A simple derivation of the Hansen-Bliek-Rohn-Ning-Kearfott enclosure for linear interval equations

Arnold Neumaier

Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Wien
Austria
email: neum@cma.univie.ac.at
WWW: http://solon.cma.univie.ac.at/~neum/

Abstract. Recently, Ning & Kearfott derived a formula for the interval enclosure of the solution set of linear systems of equations with uncertain data ranging in intervals, in the case when the coefficient matrix is an H-matrix. The enclosure is optimal when the midpoint matrix is diagonal, and when the midpoint is the identity, it reduces to the optimal method for enclosing preconditioned systems found by Hansen and Bliek and simplified by Rohn.

An elementary proof of this formula is given using only simple properties of H-matrices and Schur complements. The new proof gives additional insight into why the theorem is true. It is also shown how to preserve rigor in the enclosure when finite precision arithmetic is used.

revised version, November 1998

Keywords: linear interval equations, optimal enclosure, H-matrix, Schur complement, finite precision arithmetic, directed rounding

1991 MSC Classification: 65G10

1 Introduction

An interesting and unexpected result by Hansen [2] gave an explicit formula for the the interval hull of the solution set of a linear system of interval equations in the special case that the midpoint matrix is the identity. Since such matrices arise by preconditioning more general linear interval equations, this result is of considerable interest. The result was independently found by Bliek [1] in Chapter 4.4 of his unpublished Ph.D. thesis, but both Hansen's and Bliek's proofs were not completely rigorous. Rohn [5] gave an impeccable proof of the result and simplified the formula, reducing the amount of work needed to compute the interval hull. Recently, Ning & Kearfott [4] generalized Rohn's result to a formula giving an enclosure of the solution set of linear interval equations whose coefficient matrix is an H-matrix, and showed that the enclosure is optimal when the midpoint matrix is diagonal.

In this paper, we give a more direct and elementary proof of this formula, together with implementation details in finite precision arithmetic.

In the following, concepts and notation are as in Neumaier [3], except that interval quantities are bold face. In particular, inequalities, absolute values and operations $\inf, \sup, \max, \min$ are interpreted componentwise, $\square S = [\inf S, \sup S]$ is the interval hull of a bounded set $S \subset \mathbb{R}^n$, and

$$\Sigma(\mathbf{A}, \mathbf{b}) = \{x \in \mathbb{R}^n \mid Ax = b \text{ for some } A \in \mathbf{A}, b \in \mathbf{b}\}$$

denotes the solution set of a system of linear interval equations with coefficient matrix $\mathbf{A}$ and right hand side $\mathbf{b}$.

A square interval matrix $\mathbf{A} \in \mathbb{IR}^{n\times n}$ is called an H-matrix iff there is a vector $v > 0$ with $\langle \mathbf{A} \rangle v > 0$. Here $\langle \mathbf{A} \rangle$ is the comparison matrix with entries

$$\langle \mathbf{A} \rangle_{ii} = \langle \mathbf{A}_{ii} \rangle = \min\{|\alpha| \mid \alpha \in \mathbf{A}_{ii}\},$$

$$\langle \mathbf{A} \rangle_{ik} = -|\mathbf{A}_{ik}| = -\max\{|\alpha| \mid \alpha \in \mathbf{A}_{ik}\} \quad \text{for } k \neq i.$$

Among the many properties of H-matrices discussed in Neumaier [3, Section 3.6-3.7], we need the fact that every matrix A contained in an interval H-matrix $\mathbf{A}$ is nonsingular, the comparison matrix $\langle \mathbf{A} \rangle$ is an H-matrix, too (even an M-Matrix), and

$$|\mathbf{A}^{-1}| \leq \langle \mathbf{A} \rangle^{-1}. \tag{1}$$

We write $A^{(i)}$ for the submatrix of a square matrix A obtained by dropping the ith row $A_{i:}$ and the ith column $A_{:i}$, and $b^{(i)}$ for the subvector of a vector b obtained by dropping the ith component b_i. Similarly, $A^{(i)}_{i:}$ and $A^{(i)}_{:i}$ denote the ith row and ith column of A with the entry A_{ii} dropped.

2 The Ning-Kearfott Theorem

Ning & Kearfott [4] proved the following theorem by reducing it to Rohn's version of Hansen's result.

Theorem 2.1 *Let* $\mathbf{A} \in \mathbb{IR}^{n\times n}$ *be an H-matrix,* $\mathbf{b} \in \mathbb{IR}^n$ *a right hand side,*

$$u = \langle \mathbf{A} \rangle^{-1} |\mathbf{b}|, \quad d_i = (\langle \mathbf{A} \rangle^{-1})_{ii} \tag{2}$$

and

$$\alpha_i = \langle \mathbf{A}_{ii} \rangle - 1/d_i, \quad \beta_i = u_i/d_i - |\mathbf{b}_i|. \tag{3}$$

Then $\square\Sigma(\mathbf{A},\mathbf{b})$ *is contained in the vector* $\mathbf{x}$ *with components*

$$\mathbf{x}_i = \frac{\mathbf{b}_i + [-\beta_i, \beta_i]}{\mathbf{A}_{ii} + [-\alpha_i, \alpha_i]}. \tag{4}$$

Moreover, if the midpoint of $\mathbf{A}$ *is diagonal then* $\square\Sigma(\mathbf{A},\mathbf{b}) = \mathbf{x}$.

The new proof is based on two simple observations.

Lemma 2.2 *If* $A^{(i)}$ *is nonsingular and* $Ax = b$ *then*

$$\left(A_{ii} - A^{(i)}_{i:}(A^{(i)})^{-1}A^{(i)}_{:i}\right) x_i = b_i - A^{(i)}_{i:}(A^{(i)})^{-1}b^{(i)}. \tag{5}$$

Proof. We separate the ith variable of x and the ith equation of $Ax = b$ and find

$$A^{(i)}x^{(i)} + A^{(i)}_{:i}x_i = b^{(i)}, \tag{6}$$

$$A^{(i)}_{i:}x^{(i)} + A_{ii}x_i = b_i. \tag{7}$$

Solving (6) for $x^{(i)}$ gives $x^{(i)} = (A^{(i)})^{-1}(b^{(i)} - A^{(i)}_{:i}x_i)$, and insertion into (7) and simplification gives (5). $\square$

Lemma 2.3 *Let A be an H-matrix. Then $A^{(i)}$ is an H-matrix, too.*

Proof. By definition, A is an H-matrix iff there is a vector $v > 0$ with $\langle A\rangle v > 0$. Dropping the ith row and separating the contribution of v_i gives $\langle A^{(i)}\rangle v^{(i)} > |A_{:i}^{(i)}|v_i \geq 0$. Therefore $\langle A^{(i)}\rangle v^{(i)} > 0$, and $A^{(i)}$ is an H-matrix. $\square$

Proof of Theorem 2.1. By Lemma 2.3, we may apply Lemma 2.2 to an arbitrary $A \in \mathbf{A}$ and $b \in \mathbf{b}$, and find for the solution x of $Ax = b$ the relation (5), i.e.,

$$(A_{ii} - \tilde{\alpha}_i)x_i = b_i - \tilde{\beta}_i, \tag{8}$$

where

$$\tilde{\alpha}_i = A_{i:}^{(i)}(A^{(i)})^{-1}A_{:i}^{(i)}, \qquad \tilde{\beta}_i = A_{i:}^{(i)}(A^{(i)})^{-1}b^{(i)}. \tag{9}$$

By (1) and the rules for absolute values we find

$$|\tilde{\alpha}_i| \leq |\mathbf{A}_{i:}^{(i)}|\langle \mathbf{A}^{(i)}\rangle^{-1}|\mathbf{A}_{:i}^{(i)}| =: \alpha_i, \tag{10}$$

$$|\tilde{\beta}_i| \leq |\mathbf{A}_{i:}^{(i)}|\langle \mathbf{A}^{(i)}\rangle^{-1}|\mathbf{b}^{(i)}| =: \beta_i. \tag{11}$$

We first verify that these definitions for α_i and β_i agree with the expressions given in the theorem.

If we apply Lemma 2.2 to the equation $\langle \mathbf{A}\rangle c = I_{:i}$ in place of $Ax = b$, where $I_{:i}$ is the ith column of the identity matrix, we find

$$(\langle \mathbf{A}\rangle_{ii} - \alpha_i)c_i = 1. \tag{12}$$

Since $c_i = (\langle \mathbf{A}\rangle^{-1})_{ii} = d_i$, this shows that α_i has the value claimed in (3). And if we apply Lemma 2.2 to the equation $\langle \mathbf{A}\rangle u = |\mathbf{b}|$ in place of $Ax = b$, we find

$$(\langle \mathbf{A}\rangle_{ii} - \alpha_i)u_i = |\mathbf{b}_i| + \beta_i. \tag{13}$$

(The plus sign is due to a minus sign in the comparison matrix!) By solving (13) for β_i we see (using (12)) that β_i also has the values claimed in (3).

To prove (4), we first note that relation (12) implies $\langle \mathbf{A}\rangle_{ii} > \alpha_i$ since $c_i = (\langle \mathbf{A}\rangle^{-1})_{ii} > 0$. Therefore, the denominator in (4) cannot contain zero, and (4) follows from (8) – (11) since

$$x_i = \frac{b_i - \tilde{\beta}_i}{A_{ii} - \tilde{\alpha}_i} \in \frac{\mathbf{b}_i + [-\beta_i, \beta_i]}{\mathbf{A}_{ii} + [-\alpha_i, \alpha_i]} = \mathbf{x}_i. \tag{14}$$

For the proof of optimality we first note that we may multiply the rows of the linear systems by factors ± 1 to make all $\operatorname{mid} \mathbf{b}_k \geq 0$, and then multiply the columns of $\mathbf{A}$ and the corresponding variables by factors ± 1 to make all $\operatorname{mid} \mathbf{A}_{kk} \geq 0$. This choice of signs guarantees

$$\sup \mathbf{b} = |\mathbf{b}|, \tag{15}$$

$$\langle \mathbf{A} \rangle_{kk} = \underline{\mathbf{A}}_{kk} > 0 \quad \text{for all } k, \tag{16}$$

since the diagonal elements of H-matrices are nonzero.

If $\operatorname{mid} \mathbf{A}$ is diagonal, $\mathbf{A}$ contains the matrix $A := \langle \mathbf{A} \rangle$ and the matrices obtained from A by making the entries of the ith row and/or column nonnegative without changing their absolute value. This gives four choices for $A \in \mathbf{A}$ with

$$A^{(i)} = \langle \mathbf{A}^{(i)} \rangle, \quad A^{(i)}_{i:} = \pm |\mathbf{A}^{(i)}_{i:}|, \quad A^{(i)}_{:i} = \pm |\mathbf{A}^{(i)}_{:i}|$$

for independent choices of the signs, and this remains valid no matter how we change $A_{ii} \in \mathbf{A}_{ii}$. If we now choose $b \in \mathbf{b}$ such that $b^{(i)} = |\mathbf{b}^{(i)}|$ which is possible by (15), we see from (9)–(11) that $\tilde{\alpha}_i = \pm \alpha_i, \tilde{\beta}_i = \pm \beta_i$ with arbitrary and independent choices of the sign. Since we may still give A_{ii}, b_i any values in $\mathbf{A}_{ii}, \mathbf{b}_i$, the extremes in (14) are attained for suitable $A \in \mathbf{A}, b \in \mathbf{b}$. $\square$

Remark. It is easy to see that the Ning-Kearfott enclosure is always at least as good as that provided by the Krawczyk inverse (Theorem 3.7.8 in [3]). However, if $\operatorname{mid} \mathbf{A}$ is not diagonal, one often gets a further improvement by refining the Ning-Kearfott enclosure by a few Gauss-Seidel iterations. Since the latter only costs $O(n^2)$ operations, it should be applied generically in applications where it is not apriori known that (apart from rounding errors) $\operatorname{mid} \mathbf{A}$ is diagonal.

3 Finite Precision Calculations

In order that the enclosure given by the Ning-Kearfott Theorem 2.1 remains rigorous in finite precision arithmetic, care is needed in the calculation of the quantities defined there. In particular, we need a rigorous upper bound B for $\langle \mathbf{A} \rangle^{-1}$, in order to obtain rigorous upper bounds for α_i and β_i that render the subsequent interval calculations rigorous.

All platforms where interval arithmetic can be implemented have (hardware or emulated software) routines that set the rounding mode for subsequent operations to either upward rounding or downward rounding. For example, the MATLAB toolbox INTLAB of RUMP [6] provides for this task two routines `SetRoundDown` and `SetRoundUp`.

Because $\langle \mathbf{A} \rangle^{-1}$ is nonnegative by (1), a tight upper bound B for $\langle \mathbf{A} \rangle^{-1}$ can be constructed from an approximation $\tilde{B}$ to $\langle \mathbf{A} \rangle^{-1}$ and vectors v, w satisfying

$$I - \langle \mathbf{A} \rangle \tilde{B} \leq \langle \mathbf{A} \rangle v w^T \tag{17}$$

by taking

$$B = \tilde{B} + v w^T.$$

(*Proof:* multiply by $\langle \mathbf{A} \rangle^{-1} \geq 0$.)

Since $\mathbf{A}$ is an H-matrix, there is a vector $v > 0$ such that $u = \langle \mathbf{A} \rangle v > 0$, and for such a vector, we can satisfy (17) with the vector w with tiny components

$$w_k = \max_i \frac{-R_{ik}}{u_i},$$

where

$$R = \langle \mathbf{A} \rangle \tilde{B} - I$$

is the residual in the computation of the approximate inverse.

To find an explicit such v, we may assume that a positive vector e is available whose components are proportional to natural scales for the components of $\mathbf{x}$. Then $v := \tilde{B}e \approx \langle \mathbf{A} \rangle^{-1} e > 0$ if we indeed have an H-matrix, and then $u = \langle \mathbf{A} \rangle v \approx e$ is positive if the approximation is good enough. Since conversely, the condition $u = \langle \mathbf{A} \rangle v > 0$ implies that $\mathbf{A}$ is an H-matrix only when $v \geq 0$, we take in the program the absolute value of the computed v to ensure the H-matrix property. In the absence of information about the relative magnitudes of the components of $\mathbf{x}$ we may simply choose the all-one vector for e, and this is the choice used in the program below.

If we do the above calculations in finite precision arithmetic, we may calculate $\tilde{B}$ and v with arbitrary rounding, but u and R must underestimate the exact quantities, while w, B and the quantities calculated in (2) and (3) must overestimate the exact quantities. This is easily achieved by a proper setting of the rounding modes.

By looking at the operations involved, it is now easy to verify that the MATLAB program

```
n = dim(A);                                           % dimension of A
dA = diag(A);    % save diagonal entries before overwriting A
A = compmat(A);        % A now contains the comparison matrix
B = inv(A);        % approximate inverse of comparison matrix
v = abs(B*ones(n,1));            % enforce nonnegativity of v
SetRoundDown;
u = A*v;           % is approximately equal to all-one vector
if ~all(min(u)>0),                      % check positivity of u
  % A is numerically not an H-matrix
  x = midrad(0,inf+zeros(n,1));
else
  dAc = diag(A); % save diagonal entries before overwriting A
  A = A*B-eye(n);        % A now contains the residual matrix
  SetRoundUp;
  w = zeros(1,n);
  for i=1:n,
    w = max(w,(-A(i,:))/u(i));
  end;
  B = B+v*w;              % rigorous upper bound for exact B
  u = B*abs(b);
  d = diag(B);
  alpha = dAc+(-1)./d;              % ensures upward rounding
  beta = u./d-abs(b);
  x = (b+midrad(0,beta))./(dA+midrad(0,alpha));
end;
```

(to be used together with the INTLAB toolbox of RUMP [6]) produces the desired rigorous bounds and hence a rigorous enclosure. It should not be difficult to rewrite this to equivalent code in other programming environments.

To calculate the Ning-Kearfott enclosure in the above way we need $2n^3 + O(n^2)$ operations (counting additions and multiplications separately) for the approximate inversion and as many operations for the calculation of the residual, plus $O(n^2)$ operations for the remaining calculations, giving a total of

$4n^3 + O(n^2)$ operations. Using a running error analysis for the inversion process, slightly suboptimal bounds for the residual can be obtained with $O(n^2)$ operations, thus halving the work in high dimensions; see, e.g., Spellucci & Krier [7].

References

[1] C. Bliek, Computer methods for design automation, Ph.D. Thesis, Dept. of Ocean Engineering, Massachusetts Institute of Technology, 1992.

[2] E. Hansen, Bounding the solution of interval linear equations, SIAM J. Numer. Anal. 29 (1992), 1493-1503.

[3] A. Neumaier, Interval methods for systems of equations, Cambridge Univ. Press, Cambridge 1990.

[4] S. Ning and R. B. Kearfott, A comparison of some methods for solving linear interval equations, SIAM J. Numer. Anal. 34 (1997), 1289-1305.

[5] J. Rohn, Cheap and tight bounds: The recent result by E. Hansen can be made more efficient, Interval Computations 4 (1993), 13-21.

[6] S.M. Rump, INTLAB – INTerval LABoratory, Reliable Computing, to appear.

[7] P. Spellucci and N. Krier, Untersuchungen der Grenzgenauigkeit von Algorithmen zur Auflösung linearer Gleichungssysteme mit Fehlererfassung, pp. 288-297 in: Interval Mathematics (K. Nickel, ed.), Lecture Notes in Computer Science 29, Springer, Berlin 1975.