Pivot Tightening for Direct Methods for Solving Systems of Linear Interval Equations

Jürgen Garloff

University of Applied Sciences / HTWG Konstanz
Faculty of Computer Science

and

University of Konstanz
Department of Mathematics and Statistics
Systems of linear interval equations \([A]x = [b]\)

**Solution set** \(\Sigma ([A], [b]) := \{x \in \mathbb{R}^n | Ax = b, A \in [A], b \in [b]\}\)

General assumption: All \(A \in [A]\) are nonsingular.

Then: \(\Sigma ([A], [b])\) is compact, connected, and convex in each orthant.
Neumaier system

\[
\begin{bmatrix}
3.5 & [0, 2] & [0, 2] \\
[0, 2] & 3.5 & [0, 2] \\
[0, 2] & [0, 2] & 3.5
\end{bmatrix}
\begin{bmatrix}
x
\end{bmatrix}
= 
\begin{bmatrix}
[−1, 1] \\
[−1, 1] \\
[−1, 1]
\end{bmatrix}
\]
(Interval) Hull of the solution set

\[ [A]^H [b] := \bigcap_{\Sigma([A],[b]) \subseteq \{z\}} \quad [z] \in \mathbb{R}^n \]
Interval Gaussian Elimination (abbr. IGE, without pivoting)

To transform a given \([A] \in \mathbb{IR}^{n \times n}\) on triangular form proceed as follows: define \([A]^{(k)} = ([a_{ij}]^{(k)}) \in \mathbb{IR}^{n \times n}, k = 1, \ldots, n,\) by

\[
[a_{ij}]^{(k+1)} = \begin{cases} 
[a_{ij}]^{(k)}, & i = 1, \ldots, k, j = 1, \ldots, n, \\
[a_{ij}]^{(k)} - \frac{[a_{ik}]^{(k)}[a_{kj}]^{(k)}}{[a_{kk}]^{(k)}}, & i = k + 1, \ldots, n, j = k + 1, \ldots, n, \\
0 & \text{otherwise.}
\end{cases}
\]

Note that no rows or columns are permuted. The algorithm is feasible if and only if \(0 \notin [a_{kk}]^{(k)}, k = 1, \ldots, n.\)
Further outline

- interval pivot tightening for Gaussian elimination

- application to
  - inverse nonnegative matrices
  - inverse $M$-matrices

- the interval Cholesky method and pivot tightening

- application to
  - positive definite matrices
  - positive definite Toeplitz matrices
Notation

$\mathbb{IR}$: set of the compact, nonempty real intervals $[a] = [a, \bar{a}], \ a \leq \bar{a}$,

$\mathbb{IR}^n$: set of $n$-vectors with components from $\mathbb{IR}$. interval vectors

$\mathbb{IR}^{n \times n}$: set of $n$-by-$n$ matrices with components from $\mathbb{IR}$. interval matrices

Elements from $\mathbb{IR}^n$ and $\mathbb{IR}^{n \times n}$ may be regarded as vector intervals and matrix intervals, respectively, w.r.t. the usual entrywise partial ordering, e.g.,

$$[A] = ([a_{ij}])_{i,j=1}^n = ([a_{ij}, \bar{a}_{ij}])_{i,j=1}^n = [A, \bar{A}], \quad \text{where } A = (a_{ij})_{i,j=1}^n, \ \bar{A} = (\bar{a}_{ij})_{i,j=1}^n.$$

A vertex matrix of $[A]$ is a matrix $A = (a_{ij})_{i,j=1}^n$ with $a_{ij} \in \{a_{ij}, \bar{a}_{ij}\}$, $i,j = 1, \ldots, n$.

A degenerate interval (vector, matrix) is identified with the (only) real number (vector, matrix) which it contains.
**Interval pivot tightening**

IGE breaks down when \([a_{kk}]^{(k)}\) (termed *ordinary* interval pivot henceforth) contains zero. This can occur even though all matrices \(A \in [A]\) have nonvanishing leading principal minors. For ordinary GE, the pivot \(a_{kk}^{(k)}\) can be represented as

\[
a_{kk}^{(k)} = \frac{\det A[\{1, 2, \ldots, k\}]}{\det A[\{1, 2, \ldots, k - 1\}]}.
\]

This property is used by J. Mayer (2002) to tighten the ordinary interval pivot \([a_{kk}]^{(k)}\): Let \(D_k\) denote an enclosure for \(\det A[\{1, 2, \ldots, k\}]\) for all \(A \in [A]\), where \(D_0 := 1\). If \(0 \notin D_k, k = 1, 2, \ldots, n\), then \([a_{kk}]^{(k)}\) is replaced by

\[
[a_{kk}]^{(k)} \cap \frac{D_k}{D_{k-1}}.
\]
Aim:
Finding the range of
\[
\frac{\det A[\{1, 2, \ldots, k\}]}{\det A[\{1, 2, \ldots, k-1\}]}
\]
over \([A]\).

w.l.o.g. \[ p(A) := \frac{\det A}{\det A'}, \]

where \[ A' := A[\{1, 2, \ldots, n-1\}]. \]
Inverse nonnegative matrices

**Def.**  $A$ is *inverse nonnegative* if $0 \leq A^{-1}$.

**Proposition 1** (Kuttler, 1971). Let $[A] = [A, \overline{A}]$ be a matrix interval and $\underline{A}$ and $\overline{A}$ be inverse nonnegative. Then $[A]$ is inverse nonnegative and $\underline{A}^{-1} \leq \overline{A}^{-1}$.

**Theorem 2** (Beeck, 1975). If $[A] \in \mathbb{IR}^{n \times n}$ is inverse nonnegative, then

$$[A]^H[b] = \begin{cases} 
[\underline{A}^{-1}b, \underline{A}^{-1}\overline{b}] & \text{if } 0 \leq [b], \\
[\underline{A}^{-1}b, \underline{A}^{-1}\overline{b}] & \text{if } 0 \in [b], \\
[\underline{A}^{-1}b, \overline{A}^{-1}\overline{b}] & \text{if } [b] \leq 0.
\end{cases}$$

In the general case, one has to solve at most $2n$ linear systems to find $\inf([A]^H[b])$ and similarly $\sup([A]^H[b])$. 
We assume now that each $A \in [A]$ has the property that all its leading principal submatrices are inverse nonnegative or equivalently that $A$ allows a factorization $A = LDLU$, where $L$ and $U$ are inverse nonnegative lower and upper triangular matrices, respectively, whose diagonal entries are all one, and $D$ is a diagonal matrix with positive diagonal entries. According to Prop. 1 the condition on $[A] = [A, \overline{A}]$ is fulfilled if all leading principal submatrices of $A$ and $\overline{A}$ are inverse nonnegative.

**Theorem 3.** *If all leading principal submatrices of $[A] = [A, \overline{A}]$ are inverse nonnegative then* 

$$p([A]) = [p(A), p(\overline{A})].$$  \hspace{1cm} (1)

*Proof.* For $\underline{A} \leq A \leq \overline{A}$ it follows from Prop. 1 that 

$$(\overline{A}^{-1})_{nn} \leq (A^{-1})_{nn} \leq (A^{-1})_{nn}.$$ 

Formula (1) is now a consequence of $(A^{-1})_{nn} = \frac{1}{p(A)}$. \hfill $\square$
Example


Since \( A \), \( \overline{A} \) are inverse nonnegative, \( [A] \) contains only inverse nonnegative matrices. IGE results in the following interval containing zero

\[ [a_{33}]^{(3)} = [-.112\ldots, 4.311\ldots] \]

and breaks down.

Application of Theorem 3 results in \([\frac{6}{7}, 4]\).
Inverse $M$-matrices

**Def.**: $A \in \mathbb{R}^{n \times n}$ is called an *inverse $M$-matrix* if $A^{-1}$ is an $M$-matrix, i.e., if $0 \leq A$ and the off-diagonal entries of $A^{-1}$ are nonpositive.

Each leading principal submatrix of an inverse $M$-matrix is likewise an inverse $M$-matrix and all its principal minors are positive.

**Proposition 4** (Johnson and Smith, 2002). A *matrix interval* is an inverse $M$-matrix interval if and only if all its vertex matrices are inverse $M$-matrices.
**Theorem 5.** If \([A] \in \mathbb{IR}^{n \times n}\) is an inverse \(M\)-matrix interval then \(p([A]) = [p(A^l), p(A^u)]\), where

\[
A^l := \begin{pmatrix}
\underline{a}_{11} & \cdots & \underline{a}_{1,n-1} & \underline{a}_{1n} \\
\vdots & \ddots & \vdots & \vdots \\
\underline{a}_{n-1,1} & \cdots & \underline{a}_{n-1,n-1} & \underline{a}_{n-1,n} \\
\underline{a}_{n1} & \cdots & \underline{a}_{n,n-1} & \underline{a}_{nn}
\end{pmatrix}, \quad A^u := \begin{pmatrix}
\overline{a}_{11} & \cdots & \overline{a}_{1,n-1} & \underline{a}_{1n} \\
\vdots & \ddots & \vdots & \vdots \\
\overline{a}_{n-1,1} & \cdots & \overline{a}_{n-1,n-1} & \underline{a}_{n-1,n} \\
\underline{a}_{n1} & \cdots & \underline{a}_{n,n-1} & \overline{a}_{nn}
\end{pmatrix}.
\]
Example

Let \( [A] := \begin{pmatrix}
[1, 4] & [\frac{1}{2}, \sqrt{2}] & \frac{\sqrt{2}}{2} \\
\frac{\sqrt{2}}{2} & 1 & [\frac{1}{2}, \sqrt{2}] \\
\frac{1}{2} & \frac{\sqrt{2}}{2} & 1
\end{pmatrix} \).

It is easily checked that all eight vertex matrices are inverse \( M \)-matrices and by Proposition 4 it follows that \( [A] \) is an inverse \( M \)-matrix interval. IGE results in the following interval containing zero

\[
[a_{33}]^{(3)} = [-\frac{4\sqrt{2} + 1}{64}, 1 - \frac{\sqrt{2}}{16}]
\]

and breaks down. By Theorem 5 this pivot can be tightened to

\[
[\frac{1}{2}, 1 - \frac{\sqrt{2}}{4}] = [0.5, 0.6464 \ldots].
\]
Similar results hold for *totally nonnegative matrices*, i.e., matrices having all their minors nonnegative, cf.


Therein also results on Toeplitz systems of equations and Neville elimination can be found.
Symmetric solution set

Let $[A]$ be symmetric, i.e.,

$[A]^T = [A],$

and

$[A]_s := \{ A \in [A] \mid A = A^T \}$  \hspace{1em} (symmetric part of $[A]$).

Symmetric solution set:

$\Sigma_{sym} ([A], [b]) := \{ x \in \mathbb{R}^n \mid Ax = b, A \in [A]_s, b \in [b] \}$
Symmetric solution set – example

\[
\begin{pmatrix}
[1, 3] & [0, 1] \\
[0, 1] & [-4, 1]
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
= 
\begin{pmatrix}
[1, 4] \\
[-2, 1]
\end{pmatrix}
\]

Given \([A] = [A]^T \in \mathbb{IR}^{n \times n}\), define the lower triangular matrix \([L]\) with \([A] \subseteq [L][L]^T\) by

\[
[l_{jj}] = \left( [a_{jj}] - \sum_{k=1}^{j-1} [l_{jk}]^2 \right)^{\frac{1}{2}}, \\
\]  

\[
[l_{ij}] = \left( [a_{ij}] - \sum_{k=1}^{j-1} [l_{ik}][l_{jk}] \right) \big/ [l_{jj}], \quad i = j + 1, \ldots, n.
\]

Here \([a]^\frac{1}{2} := \left\{ a^\frac{1}{2} \mid a \in [a] \right\}\) for \(0 \leq a\).

The algorithm is feasible if and only if \(0 < l_{ii}, \ i = 1, \ldots, n\).
Interval pivot tightening

The interval Cholesky method breaks down when $0 \not< l_{ii}$. This can occur even though all matrices $A \in [A]_s$ are positive definite. For the ordinary Cholesky decomposition, $l_{ii}$ can be represented as

$$l_{ii} = \left( \frac{\det A[\{1, 2, \ldots, i\}]}{\det A[\{1, 2, \ldots, i-1\}]} \right)^{\frac{1}{2}}.$$

An easy upper bound: By Fischer’s inequality

$$\ell_{jj} \leq \frac{1}{a_{jj}^2}, \quad j = 1, \ldots, n.$$
Positive definite interval matrices

Let $A_c := \frac{1}{2}(A + \bar{A})$ \textit{midpoint matrix}
and $\triangle A := \frac{1}{2}(\bar{A} - A)$ \textit{radius matrix};
then $[A] = [A_c - \triangle A, A_c + \triangle A]$.

Consider the following vertex matrices (of cardinality of at most $2^{2n-1}$)

$$A_{yz} := A_c - \text{diag}(y_1, \ldots, y_n) \cdot \triangle A \cdot \text{diag}(z_1, \ldots, z_n), \quad y, z \in \{-1, 1\}^n;$$
in particular, the matrices $A_{zz}$, $z \in \{-1, 1\}^n$, form a subset of the vertex matrices of cardinality of at most $2^{n-1}$. 
Proposition 6 (Bialas and G. 1984, Rohn 1994). Let $[A] = [A]^T \in \mathbb{IR}^{n \times n}$. Then $[A]$ is positive definite, i.e., $[A]_s$ contains only positive definite matrices if and only if all matrices $A_{zz}$, $z \in \{-1, 1\}^n$, are positive definite.

Conjecture:

For a symmetric positive definite interval matrix $[A]$ the minimum value of $p = \det A / \det A'$ over $[A]_s$ is attained at a matrix $A_{zz}$. 
Let \( A \in \mathbb{R}^{n \times n} \) be symmetric and order its eigenvalues as \( \lambda_1(A) \leq \ldots \leq \lambda_n(A) \). By the interlacing of the eigenvalues of \( A \) and \( A' = A[\{1, \ldots, n-1\}] \), we obtain

\[
\lambda_1(A) \leq p(A) = \frac{\text{det}A}{\text{det}A'}.
\]

For \( [A] = [A]^T \in \mathbb{H}^{n \times n} \)

\[
\min \{ \lambda_1(A) \mid A \in [A]_s \} = \min \{ \lambda_1(A_{zz}) \mid z \in \{-1, 1\}^n \}
\]

(Rohn, 1998).
Lower bounds for the smallest eigenvalue of a symmetric pd matrix

Let \( A = \begin{pmatrix} A' & b \\ b^T & c \end{pmatrix} \) and denote by \( \beta_{n-1} \) a lower bound for \( \lambda_1(A') \). Then

Dembo (1988):

\[
\beta_n = \frac{1}{2} \left( c + \beta_{n-1} - \sqrt{(c - \beta_{n-1})^2 + 4b^Tb} \right) \leq \lambda_1(A)
\]

(may not be positive);

Ma and Zarowski (1998):

\[
0 < \tilde{\beta}_n = \frac{1}{2} \left( c + \beta_{n-1} - \sqrt{(c - \beta_{n-1})^2 + 4\beta_{n-1}b^T(A')^{-1}b} \right) \leq \lambda_1(A).
\]
Example (cont’d)

\[ [A] = \begin{pmatrix} [4,5] & [-3,-2] & 1 \\ [-3,-2] & 4 & [-3,-2] \\ 1 & [-3,-2] & [4,5] \end{pmatrix} \] is pd but the interval Cholesky method breaks down due to

\[ l_{33}^2 = -0.112 \ldots. \]

The matrices \( A_{zz} \):

\[
\begin{pmatrix}
4 & -3 & 1 \\
-3 & 4 & -3 \\
1 & -3 & 4
\end{pmatrix}
\begin{pmatrix}
4 & -3 & 1 \\
-3 & 4 & -2 \\
1 & -2 & 4
\end{pmatrix}
\begin{pmatrix}
4 & -2 & 1 \\
-2 & 4 & -3 \\
1 & -3 & 4
\end{pmatrix}
\begin{pmatrix}
4 & -2 & 1 \\
-2 & 4 & -2 \\
1 & -2 & 4
\end{pmatrix}
\]

Dembo bound: -1 -0.192… -0.316… 0.550…
Ma and Zarowski bound: 0.177… 0.658… 0.619…

Thus \( l_{33}^2 \) can be improved by 0.177….
Positive definite Toeplitz matrices

\[ T = T(a_1, \ldots, a_n) = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_2 \\ a_n & \cdots & a_2 & a_1 \end{pmatrix} \]

**Proposition 7** (Mukherjee and Maiti, 1988).

\[ l_{nn} \leq \ldots \leq l_{22} \leq l_{11} \leq a_1^2. \]

Practical application to \([l_{jj}] = [l_{jj}, \bar{l}_{jj}]\):

If \( \bar{l}_{j+1,j+1} > \bar{l}_{j,j} \) then put \( \bar{l}_{j+1,j+1} := \bar{l}_{j,j} \);

if \( l_{j+1,j+1} > l_{j,j} \) then put \( l_{j,j} := l_{j+1,j+1} \).

In the second case the entries below \([l_{jj}]\) should be recomputed; this may tighten \([l_{j+1,j+1}]\), too.
**Example:** $[T] = T(16, [9, 10], [4, 5], [1, 2], [0, 1])$

<table>
<thead>
<tr>
<th>diagonal entry</th>
<th>only monotonicity</th>
<th>use of matrices $A_{zz}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[l_{22}]$</td>
<td>[3.1224, 3.3072]</td>
<td></td>
</tr>
<tr>
<td>$[l_{33}]$</td>
<td>[2.8770, 3.4416]</td>
<td>[3.0382, 3.4416]</td>
</tr>
<tr>
<td></td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td></td>
<td>3.3072</td>
<td>3.3072</td>
</tr>
<tr>
<td>$[l_{44}]$</td>
<td>[2.2309, 3.6190]</td>
<td>[3.0110, 3.6417]</td>
</tr>
<tr>
<td></td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td></td>
<td>3.3072</td>
<td>3.3072</td>
</tr>
<tr>
<td>$[l_{55}]$</td>
<td>[−7.5673, ]</td>
<td>[2.9638, 3.8438]</td>
</tr>
<tr>
<td></td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.3072</td>
</tr>
</tbody>
</table>
Example (cont’d) $[T] = T(16, [9, 10], [4, 5], [1, 2], [0, 1])$

The entries are nonnegative and exhibit a monotone and convex decay.

Results by Berenhaut and Bandyopadhyay, 2005:

For each $T \in [T]$, positivity and monotonicity are maintained through the Cholesky decomposition $T = LL^T$, where the lower triangular matrix $L = (l_{ij})$ satisfies

\[
\begin{align*}
l_{ij} & \geq 0, \quad 1 \leq j \leq i \leq n; \\
l_{ij} & \geq l_{i+1,j}, \quad 1 \leq j \leq i \leq n - 1; \\
l_{jj} & \geq \sqrt{(a_1 - a_2) + a_j(a_{j-1} - a_j)} \quad (\geq 0), \quad 1 \leq j \leq n; \\
l_{ij} & \geq \left( a_{i-j+1} - a_{i-j+2} + a_i(a_{j-1} - a_j) \right) / l_{jj}, \quad 1 \leq j \leq i \leq n.
\end{align*}
\]