SOLVING INTERVAL LINEAR SYSTEMS

by

J. Rohn, Prague

Dedicated to Prof. K. Nickel on his 60th birthday

1. Introduction and notations

In this paper, we present a general method for solving an interval linear system $A^I x = b^I$ with a square regular interval matrix A^I. The results are based on a link between interval linear algebra and the theory of linear complementarity, established in theorem 1 below. The paper is written rather briefly, with the proofs omitted. Emphasis is laid on computational aspects; several further theoretical consequences are not mentioned here.

Notations used are basically the same as in [11]. We shall deal with nxn real or interval matrices; for the sake of brevity, let $N = \{1,2,\dots,n\}$. The ij-th coefficient of a matrix A is denoted by A_{ij}, the i-th row by $A_{i.}$, the j-th column by $A_{.j}$ $(i,j \in N)$. The matrix $|A|$ is defined by $|A|_{ij} = |A_{ij}|$ for each i,j (similarly for vectors). $\rho(A)$ denotes the spectral radius of A. An nxn interval matrix A^I and an interval n-vector b^I are defined by $A^I = [A_c - \Delta, A_c + \Delta] = \{A;\ A_c - \Delta \le A \le A_c + \Delta\}$, $b^I = [b_c - \delta, b_c + \delta] = \{b;\ b_c - \delta \le b \le b_c + \delta\}$, where $\Delta \ge 0$, $\delta \ge 0$.

AMS Subject Classifications: 65G10, 90C33

Let $e = (1,1,\dots,1)^T \in R^n$ and let $f = -e$. We denote $Y = \{y \in R^n;\ |y| = e\}$, so that Y has 2^n elements, $e \in Y$ and $f \in Y$. For each $y \in Y$, let T_y be the diagonal matrix whose diagonal is formed by the vector y; thus $T_e = E$ (the unit matrix) and $T_f = -E$. The k-th column of E is denoted by e_k. For each $y,z \in Y$, we introduce an important notation

$$A_{yz} = A_c - T_y \Delta T_z$$
$$b_y = b_c + T_y \delta ,$$

implying $A_{yz} \in A^I$ and $b_y \in b^I$. For $x \in R^n$, we define a vector sgn x by

$$(\operatorname{sgn} x)_i = \begin{cases} 1 & \text{if } x_i \geq 0 \\ -1 & \text{if } x_i < 0 \end{cases} \qquad (i \in N),$$

hence sgn $x \in Y$ and if $z = \operatorname{sgn} x$, then $|x| = T_z x$.

Let X be a compact (especially, finite) set in R^n. We introduce n-vectors min X, max X by

$$(\min X)_i = \min\{x_i;\ x \in X\}$$
$$(\max X)_i = \max\{x_i;\ x \in X\} \qquad (i \in N),$$

so that $[\min X, \max X]$ is the narrowest interval containing X. For each $x \in R^n$, we define x^+, x^- by $x^+ = \max\{x,0\}$, $x^- = \max\{-x,0\}$; then $x^+ \geq 0$, $x^- \geq 0$, $x = x^+ - x^-$ and $|x| = x^+ + x^-$. The convex hull of X is denoted by Conv X.

2. Basic results

Consider an interval linear system

$$A^I x = b^I \qquad (1)$$

with an nxn interval matrix $A^I = [A_c - \Delta, A_c + \Delta]$ and a right-hand side interval n-vector $b^I = [b_c - \delta, b_c + \delta]$. A^I is assumed to be regular (i.e. each $A \in A^I$ is regular). The set $X = \{x;\ Ax = b,\ A \in A^I,\ b \in b^I\}$ is called the solution set of (1), the interval vector $x^I = [\underline{x}, \overline{x}]$, where $\underline{x} = \min X$, $\overline{x} = \max X$, is called the interval solution to (1).

Oettli and Prager [9] described the solution set by $X = \{x;\ |A_c x - b_c| \leq \Delta |x| + \delta\}$. Our approach is based on an observation leading to a conjecture justified later in theorem 2 that each extremal point of Conv X satisfies the equation

$$|A_c x - b_c| = \Delta |x| + \delta \ . \qquad (2)$$

In order to bring this equation to a more usual form, put $y = \operatorname{sgn}(A_c x - b_c)$, so that $|A_c x - b_c| = T_y(A_c x - b_c)$, and substitute $x = x^+ - x^-$, $|x| = x^+ + x^-$. Then we get

$$x^+ = A_y x^- + w_y \ , \qquad (3)$$

where

$$A_y = A_{ye}^{-1} A_{yf}$$
$$w_y = A_{ye}^{-1} b_y .$$

Thus we have replaced the equation (2) by 2^n equations of the type (3) (for all possible $y \in Y$). The equation (3) is nothing else than a linear complementarity problem [4]. We shall show that under the regularity assumption each equation (3) has exactly one solution. To this end, we need several preliminary results. First, we state this theorem:

<u>Theorem 0.</u> Let for each $y \in Y$ the equation

$$A_{ye}x^1 - A_{yf}x^2 = b_y \qquad (4)$$

have a nonnegative solution x_y^1, x_y^2. Then for each $A \in A^I$ and $b \in b^I$, the equation $Ax = b$ has a solution belonging to $\mathrm{Conv}\left\{x_y^1 - x_y^2;\ y \in Y\right\}$.

The proof of this theorem can be drawn from the proof of theorem 1 in [10]. Notice that premultiplying (3) by A_{ye} yields $A_{ye}x^+ - A_{yf}x^- = b_y$, hence (3) is of the form (4). Second, we shall need this lemma whose proof is straightforward:

<u>Lemma 1.</u> Let A be a regular nxn matrix and let D_j be a matrix whose all rows except the j-th are zero. Let $\alpha = 1 + (D_jA^{-1})_{jj}$. Then we have:

(i) if $\alpha > 0$, then $(A + D_j)^{-1} = A^{-1} - \frac{1}{\alpha} A^{-1}D_jA^{-1}$,

(ii) if $\alpha \leq 0$, then $A + tD_j$ is singular for some $t \in (0,1]$.

Third, we shall utilize the result by Samelson, Thrall and Wesler [12] in the form given e.g. in [4, p. 90]: An equation $x^+ = Ax^- + w$ has exactly one solution for each right-hand side w if and only if A is a P-matrix. (Recall that A is called a P-matrix if all its principal minors are positive.) Now we are able to state our first basic result.

<u>Theorem 1.</u> A^I is regular if and only if A_y is a P-matrix for each $y \in Y$.

The "if" part of the proof follows from the above result by Samelson et al. and from theorem 0 (applied to systems $A^Ix = e_j$, $j \in N$). In the "only if" part of the proof, we first use lemma 1 to prove (by induction on j) that all leading principal minors of $A_{ef}A_{ee}^{-1}$ are positive, then (using permutations) that all its principal minors are positive. Next, employing the transpose of A^I and its premultiplying by T_y, we conclude with the P-property of A_y $(y \in Y)$.

Now, using theorems 0 and 1 and the result by Samelson et al., we obtain the second basic result:

<u>Theorem 2.</u> Let A^I be regular. Then for each $y \in Y$, the equation (3) has exactly one solution x_y. Moreover, we have $x_y \in X$ for each $y \in Y$ and Conv X = Conv $\{x_y;\ y \in Y\}$; especially,

$$\begin{aligned} \underline{x} &= \min \{x_y;\ y \in Y\} \\ \overline{x} &= \max \{x_y;\ y \in Y\} . \end{aligned} \tag{5}$$

For computational purposes, it is useful to rearrange the equation (3) to some equivalent forms. Returning from x^+, x^- back to x, $|x|$, we can bring (3) to the form

$$x = D_y|x| + d_y, \tag{6}$$

where

$$\begin{aligned} D_y &= A_c^{-1} T_y \Delta \\ d_y &= A_c^{-1} b_y , \end{aligned}$$

thus obtaining a fixed-point equation. Further, letting $z = \operatorname{sgn} x$ and substituting $|x| = T_z x$, (6) gives an equivalent system

$$\begin{aligned} A_{yz} x &= b_y \\ T_z x &\geq 0 . \end{aligned} \tag{7}$$

Due to theorem 2, both (6) and (7) have a unique solution x_y.

According to (5), for each $i \in N$ we have $\underline{x}_i = (x_y)_i$ for some $y \in Y$ (similarly for $\overline{x}_i$). Our third basic result specifies some property of this y; its proof follows from lemma 1 applied to system (7):

Theorem 3. Let A^I be regular. Then for each $i \in N$ we have:

(i) $\underline{x}_i = (x_y)_i$ for some $y \in Y$ satisfying $(A_{yz}^{-1}T_y)_{i.} \leq 0^T$, where $z = \operatorname{sgn} x_y$,

(ii) $\overline{x}_i = (x_y)_i$ for some $y \in Y$ satisfying $(A_{yz}^{-1}T_y)_{i.} \geq 0^T$, where $z = \operatorname{sgn} x_y$.

3. Computation of x_y's

In view of theorem 1, we may use the standard algorithms by Cottle and Dantzig [3] or Lemke [6] to solve the linear complementarity problem (3). However, due to specific features of our problem, we shall get a far simpler algorithm for computing x_y when employing the result by Murty [8] for solving the system (7). Notice that the condition $T_z x \geq 0$ is equivalent to $z_j x_j \geq 0$ for each $j \in N$. The following "sign accord algorithm" solves the system $A_{yz}x = b_y$ for different z until this condition is met:

Algorithm 1 (computing x_y for a given $y \in Y$).

Step 0. Set $z = e$.

Step 1. Solve $A_{yz}x = b_y$.

Step 2. If $T_z x \geq 0$, terminate with $x_y = x$.

Step 3. Otherwise compute $k = \min\{j;\ z_j x_j < 0\}$.

Step 4. Set $z_k = -z_k$ and go to step 1.

Using the Murty's result combined with theorem 1, we get this theorem:

<u>Theorem 4.</u> Let A^I be regular. Then the algorithm is finite for each $y \in Y$.

The algorithm proved to perform quite satisfactorily on test examples, taking an average of 1 to 2 returns to step 1 before stopping in step 2. Nevertheless, for large-size examples it may be useful to modify it in such a way as to avoid solving a new system at each return from step 4 to step 1. This modification is possible due to the fact that the new system differs only in the k-th column from the old one. The computations can be performed in a tableau of the form

p	A	x	z

where x and z are as above, A is an $n \times n$ matrix which is at each step equal to A_{yz}^{-1} and p is an additional pivoting column.

<u>Algorithm 2</u> (computing x_y – tableau form).

<u>Step 0.</u> Start with $A = A_{ye}^{-1}$, $x = Ab_y$, $z = e$, $\Delta' = T_y \Delta$.

<u>Step 1.</u> If $T_z x \geqslant 0$, terminate with $x_y = x$.

<u>Step 2.</u> Otherwise compute $k = \min \{ j;\ z_j x_j < 0 \}$.

<u>Step 3.</u> Compute $p = 2 z_k A (\Delta')_{.k} + e_k$. If $p_k \leqslant 0$, stop! A^I is singular. Otherwise proceed to step 4.

<u>Step 4.</u> Perform the Gaussian elimination with the rows of the tableau (except the last column z, which remains unchanged) with pivot p_k so that p become e_k.

<u>Step 5.</u> Set $z_k = -z_k$ and go to step 1.

The finiteness of this algorithm follows from the fact that it generates the same sequence of vectors x, z as algorithm 1.

Another algorithms, this time of iterative type and under additional assumptions, are based on solving the fixed-point equation (6). First, we may use iterations of Jacobi type:

$$x_y^o = d_y$$
$$x_y^m = D_y |x_y^{m-1}| + d_y \qquad (m = 1,2,\ldots).$$

If $\rho(|D_y|) < 1$, then $x_y^m \rightarrow x_y$. Since $\rho(|D_y|) \leq \rho(D)$, where $D = |A_c^{-1}|\Delta$, the convergence for each $y \in Y$ can be assured by the assumption

$$\rho(D) < 1 , \qquad (8)$$

as it was done in [11] .(Note: if $D_{jj} \geqslant 1$ for some $j \in N$, then A^I is singular.) Second, we may use iterations of the Gauss-Seidel type. Let $D_y = L_y + U_y$, where L_y is a lower triangular matrix with zero diagonal elements and U_y is an upper triangular matrix. Then the Gauss-Seidel iterations have the form

$$\tilde{x}_y^o = d_y$$
$$\tilde{x}_y^m = L_y |\tilde{x}_y^m| + U_y |\tilde{x}_y^{m-1}| + d_y \qquad (m = 1,2,\ldots).$$

Let $D = L + U$ be an analogous decomposition of D into triangular matrices. Then, if D is an irreducible matrix (e.g. $D > 0$ or $\Delta > 0$) satisfying (8), then $\tilde{x}_y^m \rightarrow x_y$ and we have $\rho((E - L)^{-1}U) < \rho(D)$ showing that the Gauss-Seidel method can be expected to converge faster than that of Jacobi.

Under more strong assumptions, x_y may be also computed when using a variational inequalities technique [5, pp. 15-17] or a linear programming technique similar to that of Mangasarian [7] .

A detailed investigation and comparison of the methods described was performed by Baumann in his thesis [1] .

4. Interval solution

Being able to compute the x_y's, we can evaluate the interval solution $[\underline{x}, \overline{x}]$ according to (5). We shall show here that in most practical cases there is no need for computing all the 2^n vectors x_y. The clue to such a reduction is contained in theorem 3. Recall that the condition $(A_{yz}^{-1}T_y)_{i.} \geqslant 0^T$ is equivalent to $(A_{yz}^{-1})_{ij}y_j \geqslant 0$ for each $j \in N$. To be able to utilize this result in an applicable manner, assume we are given an estimation of the interval inverse of A^I, i.e. an interval matrix $\tilde{B}^I = [\underset{\sim}{B}, \tilde{B}]$ such that $A^{-1} \in \tilde{B}^I$ for each $A \in A^I$. Now for each $i \in N$ define

$$Y_i = \{y;\ y_j = 1\ (\underset{\sim}{B}_{ij} > 0),\ y_j = -1\ (\tilde{B}_{ij} < 0),\ |y_j| = 1\ (\text{otherwise})\}$$

and let

$$Y_o = \bigcup_{i=1}^{n} (Y_i \cup (-Y_i))$$

(here, $-Y_i = \{-y;\ y \in Y_i\}$). Then theorem 3 gives this result.

<u>Theorem 5.</u> Under the above notations, we have

$$\underline{x}_i = \min \{(x_y)_i;\ y \in -Y_i\}$$
$$\overline{x}_i = \max \{(x_y)_i;\ y \in Y_i\} \qquad (i \in N),$$

hence also

$$\underline{x} = \min \{x_y;\ y \in Y_o\}$$
$$\overline{x} = \max \{x_y;\ y \in Y_o\}\ .$$

To obtain an estimation of the interval inverse is seemingly a difficult problem in the general case; but there are several important classes of regular interval matrices for which it can be done easily. First, consider interval matrices satisfying (8) (a condition met by all the published examples, as far as we know).

Here, we can simply put

$$\underset{\sim}{B} = A_c^{-1} - C|A_c^{-1}|$$
$$\tilde{B} = A_c^{-1} + C|A_c^{-1}| \quad ,$$

where $C = D(E - D)^{-1}$ (cf. [11]).

Next, consider interval matrices satisfying $|A^{-1}| > 0$ for each $A \in A^I$; we call them inverse-stable matrices. In this case, each Y_i $(i \in N)$ consists of a single vector, namely $Y_i = \{y(i)\}$, where

$$y(i) = \text{sgn}\, (A_c^{-1})_{i.} \quad ,$$

hence at most 2n vectors x_y are to be computed. An interval matrix satisfying (8) is inverse-stable if $C|A_c^{-1}| < |A_c^{-1}|$ holds.

As a special case, consider positively invertible interval matrices ($A^{-1} > 0$ for each $A \in A^I$). Then $Y_i = \{e\}$ for each $i \in N$, hence $\underline{x} = x_f$ and $\overline{x} = x_e$. Thus we can either twice use algorithm 1 (which is similar to Beeck's algorithm [2] in this case) or compute $\underline{x}$, $\overline{x}$ by Jacobi (Gauss-Seidel) iterations using the fact that (8) always holds for positively invertible matrices (a consequence of the Perron-Frobenius theorem).

5. Regularity and invertibility

In this section, we give some results concerning interval matrices; these results can be obtained from theorems 0-3 when applying them to systems $A^I x = e_j$ $(j \in N)$.

For a regular interval matrix A^I, the inverse interval matrix $B^I = [\underline{B}, \overline{B}]$ is defined by

$$\underline{B} = \min \{ A^{-1};\ A \in A^I \}$$
$$\overline{B} = \max \{ A^{-1};\ A \in A^I \} \ .$$

Theorem 1 gives no hint how to compute B^I; such an information is contained in the following, more complex, theorem:

<u>Theorem 6.</u> A^I is regular if and only if for each $y \in Y$ the matrix equation

$$B = D_y|B| + A_c^{-1}$$

has a solution B_y. If this condition is met, then B_y is unique for each $y \in Y$. Moreover, for each $A \in A^I$ there exist nonnegative diagonal matrices L_y $(y \in Y)$ satisfying $\sum_{y \in Y} L_y = E$ such that

$$A^{-1} = \sum_{y \in Y} B_y L_y$$

holds; especially, we have

$$\underline{B} = \min\{B_y;\ y \in Y\}$$
$$\overline{B} = \max\{B_y;\ y \in Y\}\ . \tag{9}$$

Assume that $\rho(|D_y|) < 1$; then the sequence $\{B_y^m\}_{m=0}^{\infty}$ given by

$$B_y^0 = A_c^{-1}$$
$$B_y^m = D_y|B_y^{m-1}| + A_c^{-1} \qquad (m = 1,2,\ldots)$$

tends to B_y (hence, $\max_{y \in Y} \rho(|D_y|) < 1$ is a sufficient regularity condition generalizing (8)). For practical computations, Y may be replaced in (9) by Y_o constructed in the same way as in section 4 (cf. [11]).

We have also this general result:

<u>Theorem 7.</u> Let A^I be regular and let $i,j \in N$. Then, we have:

(i) $\underline{B}_{ij} = (A_{yz}^{-1})_{ij}$ for some $y,z \in Y$ satisfying $(A_{yz}^{-1}T_y)_{i.} \leq 0^T$, $(T_z A_{yz}^{-1})_{.j} \geq 0$,

(ii) $\overline{B}_{ij} = (A_{yz}^{-1})_{ij}$ for some $y,z \in Y$ satisfying $(A_{yz}^{-1}T_y)_{i.} \geq 0^T$, $(T_z A_{yz}^{-1})_{.j} \geq 0$.

This theorem is of special use for inverse-stable matrices. In this case, letting

$$y(i) = \operatorname{sgn}\,(A_c^{-1})_{i.}$$
$$z(j) = \operatorname{sgn}\,(A_c^{-1})_{.j} \qquad (i,j \in N),$$

we obtain simple formulae

$$\underline{B}_{ij} = (A^{-1}_{-y(i),z(j)})_{ij}$$
$$\overline{B}_{ij} = (A^{-1}_{y(i),z(j)})_{ij} \qquad (i,j \in N),$$

enabling us to compute B^I step by step.

We conclude this section with three theoretical results characterizing regularity, inverse-stability and positive invertibility by necessary and sufficient conditions.

Theorem 8. A^I is regular if and only if for each $y \in Y$ and each $j \in N$ there exists a $z \in Y$ such that $(T_z A^{-1}_{yz})_{.j} \geqslant 0$.

Theorem 9. A^I is inverse-stable if and only if $|A^{-1}_{yz} - A_c^{-1}| < |A_c^{-1}|$ for each $y, z \in Y$.

Theorem 10. A regular interval matrix A^I is positively invertible if and only if $A^{-1}_{ef} > 0$.

6. Acknowledgment

I am very indebted to Dr. J. Mandel for drawing my attention to the literature of the linear complementarity problem, where I found the papers [12], [8] that enabled me to reach the general results.

References

[1] M. Baumann, Solutions of Systems of Interval Linear Equations, Thesis, Charles University, Prague 1983

[2] H. Beeck, Zur Problematik der Hüllenbestimmung von Intervallgleichungssystemen, in: Interval Mathematics (K. Nickel, Ed.), Springer-Verlag, Berlin-Heidelberg-New York 1975, 150-159

[3] R. Cottle, G. Dantzig, Complementary Pivot Theory of Mathematical Programming, Linear Algebra and Its Applications 1(1968), 103-125

[4] R. Cottle, J.-S. Pang, On Solving Linear Complementarity Problems as Linear Programs, Mathematical Programming Study 7(1978), 88-107

[5] R. Glowinski, J.-L. Lions, R. Trémolières, Numerical Analysis of Variational Inequalities, North-Holland, Amsterdam 1981

[6] C. Lemke, Bimatrix Equilibrium Points and Mathematical Programming, Management Science 11(1965), 681-689

[7] O. Mangasarian, Characterization of Linear Complementarity Problems as Linear Programs, Mathematical Programming Study 7(1978), 74-87

[8] K. Murty, Note on a Bard-Type Scheme for Solving the Complementarity Problem, Opsearch 11(1974), 123-130

[9] W. Oettli, W. Prager, Compatibility of Approximate Solution of Linear Equations with Given Error Bounds for Coefficients and Right-Hand Sides, Numerische Mathematik 6(1964), 405-409

[10] J. Rohn, On the Interval Hull of the Solution Set of an Interval Linear System, Freiburger Intervall-Berichte 81/5, 47-57

[11] J. Rohn, An Algorithm for Solving Interval Linear Systems and Inverting Interval Matrices, Freiburger Intervall-Berichte 82/5, 23-36

[12] H. Samelson, R. Thrall, O. Wesler, A Partition Theorem for Euclidean n-Space, Proceedings of the American Mathematical Society 9(1958), 805-807.

Author's address: J. Rohn, Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 11800 Prague, Czechoslovakia.