x]_j^G\Big) / [a]_{ii}^{(n)}, \quad i = n(-1)1,$$

where $\sum_{j=n+1}^{n} \ldots := 0$. Note that $[x]^G$ is defined without permuting rows or columns. The construction of $[x]^G$ is called the *interval Gaussian algorithm*. It is feasible for any $[b] \in \mathbf{IR}^n$ if and only if $0 \notin [a]_{kk}^{(k)}$, $k = 1, \ldots, n$.

This can be guaranteed for the following two classes of matrices.

Theorem 3.1. *Let* $[b] \in \mathbf{IR}^n$ *and let* $[A] \in \mathbf{IR}^{n\times n}$ *be an H-matrix. Then* $[x]^G$ *exists.*

P r o o f. See [1]. □

Corollary 3.2. *Let* $[b] \in \mathbf{IR}^n$, $[A] \in \mathbf{IR}^{n\times n}$. *If* $\langle [A] \rangle$ *is irreducibly diagonally dominant then* $[A]$ *is an H-matrix, hence* $[x]^G$ *exists.*

P r o o f. See [5] or [7]. □

If $\langle [A] \rangle$ is weakened to an irreducible and diagonally dominant matrix, Corollary 3.2 needs no longer be true as was pointed out in Section 1 by a singular point matrix. But even if $[A]$ is a nonsingular interval matrix satisfying $\langle [A] \rangle e = 0$, the feasibility of the interval Gaussian algorithm may fail as the example

$$[A] := \begin{pmatrix} 2 & 1 & -1 \\ [-1,1] & 2 & -1 \\ [-1,1] & -1 & 2 \end{pmatrix}$$

shows; see [5] for details. In the same paper the following sufficient criterion was proved.

Theorem 3.3. *Let* $[A] \in \mathbf{IR}^{n\times n}$, $n \geqslant 2$, *and let* $[b] \in \mathbf{IR}^n$. *Assume that* $\langle [A] \rangle$ *is irreducible and diagonally dominant. If*

$$\text{sign}(\check{a}_{ij}) \cdot \text{sign}(\check{a}_{ik}) \cdot \text{sign}(\check{a}_{kj}) = \begin{cases} \text{sign}(\check{a}_{kk}), & \text{if } i \neq j \\ -\text{sign}(\check{a}_{kk}), & \text{if } i = j \end{cases} \tag{13}$$

for at least one triple (i,j,k) *satisfying* $k < i, j$, *then* $[x]^G$ *exists.*

Unfortunately, criterion (13) is not a necessary one as the example

$$[A] = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 2 & 2 \\ 2 & 0 & [2,4] \end{pmatrix} \tag{14}$$

shows. Difficulties arise if $\check{a}_{ij} = 0$. These zeros may vanish in a later stage of the algorithm. Therefore we will associate with $[A]$ an extended sign matrix S' which takes this change into account.

Definition 3.4. Let $[A] \in \mathbf{IR}^{n\times n}$ and assume that the matrices $[A]^{(k)}$, $k = 1, \ldots, n$, of the interval Gaussian algorithm exist. Define the sign matrices $S, S^{(k)} \in \mathbf{R}^{n\times n}$ by

$$s_{ij} := \text{sign}(\check{a}_{ij})$$
$$s_{ij}^{(k)} := \text{sign}(\text{mid}([a]_{ij}^{(k)})).$$

Then the extended sign matrix S' is defined by the process

$$
\begin{aligned}
&S' := S\\
&\quad \text{for } k := 1 \text{ to } n-1 \text{ do}\\
&\quad\quad \text{for } i := k+1 \text{ to } n \text{ do}\\
&\quad\quad\quad \text{for } j := k+1 \text{ to } n \text{ do}\\
&\quad\quad\quad\quad \text{if } s'_{ij} = 0 \text{ then } s'_{ij} := -s'_{ik}s'_{kk}s'_{kj}\,.
\end{aligned}
$$

S' has been defined in view of (6) and (10) in Lemma 2.1. How it is related to $S^{(k)}$ in special cases will be explained later on. With its definition we are able to formulate the following necessary and sufficient criterion for the existence of $[x]^G$.

Theorem 3.5. *Let $[A] \in \mathbf{IR}^{n\times n}$ and let $\langle [A]\rangle$ be irreducible satisfying $\langle [A]\rangle e = 0$. Then $[x]^G$ exists for any vector $[b] \in \mathbf{IR}^n$ if and only if*

$$s'_{ij}s'_{ik}s'_{kk}s'_{kj} = \begin{cases} 1, & \text{if } i \neq j \\ -1, & \text{if } i = j \end{cases} \tag{15}$$

holds for at least one triple (i, j, k) with $k < i, j$ and with S' as in Definition 3.4.

To prove Theorem 3.5 we need some auxiliary results.

Definition 3.6. Let $[A] \in \mathbf{IR}^{n\times n}$ and let $[A]^{(k)}$ exist. Then the matrix $\pi_k([A]) \in \mathbf{R}^{(n-k+1)\times(n-k+1)}$ is defined by

$$\pi_k([A]) := \begin{pmatrix} [a]_{kk}^{(k)} & \ldots & [a]_{kn}^{(k)} \\ \vdots & \ldots & \\ [a]_{nk}^{(k)} & \ldots & [a]_{nn}^{(k)} \end{pmatrix}$$

Lemma 3.7. *Let $[A] \in \mathbf{IR}^{n\times n}$, $n \geqslant 2$, $0 \notin [a]_{11}$. Then the following assertions hold:*

(a) $\langle \pi_2([A])\rangle \geqslant \pi_2(\langle [A]\rangle)$,

(b) $\langle [A]^{(2)}\rangle \geqslant \langle [A]\rangle^{(2)}$,

(c) $\langle [A] \rangle e \begin{Bmatrix} \geqslant \\ > \\ = \end{Bmatrix} 0$ *implies* $\langle [A]^{(2)} \rangle e \geqslant \langle [A] \rangle^{(2)} e \begin{Bmatrix} \geqslant \\ > \\ = \end{Bmatrix} 0.$

(d) *If* $\langle [A] \rangle e \geqslant 0$ *and if* $(\langle [A] \rangle e)_i > 0$ *for some index* $i \in \{1, \ldots, n\}$ *then* $(\langle [A]^{(2)} \rangle e)_i > 0$.

Proof. (a)–(c) are proved in [5].

(d) follows by considering the i-th component of the inequality

$$\langle [A]^{(2)} \rangle e \geqslant \langle [A] \rangle^{(2)} e = L_{\langle [A] \rangle} (\langle [A] \rangle e) \geqslant I \cdot \langle [A] \rangle e. \tag{16}$$

This inequality holds with

$$L_{\langle [A] \rangle} := \begin{pmatrix} 1 & & \\ \frac{|[a]_{21}|}{\langle [a]_{11} \rangle} & 1 & 0 \\ \vdots & & \ddots \\ \frac{|[a]_{n1}|}{\langle [a]_{11} \rangle} & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

using (b) and $\langle [A] \rangle^{(2)} = L_{\langle [A] \rangle} \cdot \langle [A] \rangle$. □

Applying Lemma 3.7 to the matrices $\pi_l([A])$, $l = 1, \ldots, k-1$, one can immediately prove the following lemma.

Lemma 3.8. *Let* $[A] \in \mathbb{IR}^{n \times n}$ *and let* $[A]^{(k)}$ *exist. Then the following assertions hold.*

(a) $\langle [A]^{(k)} \rangle \geqslant \langle [A] \rangle^{(k)}$,

(b) $\langle [A] \rangle e \begin{Bmatrix} \geqslant \\ > \\ = \end{Bmatrix} 0$ *implies* $\langle [A]^{(k)} \rangle e \geqslant \langle [A] \rangle^{(k)} e \begin{Bmatrix} \geqslant \\ > \\ = \end{Bmatrix} 0.$

(c) *If* $\langle [A] \rangle e \geqslant 0$ *and if* $(\langle [A]^{(l)} \rangle e)_i > 0$ *for some integer* $l < k$ *and for some index* $i \in \{1, \ldots, n\}$ *then* $(\langle [A]^{(k)} \rangle e)_i > 0$.

Lemma 3.9. *Let* $[A] \in \mathbb{IR}^{n \times n}$ *and let* $\langle [A] \rangle$ *be irreducible with* $\langle [A] \rangle e \geqslant 0$. *Then we get*

(a) $\langle [a]_{ii} \rangle > 0, \; i = 1, \ldots, n;$

(b) $\pi_2(\langle [A] \rangle)$ *is irreducible provided* $[A]$ *has at least 3 rows.*

Proof. See [5].

□

Lemma 3.10. *Let $[A] \in \mathbf{IR}^{n\times n}$ and let $\langle [A] \rangle$ be irreducible with $\langle [A] \rangle e = 0$. Then*

(a) $\langle [A] \rangle^{(k)}$, $k = 1, \ldots, n$, *exists*, $\langle [A] \rangle^{(k)} e = 0$, $\pi_k(\langle [A] \rangle)$, $k = 1, \ldots, n-1$, *is irreducible,*

(b) $[A]^{(k)}$, $k = 1, \ldots, n$, *exists,*

(c) $0 \in [a]_{nn}^{(n)}$ *(i.e., $[x]^G$ does not exist for any vector $[b] \in \mathbf{IR}^n$) if and only if*

$$\langle [A]^{(k)} \rangle = \langle [A] \rangle^{(k)}, \quad k = 1, \ldots, n. \tag{17}$$

Proof. The assertions are trivial if $n = 1$. Therefore we assume $n \geqslant 2$.

(a) We prove (a) by induction. By assumption, the assertions are true for $k = 1$. Assume now that they hold for some positive integer $k < n$. Then $\langle [A] \rangle^{(k)}$ exists and $\pi_k(\langle [A] \rangle^{(k)})$ is irreducible satisfying

$$\left(\pi_k(\langle [A] \rangle^{(k)})e\right)_{i-k+1} = (\langle [A] \rangle^{(k)} e)_i = 0, \quad i = k, \ldots, n,$$

where $e = (1, \ldots, 1)^T$ is chosen from $\mathbf{R}^{n-k+1}$ on the left hand side. Replacing $[A]$ in Lemma 3.9 by $\pi_k(\langle [A] \rangle^{(k)})$ yields $(\langle [A] \rangle^{(k)})_{kk} > 0$, hence $\langle [A] \rangle^{(k+1)}$ exists. Since $\pi_{k+1}(\langle [A] \rangle) = \pi_2(\pi_k(\langle [A] \rangle^{(k)}))$ holds, part (b) of the same lemma shows that $\pi_{k+1}(\langle [A] \rangle)$ is irreducible provided $\pi_k(\langle [A] \rangle^{(k)})$ has at least 3 rows, i.e. $k + 1 < n$. Lemma 3.8 (b) (case '=') completes the proof.

(b) Let $k < n$ and let $[A]^{(k)}$ exist. (This is certainly true for $k = 1$.) Since $\langle [A] \rangle^{(k+1)}$ exists by (a) we get $(\langle [A] \rangle^{(k+1)})_{kk} > 0$, and Lemma 3.8(a) implies

$$\langle [A]_{kk}^{(k)} \rangle \geqslant (\langle [A] \rangle^{(k)})_{kk} > 0.$$

Hence $[A]^{(k+1)}$ exists, and the proof is completed by induction.

(c) '$\Longrightarrow$' Let $0 \in [a]_{nn}^{(n)}$ and assume that (17) is false. Then by Lemma 3.8(a) there is an integer $k > 1$ and there are indices $i_0, j_0 \in \{k, k+1, \ldots, n\}$ such that

$$\langle [A]^{(l)} \rangle = \langle [A] \rangle^{(l)}, \quad l = 1, \ldots, k-1 \tag{18}$$

and

$$\langle [A]^{(k)} \rangle_{i_0 j_0} > (\langle [A] \rangle^{(k)})_{i_0 j_0} \tag{19}$$

hold. If $n = 2$, (18) yields $k = 2 = i_0 = j_0$, hence $0 \notin [a]_{nn}^{(n)}$ which contradicts our assumption. Therefore let $n \geqslant 3$. We slightly blow up the degenerate non-zero entries of $[A]$ such that the resulting matrix $[B]$ possesses the following properties

$$[B] \supseteq [A], \tag{20}$$

$$\langle [B] \rangle = \langle [A] \rangle, \tag{21}$$

$$[b]_{ij} \neq 0 \Longrightarrow d([b]_{ij}) > 0. \tag{22}$$

Property (21) and (b) guarantee the existence of $[B]^{(l)}, l = 1, \ldots, n$; (20) implies $[B]^{(l)} \supseteq [A]^{(l)}, l = 1, \ldots, n$, hence $\langle [B]^{(l)} \rangle \leqslant \langle [A]^{(l)} \rangle, l = 1, \ldots, n$. Property (21), once more, yields $\langle [B] \rangle^{(l)} = \langle [A] \rangle^{(l)}$, $l = 1, \ldots, n$. Together with Lemma 3.8(a) we thus get

$$\text{(23)} \qquad \langle [A] \rangle^{(l)} = \langle [B] \rangle^{(l)} \leqslant \langle [B]^{(l)} \rangle \leqslant \langle [A]^{(l)} \rangle, \; l = 1, \ldots, n,$$

therefore (18) yields

$$\text{(24)} \qquad \langle [A] \rangle^{(l)} = \langle [B] \rangle^{(l)} = \langle [B]^{(l)} \rangle = \langle [A]^{(l)} \rangle, \; l = 1, \ldots, k-1.$$

By continuity of the interval arithmetic operations we can choose $[B]$ such that, in addition to (20)–(22), the condition

$$\text{(25)} \qquad (\langle [B]^{(k)} \rangle)_{i_0 j_0} > (\langle [B] \rangle^{(k)})_{i_0 j_0} = (\langle [A] \rangle^{(k)})_{i_0 j_0}$$

holds with i_0, j_0 from (19). From (a), (23) and (25) we get

$$\text{(26)} \qquad 0 = \langle [A] \rangle^{(k)} e = \langle [B] \rangle^{(k)} e \leqslant \langle [B]^{(k)} \rangle e$$

with strict inequality for the i_0-th component.

We will show now that $\langle \pi_k([B]) \rangle$ is irreducibly diagonally dominant. To this end we remark that $(\langle [B] \rangle^{(2)})_{ij} \neq 0$ implies $((\langle [B] \rangle^{(1)})_{ij} \neq 0)$ or $((\langle [B] \rangle^{(1)})_{i1} \neq 0$ and $(\langle [B] \rangle^{(1)})_{1j} \neq 0)$, hence $([b]_{ij} \neq 0)$ or $([b]_{i1} \neq 0$ and $[b]_{1j} \neq 0)$. Taking

$$\text{(27)} \qquad \operatorname{rad}([b]_{ij}^{(2)}) = \operatorname{rad}([b]_{ij}) + \operatorname{rad}\left(\frac{[b]_{i1}[b]_{1j}}{[b]_{11}} \right)$$

into account which is true by (2), we see that $(\langle [B] \rangle^{(2)})_{ij} \neq 0$ implies $[b]_{ij}^{(2)} \neq 0$. Thus, since $\pi_2(\langle [B] \rangle)$ is irreducible by (a), $\langle \pi_2([B]) \rangle$ is irreducible, as well. Moreover, by (27), property (22) holds also for $[B]^{(2)}$ instead of $[B]$. An inductive argument shows that $\langle \pi_k([B]) \rangle$ is irreducible, thus, by (26), $\pi_k([B])$ is irreducibly diagonally dominant. Therefore, Corollary 3.2 yields $0 \neq [b]_{nn}^{(n)} \supseteq [a]_{nn}^{(n)}$ in contrast to our assumption. This proves the assertion.

'$\Longleftarrow$' By choosing $k = n$ in (a) and (17), we get $\langle [a]_{nn}^{(n)} \rangle = (\langle [A] \rangle^{(n)})_{nn} = (\langle [A] \rangle^{(n)} e)_n = 0$, hence $0 \in [a]_{nn}^{(n)}$. □

With the matrices $S^{(k)}, S'$ from Definition 3.4 we can state our final lemma.

Lemma 3.11. *Let $[A] \in \mathbb{IR}^{n \times n}$ and let $\langle [A] \rangle$ be irreducible with $\langle [A] \rangle e = 0$. If $\langle [A]^{(l)} \rangle = \langle [A] \rangle^{(l)}$, $l = 1, \ldots, k$, then the following assertions are true.*

(a) $s_{ii}^{(l)} = s_{ii}^{(1)} = s_{ii}^{'}$, $l = 1, \ldots, k$, $i = 1, \ldots, n$, *with the exception of* $k = n = l = i$, *where* $s_{nn}^{(n)} = s_{nn}^{(1)} = s_{nn}^{'}$ *or* $s_{nn}^{(n)} = 0$ *are possible.*

(b) *For any indices* i, j *with* $j \geqslant \min\{l, i\}, j \neq i$, *we have*

$$s_{ij}^{(l)} = 0, \ l = 1, \ldots, k \tag{28}$$

or there is an integer $l_0 = l_0(i,j) \leqslant k$ *such that*

$$s_{ij}^{(l)} = \begin{cases} 0, & l = 1, \ldots, l_0 - 1 \\ s_{ij}^{'}, & l = l_0, l_0 + 1, \ldots, k \end{cases}.$$

If $i \leqslant l$ *or if* $j \leqslant l$ *in* (28), *then*

$$s_{ij}^{(l)} = s_{ij}^{'} = 0. \tag{30}$$

Proof. (a) The assumption, Lemma 3.10(a) and Lemma 3.9(a) applied to $\pi_l(\langle [A] \rangle)$ guarantee

$$\langle [a]_{ii}^{(l)} \rangle = (\langle [A] \rangle^{(l)})_{ii} > 0, \ l = 1, \ldots, k-1, \ i = 1, \ldots, n. \tag{31}$$

Inequality (31) remains true for $l = k$ if $k < n$ or if ($k = n$ and $i < n$). Hence $[a]_{ii}^{(l)} \neq 0$ in these cases, and (10) proves the assertion. In the case $k = n = l = i$, Lemma 3.10(c) implies $0 \in [a]_{nn}^{(n)}$. If $[a]_{nn}^{(n)} = 0$ then $s_{nn}^{(n)} = 0$; otherwise (10) completes the proof.

(b) The assertion follows by (6) and by the definition of $S^{'}$. □

Proof of Theorem 3.5. '$\Longrightarrow$' Let $[x]^G$ exist for any vector $[b] \in \mathbf{IR}^n$. Then $0 \notin [a]_{nn}^{(n)}$, hence by Lemma 3.10(c) and Lemma 3.8(a), there is an integer $k < n$ such that $(\langle [A]^{(k+1)} \rangle)_{i_0 j_0} > \langle [A] \rangle_{i_0 j_0}^{(k+1)}$ for some indices i_0, j_0. Let k be as small as possible, i.e., $\langle [A]^{(l)} \rangle = \langle [A] \rangle^{(l)}, l = 1, \ldots, k$. Then $i_0, j_0 > k$, and Lemma 2.1 implies

$$s_{i_0 j_0}^{(k)} s_{i_0 k}^{(k)} s_{kk}^{(k)} s_{k j_0}^{(k)} = \begin{cases} 1, & \text{if } i_0 \neq j_0 \\ -1, & \text{if } i_0 = j_0 \end{cases}.$$

In particular, none of these four elements of $S^{(k)}$ can be zero. By Lemma 3.11(b) they must therefore be equal to the corresponding elements of $S^{'}$ which proves the assertion with the triple $(i, j, k) = (i_0, j_0, k)$.

'$\Longleftarrow$' Let (15) be true and assume that $[x]^G$ does not exist for any vector $[b] \in \mathbf{IR}^n$. By Lemma 3.10(c) we get

$$\langle [A]^{(l)} \rangle = \langle [A] \rangle^{(l)}, l = 1, \ldots, n. \tag{32}$$

We first show

$$s_{ik}^{(k)} = s_{ik}^{'}, \; s_{kk}^{(k)} = s_{kk}^{'}, \quad s_{kj}^{(k)} = s_{kj}^{'}$$

for i, j, k from (15). To this end, we remark that none of the four factors in (15) can be zero. Since $i, j > k$ there, we have $k < n$. Hence by Lemma 3.11(a), $s_{kk}^{(k)} = s_{kk}^{'}$. If (29) holds for $s_{ik}^{(k)}$ then trivially $s_{ik}^{(k)} = s_{ik}^{'}$. Otherwise, (30) with $j = l = k$ implies $s_{ik}^{(k)} = s_{ik}^{'} = 0$ which contradicts (15). Analogously one gets $s_{kj}^{(k)} = s_{kj}^{'}$.

We now show $s_{ij}^{(k)} = s_{ij}^{'}$. This is obviously true by Lemma 3.11(a) for $i = j$, and by Lemma 3.11(b) for $i \neq j$ provided (29) holds. In both these cases, (15) is fulfilled with $s^{'}$ replaced by $s^{(k)}$. Hence by Lemma 2.1 we obtain $\langle [A]^{(k+1)} \rangle_{ij} > (\langle [A] \rangle^{(k+1)})_{ij}$ contradicting (32). Therefore we must have $i \neq j$ with (29) being false. By (28), we get $s_{ij}^{(k)} = 0$. Since

$$s_{ik}^{(k)} s_{kk}^{(k)} s_{kj}^{(k)} = s_{ik}^{'} s_{kk}^{'} s_{kj}^{'} \neq 0 \tag{33}$$

the definition of $S^{'}$ yields $s_{ij}^{(k+1)} = -s_{ik}^{(k)} s_{kk}^{(k)} s_{kj}^{(k)} = s_{ij}^{'}$ which results in $s_{ij}^{'} s_{ik}^{'} s_{kk}^{'} s_{kj}^{'} = -1$ by virtue of (33). This contradicts (15). Therefore $[x]^G$ exists for any vector $[b] \in \mathbb{IR}^n$. □

We illustrate Theorem 3.5 with the matrix $[A]$ of (14). Here we have

$$S^{'} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix},$$

therefore (15) holds for $(i, j, k) = (3, 3, 2)$, and $[x]^G$ exists.

Replacing $[a]_{33} = [2, 4]$ by $[a]_{33} = [-4, -2]$ results in

$$S^{'} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}$$

and (15) does not hold as one easily checks. Therefore, $[x]^G$ does not exist.

Combining Corollary 3.2 with Theorem 3.5 we immediately get the main result of our paper.

Theorem 3.12. *Let* $[b] \in \mathbb{IR}^n$, $[A] \in \mathbb{IR}^{n \times n}$. *Let* $\langle [A] \rangle$ *be irreducible and diagonally dominant. Then* $[x]^G$ *exists if and only if* $\langle [A] \rangle$ *is irreducibly diagonally dominant or if*

$$s_{ij}^{'} s_{ik}^{'} s_{kk}^{'} s_{kj}^{'} = \begin{cases} 1, & \text{if } i \neq j \\ -1, & \text{if } i = j \end{cases}$$

holds for at least one triple (i,j,k) with $k < i,j$ and with S' as in Definition 3.4.

We cannot dispense in Theorem 3.12 with the passage 'if $\langle[A]\rangle$ is irreducibly diagonally dominant', as the example $[A] = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ illustrates. Here, (15) is apparently not fulfilled although $[x]^G$ exists. Note also that no nonsingularity of $[A]$ is assumed in Theorem 3.12 .

If in this theorem $[A]$ is not an H-matrix then

$$\langle[a_{ii}]\rangle - \sum_{\substack{j=1 \\ j\neq i}}^{n} |[a_{ij}]| = 0, \; i = 1,\dots,n. \tag{34}$$

Let in this case condition (15) be fulfilled. Then the feasibility of the interval Gaussian algorithm is guaranteed for all interval matrices in some neighbourhood[2] of $[A]$ by the continuity of the interval arithmetic. This is, in particular, true for all non-H-matrices which are sufficiently close to $[A]$, a remark which is important in view of rounding errors on a computer.

The equation (34) can easily be checked in practice (without rounding errors!) when using programming languagues like PASCAL-XSC (cf. [6]) with its exact scalar product facility—provided that $[A]$ is representable by machine numbers, of course.

We will show now that Theorem 3.3 follows at once from Theorem 3.12.

Since in Theorem 3.3 the matrix $\langle[A]\rangle$ is assumed to be irreducible, none of its rows can contain only zeros. Therefore, $\langle[A]\rangle e \geqslant 0$ implies $\langle[a]_{kk}\rangle > 0$, hence $\mathrm{sign}(\check{a}_{kk}) \neq 0$. This shows that no factor in (13) can be zero. Thus it coincides with the corresponding element of S' so that (13) implies (15) with the same triple (i,j,k). The assertion of Theorem 3.3 now follows from Theorem 3.12.

If no element of $\check{A}$ is zero, $\langle[A]\rangle$ is irreducible and $S^{(1)} = S'$. This is the base of the following corollary.

Corollary 3.13. *Let $[A] \in \mathbb{IR}^{n\times n}$ with $\check{a}_{ij} \neq 0$ for any $i,j \in \{1,\dots,n\}$ and let $\langle[A]\rangle$ be diagonally dominant. Then $[x]^G$ exists for any vector $[b] \in \mathbb{IR}^n$, if and only if $\langle[A]\rangle$ is irreducibly diagonally dominant or if*

$$\mathrm{sign}(\check{a}_{ij})\,\mathrm{sign}(\check{a}_{ik})\,\mathrm{sign}(\check{a}_{kk})\,\mathrm{sign}(\check{a}_{kj}) = \begin{cases} 1, & \text{if } i \neq j \\ -1, & \text{if } i = j \end{cases}$$

holds for at least one triple (i,j,k) with $k < i,j$.

---

[2] We use the standard topology in $\mathbb{IR}^{n\times n}$—cf. e.g. [2]

We conclude our paper with a numerical example which illustrates the theory. Let $[A] \in \mathbb{IR}^{5\times 5}$ be defined by its bounds

$$\overline{A} := \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{16} & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \underline{a}_{ij} := \begin{cases} \overline{a}_{ij} & \text{if } i = j \\ \frac{1}{2}\overline{a}_{ij} & \text{if } i \neq j \end{cases},$$

and choose $[b] := e \in \mathbb{IR}^5$. Then

$$[x]^G \subseteq \begin{pmatrix} [-0.6135, & 2.5378] \\ [-0.5399, & 2.5213] \\ [-0.5403, & 2.5510] \\ [-0.7533, & 2.4633] \\ [-0.3091, & 3.2574] \end{pmatrix} \tag{35}$$

where the numbers result from the interval Gaussian algorithm executed in PASCAL-XSC on an HP Apollo 720 Workstation, taking rounding errors into account. We expect that our numerical results overestimate $[x]^G$ at most by one unit of the last place being shown, since the dimension of the problem is not very high and since we applied the exact scalar product of PASCAL-XSC as often as it was possible.

Note that in the example above the matrix $[A]$ contains only nonnegative matrices with row sums and column sums at most one. Such matrices belong to the class of double substochastic matrices as defined in [4], p. 104. The upper bound $\overline{A}$ has row sums and column sums equal to one and is thus a double stochastic matrix (see again [4]). It is obvious that such matrices belong to the class mentioned in Theorem 3.12, provided $\overline{A}$ is irreducible with diagonal entries being not less than one half.

The matrix $[A]$ is not an H-matrix since $\langle [A] \rangle e = 0$. Corollary 3.13 applies with $(i, j, k) = (2, 3, 1)$, e.g. Thus the interval Gaussian algorithm is here feasible in theory. That it is also in practice, is shown by (35).

References

[1] *G. Alefeld*: Über die Durchführbarkeit des Gaußschen Algorithmus bei Gleichungen mit Intervallen als Koeffizienten, Computing Suppl. *1* (1977), 15–19.

[2] *G. Alefeld and J. Herzberger*: Introduction to Interval Computations, Academic Press, New York, 1983.

[3] *H. Bauch, K.-U. Jahn, D. Oelschlägel, H. Süsse and V. Wiebigke*: Intervallmathematik, BSB B.G. Teubner Verlagsgesellschaft, 1987.

[4] *A. Berman and R. J. Plemmons*: Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.

[5] *A. Frommer and G. Mayer*: A new criterion to guarantee the feasibility of the interval Gaussian algorithm, SIAM J. Matrix Anal. Appl., in press.

[6] *R. Klatte, U. Kulisch, M. Neaga, D. Ratz and Ch. Ullrich*: PASCAL-XSC, Sprachbeschreibung mit Beispielen, Springer, Berlin, 1991.

[7] *G. Mayer*: Old and new aspects of the interval Gaussian algorithm, Computer Arithmetic, Scientific Computation and Mathematical Modelling (E. Kaucher, S.M. Markov, G. Mayer, eds.), IMACS Annals on Computing and Applied Mathematics 12, Baltzer, Basel, 1991, pp. 329–349.

[8] *R.E. Moore*: Interval Analysis, Prentice Hall, Englewood Cliffs, N.J., 1966.

[9] *A. Neumaier*: New techniques for the analysis of linear interval equations, Linear Algebra Appl. *58* (1984), 273–325.

[10] *A. Neumaier*: Interval Methods for Systems of Equations, Cambridge University Press, Cambridge, 1990.

[11] *K. Reichmann*: Abbruch beim Intervall-Gauss-Algorithmus, Computing *22* (1979), 355–361.

[12] *R.S. Varga*: Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1963.

Authors' address: *Günter Mayer*, *Lars Pieper*, Institut für Angewandte Mathematik, Universität Karlsruhe (TH), Kaiserstr. 12, D-7500 Karlsruhe 1, Deutschland.